Essays in Economics of Education: Free Primary Education, Birth Order and Human Capital Development in Lesotho

by

RAMAELE E. MOSHOESHOE

Thesis Presented for the Degree of

DOCTOR OF PHILOSOPHY

in the School of Economics
UNIVERSITY OF CAPE TOWN

14th August 2015

Supervisor:
A/Professor Cally ARDINGTON

Co-Supervisor:
A/Professor Patrizio PIRAINO

Declaration

I declare that this thesis is my own work, except where acknowledged in the text. I further declare that this thesis has not been submitted for a degree at any other university.

Ramaele Moshoeshoe

____
Signature

____
Date

Abstract

Given the the low levels of educational standards in the developing world, the World Education Forum adopted the Dakar Framework for Action (DFA) in 2000, calling for *quality* 'Education for All' children of school-going age. Heeding to this call, many sub-Saharan African countries instituted Free Primary Education (FPE) policies. Lesotho instituted the FPE programme in 2000 on a grade-by-grade basis; first abolishing school fees in grade one, and then in successive higher grades each following year.

This thesis consists of a short introductory chapter, three self-contained analytical chapters which empirically evaluate the importance of the FPE policy and family factors on education in Lesotho, and the summary chapter. It first examines the effect of the FPE policy on primary school enrolment in Chapter 2 using household level data for before and after the policy. A difference-in-differences strategy is employed to tease out the FPE effect. This exploits the variations in enrolment rates over time and across grade-groups (i.e. grades covered versus those not-yet covered) created by the implementation strategy of the programme. The findings demonstrate that the policy significantly increased enrolment of primary school-age children by at least 9.3 percentage points (or 13.2 percent). There is also evidence that this policy disproportionately raised enrolment levels of children from poor households and that of boys (the historically disadvantaged group), thereby bridging the gender- and wealth-related educational (enrolment) inequalities.

In Chapter 3, the thesis draws on grade six pupils' standardised maths and (English) reading test scores from 2000 and 2007 to analyse changes in educational achievement and educational inequality, and the determinants of such changes. The analysis of the data shows that educational achievement increased significantly for both low- and high-ability pupils over the period of analysis. Nonetheless, this increase in achievement was accompanied by a significant rise in educational inequality, especially in reading test scores. The analysis further shows that these changes are statistically related to policy measures taken under the FPE programme. In particular, the results show that pupil-teacher ratio is negatively correlated with changes in reading performance of low-ability pupils, while teacher effort (i.e. subject-testing frequency and teaching hours per week) and grade repetition have a positive influence on changes in educational achievement. These results suggest that the fall in pupil

teacher ratio between 2000 and 2007 has helped increase educational achievement.

The analysis, however, reveals that much of the increase in educational achievement and educational inequality is unexplained by both school and pupils' family characteristics, which suggests that there could be other unobserved family and school factors that influence achievement and inequality. Therefore, in Chapter 4 of the thesis I shift focus from FPE policy effects and look at the impact of family factors on human capital accumulation. Specifically, Chapter 4 examines the effect of a child's order of birth on human capital development (i.e. enrolment, educational attainment, and schooling progression) using family-fixed effects models. Birth order has received much attention in the economics and psychology literature. Contrary to much of the evidence from developing countries, I find that birth order has a strong negative effect on human capital development. These birth order effects are pronounced in large families, and families with first-born girls, thereby revealing the strong girls' education preferences in Lesotho. Turning to potential pathways of these effects, I find that birth order effects are not propagated through family wealth, but mainly though birth- (or child-) spacing. These results suggest that there are some intra-household inefficiencies that could explain the changes in educational achievement and inequality.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Cally Ardington and Patrizio Piraino, for their advice, encouragement and guidance throughout this PhD project. They both painstakingly read a countless number of drafts of this thesis, and I have enormously benefited from their critiques and discussions. The first analytical chapter of this thesis benefited particularly from Cally Ardington, who treated it like her own baby from inception. Martin Wittenberg lured me into applied microeconometrics research through his lectures. I also thank him for giving me the opportunity to be his teaching assistant, which helped to hone my econometrics understanding.

I am also thankful for the comments I received from participants at the African Economic Research Consortium (AERC) biannual workshops in Arusha and Nairobi in 2013, the University of Cape Town (UCT) School of Economics brown bag seminars, and the Southern Africa Labour and Development Research Unit (SALDRU) seminars.

The students in the PhD programme in the School of Economics, UCT, have been a source of inspiration, support and learning. I am particularly indebted to David Khaoya, Grieve Chelwa, Mduduzi Mthembu, and Nicholas Masiyandima for countless discussions that shaped this project, and for the laughs that kept me going.

This thesis project was made possible by the generous scholarship from the AERC, which sponsored the first four years of my PhD journey, and the Carnegie scholarship, which enabled me to stay for one more year to successfully complete this project.

Finally, I extend special thanks to my wife Reley 'Maleseli Moshoeshoe for sharing my exciting and tough moments. Thanks for always believing in me. My mother, 'Mamonyane Moshoeshoe, and my sister, 'Malefu Moshoeshoe, have also been a constant source of support and motivation throughout this journey. I am greatly thankful.

Dedication

To my wife Reley 'Maleseli Moshoeshoe, our daughter Leseli, my mother 'Mamonyane Moshoeshoe, and in memory of my father, Matela Moshoeshoe.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Economic theory predicts and many empirical studies confirm that quality education is important for individuals' health and lifetime earnings as well as economic growth (Becker, 1964; Hanushek and Woessmann, 2012). Thus, the poor education in much of the developing world will likely retard growth and development in these countries.

In 2000, the World Education Forum adopted the Dakar Framework for Action (DFA), which called for *complete, free and compulsory quality education*, and *redress of educational inequalities* (UNESCO, 2000, p.8). This was the world's response to addressing the problem of what Pritchett (2004) calls 'lack of education' (that is, low levels of school enrolment, educational attainment, and educational achievement) in the developing world. For example, at the time, sub-Saharan Africa (SSA) was trailing all other regions in all education indicators with a net enrolment rate of 56 percent, a 13 percent repetition rate, and high delayed school entry estimated at 70 percent (see Pritchett, 2004).

It is often said that education is like a three-legged pot; it requires equal contribution from parents, teachers, and the community. This implies that a child's development outcomes are influenced by her family, community, and school experiences. Thus, public policies that increase the supply of schooling, and aim to change a community's and family's preferences towards education might help to improve education.

Since the adoption of the DFA, several sub-Saharan African countries[1] have instituted free primary education (FPE) programmes, which involve the abolition of all user fees in public schools to encourage school enrolment as a first step towards addressing the 'lack of education' problem.

Lesotho implemented the FPE policy sequentially from grade one up to grade seven between 2000 and 2006. Today, primary schooling is free of user-fees. Under the FPE programme, other private schooling costs such as stationery and textbooks' costs have largely

[1] These include Burundi, Cameroon, Ghana, Kenya, Lesotho, Mozambique, Namibia, Rwanda, Swaziland, Tanzania, and Zambia.

been eliminated. Moreover, since the government's primary objective was to put all children of primary school-going age in school, the abolishment of user-fees was accompanied by an aggressive school building programme and school feeding programme. For example, between 2002 and 2011, the number of schools in Lesotho increased by 10 percent, from 1333 schools in 2002 to 1468 schools in 2011 (MOET, 2011). Much of the increase in the number of schools during this period occurred between 2002 and 2006. The number of schools in this period, 2002 to 2006, increased by 9 percent.

As of 2011, 40 percent of primary schools were in the mountains region (which is a largely rural region), versus 36 percent in the lowlands (a largely urban) region. The remaining 24 percent of schools were also in the mainly rural regions; the Senqu River Vally region with 4 percent, and the Foothills region with 36 percent. Data from the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) shows that the share of urban schools dropped by 1.5 percentage points from 35.1 percent to 33.6 between 2000 and 2007. This was a statistically significant drop (increase) in the number of urban (rural) schools over this period. This indicates that, there were more schools built in rural areas than in urban areas where relatively more schools existed before the programme.

The school building programme is aimed at increasing the schools' capacity to handle the influx of students attracted by the abolishment of user-fees. The school feeding programme on the other hand is aimed at retaining children within the schooling system, especially children from poor families who are at higher risk of dropping out of school because of hunger. In order to mitigate against the potential decline in educational quality, the government also increased the number of teachers in primary schools. Between 2002 to 2011, the number of teachers increased by 28 percent, from 8908 teachers in 2002 to 11378 teachers in 2011 (MOET, 2011). Similar to the increase in the number of schools, the biggest increase in the number of teachers hired happened between 2002 and 2006. During this time, the number of teachers increased by 21 percent, which accounts for about 75 percent of the increase between 2002 and 2011. This shows that, at the time the FPE was being rolled out to higher grades, the number of schools and teachers were also increasing.

To my knowledge, no other sub-Saharan African country followed the same strategy as Lesotho in implementing the FPE programme. For example, in Malawi, school fees were first abolished on a grade-by-grade basis for the first three grades, starting with grade one from 1991 to 1993, and then simultaneously eliminated for all remaining grades in 1994. In Uganda and Kenya, user fees were simultaneously eliminated for all primary school grades in 1997 and 2003, respectively. The effects of this policy on education in Lesotho have not been studied before, and may differ from those found in other countries.

As mentioned earlier, taking children to school is just one way of addressing the lack of education in the developing world. Improving education does not only require changing demand and supply constraints, but also requires deep understanding of other family, community, and school environment factors that impact on a child's development. Knowing which family factors affect child outcomes, and how they do so, is equally important for addressing the lack of education problem. There are a plethora of important family factors that have an influence on education, but the child's order of birth has been a recurring theme in both the economics of education and psychology literatures. That birth order has an influence on human capital accumulation is undisputed in both literatures (Black et al., 2005; De Haan, 2010; De Haan et al., 2014; Tenikue and Verheyden, 2010; Sulloway, 2007), but the important question of how it impacts a child's outcomes remains open.

Therefore, the big question that forms the basis of this thesis is, what is the effect of public policies and the family environment on human capital development? Specifically, the thesis addresses the following questions:

1. What is the effect of the FPE programme on school enrolment?

2. What are the implications of FPE policy on educational quality?

3. What is the effect of birth order on education?

Looking at the effects of both government policy (i.e. the FPE) and family environment (i.e. birth order) on education acknowledges the fact that both family and school factors, and government policy are essential in human capital accumulation.

This thesis is related to several strands of the literature. First, it relates to the literature that analyses changes in schooling and educational achievement in developing countries over time, sometimes following the introduction of large public policy reforms (Jopo et al., 2011b; Carnoy and Rothstein, 2013; Lekhetho, 2013; Rodrigues et al., 2013; Taylor and Spaull, 2013). Most of this literature, however, largely use measures of location (that is, mean, median, and variance) to analyse changes in educational quality (or achievement) and schooling enrolment (see Jopo et al., 2011b; Carnoy and Rothstein, 2013; Taylor and Spaull, 2013). For example, Carnoy and Rothstein (2013) use changes in PISA (Program for International Student Assessment) test scores since 2000 to assess changes in the U.S.'s educational quality, while Jopo et al. (2011b) look at changes in SACMEQ average test scores for Lesotho to infer changes in education quality since the introduction of FPE.

However, the use of this traditional summary statistic (i.e. the mean) leaves much of the information inherent in a test scores' distribution untapped and thus fails to reveal much of education quality dynamics. As a consequence, it is unclear whether the increase in average achievement in some countries, for example, is due to improved student performance levels at the upper and/or lower tails of the distribution. On the one hand, if the increase in average achievement is due to an increase in the performance of high-ability students, then we are faced with a scenario of improved average education quality and a rise in education inequality. If, on the other hand, the increase in average scores is a consequence of increased performance of low-ability students, then we have a case of a reduction in education inequality and improved education quality. Knowing which scenario prevails is essential for education policy in order to ensure that no child is left behind, not only outside the schooling system but also within the schooling system.

Furthermore, studies that go beyond averages in analysing changes in education (for example, Rodrigues et al., 2013) fall short of analysing the determinants of such changes. In this thesis, I examine the distributional changes in educational quality (or achievement), and the determinants of such changes, following the introduction of the FPE programme.

Second, this thesis speaks to the literature on the effects of large public policy reforms on school enrolment and educational achievement. For example, Schultz (2004) evaluates the effect of the PROGRESA programme in Mexico, which provides rural mothers with education grants, on school enrolment. In China, Chyi and Zhou (2014) examine the impact of public policy reforms targeting the rural poor (that is, tuition waivers, free textbooks, and living stipends for children). Both in China and Mexico, public policy reforms implemented in the respective countries had a significant positive effect on school enrolment.

I mentioned earlier that a number of sub-Saharan African countries have implemented FPE programmes. But there are few studies, concentrated on a few countries, that attempt to tease out the effect of these programmes on education in sub-Saharan Africa (see Deininger, 2003; Grogan, 2009; Lucas and Mbiti, 2012). For example, Deininger (2003), Nishimura et al. (2008), and Grogan (2009) assess the effects of the Ugandan FPE programme on schooling. Al-Samarrai and Zaman (2007), Lucas and Mbiti (2012), and Hoogeveen and Rossi (2013) evaluate the effects of FPE on education in Malawi, Kenya and Tanzania, respectively. These papers find positive effects of the FPE policy on schooling.

However, the results of some papers are potentially biased because of the methods they used. For instance, Deininger (2003) examines the impact of FPE on enrolment in Uganda by estimating Ordinary Least Squares (OLS) regressions of primary and secondary enrolment, respectively, on a vector of controls plus a dummy for post-FPE period which is meant to pick the effect. The results indicate that after the programme, primary school enrolment increased more than secondary school enrolment, and this is attributed to the FPE policy.

But we know that primary school enrolment rates are higher than secondary school enrolment rates in many developing countries, and that drop-out rates are differentially high among girls (Pritchett, 2004). This suggests that secondary school enrolment is potentially a bad counterfactual for primary school enrolment. Thus, it is hard to infer and/or quantify the impact of FPE from the analysis. Therefore, Lucas and Mbiti (2012) lament that, in spite of the widespread implementation of FPE programmes in SSA, there is a dearth of credible research evaluating their casual effect on child schooling. Furthermore, there is little knowledge on the implications of this policy for changes in educational achievement over time. This thesis adds to this growing literature while addressing some identification issues that have plagued earlier studies as highlighted by Lucas and Mbiti (2012). Given that resources are limited, it is always important to know whether any given public policy achieves the desired outcomes.

Third, this thesis relates to the literature that examines the influence of the family environment on human capital development (Leibowitz, 1974; Cunha et al., 2006). Specifically, it contributes to the literature that evaluates the effect of a child's unique characteristic, the order of birth, on human capital accumulation (Black et al., 2005; De Haan, 2010; Jayachandran and Kuziemko, 2011; De Haan et al., 2014). Even though a child's order of birth is biologically determined, parents can actively or passively choose to create different home environments for children of different birth orders, which will then affect their cognitive and non-cognitive skills' development. On the one hand, parents may enforce stricter disciplinary rules on the first-borns than later-borns, and thus lead to different educational outcomes between siblings. This constitutes passive differential investments on children. On the other hand, parents may purposefully invest relatively more on later-borns and/or boys because of cultural and/or economic pressures.

The available evidence on birth order effects on educational attainment shows a consistent divide between the developing world and the developed world: there tend to be negative birth order effects in developed countries, while there is evidence of positive birth order effects from developing countries (De Haan et al., 2014). Therefore, birth order effects on child outcomes appear to be context-specific, as it relates to countries' levels of development. The literature on birth order effects in the sub-Saharan African context is very sparse (see Tenikue and Verheyden, 2010) and the available evidence may not be generalised to other countries in the region. This points to the value of this thesis, which aims to provide new evidence on this subject from another sub-Saharan African country.

Lastly, the thesis contributes to the literature on child gender and parental investments. According to this literature, parents discriminate against girls in terms of investments in human capital development. For example, in India, parents invest more on boys than girls in terms of breastfeeding and provision of medical care, and this preferential treatment increases with birth order (Jayachandran and Kuziemko, 2011). These early better investments in the health of boys relative to girls, and younger children relative to earlier borns may translate into better human capital accumulation capabilities for boys and young children relative to their female and older siblings, respectively, (Cunha and Heckman, 2009). Furthermore, parents in most developing countries are more likely send boys to schools than girls (Morduch, 2000).

The fundamental question that guides this thesis is, what is the effect of public policy reforms and the family environment on human capital accumulation? With the following three self-contained analytical chapters, I attempt to answer this fundamental question, and also contribute the literature by filling in the research gaps highlighted above.

Lesotho presents an interesting case study for addressing this important question for the following reasons. Firstly, because of the unique way in which the FPE policy was introduced in Lesotho, one is able to analyse the causal effect of the policy on education using a relatively clean identification strategy. In countries where the FPE policy was introduced at once for all primary school grades, it have been difficult to tease out the potential effects of the policy on school enrolment and educational achievement (Lucas and Mbiti, 2012). This has been a significant impediment to our understanding of the effectiveness of the FPE programmes in different contexts.

Secondly, the availability of educational achievement data for before and after the FPE policy implementation makes it possible to analyse the dynamics of educational achievement since the introduction of the policy. Thirdly, being a less developed country with a relatively low fertility rate, Lesotho presents a perfect testing ground for the hypothesis that birth order effects on human capital development depend on a country's level of development. That is, given that Lesotho is a developing country, are birth order effects on educational attainment positive? This will not only tell us how birth order as a family environment factor affects child development, but also whether context other than a country's level of development determines such an influence. Lastly, I am not aware of any other study in Lesotho that engages the questions raised and/or address them in the way I do in this thesis. This makes this thesis unique both in the questions addressed and the approaches taken to address them.

To offer a preview, in Chapter 2, I examine the effect of the FPE policy on child schooling in Lesotho, thereby contributing to the scant literature evaluating the impact of this policy on education in sub-Saharan Africa. The analysis uses data from two cross-sectional household surveys, the 1994/5 Household Budget Survey data collected before the policy was introduced, and the Lesotho Core Welfare Indicators Questionnaire (CWIQ) survey data collected in 2002 after the introduction of the policy. To identify the causal impact of FPE on school enrolment, I use the difference-in-differences method. I estimate the lower- and upper-bound effects of the policy. The results show that FPE has significantly raised enrolment of primary school-age children, and more so that of boys (the historically disadvantaged group) and children from poor backgrounds. The most important implication of these results is that it is possible that, if the FPE policy can be instituted in those sub-Saharan African countries still lagging behind in education indicators, the 'Education for All' goal can be achieved.

While getting children to school is important in and of itself, schooling is not education, but rather an essential input in the production of education (Behrman, 2010). The next challenge is to make sure that children learn while in school. Indeed the ultimate goal of FPE policies, as contained in the DFA call, is to improve education quality (or achievement)[2], which is a much bigger problem in SSA. For example, Beatty and Pritchett (2012) estimate that it will take 150 years (134 years) for a median Southern and Eastern African country (in terms of SACMEQ[3] tests performance) to reach OECD reading (maths) levels.

There are arguably several factors contributing to this dire state of educational quality in SSA, but a common presumption in the current literature is that it has been exacerbated by the recent push to meet the universal primary education Millennium Development Goal (MDG) by 2015 (see for example, Colclough et al., 2008; Zuze and Leibbrandt, 2011). Although there is some evidence showing negative effects of FPE on achievement immediately after the introduction of the policy (see Lucas and Mbiti, 2012; Hoogeveen and Rossi, 2013), we have very little knowledge of how educational achievement and educational inequality have evolved since the policy was introduced. In particular, the questions of whether educational achievement falls (or increases) for all students (both low- and high-ability), and whether such a change in achievement comes with an increase (or decrease) in educational inequality have not been answered. Not only does educational inequality explain income inequality, it also correlates with health inequality (Deaton, 2003; Alves, 2012; Ferreira and Gignoux, 2014).

Chapter 3 of the thesis tries to answer these questions in the case of Lesotho. Using data on grade six pupils' standardised maths and reading test scores from SACMEQ, I first employ the relative distribution method of Handcock and Morris (1998, 1999) to analyse the distributional changes in achievement between 2000 and 2007, following Rodrigues et al. (2013). While the relative distribution method is good in both quantifying the changes in achievement and showing them graphically without any distributional assumptions, it has a limitation of not being able to apportion these changes to individual characteristics. Therefore, I apply the Recentered Influence Function regression decomposition by Firpo et al. (2007, 2009) and Fortin et al. (2011) to analyse the determinants of such changes. The results from this analysis reveal that the increase in educational achievement in Lesotho, between 2000 and 2007, was driven by the rise in the performance of both low-ability and high-ability students. I also find that educational inequality increased during this period. The results show that pupil-teacher ratio, grade repetition, "teacher effort" (that is, the number of teaching hours per week and frequency of testing students), and students' socioeconomic status explain some of the changes in educational quality and inequality. However, much of these changes in educational quality and inequality are not explained by observed school and child characteristics, suggesting that there could be some intangible school and family factors influencing these changes.

[2] I use educational achievement and education quality interchangeably throughout this thesis.
[3] Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ).

Even though changes in educational achievement and educational inequality in many sub-Saharan African countries have been linked to the changes in schools' resources and characteristics due to the FPE policy, the result that these school characteristics explain very little of the changes in achievement is not surprising in the literature. For example, Glewwe et al. (2011) review the literature that estimates the effects of school resources on students' achievement. From a list of studies they review, they find that there is little evidence that school resources have a positive impact on students' achievement.

There are at least three important implications of the results from chapter 3 of the thesis. First, Lesotho is one of the first countries to have sequentially implemented the FPE programme in sub-Saharan Africa. The fact that educational achievement increased since the policy was introduced, and that this increase is positively associated with educational policy variables (for instance, PTR and instructional time) suggests that the sequential implementation strategy potentially helps to minimise the disruptions in the system and reduce potential negative quality effects of the programme on quality. This is possibly a good policy lesson to those countries that are yet to institute FPE programmes.

Second, the increase in educational achievement inequality will have negative implications on health and earnings inequalities (Becker and Tomes, 1986a; Alves, 2012; Ferreira and Gignoux, 2014). The Lesotho government thus should implement policies to address educational inequality in order to reduce the high level of income inequality. More generally, the evidence suggests that for countries to effectively address the income inequality problem, distributional changes in educational quality have to be closely monitored, such that any emerging inequalities can be bridged in time.

Third, the result that grade repetition has a significant influence on achievement seems to support the evidence found by Foureaux Koppensteiner (2014) in Brazil (see also Rodrigues et al., 2013) which indicates that relaxing the retention policy reduces students' performance. Therefore, the finding suggests that the automatic grade promotion policy that is now being implemented in Lesotho could hurt educational quality.

In Chapter 4, I shift focus and evaluate the effect of family factors on the household distribution of education. Specifically, I examine the effect of birth order on children's educational attainment in Lesotho. "Childhood is the province of the family" (Heckman and Mosso, 2014, p.3). Therefore, much of schooling decisions and education happen within the household. In as much as taking children to school can help increase education attainment and reduce educational inequality, there are childhood factors that influence intra-household distribution of education among siblings, and ultimately the level of educational achievement and inequality within a country. I use the 2006 census data for children aged six to twelve years to estimate birth order effects on educational attainment. To deal with potential endogeneity of birth order resulting from its correlation with family size and other unobserved family factors, I use family fixed effects models.

Apart from this thesis being the first to estimate birth order effects on human capital development in Lesotho, it adds to the literature in three other ways. First, I use different measures of educational attainment, instead of one as is common in the literature. I estimate birth order effects on a short-run measure of education (that is, enrolment), and two long-run measures of education (that is, completed years of education, and *schooling progression* or age-adjusted schooling). Second, I examine two possible sources of heterogeneity in birth order effects: differences due to families of different sizes, and gender bias (that is, gender of the first-born). Lastly, I investigate whether family wealth and child-spacing can explain the observed birth order effects.

Different from the available evidence based on many developing countries data, I find large and significant negative birth order effect on child educational attainment. The birth order effects differ by family size and gender. Specifically, the negative birth order effects are more pronounced in large families, and in families of first-born girls. I also find strong evidence that this negative birth order effect is transmitted through birth-spacing, and not family wealth, contrary to earlier evidence from other developing countries (for example, Ecuador and some sub-Saharan countries) which shows that wealth is the underlying causal mechanism behind the positive, not negative, birth order effect on attainment (Tenikue and Verheyden, 2010; De Haan et al., 2014).

Surprisingly, the results affirm the negative birth order effects found in developed countries (for example, Norway and the United States) (Black et al., 2005; De Haan, 2010). The difference with the latter evidence is that, in the United States, for example, De Haan (2010) finds that family wealth, and not birth-spacing, explains the negative birth order effect. Therefore, I tentatively conclude that these findings are consistent with the *confluence* model's predictions, even though the hypothesis that first-borns do better because of being brought up under tougher parental disciplinary rules than their younger siblings cannot be ruled out.

The implications of the negative birth order results are twofold. One is that, the theories that try to explain birth order effects in developing country contexts cannot be generalised as yet. We still have to identify those conditions or contexts that fit the negative birth order effects and those that fit the positive birth order effects because it is evident that the level of development cannot be the only factor. Second, for policy direction, the results suggest that, in order to attenuate these birth order effects and increase educational attainment of later-borns, it is essential that the government designs policies that will improve school participation of later-borns, especially for boys and those from larger families. Without speculating on the general equilibrium effects, one such policy intervention could be the introduction of conditional cash transfers to larger families, particularly those with first-born boys.

Finally, in Chapter 5 I summarise the main results and conclusions presented in the preceding chapters.

Chapter 2

The Effect of Free Primary Education on Enrolment in Lesotho

2.1 Introduction

Since the adoption of the DFA by the World Education Forum in 2000, a number of sub-Saharan African countries have implemented free primary education (FPE) programmes, aiming to boost enrolment levels by abolishing all user-fees. Malawi and Uganda were among the first countries to introduce this policy in the mid-1990s. The policy, however, has been implemented differently in different countries. For example, in Malawi, school fees were first abolished on a grade-by-grade basis for the first three grades, starting with grade one from 1991 to 1993, and then simultaneously eliminated for all remaining grades in 1994. In Uganda and Kenya, user fees were simultaneously eliminated for all primary school grades in 1997 and 2003, respectively. In this thesis, I will refer to the FPE policy in these countries as the 'big bang' FPE because user-fees were eliminated for all grades in one year. In Lesotho, however, fee-elimination was phased-in grade-by-grade, starting with grade one in 2000, until it covered the entire primary schooling system in 2006.

To date, there is only a handful of studies that attempt to quantify the causal effects of FPE programmes, and almost all of them have concentrated on a few countries where the policy was implemented in a big bang fashion. Deininger (2003), for example, examines the impact of FPE on enrolment in Uganda by estimating Ordinary Least Squares (OLS) regressions of primary and secondary enrolment, respectively, on a vector of controls plus a dummy for post-FPE period which is meant to pick the effect. The results indicate that after the programme, primary-school children were 89 percent more likely to be enrolled, while secondary school enrolment increased by 62 percent during the same period (see Table 3 of

Deininger, 2003, p.299)[1]. The author then interprets the large increase in primary school enrolment as the impact of the FPE policy.

However, we know that primary school enrolment rates are higher than secondary school enrolment rates in many developing countries, and that drop-out rates are differentially high among girls (Pritchett, 2004). The fact that secondary school enrolment, which is used as a control, also increased significantly after the policy, could be attributed to part of the FPE policy in Uganda which specifically targeted girls' education by campaigning against early marriages (see Deininger, 2003 p.294). This suggests secondary school enrolment is potentially a bad counterfactual for primary school enrolment. Thus, it is hard to infer and/or quantify the impact of FPE from the analysis.

Grogan (2009) also assesses the Ugandan FPE programme, but instead of looking at its enrolment effects, she examines its effect on the age at which a child enters school. She specifically examines the FPE effect on the probability that a child begins school before age nine, which is the age associated with a lower probability of completing at least seven years of schooling. Exploiting the discontinuity created by the policy in the relation between child age and the probability of enrolment, she applies a regression discontinuity strategy, and discovers that FPE increased the probability of on time enrolment by 3 percent with a much higher effect (5 percent) for girls (see also Nishimura et al., 2008). This evidence is suggestive of sustained enrolment benefits of FPE for those who enrol on time. But given large fractions of overage children attracted by fee waivers, and their high drop out rates, it is hard to infer the longer term overall enrolment effects of FPE from the results.

In the case of Malawi, Al-Samarrai and Zaman (2007) document increases in primary and secondary school enrolment between 1990 and 1997, and also present suggestive evidence that girls and children from poor backgrounds benefited the most from FPE. They, however, do not identify the causal effect of FPE on enrolment. Lucas and Mbiti (2012) apply a difference-in-differences strategy, exploiting differences in pre-programme primary school drop out rates (and post-programme completion rates) across districts in Kenya, to tease out the FPE effect on access. They find a positive effect of FPE on the primary school completion rate, and a much higher effect on that of children from disadvantaged backgrounds.

Indeed the 2013/4 Education for All Global Monitoring report attributes the fall by half in the global number of out-of-school children between 1999 and 2011 to FPE programmes (UNESCO, 2014). But the question of how much of this drop is due to the causal effect of FPE, unconfounded with improving educational tastes over time, still begs for an answer. Moreover, the Global Monitoring report also reveals that universal primary education is unlikely to be met by 2015, and that sub-Saharan Africa still lags far behind all other

[1]Note, however, that Deininger (2003, p. 298), reports different figures that cannot be read from Table 3 in page 299. For instance, the author reports that primary school enrolment increased by 60 percent, while the time dummy coefficient is negative and statistically insignificant for secondary school enrolment.

regions with an estimated 22 percent of primary school-age children still out of school in 2011. Progress is even more abysmal for children from disadvantaged backgrounds. This further points to the value of learning more about FPE effects on enrolment from different contexts.

This chapter contributes to this growing literature, taking the FPE programme in Lesotho as a case study. I specifically examine the causal effect of FPE on overall enrolment, and whether the policy has bridged gender- and wealth-related educational inequalities. Lesotho presents an appropriate setting for evaluating FPE effect on enrolment while, at the same time, credibly addressing identification concerns. I take advantage of the unique phasing-in approach to school-fees elimination to identify the FPE effect on primary school enrolment, three years after its implementation. Looking at the FPE effect on primary school enrolment after three years gives an indication of whether the short-term (or immediate) effects on enrolment of the policy reported in other countries will likely be sustained or whether such gains might be reversed as a result of school drop out of primary school, for example.

The main outcome variable is net enrolment rate (NER), defined as the proportion of official primary school-age children who are currently in school. In this study, official primary school-age children consist of all those aged 6 to 12 years old. By comparing changes in enrolment in grades covered by FPE to changes in enrolment in grades not-yet-covered, I am able to wipe out the common trends in education demand over time and hence credibly identify the FPE effects. In line with previous studies elsewhere, I show that the removal of school fees significantly increased net enrolment of the primary school-age population by at least 9.3 percentage points (or 13.2 percent). More importantly, the results reveal that FPE led to a significant increase in enrolment of boys (the historically disadvantaged group in Lesotho)[2] relative to that of girls, and of children from poor backgrounds.

The chapter proceeds as follows. In Section 2, I present Lesotho's institutional context and the FPE policy background. Section 3 presents the theoretical framework, while Section 4 discusses the data used. Section 5 presents the descriptive evidence of FPE effects. Section 6 explains the identification strategy and Section 7 discusses the main results. Section 8 concludes the paper.

2.2 The institutional context and policy background

2.2.1 The institutional context

Lesotho education follows a 7-3-2-4 system with seven years of primary education, three years of junior secondary education, two years of senior secondary education, and four years

[2] In Lesotho, boys look after their families' livestock, while girls do household chores. This, therefore, disadvantages boys relative to girls when it comes to schooling.

of university education. The official age of entry into primary schooling is six years (or five years for children born on or before the 30$^{\text{th}}$ of June) such that by age twelve, children should be in grade seven completing their primary school. This implies that the official primary school-age is 6 to 12 years.

Unlike many other countries, most primary schools in Lesotho (about 80%) are owned and controlled by various churches (see Table A.1), and these churches are represented in the national education advisory board by their appointed education secretaries (Ambrose, 2007). Non-religious private schools constitute only one percent, and are not covered by FPE. Notwithstanding this co-ownership structure, all schools follow the same national curriculum provided by the Ministry of Education and Training (MOET). Additionally, the government has an overall authority in pronouncing education policies, management and regulation of education, training of teachers, and teachers' appointments, dismissals, and deployments.

With regard to students' progression within the system, the *de jure* government policy since 1967 is that of automatic grade promotion (see Ambrose, 2007), but *de facto*, schools still practice grade retention. This, coupled with delayed school enrolment, implies that, in any given grade, one is most likely to find students of different cohorts. Beyond physical and monetary costs, there are no regulatory restrictions with regard to school choice in Lesotho. Parents are thus free to choose schools to which they prefer to take their children.

Primary schools' funding mainly comes from the government through payment of teacher salaries. At least until the year 2000, primary education was not free in Lesotho. Thus, another way that schools used to raise funds was through charging school fees. The situation changed in 2000 when the government introduced the FPE programme, starting with standard 1 (or grade 1). I now turn to the background of this policy.

2.2.2 Policy background

As a signatory to the World Declaration on 'Education for All', and the United Nations' Millennium Development Goals (MDGs), Lesotho has pledged to ensure that primary education is free and available to all (UNESCO, 2000). Therefore, in 2000, the country instituted the FPE programme. However, as mentioned earlier, Lesotho's implementation strategy was different from that followed by other African countries. School fees were eliminated sequentially on a grade-by-grade basis starting with grade one in 2000; such that by 2002, the first three grades were covered, and all seven primary school grades were under FPE by 2006[3]. The main reason for this implementation strategy was to cushion FPE's financial impact on the public budget (Urwick, 2011).

The FPE policy addresses both the demand and supply side constraints to schooling. On the demand side, the policy involves elimination of some private schooling costs, such

[3] Figure A.1 shows the phasing-in of the FPE programme.

as school fees, stationery and textbooks' costs. On the supply side, the government builds additional classrooms in existing schools, and new government schools where none existed before. The latter helps to reduce the distance to schools, and hence transportation costs.

In addition, the government provides annual capitation grants, furniture and teaching materials to all schools, including church/private schools (Jopo et al., 2011a; Lekhetho, 2013). As of 2000, these capitation grants were calculated as follows. The book rental and maintenance fees were set at 15 maloti (M15)[4] and M5 per pupil per year, respectively; pupils' stationery per class per year was M11.51, while the annual teaching materials fee per teacher was set at M37 (Jopo et al., 2011a). These fees have been increasing but not on an annual basis. The school feeding programme has been integrated into the FPE programme to further encourage school attendance. It is necessary to mention, though, that prior to the implementation of FPE, schools were already providing lunch meals to pupils. These costs, however, were included in the school fees. Under FPE, all pupils in covered grades are provided with free meals.

In the first year of the policy, there was a 75 percent surge in the number of new school entrants in grade one (MOET, 2011), and a 33 percent increase in primary school net enrolment rate[5]. Urwick (2011) also reports that gross enrolment rate (GER)[6] increased by 34 percentage points from 107% to 141% between 1996 and 2006. The author then posits that this increase in GER was due to high grade repetition rate and late school entry, especially after FPE introduction. These figures, however, cannot be given a causal interpretation.

2.3 Theoretical framework

I use the canonical model of household behaviour by Becker (1965) where a household is both a consumer and a producer. I assume that a household combines time (for example, schooling time, child care time, etc.) with market goods (for example, books, school materials, school uniform, etc.) through a production function to produce basic commodities such as child quality. Subject to the schooling production function, income, and time constraints, the household then maximises its utility and chooses the best combination of these commodities, including child quality (see Thomas et al., 1990; Schultz, 2004; Mani et al., 2013).

In the presence of non-market goods, such as parental time spent on child care, a conditional demand function for schooling is one of the solutions to the household maximisation problem explained above. Therefore, following Thomas et al. (1990), Schultz (2004), and

[4]A Lesotho loti is pegged 1-to-1 to a South African rand (ZAR). 1 USD equals 12.54 ZAR as of 29/07/2015.

[5]Data available at http://data.uis.unesco.org/. UNESCO Institute of Statistics.

[6]GER is the ratio of the total number of primary school students, regardless of their age, to the total number of official primary school-age population.

Mani et al. (2013), I model determinants of schooling using the conditional demand function for schooling given as:

$$Y_i = f\left(P_i^s, \mathbf{P}_i^c, w_i, \mathbf{X}_i, z_i, \eta_i, \mu_i\right), \tag{2.1}$$

where Y_i is enrolment by child i, P_i^s is the price of schooling faced by child i, $\mathbf{P}_i^c$ is a vector of prices for all other commodities, w_i is the price of leisure, z_i is child i's household wealth, and $\mathbf{X}_i$ is a vector of household and individual characteristics (age, gender, and household demographic composition). η_i is child i's innate ability which influences parental allocation of schooling inputs, while μ_i captures all child rearing and caring practices in student i's home which also affect schooling.

I think of child schooling as measured by a short term measure like enrolment as one dimension of child quality. Child height and the child survival rate are other dimensions of child quality which can also influence child schooling (see Thomas et al. 1990), but which I do not consider in this study.

According to Becker (1965), the full price of schooling or investment in human capital is the sum of direct costs (for example, school fees, transportation, school uniforms, textbooks and materials) and indirect costs such as the forgone child wage. Therefore, a household will invest in a child's education if the marginal utility (or benefit) of an extra year of schooling exceeds the marginal cost. The implementation of FPE reduced the price of schooling and that of child feeding (due to free lunch meals at school) faced by households. Consequently, to the extent that schooling is a normal consumption good, the introduction of FPE should increase the household's demand for schooling, *ceteris paribus* (Schultz, 2004).

2.4 Data sources and summary statistics

I use data from two cross-sectional household surveys conducted by the Lesotho Bureau of Statistics. The first is the 1994/95 Household Budget Survey (HBS) which collected information from a sample of 3,784 households (with 25,272 individuals and 4,654 primary school-age children) interviewed roughly five years prior to the introduction of FPE policy (Bureau of Statistics, 2006). The second is the Lesotho Core Welfare Indicators Questionnaire (CWIQ) survey conducted in 2002 – three years after FPE introduction – collecting information from 4,954 households totalling 20,031 individuals, 4,175 of whom are primary-school-age children (Bureau of Statistics, 2002)[7].

The main variable of interest from these data is enrolment of primary school-age children

[7]These are the only surveys that are close enough to the cut-off point, 2000, and have comparable information on individuals' current educational status and grade level. For instance, individuals are asked about their *highest level of education*, which is the highest grade completed in both surveys.

in 1995 and 2002. In line with conventional practice in the literature, I focus on net enrolment rate (NER) rather than gross enrolment rate (GER) as the schooling access measure of interest. The GER may exceed 100 percent due to high participation of over-age children and/or under-age children. Therefore, Bold et al. (2015) argue that increases in GER are not necessarily a positive development, and that using NER is more appropriate when one is interested in measuring demand for schooling among comparable age-groups of children across time periods.

To improve the precision of our estimate of the effect of FPE, I need to control for other factors that potentially have an influence on enrolment. Therefore, student characteristics, such as age and gender, and parental (or household head's) education proxied by years of schooling, are important. Controlling for age is important for two reasons. Firstly, there is an official age of entering the school system, which means that older children are more likely to enrol than younger ones. Secondly, the opportunity costs of sending children to school increase with age as they become more primed for the child labour market. The prevalence of specific-gender preferences (for example boys being preferred over girls in many societies) when it comes to education makes it essential to control for gender in the schooling equation. With regard to parental education, Behrman (2010) notes that it increases child's education through home teaching, especially during pre-school and school-age stages. Mani et al. (2013) further posit that parental education affects child's education via parental choice of schooling inputs.

Following Mani et al. (2013), I also control for differences in households' demographic composition using the number of adult males and females (i.e. those aged 18 and above) in the household, and for the household's experience in child caring and life cycle position using the age of the household's head.

Lastly, I control for household size, and wealth proxied by the household's socio-economic status index, which measures the long-term position of the household[8]. This wealth index is constructed using principal components analysis based on a number of household wealth-indicator variables which include household amenities, such as type of house roofing and wall materials, sources of drinking water and fuel, distance to public services (e.g. clinic, etc.) and household ownership of assets like a TV, refrigerator, etc. To ensure comparability of the index over time, I first pooled the data to ensure that the same weights on respective wealth indicators are applied on both surveys. Moreover, I use only those wealth indicators that are common to both surveys. I then normalised the index to have a minimum value of zero for ease of interpretation.

Table 2.1 shows summary statistics for the outcome and control variables, and in the notes, it mentions all wealth indicator variables. The reported statistics are for a sample of

[8]Note that parental education is also used to proxy household wealth, specifically that part of wealth not captured by the wealth index.

all primary school-age children. We can see from the the table that enrolment significantly increased between 1995 and 2002. For example, in 1995, enrolment among primary school-age children was 69 percent, and it increased by 15 percentage points in 2002 to stand at 84 percent. This increase in enrolment is both statistically and economically significant. The table further shows that household wealth dropped, and that the number of adult males and females in the household declined, in line with a one member decrease in household size over the period. Next, I take a closer look at what drives much of the change in enrolment.

Table 2.1: Summary statistics

Variable	1995			2002			|t|-statistic
	mean	sd	N	mean	sd	N	
Enrolment	0.707	0.463	4631	0.839	0.367	4149	14.948***
Wealth index	5.433	0.041	4415	5.115	0.039	4175	5.660***
Child Age	9.094	1.995	4654	9.038	2.061	4175	1.292
Child Sex (Male)	0.514	0.500	4654	0.503	0.500	4175	0.967
Household Head's educ	4.333	3.911	4652	4.388	3.935	4175	-0.656
Household Head's age	49.47	13.32	4557	49.01	14.09	4122	1.583
Adult males	1.501	1.043	4654	1.239	0.928	4175	12.400***
Adult females	1.754	1.006	4654	1.508	0.805	4175	12.603***
Household size	7.314	2.716	4654	6.380	2.321	4175	17.260***

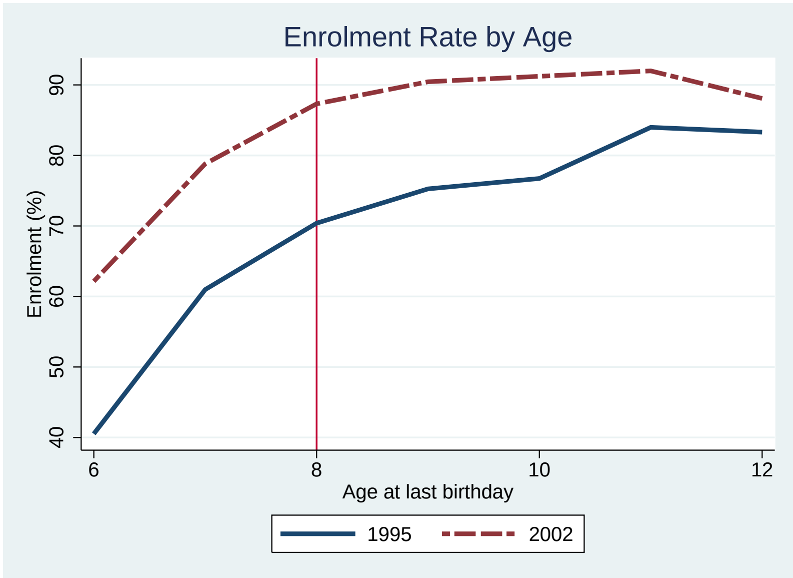
Source: Own calculations using 1995 HBS and 2002 CWIQ surveys. *Notes*: The sample is all primary school-age children. All the figures are unweighted because both surveys do not provide sampling weights. * significant at 10%, ** significant at 5%, *** significant at 1%. The construction of the wealth index is based on the following variables: roofing and wall materials; number of rooms in the house; source of drinking water; source of fuel for cooking, heating, and lighting; distance to nearest water source, public transport, health centre, primary school, and local shop or food market; refrigerator; radio, bicycle, wheel barrow, car, land, cattle, sheep, goats, horses, donkeys, sewing machine, and motor vehicle.

2.5 Descriptive evidence on enrolment changes since FPE

In this section, I provide suggestive evidence on the effect of FPE on school enrolment. Figure 2.1 plots enrolment rates by primary school-appropriate age, from age 6 to age 12. For age 6, for example, NER is defined as the proportion of six year old pupils to the total number of 6 year olds in the population. The figure clearly shows that enrolment increases with age, and this age gradient is high between ages 6 and 8. For instance, in 1995, NER at age 6 was about 30 percent, then almost doubled to 58 percent at age 7, and further increased to about 65 percent at age 8. Similarly, in 2002, NER for 6 year olds was about 60 percent, increased to about 80 percent by age 7, and then stabilised at 85 percent from age 8.

Furthermore, the figure shows that NER increased significantly for all primary school ages, but evidently more so for the 6 to 8 year old group. For example, NER of the 6 year olds more than doubled between 1995 and 2002, jumping from 30 percent to more than 60 percent. NER of the 7 year olds also increased by an impressive 22 percentage points from 58 percent to 80 percent in 2002. A similar increase is observed for the 8 year olds during the same period. However, the increase in NER of the 9 to 12 year olds over the same period was relatively small.

Figure 2.1: Pre- and Post-FPE enrolment rates by age



Source: Own calculations based on 1995 HBS and 2002 CWIQ data. *Notes*: Enrolment is a dummy equal to 1 if a child is aged between 6 and 12 years is enrolled in primary school.

I now turn to Table 2.2 which reports NERs by gender, age, grade and household wealth, before and after FPE. To aid the discussion, it is necessary to outline some definitions. NER is defined for all primary school-age children excluding those enrolled in secondary schools[9]. In anticipation of the discussion later on in the chapter, I report net enrolment rates by *age-group* (i.e. the 6-to-8 year old group versus the 9-to-12 year old group), and by *grade-group* (i.e. primary school children with at most grade 2 versus those with grade 3 to 7). Thus, I define the *age-group NER* as the total number of children in the age-group (say 6 to 8 years old group) enrolled in primary school divided by the total number of children in that age group expressed as a percentage, and *grade-group NER* is the proportion of all primary school-age pupils in the grade-group (say all 6-to-12 year olds in grade-group '*Grade* $<$ *3*' and enrolled) to the total number of primary school-age children who have completed at most grade 2 (i.e. those with Grade less than 3).

Table 2.2 shows that overall NER increased significantly by about 13 percentage points after the introduction of FPE. This is consistent with our expectations and is similar to

[9]Some primary school-age children who enrolled before the entry age of 6 and/or those who were fast-tracked are in secondary schools.

results found in Kenya and Uganda (Deininger, 2003; Grogan, 2009; Bold et al., 2015). Male children, who historically have lower enrolment rates compared to their female counterparts, appear to have been the biggest beneficiaries of the FPE programme, even though their enrolment rate is still low relative to that of females. For example, after FPE, net enrolment for boys increased by 16 percentage points from 65.4 percent in 1995 to 81.3 percent in 2002, while that of girls increased by 10.4 percentage points over the same period.

With regard to enrolment patterns by age-group, the table shows that the NER for all age-groups increased over the period, but much more so for the 6-to-8 years old group. For instance, while the NER of the 9-to-12 year olds increased by 10.8 percentage points, that of the 6-to-8 year olds increased by about 17.8 percentage points after FPE.

In anticipation of the discussion later on the chapter, the table also presents changes in NER by grade-groups; the grade-group '*Grade < 3*' (i.e. the group whose highest grade completed is less than 3), and grade-group '*Grade 3 to 7*' (i.e. the group whose highest grade completed is between 3 and 7). For the grade-group '*Grade < 3*', NER is defined as the total number of primary school-age pupils with less than grade 3 divided by the total number of primary school-age children who have completed at most grade 2, expressed as a percentage. Given that the FPE has no age restrictions, changes in grade-group NER are a broad measure of FPE effects on enrolment. We can see from the table that the NER of children with grades 3 to 7 remained almost constant at 96 percent over the period, while that of children with less than grade 3 (i.e. those eligible for FPE in 2002 irrespective of age) increased by an impressive 21.3 percentage points to stand at 79.3 percent after FPE. From this perspective, irrespective of how the FPE target group is defined (either by grade-appropriate age or by highest grade completed), there is a strong *prima facie* evidence that FPE raised enrolment rates between 1995 and 2002.[10]

Furthermore, FPE appears to have benefited the poor the most relative to the rich (see Table 2.2 and Figure 2.2). Firstly, Table 2.2 and Figure 2.2 show a clear evidence of a wealth gradient in enrolment in both periods. Table 2.2 shows that, in 1995, less than half of primary school-age children in the bottom wealth quintile (i.e. 48.5 percent) were enrolled, whereas almost three quarters (i.e. 74.4 percent) of those in the third quintile and 86.6 percent of those in the wealthiest quintile were enrolled. Similarly, in 2002, about 16 percent of primary school-age children in the bottom quintile were not enrolled, whereas only 6 percent of those in the wealthiest quintile were not enrolled.

Figure 2.2 clearly shows that the wealth gradient in 1995 was much larger between the bottom wealth quintile and the third wealth quintile than it was between the third quintile and the fifth quintile. In 2002, the wealth gradient was almost constant through out the

[10]Note, however, that those with at least grade 3 have already enrolled in school at some point and are potentially a select group. Later in the chapter, I empirically evaluate whether this group is a reasonably good counterfactual for those with less than grade 3.

Table 2.2: Pre- and Post-FPE Net Enrolment Rates by groups

	Obs.		NER	
	1995	2002	1995	2002
All children	4631	4149	70.7	83.9
By Gender				
Female	2252	2058	76.2	86.6
Male	2379	2091	65.4	81.3
By Age group				
6-8 years old	1880	1765	57.5	75.3
9-12 years old	2751	2384	79.6	90.4
By Highest Grade				
Grade < 3	3042	2980	58.0	79.3
Grade 3 to 7	1428	1131	96.6	96.2
By Wealth				
Quintile 1	893	876	48.5	74.1
Quintile 2	890	930	65.4	80.0
Quintile 3	883	951	74.4	85.8
Quintile 4	911	751	78.6	89.5
Quintile 5	815	641	86.6	93.9

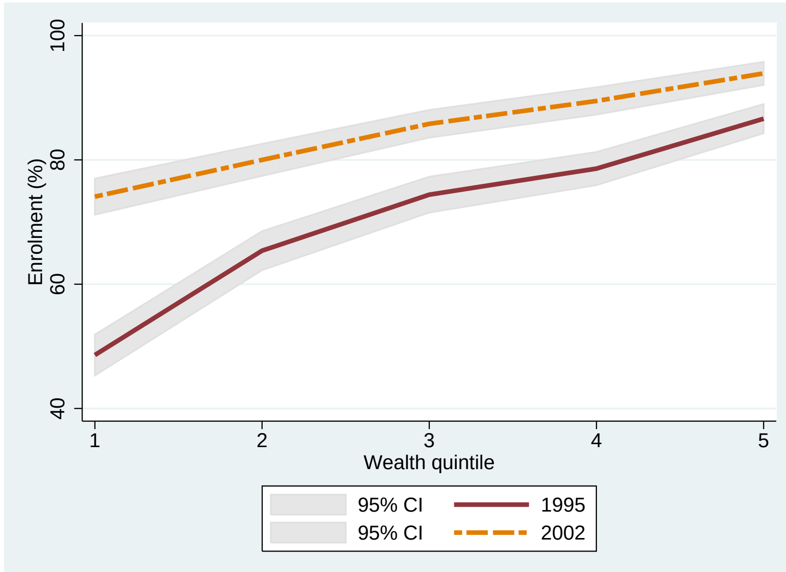
Source: Own calculations based on 1995 HBS and 2002 CWIQ data. *Notes*: The sample is primary school-age children, excluding those enrolled in secondary school. All statistics are unweighted. The age-group NER is the proportion of pupils in the age-group (say 6-8 year old group) to the total number of children in that age-group. Grade-group NER is the proportion of all primary school-age pupils in the grade-group (say all 6-to-12 year olds in grade-group '*Grade < 3*' and enrolled) to the total number of primary school-age children who have completed at most grade 2. NERs for all other categories are calculated using the conventional definitions.

wealth quintiles, and less than it was in 1995. This suggests that FPE lessened the influence of wealth on enrolment by benefiting the poor more than the rich.

Table 2.2 shows that NER of students from poor backgrounds (i.e. those at the bottom wealth quintile) jumped by an impressive 25.6 percentage points from 48.5 percent in 1995 to 74.1 percent in 2002. Primary school-age children in the second wealth quintile were the second greatest beneficiaries, their enrolment increased by 14.6 percentage points to stand at 80 percent in 2002. The smallest increase in enrolment after FPE occurred in the fifth quintile with only an 7.3 percentage points increase. Enrolment in the third and fourth wealth quintiles increased by 11.4 and 10.9 percentage points, respectively. By and large, even though the NER increased across all wealth quintiles, the increase was much higher for those coming from poor households. As Figure 2.2 clearly shows, these enrolment gains are statistically significant across wealth quintiles.

These naïve comparisons of enrolment rates before and after the programme do not reflect the causal effect of FPE on enrolment because the observed changes in enrolment could have entirely been driven by factors unrelated to FPE; for example, changes in tastes for education over time. To attach a causal interpretation to these changes, it is necessary to address some

Figure 2.2: Pre- and Post-FPE Net Enrolment Rates by Wealth



Source: Own calculations based on 1995 HBS and 2002 CWIQ data. *Notes*: Net enrolment is calculated for each wealth quintile. Wealth quintiles are for the household asset index constructed using the principal components analysis.

of these identification problems. I turn to this issue in the next section.

2.6 Empirical strategy

2.6.1 The model

To tease out the treatment effect of the FPE programme on enrolment, I employ the difference-in-differences (DID) identification strategy (Angrist and Pischke, 2009; Blundell and Dias, 2009; Imbens and Wooldridge, 2009). The DID approach is appropriate in settings where we observe outcomes for individuals in two groups (i.e. control and treatment groups) for two or more periods (pre- and post-programme) as in panel data, and in settings where we have repeated cross-sections from the same population, pre- and post-programme, for individuals in the control and treatment groups.

I use data from repeated cross-sections. In this case, the DID strategy is as follows. Suppose child i belongs to a group $G_i \in \{0, 1\}$ (where $G_i = 1$ is the treatment group), and is observed in period $T_i \in \{0, 1\}$ (where $T_i = 1$ is post-programme period). Let $Y_i(0)$ and $Y_i(1)$ be her potential outcomes (e.g. enrolment status) before and after the programme, respectively. Therefore child i's outcome is given as

$$Y_i = \begin{cases} Y_i(0) \equiv \alpha + \gamma T_i + \lambda G_i + \varepsilon_i, & \text{if } T_i = 0 \\ Y_i(1) \equiv \alpha + \gamma T_i + \lambda G_i + \delta_{DID} + \varepsilon_i, & \text{if } T_i = 1 \end{cases} \tag{2.2}$$

where γ is the year-specific effect common to both control and treatment groups; λ is a group-specific, time-invariant coefficient; δ_{DID} is the DID effect parameter; and ε_i is an unobserved individual error term.[11] The DID estimand, δ_{DID}, is then given by

$$\begin{aligned} \delta_{DID} &= E[Y_i(1) - Y_i(0)] \\ &= (E[Y_i|G_i = 1, T_i = 1] - E[Y_i|G_i = 1, T_i = 0]) \\ &\quad -(E[Y_i|G_i = 0, T_i = 1] - E[Y_i|G_i = 0, T_i = 0]) \end{aligned} \tag{2.3}$$

This double differencing procedure removes biases due to permanent pre-existing differences between the two groups, and biases resulting from common time trends unrelated to the programme. Let D_i be the treatment dummy equal to the interaction of the group and time dummies, $D_i = T_i \cdot G_i$, which equals 1 if $T_i = 1$ and $G_i = 1$. Then equation 2.3 can simply be estimated by least squares methods using the following regression DID[12]

$$Y_i = \alpha + \gamma T_i + \lambda G_i + \delta_{DID} D_i + \mathbf{X}_i\beta + \varepsilon_i \tag{2.4}$$

where $\mathbf{X}_i$ is a vector of controls.

The DID method requires that the control group not be influenced by the programme, and be comparable to the treatment group. As mentioned earlier, the FPE programme was progressively implemented from grade 1 in year 2000 such that by 2002, it was covering grades 1 to 3. Given the school entry age of 6 years, the age-appropriate cohort for grades 1 to 3 is the 6-to-8 year olds (see Figure A.1). Thus, I define the treatment group in two ways: (i) by age-group (i.e. the age-appropriate cohort for grades 1 to 3), and (ii) by grade-group (i.e. whether a school-age child had completed at most grade 2 by year 2002).

Under the first definition (the age-group definition), the treatment group refers to all children aged 6 to 8 years old. I use all those aged 9 to 12 years old as the control group. These two groups are quite similar in terms of their characteristics as I show in Table 2.3. Notice that since the FPE programme had no age restrictions, the control group in this case inevitably includes some pupils who have received treatment because of delayed enrolment and/or grade repetition[13]. Second, there could be intra-household spillover effects such that eliminating fees a younger sibling in the household may free up household resources to be used to pay fees for an older sibling. For these reasons, I call the treatment group defined this way a *narrowly-defined* treatment group. The implication of this definition is that the

[11] A much stronger version about the individual error term for the DID strategy is that it is assumed to be independent of the G_i, and to have the same distribution over time, $\varepsilon_i \perp (G_i, T_i)$.

[12] This is the estimable counterpart of the conditional schooling demand function in equation 2.1.

[13] Figure A.2 in the Appendix shows the age composition of each grade. For example, there were more 12 year olds with at most grade 2 in 2002 than in 1995. This clearly indicates that fee-elimination induced many overage children who would have otherwise been left out of the schooling system to enrol.

DID estimand will be biased downward, and hence identify the lower bound effect. The advantage of this definition of the treatment group is that it is not possible for any child to manipulate her age in order to fall in any particular age group. After controlling for individual and household characteristics, any other allocation mechanism not explained by age is largely random.

According to the second definition (the grade-group), the treatment group is all primary school-age children who had completed at most grade 2 by year 2002 such that, if enrolled in 2002, their highest current grade should be grade 3. I call the treatment group defined this way a *broadly-defined* treatment group. Under this treatment-group definition, the control group is all primary school-age children who had completed at least grade 3 in 2002. By definition, unless they intentionally decided to repeat lower grades, these children should not have benefited from the policy in 2002. If they did repeat lower grades, this will bias upwards the FPE effect. But if they instead decided to stay out of school waiting for fee elimination in their new grades, this will bias downwards the FPE effect. Thus, one cannot tell *a priori* the direction of the bias in this case. Nevertheless, I show in Section 2.6.3 that there were no unusual spikes in the trends of repetition rates (and percentage of repeaters) by grade over time, which, if evident, would suggest the presence of some bias. I also show that there are no group compositional changes over time.

The second concern with the broadly-defined treatment group is that, the control group (i.e. those with at least grade 3) may not be a good counterfactual. This is because children with at least grade 3 have already enrolled in school at some point and are potentially a select group of high-ability children and/or from wealthy families. However, I show in the next section that the control group is a reasonably good counterfactual, and I control for family wealth in all the regressions to address this concern.

There are at least three concerns about the internal validity of the DID strategy, with respect to both group definitions. First, there may be intra-school spillover effects. For example, schools may respond to rising enrolment in lower grades (that is, grades affected by the policy) by diverting resources (e.g. teachers, classroom space, etc.) away from higher grades (that is, grades not-yet-affected by the policy). These spillover effects would lower enrolment in the control group and thus overstate the policy effect. However, the concurrent occurring of school-building and teacher-recruitment programmes would likely mitigate any intra-school spillover effects. As mentioned earlier in the introduction, the number of schools and teachers increased considerably between 2002 and 2006, by 9 percent and 21 percent, respectively. This suggests that, as demand for schooling was increasing in lower grades, resources were made available to cater for such demands. Therefore, the intra-school enrolment spillover effects, if any, should at least be very minimal. In fact, this should have more bite on the policy effect under the broad definition of the treatment group where such resource-shifting would affect children of all ages in the control and treatment grades.

Second, the school construction programme may have differentially affected the treatment and control groups. By design of the FPE programme, younger children and children who had not completed the treated grades have likely benefited the most from the school expansion programme especially in cases where only additional classes were built to cater for the increased demand in lower grades. In cases where new schools were built, this is likely to have benefited the control and treatment groups equally. In either case, it is unlikely that the school building programme negatively affected the enrolment in the control group.

Third, the school expansion and fee elimination could have interaction effects on enrolment. Similarly, if this interaction effects affect the control and treatment groups differentially, this will violate the DID strategy's assumptions. But, if these effects are similar for both control and treatment groups, then this will have no implication on the identified average treatment effect of the FPE policy. It is hard to test and/or solve this possibility. However, if the this interaction effects are positive, given that both the fee elimination and school construction programmes have at least positive influence on enrolment, then the FPE policy effect will be attenuated under either treatment group definition.

2.6.2 Assessing the plausibility of assumptions under the narrowly-defined treatment group

Given that the treatment group is defined in two different ways, it is fitting to discuss the plausibility of the identifying assumptions in relation to each treatment group definition separately. In this sub-section I discuss the identifying assumptions and their plausibility in relation to the narrowly-defined treatment-group, and relegate a similar discussion in respect of the broadly-defined treatment group to sub-section 2.6.3.

To effectively wipe out the common time trend between the groups, the main identifying assumption required is that, in the absence of the policy, the change in outcome (e.g. enrolment) of the treatment group would have been equal to the change in enrolment of the control group. In this two-period, two-group setting, however, this assumption cannot be tested. Pre-programme enrolment data corresponding to the narrowly-defined treatment group is not available for this study, otherwise it would have been possible to indirectly assess the plausibility of the common trends assumption[14].

The second most important assumption required by the DID strategy in cross-sections is that there should be no significant compositional changes in the control and treatment groups

[14]But with respect to our broadly-defined treatment group, we are able to test the plausibility of this assumption using enrolment-related data (i.e. repetition and drop-out rates, and percentage of repeaters by grade over time).

before and after the policy to ensure before-after comparability (Blundell and Dias, 2009). To assess the plausibility of this assumption, I follow Foureaux Koppensteiner (2014) and Bettendorf et al. (2012), and compare characteristics of the control and treatment groups before and after FPE in Table 2.3. The table shows the descriptive statistics of the characteristics of the control and treatment groups, and their differences in 1994/95 and 2002. Specifically, the table shows the means and differences in means of the outcome variable (enrolment), and individual and household characteristics (child age, child sex, household wealth, household size, household head's education and age, and number of adult males and females in the household) of the control and treatment groups, before and after the programme.

Columns (4) and (9) of the table show that the differences in all characteristics are quite small, the highest being 3.6 for child age. The large differences in child age are inevitable given our definition of the treatment group. The most important result is that the differences in most characteristics between the groups have remained stable over time. For example, the group differences in child age, child sex, household head's age, and number of adult males are quite similar over time. Even though there appears to be a slightly faster decline in household size and number of adult females in the household in the treatment group relative to the control group, I note that changes in all these variables normally follow a gradual process and are unlikely to have been caused by sudden jumps that could confound the results. I therefore conclude that changes in the composition of the groups with respect to their fixed effects (or unobservable characteristics) have remained unchanged over time, and hence unlikely to bias the estimation results.

To drive home this argument, I check whether the control and treatment groups are balanced in characteristics. Although this balancing property (or the unconfoundedness assumption) is not fundamental in the DID strategy, it is generally reassuring when it is satisfied (Foureaux Koppensteiner, 2014). Columns (3) and (8) of Table 2.3 show the p-values for the t-test of differences in mean characteristics of the control and treatment groups. Looking at these p-values, we can see that for all characteristics, except child age and household head's age, the hypothesis that they are balanced between groups cannot be rejected at the 5 percent level of significance.[15]

I further check the validity of this balancing property by looking at the normalized differences in characteristics following Imbens and Wooldridge (2009). The normalized difference between means is a scale-free measure of the difference in characteristics' distributions by treatment status, and is given as

$$\text{Norm-diff} \equiv \triangle_{\mathbf{X}} = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{S_0^2 + S_1^2}} \tag{2.5}$$

[15] Differences in household head's age are merely a function of differences in child age induced by the definition of the treatment group.

Table 2.3: Mean characteristics of control and treatment groups in 1994/95 and 2002

	1994/95					2002				
Variable	Treatment (1)	Control (2)	p-value (3)	Diff $(X_1 - X_0)$ (4)	Norm-diff (5)	Treatment (6)	Control (7)	p-value (8)	Diff $(X_1 - X_0)$ (9)	Norm-diff (10)
Enrolment	0.5750	0.7964	0.000	-0.2214	-0.3472	0.7530	0.9035	0.000	-0.1505	-0.2880
	(0.4945)	(0.4027)				(0.4314)	(0.2953)			
Wealth Index	5.5027	5.3859	0.1586	0.1168	0.0305	5.0741	5.1441	0.3728	-0.0670	-0.0197
	(2.7100)	(2.6962)				(2.5124)	(2.5019)			
Child Age	7.0122	10.5085	0.000	-3.4963	-2.5017	6.9399	10.5747	0.000	-3.6347	-2.6025
	(0.8146)	(1.1357)				(0.8205)	(1.1301)			
Child Sex (Male)	0.5088	0.5168	0.5913	-0.0080	-0.0113	0.5076	0.5000	0.6254	0.0076	0.0108
	(0.5001)	(0.4998)				(0.5001)	(0.5001)			
Household Head's Educ	4.3714	4.2458	0.2826	0.1257	0.0227	4.4006	4.3788	0.8601	0.0217	0.0039
	(3.9585)	(3.8849)				(3.9338)	(3.9365)			
Household Head's Age	48.6839	50.0059	0.0010	-1.3220	-0.0701	48.2342	49.5706	0.0026	-1.3364	-0.0669
	(13.440)	(13.216)				(14.433)	(13.815)			
Adult males	1.4870	1.5106	0.4478	-0.0237	-0.0161	1.2170	1.2556	0.1842	-0.0386	-0.0296
	(1.0065)	(1.0678)				(0.8946)	(0.9514)			
Adult females	1.7732	1.7409	0.2815	0.0323	0.0227	1.5065	1.5087	0.9306	-0.0022	-0.0019
	(1.0289)	(0.9895)				(0.8057)	(0.8053)			
Household size	7.3261	7.3053	0.7979	0.0208	0.0054	6.3082	6.4332	0.0857	-0.1250	-0.0381
	(2.7259)	(2.7094)				(2.3048)	(2.3321)			
Observations	1883	2771				1765	2410			

Source: Own calculations based on 1995 HBS and 2002 CWIQ data. *Notes*: The sample is all primary school-age children (i.e. $6 \leq Age \leq 12$). The treatment group is all primary school-age children aged 6 to 8 years old. All mean values are unweighted. Standard deviations are in parentheses. The reported p-values are for test of equality of means between the treatment and control groups (independent samples). Diff means difference in means by treatment status. Norm-diff means normalized differences between treatment and control means computed as $\triangle_{\mathbf{X}} = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{S_0^2 + S_1^2}}$, where S_G^2 is the sample variance of X_i in subsample with treatment $G_i = \{0, 1\}$. $G = 1$ if a child is aged between 6 and 8 years old.

where S_G^2 is the sample variance of X_i in sub-sample with treatment $G_i = \{0, 1\}$. $G = 1$ if a child is aged between 6 and 8 years. According to Imbens and Wooldridge (2009), if the normalized difference is less than one quarter in absolute values, the unconfoundedness assumption is likely satisfied, and hence estimating equation 2.4 by linear regression methods is justified. The normalized differences reported in columns (5) and (10) of Table 2.3 are less than this rule of thumb absolute value of 0.25 for all characteristics, except for child's age, This result suggests that control and treatment children are very similar in terms of their characteristics.

Given this similarity in characteristics between the treatment and control groups, it is plausible to consider the assignment of primary school-age children to control and treatment groups, based on the grade-appropriate cohort criteria, to be conditionally random (Foureaux Koppensteiner, 2014). This is further reassurance that the estimation results can only be biased downwards, as explained earlier, and not due to compositional changes.

2.6.3 Assessing the plausibility of assumptions under the broadly-defined treatment group

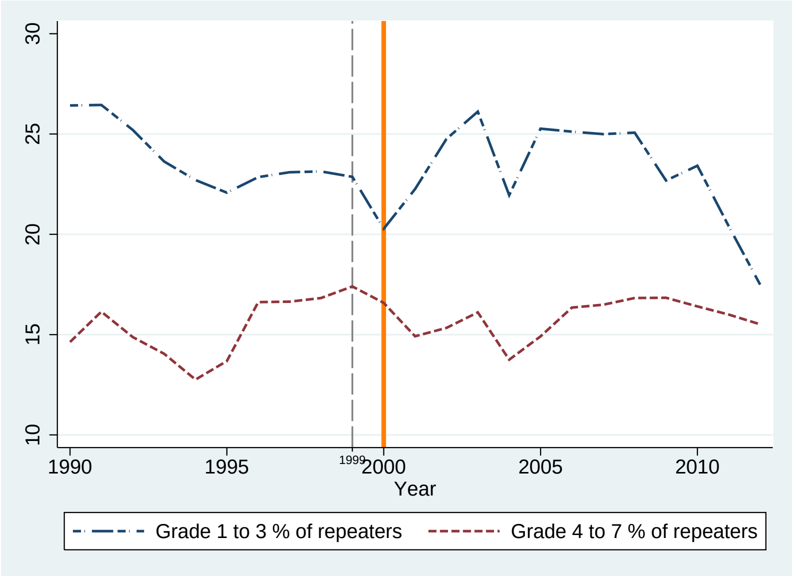
As pointed out earlier, the common trend assumption can not be tested directly. I do not have pre-reform data on enrolment by grade, which would be ideal to indirectly test the plausibility of this assumption. Fortunately, I have pre-reform national data on enrolment-related measures of schooling (i.e. the percentage of repeaters by grade, repetition rates, and drop-out rates) from UNESCO Institute of Statistics. Therefore, to assess the plausibility of the common trend assumption, I look at pre-reform trends in these measures.[16]

In Figure 2.3, I plot the average percentage of repeaters for the treatment group (i.e. grades 1 to 3 combined) and the control group (i.e. grades 4 to 7 combined) from 1990 to 2012. The long-dash vertical line in the figure marks the end of the fee-paying period, while the solid vertical line marks the start of the FPE programme. Looking at Figure 2.3, we can see that the graphs for the treated group and the control group are moving parallel to each other. Between 1990 and 1994, the proportion of repeaters in each grade-group (treatment and control) declined, then slightly increased between 1995 and 1996, and remained stable until 1999. The slightly bigger drop in the percentage of repeaters in grades one to three between 1999 and 2000 is due to the huge increase of new school entrants. The overall percentage of repeaters throughout the period remained at around 20 percent before and after FPE implementation (not shown in the figure).

A similar pattern prevails with respect to the trends in drop-out rates. Figure 2.4 presents

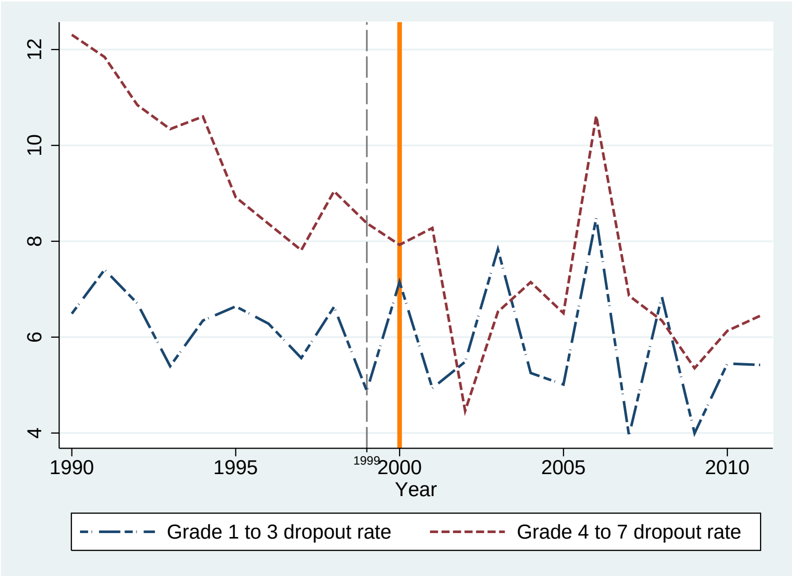
[16]The UNESCO Institute of Statistics data for each country comes from the annual education censuses. We do not report trends in repetition rates because they are the same as trends in the percentage of repeaters by grade.

Figure 2.3: Trends in the percentage of repeaters by treatment status: 1990-2012



Source: UNESCO Institute for Statistics (UIS), Data Centre, 21 August 2014. *Notes*: Percentage of repeaters is defined as the "total number of students who are enrolled in the same grade as the previous year, expressed as a percentage of total enrolment in the given grade of education" (UIS glossary).

Figure 2.4: Trends in drop-out rates by treatment status: 1990-2011



Source: UNESCO Institute for Statistics (UIS), Data Centre, 21 August 2014. *Notes*: Grade Drop-out rate is defined as "the proportion of pupils from a cohort enrolled in a given grade at a given school year who are no longer enrolled in the following year" (UIS glossary).

the trends in drop-out rates for each grade-group, that is, the treatment grade-group and the control grade-group. Similarly, the long-dash vertical line in the figure marks the end of the fee-paying period, while the solid line shows the beginning of the FPE programme. The figure indicates that, in 1991, the drop-out rate for the control group was about 12 percent, and 7 percent for the treatment group. In 1997, the drop-out rate dropped to about 8 percent for the control group, and 5.5 percent for the treatment group. It slightly increased in 1998 and then declined in 1999 for each group. Even though the gradient is slightly higher for the control group between 1990 and 2000, it appears similar for both groups between 1995 and 2000.

Both the percentage of repeaters (or the repetition rate) and the drop-out rate are closely related to enrolment in any given year. The fact that the control and treatment groups appear to have a pre-reform common trend in each of these measures gives one confidence that pre-reform enrolment rates for these groups also have a common trend. Thus, there should be little concern for self-selection of children into the treatment group by either differentially dropping out or repeating classes. This suggests that the control group, grade 4-to-7, is potentially a valid control group.

I now turn to the discussion of the compositional changes assumption. Table 2.4 shows the descriptive statistics of the control group (i.e. children in grades 4 to 7) and the treatment group (i.e. children in grades 1 to 3) in 1994/95 and 2002. Essentially, Table 2.4 presents similar information to that presented in Table 2.3, but for a broadly-defined treatment group.

I have highlighted earlier that, by design, the control group under the broadly-defined treatment group includes all children who have enrolled for sometime and excludes all children who have never enrolled. Therefore, the control group inevitably includes children from wealthier families, relative to those in the treatment group. This is because, before FPE, schools were still charging fees, and hence the group with at least grade 3 would have come from relatively wealthier households. Indeed, we can see from columns (1) and (2), and columns (6) and (7) of Table 2.4 that the wealth index for the control group is higher than that of the treatment group. However, the main focus is on whether these wealth differences between the control and treatment groups, for example, are stable over time.

Columns (4) and (9) show the mean differences between the control and treatment groups for each characteristic in 1995 and 2002. We can see from these columns that the differences in all the characteristics are small. As we have seen in Table 2.3, the largest difference between the control and treatment groups is in age of the child (about 2.5 years), which is still small, and should be expected given that children in lower grades are generally younger than those in higher grades. What is particularly worth noting from Table 2.4 is that the differences in all the characteristics are stable over time. For example, the difference between the control and treatment groups in sex composition in 1994/95 is 0.1056, and is very comparable to

Table 2.4: Mean characteristics of control and treatment groups in 1994/95 and 2002

Variable	1994/95					2002				
	Treatment (1)	Control (2)	p-value (3)	Diff $(X_1 - X_0)$ (4)	Norm-diff (5)	Treatment (6)	Control (7)	p-value (8)	Diff $(X_1 - X_0)$ (9)	Norm-diff (10)
Enrolment	0.5802	0.9657	0.0000	-0.3855	-0.7327	0.7933	0.9620	0.0000	-0.1687	-0.3766
	(0.4936)	(0.1821)				(0.4050)	(0.1913)			
Wealth Index	4.9429	6.2665	0.0000	-1.3237	-0.3590	4.7323	6.0442	0.0000	-1.3119	-0.3781
	(2.5725)	(2.6405)				(2.3926)	(2.5122)			
Child Age	8.3603	10.7367	0.0000	-2.3764	-1.0844	8.3349	10.8797	0.0000	-2.5448	-1.1543
	(1.8353)	(1.1977)				(1.8941)	(1.1280)			
Child Sex (Male)	0.5510	0.4454	0.0000	0.1056	0.1501	0.5322	0.4320	0.0000	0.1002	0.1424
	(0.4974)	(0.4972)				(0.4990)	(0.4956)			
Household Head's Educ	3.7235	5.2178	0.0000	-1.4943	-0.2727	3.9973	5.3532	0.0000	-1.3559	-0.2403
	(3.6435)	(4.0917)				(3.7245)	(4.2373)			
Household Head's Age	49.7471	49.2601	0.2592	0.4870	0.0259	49.1600	48.5417	0.2088	0.6183	0.0316
	(13.4415)	(13.1171)				(14.4192)	(13.2232)			
Adult males	1.5253	1.4552	0.0367	0.0701	0.0476	1.2369	1.2381	0.9706	-0.0012	-0.0009
	(1.0547)	(1.0274)				(0.9188)	(0.9436)			
Adult females	1.7728	1.7248	0.1391	0.0481	0.0340	1.5081	1.5039	0.8817	0.0042	0.0036
	(1.0367)	(0.9592)				(0.8021)	(0.8151)			
Household size	7.4412	7.1218	0.0003	0.3193	0.08389	6.4460	6.2139	0.0039	0.2321	0.0711
	(2.7790)	(2.6011)				(2.3320)	(2.2821)			
Observations	3042	1428				2980	1155			

Source: Own calculations based on 1995 HBS and 2002 CWIQ data. *Notes*: The sample is all primary school-age children (i.e. $6 \leq Age \leq 12$). The treatment group is all children with at most grade 2 (i.e. 'Grade<3'). All mean values are unweighted. Standard deviations are in parentheses. The reported p-values are for test of equality of means between the treatment and control groups (independent samples). Diff means difference in means by treatment status. Norm-diff means normalized differences between treatment and control means computed as $\triangle_{\mathbf{X}} = \frac{\bar{X}_1 - \bar{X}_0}{\sqrt{S_0^2 + S_1^2}}$, where S_G^2 is the sample variance of X_i in subsample with treatment $G_i = \{0, 1\}$. $G = 1$ if a child is aged between 6 and 8 years old.

0.1002 difference in 2002. Similarly, the difference in the wealth index between the control and treatment group has remained stable over time at about 1.32 points.

Furthermore, the groups appear to be balanced in all characteristics, except for wealth, household head's age and age of the child[17], as indicated by the small normalized differences in columns (5) and (10) of the table (i.e. they are smaller than 0.25 in absolute values). This evidence supports the assumption that the composition of groups, in terms of their unobservable characteristics, has remained unchanged pre- and post-FPE. This taken together with the common trends evidence is further evidence that the control group is potentially valid.

The main concern with the analysis under this broad-group definition is that, prior to FPE, the control group's enrolment rate was 96.6 percent (see column 2 of Table 2.4), with a very limited scope to increase over the period. When using the linear probability model, this forces one to attribute almost the entire increase in enrolment over time in the treatment group to FPE. To address this concern, I check the robustness of the results by running a non-linear probability model (i.e. the probit model).

2.7 Results

In this section, I provide empirical evidence on the effect of FPE on primary school enrolment in Lesotho. I first present the results when the treatment group is narrowly-defined, which gives the lower bound effect of FPE on enrolment, and then results when the treatment group is broadly-defined.

2.7.1 Lower bound Effect of FPE on Enrolment

This sub-section presents the lower bound effect of FPE on enrolment. Table 2.5 offers results on the effect of FPE on access to primary schooling by children aged 6 to 8 years old. It shows that, post-FPE, access to primary schooling increased, and this result is robust to different specifications. The results are largely in line with the descriptive evidence seen in the previous section (in Table 2.2).

Column 1 of the table presents FPE effect results with no controls. In columns 2 to 4, I, respectively, control for wealth, child gender, and household head's education, including their interactions with group, time and FPE dummies. Column 5 presents the preferred results where I control for all other individual and household characteristics thought to be important to child schooling.

[17]This indicates potential violations of the linearity assumption. I check the robustness of the results by running a probit model.

Table 2.5: The Lower-bound Effect of FPE on Primary School Enrolment

VARIABLES	(1)	(2)	(3)	(4)	(5)
FPE (D=1)	0.0709***	0.0497	0.0894***	0.0704**	0.0934**
	(0.0182)	(0.0427)	(0.0245)	(0.0283)	(0.0441)
Year (T=2002)	0.1071***	0.2439***	0.0689***	0.1573***	0.1918***
	(0.0098)	(0.0250)	(0.0120)	(0.0160)	(0.0250)
Eligible (G=1)	-0.2214***	-0.2941***	-0.2437***	-0.2673***	-0.0662*
	(0.0137)	(0.0315)	(0.0186)	(0.0208)	(0.0381)
Wealth		0.0434***			0.0370***
		(0.0028)			(0.0036)
Wealth×Eligible		0.0124***			0.0081
		(0.0048)			(0.0060)
Wealth×Year		-0.0243***			-0.0213***
		(0.0037)			(0.0045)
Wealth×FPE		0.0063			0.0040
		(0.0065)			(0.0079)
Male			-0.1268***		-0.1254***
			(0.0150)		(0.0149)
Male×Eligible			0.0414		0.0362
			(0.0273)		(0.0262)
Male×Year			0.0720***		0.0699***
			(0.0193)		(0.0191)
Male×FPE			-0.0334		-0.0332
			(0.0362)		(0.0348)
HHEduc				0.0214***	0.0063***
				(0.0017)	(0.0023)
HHEduc×Year				-0.0122***	-0.0029
				(0.0023)	(0.0029)
HHEduc×Eligible				0.0096***	0.0058
				(0.0031)	(0.0040)
HHEduc×FPE				0.0008	-0.0009
				(0.0042)	(0.0052)
Constant	0.7964***	0.5620***	0.8621***	0.7064***	-1.2882***
	(0.0077)	(0.0193)	(0.0095)	(0.0123)	(0.1388)
Observations	8,780	8,541	8,780	8,779	8,397
R-squared	0.075	0.139	0.086	0.114	0.187

Notes: All regressions are unweighted. Robust standard errors are in parentheses. *significance at 10%; **significance at 5%; ***significance at 1%. Enrolment is the dependent variable and equals 1 if a 6-12 year old is enrolled in primary school. Eligible is a group indicator equals 1 if a child is in the treatment group (i.e. if aged 6 to 8 years). FPE is the treatment indicator, the interaction of Eligible and Year dummies. The regression in column (5) controls for Age, Age squared, Household head's age, Number of adult males and females in the household, and Household size. HHEduc is Household Head's Education.

Across all specifications, except for column 2, the FPE coefficient is statistically significant and ranges between 0.070 and 0.093. Similarly, the coefficient on the year dummy is positive and statistically significant throughout, indicating a positive time effect on enrolment. We can also see that enrolment of the treatment group (i.e. the 6 to 8 year olds) is significantly lower than that of the control group (i.e. the 9 to 12 year olds). Furthermore, we can see in column 2 of the table that a unit increase in the wealth index increases enrolment by 4.3 percentage points. This positive wealth influence is also supported by the positive coefficient on household head's education in column 4 of the table. The enrolment rate of boys is revealed to be significantly lower than that of girls in column 3.

The point estimates reported in columns 1 to 4 are potentially biased in ways that we cannot tell a priori. This is because, as mentioned in section 2.6, the assignment of primary school-age children to control and treatment groups is random, conditional on their characteristics. Controlling for all child's characteristics related to treatment status and enrolment helps to net-out this potential bias. This leads us to the preferred results in column 5 of the table, where all observed individual and household characteristics have been controlled for. The estimates in column 5 reveal that FPE raised enrolment of the 6 to 8 year olds (the treated group) by about 9.3 percentage points relative to enrolment of the 9 to 12 year olds. This translates to about 13.2 percent (i.e. $(9.3/70.7) \times 100$) increase in enrolment attributable to FPE.

Consistent with the descriptive evidence in Table 2.2 and Figure 2.2, results show a significant positive wealth gradient. The coefficients on the wealth index, gender, and household head's education have remained statistically significant and relatively stable, even after controlling for all other variables. We can see in column 5 that a unit increase in the wealth index increases primary-school enrolment by about 4 percentage points, and children of educated parents are more likely to enrol relative to their counterparts with uneducated parents. There is also clear evidence of the gender gap in favour of girls. Boys' enrolment is about 13 percentage points lower than that of girls.

Furthermore, the results in Table 2.5 indicate that FPE has not benefited children from poor backgrounds more than those from rich families. For instance, we can see from column 5 of the table that the coefficient of the wealth-FPE interaction term is positive, generally implying that school enrolment by the rich, not the poor, increased after FPE. However, this coefficient is statistically insignificant, potentially due to the fact that the estimates here are lower bound estimates given the restrictive nature of the treatment group. Similarly, the coefficient for the interaction term Male–FPE is negative, contrary to what we saw in Table 2.2, but it is insignificant. This could be due to the fact that many boys and poor children who have benefited from the policy are more likely to be older than 8 years given that these groups (i.e. boys and the poor) are the one most likely to delay entry into the schooling system. I elaborate on this point further in the next section.

2.7.2 Upper bound Effect of FPE on Enrolment

Table 2.6 presents results of the effect of FPE on enrolment by primary school-age children using the broadly defined treatment group (i.e. all primary school-age children who had completed at most grade 2 in 2002). Column 5 of the table shows the upper bound FPE effect, having controlled for all observed individual and household characteristics.

By and large, the qualitative conclusions remain the same, in spite of how the treatment group is defined. The most important result from Table 2.6 is that enrolment of the primary school-age population increased by about 28.8 percentage points due to the implementation of the FPE programme[18]. Given the base enrolment rate of 70.7 percent in 1995, this translates to about 41 percent increase in enrolment due to FPE. Even though this estimated FPE effect is true for the sample at hand (because of not weighting), it is very much higher than the 33 percent increase in primary school net enrolment rate between 1999 and 2000, which is calculated using the national enrolment statistics[19], and much less than the widely cited 75 percent surge in enrolment of new entrants during the same period.

In addition, as mentioned earlier, the bias on the effect parameter under the broad definition of the treatment group could go either way: it could be positive if children intentionally repeated lower grades, or negative if those in higher grades stayed out of school waiting for FPE. Given that those who continued schooling beyond grade 3 before FPE must have come from relatively well-off households with high tastes for education, it is reasonable to believe that the majority of them could have, at worst, repeated lower grades than stay out of school. This makes it more probable that the effect could be biased upwards, and, hence, strongly suggests that this result is the upper bound causal effect of FPE that can be generalised to the population.

I speculated in Section 2.7.1 that using the narrowly-defined treatment group could explain the insignificance of the FPE effect on enrolment of children from poor households seen in Table 2.5. The results in Table 2.6 confirm these speculations. The FPE programme has clearly narrowed gender- and wealth-related schooling inequalities by significantly benefiting boys and children from poor backgrounds. This is consistent with the descriptive evidence. For instance, the negative coefficient of the Wealth-FPE interaction term implies that, post-FPE introduction, a percentage fall in household wealth leads to a further 2 percentage points increase in the FPE enrolment effect to 30.8 percentage points. That is, FPE increased enrolment of children at the lower tail of the wealth distribution the most.

[18] This result is robust to possible violations of the linearity assumption. Using the probit model, this effect is estimated at 37.6 percentage points, and is statistically significant at the 1 percent level. See Table A.2.

[19] Data available at http://data.uis.unesco.org/. UNESCO Institute of Statistics.

Table 2.6: The Upper-bound Effect of FPE on Primary School Enrolment

VARIABLES	(1)	(2)	(3)	(4)	(5)
FPE (D=1)	0.2168***	0.3273***	0.1745***	0.2526***	0.2877***
	(0.0138)	(0.0333)	(0.0189)	(0.0218)	(0.0345)
Year (T=2002)	-0.0037	-0.0035	0.0035	-0.0013	0.0037
	(0.0075)	(0.0209)	(0.0090)	(0.0132)	(0.0206)
Eligible (G=1)	-0.3855***	-0.6036***	-0.3279***	-0.4698***	-0.4922***
	(0.0102)	(0.0244)	(0.0144)	(0.0155)	(0.0262)
Wealth		0.0042**			0.0024
		(0.0020)			(0.0026)
Wealth×Eligible		0.0455***			0.0430***
		(0.0039)			(0.0048)
Wealth×Year		0.0003			-0.0004
		(0.0028)			(0.0036)
Wealth×FPE		-0.0215***			-0.0196***
		(0.0053)			(0.0063)
Male			-0.0062		-0.0087
			(0.0098)		(0.0106)
Male×Eligible			-0.1034***		-0.1123***
			(0.0203)		(0.0203)
Male×Year			-0.0168		-0.0120
			(0.0154)		(0.0162)
Male×FPE			0.0789***		0.0742***
			(0.0278)		(0.0276)
HHEduc				0.0027**	0.0010
				(0.0012)	(0.0016)
HHEduc×Year				-0.0005	-0.0008
				(0.0017)	(0.0022)
HHEduc×Eligible				0.0238***	0.0112***
				(0.0026)	(0.0032)
HHEduc×FPE				-0.0109***	-0.0030
				(0.0034)	(0.0042)
Constant	0.9657***	0.9385***	0.9684***	0.9513***	-0.7269***
	(0.0048)	(0.0150)	(0.0062)	(0.0089)	(0.1043)
Observations	8,581	8,353	8,581	8,581	8,215
R-squared	0.136	0.176	0.143	0.161	0.231

Notes: All regressions are unweighted. Robust standard errors are in parentheses. *significance at 10%; **significance at 5%; ***significance at 1%. Enrolment is the dependent variable and equals 1 if a 6-12 year old is enrolled in primary school. Eligible is a group indicator equals 1 if a child is in the treatment group (i.e. if a primary school age child has completed at most grade 2). FPE is the treatment indicator, the interaction of Eligible and Year dummies. The regression in column (5) controls for Age, Age squared, Household head's age, Number of adult males and females in the household, and Household size. HHEduc is Household Head's Education.

Additionally, the positive coefficient of the male-FPE interaction term indicates that FPE increased school enrolment of boys by as much as 36.2 percentage points relative to that of girls.

On the whole, these results are consistent with the evidence provided by Deininger (2003) in Uganda, and Lucas and Mbiti (2012) in Kenya who found that FPE increased access to schooling. What is more, they show that in an environment where there is little scope for schooling demand shifts from public schools (i.e. those covered by the FPE policy) to private schools[20]. This suggests that FPE does indeed significantly raise access to schooling. Furthermore, in line with the evidence by Deininger (2003) in Uganda and Al-Samarrai and Zaman (2007) in Malawi, this chapter's results show that the abolition of user fees mostly benefited the poor, and bridged the gender-related inequalities in schooling, thereby enhancing equity in access.

2.8 Conclusion

In this chapter, I examine the effect of FPE programme on enrolment in Lesotho. I find that this policy significantly increased the enrolment of the primary school-age population. Using the narrow definition of the treatment group, I derive a lower-bound FPE effect of 9.3 percentage points (or about 13.2 percent) increase in enrolment of the 6 to 8 year olds. The effect increases to 28.8 percentage points (or about 40.7 percent) increase in enrolment for all of the primary school-age population when a broadly defined treatment group is used. What is more, the results also show that FPE significantly increased enrolment of boys (the historically disadvantaged group in Lesotho) relative to girls and of children from poor backgrounds.

The main concern with this analysis is that the broadly defined control group's enrolment rate was already high (i.e. close to 100 percent) even before FPE, thereby making it difficult to isolate the FPE effect from the time trend. Even though I have checked the robustness of the results against this potential pitfall, it may still be considered a limitation. Another potential limitation is that, under the broadly-defined treatment group, the control group is a select group of children who have enrolled in school for at least three years and are potentially from wealthier families. I have attempted to address this problem by controlling for household wealth and household head's education level.

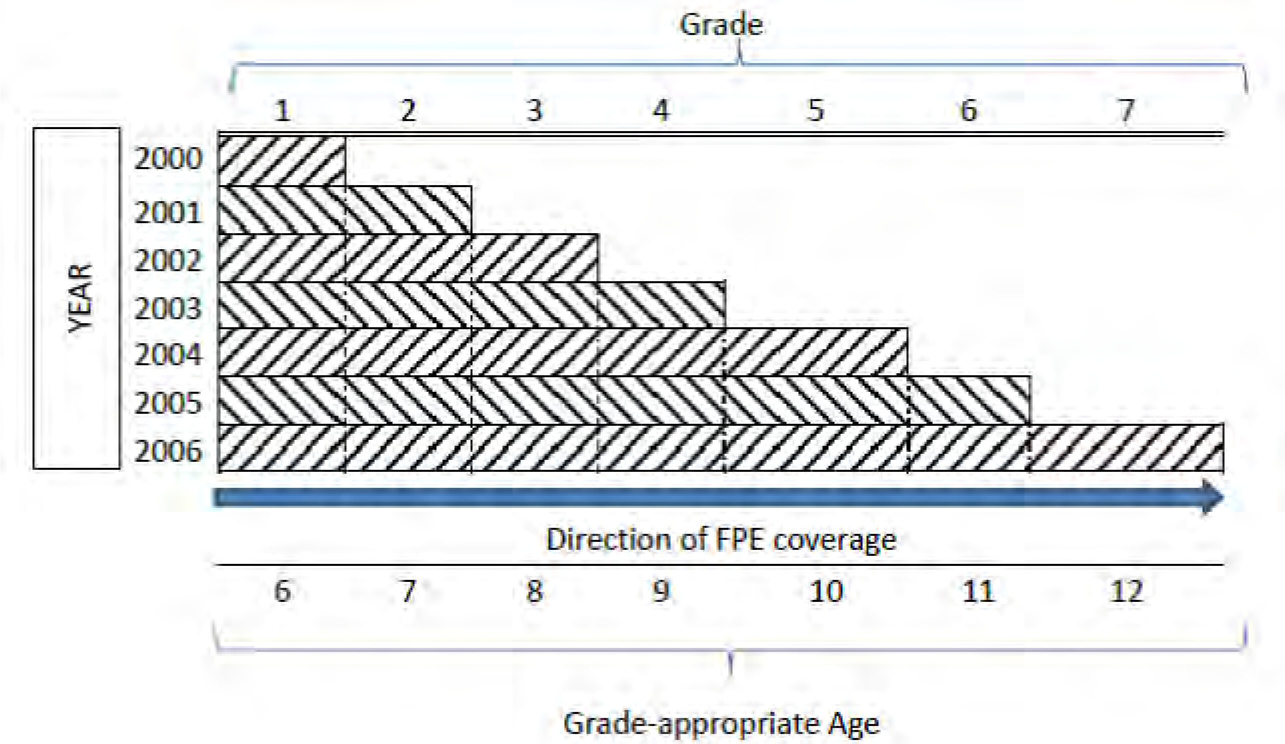
By and large, however, the results in this chapter highlight that FPE in Lesotho has largely achieved its objective of increasing enrolment and redressing inequalities in schooling access. It is therefore possible that, if instituted in those sub-Saharan African countries still

[20] In Lesotho, only one percent of all primary (non-religious) schools are private and FPE covers almost all schools.

lagging behind in education indicators, the 'Education for All' goal can be achieved. What remains unknown is the effect of this policy on educational quality. Therefore, future research should focus on this margin if we are to understand the full impact of FPE on education. In the next chapter, I make a contribution towards this margin by looking at how education quality has changed since the introduction of this policy.

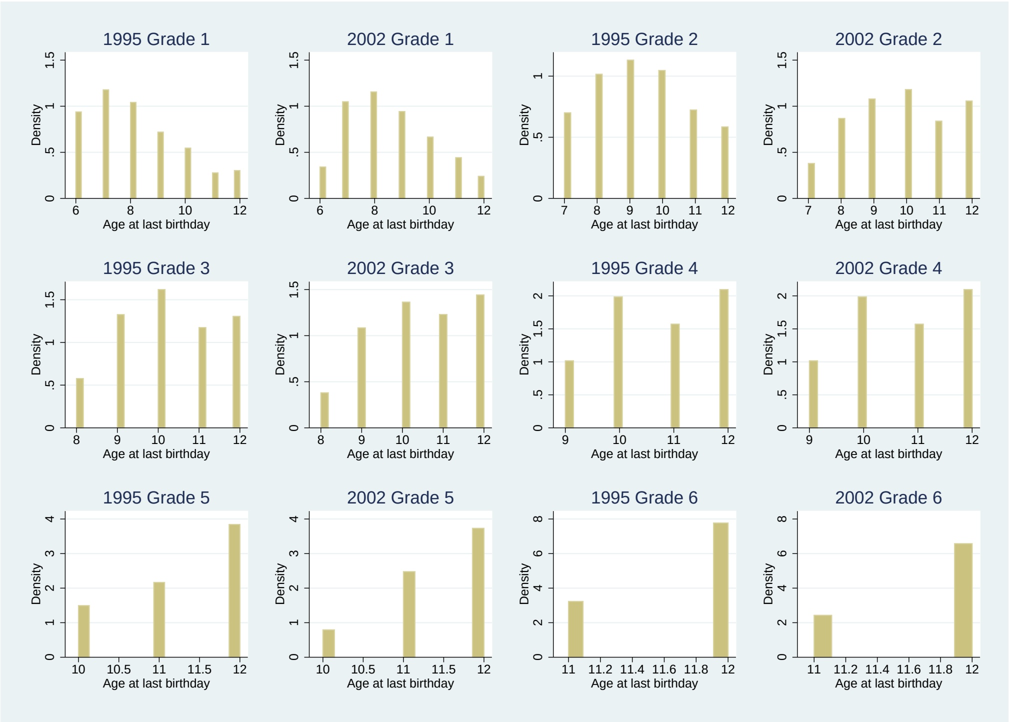
Appendix

Figure A.1: Phasing-in of the FPE programme



Source: Own representation. *Notes*: On the left-hand side (i.e. y-axis) of the figure we have years from 2000 to 2006, the top horizontal axis shows grades 1 to 7, the bottom horizontal axis (i.e. x-axis) shows grade-appropriate ages 6 to 12 years. The shaded regions of the figure indicate the grade(s) and grade-appropriate age(s) covered by FPE in each year, starting from 2000.

Figure A.2: Age distribution by highest grade completed



Source: Own calculations based on 1995 HBS and 2002 CWIQ data.

Table A.1: Primary School ownership in Lesotho

Proprietor	Share of ownership (%)		
	2000	2007	2011 (census)
African Methodist Episcopal Church	1.87	0.35	2
Anglican Church of Lesotho (A.C.L)	13.19	8.44	12
Lesotho Evangelical Church (Protestants)	38.86	38.66	33
Roman Catholic Church (R.C.C)	41.27	32	34
Community	1.39	5.31	3
Government	0.6	8.8	11
Private	0	0.59	1
Other Churches	2.82	5.85	4
Total	100	100	100

Source: Own calculations from the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ), SACMEQ II of 2000 and SACMEQ III of 2007, and MOET (2011).

Table A.2: The Effect of FPE on Primary School Enrolment

VARIABLES	(1) Enrolment	(2) Marginal Effects (dy/dx)
FPE (D=1)	2.2246***	0.3759***
	(0.2744)	(0.0459)
Year (T=2002)	0.0666	0.0113
	(0.2322)	(0.0392)
Eligible	-1.7681***	-0.2988***
	(0.1825)	(0.0303)
Wealth	0.0390	0.0066
	(0.0325)	(0.0055)
Wealth×Eligible	0.0940***	0.0159***
	(0.0349)	(0.0059)
Wealth×Year	0.0016	0.0003
	(0.0445)	(0.0075)
Wealth×FPE	-0.1152**	-0.0195**
	(0.0529)	(0.0089)
Male	-0.1747	-0.0295
	(0.1328)	(0.0224)
Male×Eligible	-0.1643	-0.0278
	(0.1432)	(0.0242)
Male×Year	-0.1077	-0.0182
	(0.1948)	(0.0329)
Male×FPE	0.3235	0.0547
	(0.2262)	(0.0382)
HHEduc	0.0230	0.0039
	(0.0214)	(0.0036)
HHEduc×Year	-0.0079	-0.0013
	(0.0300)	(0.0051)
HHEduc×Eligible	0.0148	0.0025
	(0.0230)	(0.0039)
HHEduc×FPE	-0.0262	-0.0044
	(0.0355)	(0.0060)
Constant	-4.5782***	
	(0.5057)	
Observations	7,386	7,386

Notes: Probit regression results. Other controls are Age, Age squared, Household head's age, Number of adult males and females in the household, and Household size. HHEduc is Household Head's Education. Enrolment is the dependent variable and equals 1 if a 6-12 year old is enrolled in primary school. Eligible is a group indicator equals 1 if a child is in the treatment group (i.e. if a primary school age child has completed at most grade 2). FPE is the treatment indicator, the interaction of Eligible and Year dummies. Robust standard errors are in parentheses. *significance at 10%; **significance at 5%; ***significance at 1%.

Chapter 3

Educational Achievement and Educational Inequality in Lesotho: Changes and Determinants

3.1 Introduction

Education quality (or achievement)[1] is low in much of the developing world, and is especially so in Sub-Saharan Africa (SSA). A common presumption in the current literature is that the dire state of educational quality in SSA has been exacerbated by the recent push to meet the universal primary education Millennium Development Goal (MDG) by 2015 (see, for example, Zuze and Leibbrandt, 2011; Colclough et al., 2008). Even though the FPE does increase the educational quality of those children who would have otherwise not attended school, it may reduce the educational quality of those who were already schooling. Therefore, the average effect of the FPE may be negative if the latter effect dominates the former effect.

Theoretically, this access-quality hypothesis derives from Cunha et al. (2006)'s theory of life cycle skill formation which predicts that exposing children to little and/or low quality education early in their lives will negatively affect their future skill acquisition. According to this theory, when young children are in a less stimulating environment, for example, staying with parents who are less involved in their education and/or being in overcrowded classrooms with little or no pupil-teacher contact, they are likely to acquire less education from schooling[2] which will be reflected by their poor performance in achievement tests. This hypothesis is supported by empirical evidence from some SSA countries, for example, Kenya, Malawi, Mozambique and Uganda, which saw a decline in education quality after

[1]Educational achievement and educational quality are used interchangeably throughout this chapter.

[2]Schooling is different from education. Schooling is an input into the education production function (Behrman, 2010).

the introduction of the free primary education (FPE) policy (Lee and Zuze, 2011; Lucas and Mbiti, 2012).

However, the FPE policy is not always associated with a decline in education quality. For example, Lesotho's average reading and maths performance on SACMEQ tests increased[3] by 4% and 7%, respectively, between 2000 (when FPE was first introduced) and 2007. Taylor and Spaull (2013) also argue that a decline in a country's average score does not always reflect a falling education system. The authors show, using their "effective enrolment" measure (that is, the proportion of an age-specific population achieving particular skill levels), that increased education access in Southern and Eastern Africa was accompanied by an increase in education quality.

Using changes in average test scores to infer changes in educational quality over time is common in the education literature (Carnoy and Rothstein, 2013; Jopo et al., 2011b). However, the use of this traditional summary statistic (that is, the mean) leaves much of the information inherent in a test scores' distribution untapped and thus fails to reveal much of education quality dynamics. It is unclear, therefore, whether the increase in average achievement in some countries, for example, is due to improved student performance levels at the upper and/or lower tails of the distribution. On the one hand, if the increase in average achievement is due to an increase in the performance of high-ability students, then we are faced with a scenario of improved average education quality and a rise in education inequality. If, on the other hand, the increase in average scores is a consequence of increased performance of low-ability students, then we have a case of a reduction in education inequality and improved education quality. Knowing which scenario prevails is essential for education policy in order to ensure that no child is left behind, not only outside the schooling system but also within the schooling system.

In fact, the Dakar Framework for Action by the World Education Forum does not only call for quality FPE, but also the redress of educational inequalities (UNESCO, 2000). This is because there is consensus in the literature that inequalities in educational achievement have negative implications for other socially important dimensions, such as health and earnings inequalities (Becker and Tomes, 1986b; Ferreira and Gignoux, 2014). Nonetheless, there is little evidence on the extent, changes and determinants, of educational inequalities in many developing countries, including Lesotho. Moreover, discussions on changes in educational quality hardly ever mention educational inequalities. As Rodrigues et al. (2013) argue, an increase in educational achievement should be interpreted as a positive change in education quality only when all, and not just the high-ability, students have improved performance levels.

This chapter seeks to fill this gap by going beyond the averages in analysing changes in

[3] Based on our calculations from the SACMEQ data described in the data section below.

educational achievement in Lesotho. Lesotho is an appropriate setting on which to study these changes in educational achievement and educational inequality for two reasons. Firstly, following the implementation of the FPE programme in the year 2000, the enrolment of new school entrants surged by 75% between 1999 and 2000, resulting in high pupil teacher ratios, shortages in physical infrastructure, and an increased use of unqualified teachers, especially in rural schools (Lekhetho, 2013; Urwick, 2011). But between 2000 and 2007, the average achievement of grade 6 students increased, contrary to what would be expected and what other SACMEQ countries with the same policy experienced. It is interesting to know whether this increase in performance was driven by all or just a few gifted students, and also the resulting implications of such an increase on educational inequality. Secondly, income inequality in Lesotho is among the highest in SSA, with an estimated Gini coefficient of 0.52 (see Bureau of Statistics, 2006). The fact that part of income inequality is explained by inequalities in educational achievement (see for example, Alves, 2012) points to the value of also understanding the determinants of educational inequality.

Using the relative distribution method of Handcock and Morris (1998, 1999), I find that the increase in educational achievement between 2000 and 2007 is driven by the improved performance of students across the achievement quantiles, but mostly due to that of low-ability students (that is, those below the median) and high-ability students (that is, those above the 80^{th} quantile). For students between the median and the 80^{th}, their performance appears to have remained constant. That is, although the inequality in educational achievement fell below the median, it increased above the median largely due to the increase in performance of those at the very top of the achievement distribution. Therefore, this led to an increase in both education quality and educational quality polarization (or educational inequality) over this period. Specifically, the results show that the increase in maths test scores distribution is largely driven by the *location* (mean and median) shift, and that in reading scores distribution is significantly driven by both location and *shape* (variance) changes.

These changes in educational achievement and educational inequality could be due to changes in school resources and school characteristics, which are most likely to have been affected by the implementation of the FPE policy. Glewwe et al. (2011) review the literature that estimates the effects of school resources on students' achievement. From a list of studies they review, they find that there is little evidence that school resources have a positive impact on students' achievement. For example, Eide and Showalter (1998) find that the impact of school inputs on achievement is nonlinear. They show that per pupil expenditures increase maths test scores of low ability students and not that of an average student, and that pupil-teacher ratio has no impact on achievement. Leuven et al. (2008) also find that class size has no effect on student performance. However, Case and Deaton (1999) find that school resources (proxied by district average pupil-teacher ratio) in South Africa have a positive effect on student achievement. Therefore, the results on the effects of schools resources on achievement are, at best, mixed.

Consequently, I employ the unconditional quantile regressions to uncover determinants of these distributional changes (Firpo et al., 2009). I find that much of the increase in educational attainment is unexplained. However, two important educational policy variables, pupil-teacher ratio and teacher effort, have a significant influence on the increase in reading test scores. For example, a fall in pupil-teacher ratio and an increase in teacher effort, respectively lead to a significant increase educational achievement. On the other hand, the increase in maths performance was partly driven by pupil socio-economic background, grade repetition, age, school wealth and school social capital (that is, whether the surrounding community contributes to school activities such as school building and maintenance). Thus, the results seem to support the earlier evidence that automatic grade promotion could be harmful to educational quality (Foureaux Koppensteiner, 2014).

The results further indicate that the increase in educational inequality (i.e. the increase in maths and reading variances) is strongly associated with a rise in grade repetition, and a fall in teacher effort and school social capital. However, as is the case with changes in educational achievement, much of the increase in educational inequality is left unexplained. I, therefore, conjecture that there could be some family background factors (for instance, parental sex preferences, a child's order of birth, etc) and intangible school quality factors, such as changes in management and leadership practices, that are associated with changes in educational quality and inequality in Lesotho.

The chapter proceeds as follows. The next section describes the data while Section 3 discusses the empirical strategies. Section 4 presents and analyses the findings. Section 5 then concludes the chapter.

3.2 Data and Descriptive Statistics

3.2.1 Data

I use data from SACMEQ, a consortium of education ministries from fourteen Anglophone African countries[4], policy makers, researchers, and the International Institute for Educational Planning (IIEP), with the aim of improving research capacity and technical skills of educational planners to monitor and evaluate educational quality (Ross et al., 2005; Murimba, 2005). Thus far, SACMEQ has successfully carried out three survey projects: SACMEQ I conducted in 1995-1998; SACMEQ II in 1998-2000; and SACMEQ III in 2005-2007. Lesotho has only participated in the last two surveys.

[4]These are Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe.

The target population for these surveys is all grade (or standard) six students rather than a specific age group as is the case with, for instance, PISA (Program for International Student Assessment). They use a two-stage sampling method to collect information. Firstly, schools are stratified by region (or district), and within each region, schools are sampled using the probability proportional to size lottery method. The size of each school is determined from the previous year's administrative records. Secondly, a simple random sample of grade six students within each selected school is chosen (Ross et al., 2005). SACMEQ II data is from 177 primary schools and 3155 students, while SACMEQ III data comes from 182 primary schools and 4240 students. These surveys are nationally (but not regionally) representative.

These data have rich background information on students', teachers' and schools' characteristics, and standardised (English) reading and mathematics test scores for students and teachers. To get these standardised test scores, reading (maths) test questions were classified into eight levels of increasing difficulty, from pre-reading (pre-numeracy) to critical reading (abstract problem solving). Exploiting this test structure, the raw test scores were then transformed using Rasch Item Response Theory (IRT) to reflect each student's and teacher's competency level[5]. These scores were further transformed to have an international mean of 500 and a standard deviation of 100. I use these transformed test scores as the measure of educational achievement.

I use students' family background information from the student questionnaire to construct the pupil socio-economic status (SES) index. First, I selected the relevant variables (e.g. parental education) based on Shanks and Robinson (2013) 'framework for understanding child and educational outcomes'. Second, I pooled data from both years, and then employed the principal components analysis (PCA) method on a set of variables common to both years to get the index.[6] Pooling the data, and using common items in both years to construct the index ensures comparability of the index over time.

Parental education is a categorical dummy of the highest level of education completed: it is coded into seven categories from 'no education' to 'completed university education'. In 2000 and 2007, respectively, I have 39% and 20% missing father education, and 19% and 11% missing mother education[7]. Therefore, I use this detailed categorisation of parental education to convert it into completed years of education, assuming no grade repetition. The main motive is to impute the missing values. I then apply the predictive mean matching method to impute these missing values based on all other students' covariates. Doing so,

[5]The IRT assumes that the distribution of latent ability is normally distributed and estimates the probability that an individual gives the correct response to an item (or question) conditional on his/her cognitive ability and item difficulty using a logit model. Combining these probabilities with the observed raw test scores distribution helps to back out the distribution of cognitive ability (Ferreira and Gignoux, 2014).

[6]A complete list of these covariates used to construct the SES index is shown in Table A.1. I use the ordinary (or centered) PCA method to construct the index given that not all the variables are assets.

[7]All children without parents and/or knowledge of their parents' education did not report parental education.

however, does not affect the distribution of the SES index (see Figure A.1), but helps to keep all the students in the data.

Table 3.1 gives a complete list of the covariates, including the SES index, and their summary statistics. There 3155 and 4240 students in 2000 and 2007, respectively. But there are about 11 students and one student with missing maths scores in 2000 and 2007, in that order. Between 2000 and 2007, the average class size slightly increased from 45 pupils to 46 pupils, but the average Pupil-Teacher Ratio (PTR) dropped from 54 pupils per teacher to 42 pupils per teacher. This is reflective of the increased number of teachers and schools built under the free primary education (FPE) programme over the same period.

The average number of classrooms remained almost constant at about 1.5 ($\approx$ 2 classrooms), while teacher years of experience dropped by about one quarter, which indicates that there was an increase in the number of new and less experienced teachers. Nonetheless, the average level of teacher qualification increased, which suggests that the new teachers were relatively more educated.

Additionally, the number of students with textbooks increased between 2000 and 2007. The percentage of students with own maths textbooks increased by 10 percentage points from 46 percent to 56 percent, while that of students with reading textbooks increased by about a percentage point from 55 percent to 56 percent.

Table 3.1: Summary Statistics

Variable	2000			2007			
	Mean	Std.Dev	N	Mean	Std.Dev	N	$\lvert t\rvert$-value for mean equality
Pupil, and subject-specific characteristics							
Reading Score	451.2	57.94	3155	467	70.21	4240	7.754
Maths score	447.2	60.36	3144	476.9	67.27	4239	17.903
Gender (female)	0.556	0.497	3146	0.546	0.498	4240	0.2
Pupil age in months	169.63	22.15	3155	168.02	21.23	4240	3.147
Own reading textbook	0.553	0.497	3155	0.559	0.497	4240	1.572
Own maths textbook	0.456	0.498	3155	0.563	0.496	4240	10.445
Share reading textbook	0.326	0.469	3155	0.369	0.483	4240	2.708
Share maths textbook	0.431	0.495	3155	0.362	0.481	4240	6.987
Once repeated a class	0.608	0.488	3155	0.517	0.500	4240	8.381
SES index	-0.350	0.034	3155	-0.098	0.035	4240	5.156
Speak English at home	0.707	0.455	3155	0.762	0.426	4240	3.846
Pupil Problem Index	-0.146	1.545	3155	0.109	2.503	4240	5.390
School characteristics							

Continued on next page

Table 3.1: *Continued from previous page*

Variable	2000			2007			
	Mean	Std.Dev	N	Mean	Std.Dev	N	$\|t\|$-value for mean equality
Class size	44.90	18.09	177	46.16	22.28	174	0.742
Pupil-teacher ratio	53.85	177	18.49	41.80	16.39	174	23.792
Grade enrolment	70.84	54.55	177	70.19	52.91	174	1.204
School enrollment	616.5	384.0	177	493.5	325.1	174	14.745
Number of grade 6 class rooms	1.531	0.963	177	1.516	0.974	174	0.456
School location (Urban)	0.351	0.477	177	0.336	0.472	174	2.467
School Building Condition (poor)	0.66	0.008	177	0.512	0.008	174	13.419
School asset index	6.31	0.039	177	5.58	0.032	174	14.544
School days lost per year	4.66	5.451	177	0.84	3.040	174	35.474
Sch. distance from services	28.54	36.337	177	26.54	29.778	174	2.526
Teacher Problem Index	0.091	1.497	177	-0.068	1.899	174	4.029
Maths teacher characteristics							
Test score	739.4	70.67	177	738.8	68.59	174	1.675
Gender (female)	0.763	0.425	177	0.684	0.465	174	4.541
Age	40.96	9.136	177	38.04	9.595	174	11.945
Years of professional training	2.723	1.264	177	2.655	1.542	174	5.977
Years of experience	16.33	9.942	177	12.40	9.119	174	15.247
Teaching hours per week	23.10	6.839	177	19.71	8.954	174	17.475
Test frequency	1.393	0.775	177	1.391	0.703	174	5.736
Qualification (primary)	0.512	0.500	177	0.358	0.479	174	12.391
Qualification (junior secondary)	0.112	0.315	177	0.0319	0.179	174	13.351
Qualification (Senior secondary)	0.157	0.364	177	0.208	0.406	174	8.459
Qualification (A-level)	0.161	0.368	177	0.263	0.440	174	9.159
Qualification (Tertiary)	0.0582	0.234	177	0.139	0.346	174	9.325
Reading teacher characteristics							
Test score	722.0	60.19	177	721.3	57.77	174	17.903
Gender (female)	0.751	0.433	177	0.722	0.448	174	4.541
Age	41.09	9.186	177	39.32	10.46	174	6.251
Years of professional training	2.736	1.260	177	2.821	1.478	174	1.096
Years of experience	16.58	9.954	177	12.87	9.247	174	13.328
Teaching hours per week	22.99	6.847	177	19.11	8.929	174	20.001
Test frequency	1.575	0.688	177	1.408	0.703	174	13.168

Continued on next page

Table 3.1: *Continued from previous page*

Variable	2000			2007			
	Mean	Std.Dev	N	Mean	Std.Dev	N	$\|t\|$-value for mean equality
Qualification (primary)	0.509	0.500	177	0.330	0.470	174	14.158
Qualification (junior secondary)	0.122	0.327	177	0.0419	0.200	174	12.886
Qualification (Senior secondary)	0.157	0.364	177	0.208	0.406	174	5.75
Qualification (A-level)	0.161	0.368	177	0.263	0.440	174	12.553
Qualification (Tertiary)	0.0582	0.234	177	0.139	0.346	174	10.839

Source: SACMEQ II and SACMEQ III data. *Notes*: The school asset index is a count index provided with the data, and is calculated as the sum of assets such as school library, staff room, electricity, radio, TV, VCR, computer, etc, that the school owns. The SES index is the pupil's socio-economic status index calculated using the PCA. I use the ordinary PCA index because most household items used are not assets, and it is not clear whether having a wooden floor is better that a cement floor, for example. Pupil- and Teacher-Problem indices are composite indices made of pupil and teacher problem behaviours such as drug abuse, alcohol abuse, late school arrivals, unjustified absenteeism, bullying of pupils and/or teachers, etc.

3.2.2 Descriptive analysis of test scores distributions

It is essential to have a clear picture of educational achievement dynamics before diving into the more technical analysis. Table 3.2 gives a glimpse of the evolution of education quality using traditional summary measures; the mean, median, standard deviation, and the coefficient of variation. The first two columns show the summary statistics, respectively, for reading and maths scores in 2000, while the last two columns present the same information for 2007.

We can see from the table that there has been an upward shift in both the location (mean and median) and the shape (variance or coefficient of variation) of maths and reading test scores. On average, maths test scores increased by 30 points, while reading test scores increased by 17 points between 2000 and 2007. The extent of inequality is higher in the maths test scores' distribution with a coefficient of variation of 13.50 compared to 12.84 for reading scores in 2000. However, the increase in inequality is higher in reading than in maths scores. For instance, the reading scores coefficient of variation increased by 2.2 points, while that of maths scores increased by 0.5 points.

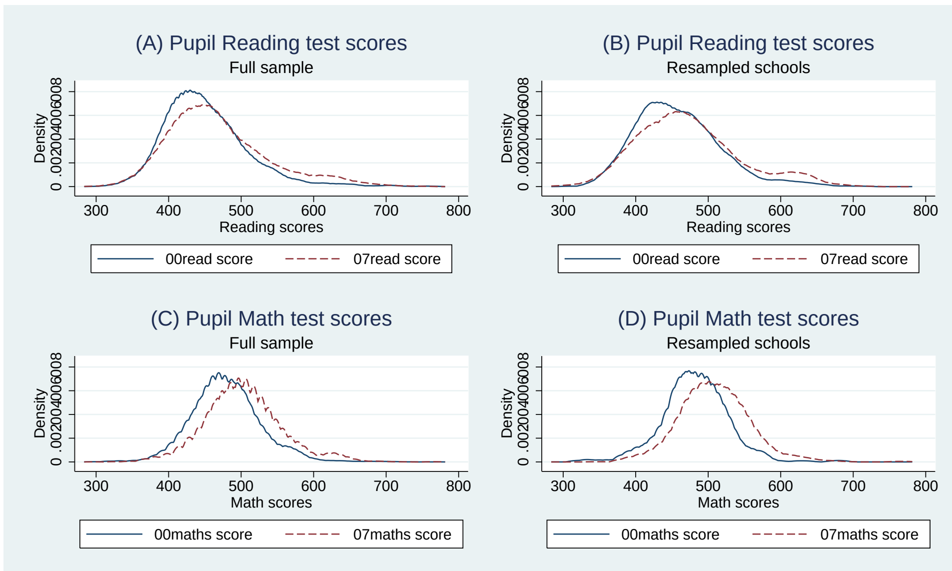
These changes are much clearer in Figure 3.1. This figure plots the distributions of reading and maths scores for the full sample of schools (panels A and C) and the resampled

Table 3.2: Grade 6 Reading and Maths Test Scores in Lesotho: 2000 - 2007

Summary measures	2000 Test scores		2007 Test scores	
	Reading	Maths	Reading	Maths
Mean	451.226	447.178	467.869	476.907
Median	445.318	440.489	455.524	468.619
Standard deviation (Sd)	57.936	60.363	70.211	67.273
Coefficient of Variation (CV)	12.84	13.50	15.01	14.11
Sample size (N)	3155	3144	4240	4239

Source: Own calculations from SACMEQ data.

Figure 3.1: Distribution of grade pupil test scores: 2000–2007



Source: Own calculations based on SACMEQ II and SACMEQ III. *Notes*: Kernel densities of test scores for a full sample of schools and resampled schools. Resampled schools is a panel of 42 schools included in both surveys.

schools (panels B and D)[8]. Panels A and B show the reading scores' distributions. They clearly show a positive shift in the location of the reading distribution with a change in shape. This, thus, shows that the increase in students' reading performance between 2000 and 2007 was accompanied by an increase in educational inequality. Panels C and D of Figure 3.1 show the distributions of maths scores. Looking at these panels (panels C and D), we can see the positive location shift in the distribution of maths scores, with no obvious change in distributional shape. Appendix Figure A.2 shows that the 2007 test scores first order stochastically dominate the 2000 test scores, especially in maths scores.

I now analyse these test scores' changes by socio-economic status (or wealth). Table 3.3a gives the changes in average reading achievement by wealth quintiles, while Table 3.3b provides similar information for maths achievement. Column 8 in each table shows the percentage change in average scores by quintile. It is evident from these tables that there

[8]The 'resampled' schools (42 in all) are a panel of schools that have been included in both surveys.

Table 3.3: Pupil test scores by SES quintiles

(a) Pupil Reading scores by SES quintiles

Quintiles of SES index	2000				2007					
	Mean	Sd.	C.V	N	Mean	Sd.	C.V	N	△mean [2007-2000] (%)	△CV [2007-2000]
1	440	47	10.68	716	447	55	12.30	1024	1.59	1.62
2	445	54	12.13	726	453	62	13.69	807	1.80	1.56
3	449	54	12.03	669	459	63	13.73	772	2.23	1.70
4	457	60	13.13	606	473	72	15.22	761	3.50	2.09
5	470	72	15.32	438	500	80	16.00	876	6.38	0.68
Overall	451	58	12.84	3155	468	70	15.01	4240	3.77	2.17

(b) Pupil Maths scores by SES quintiles

Quintiles of SES index	2000				2007					
	Mean	Sd.	C.V	N	Mean	Sd.	C.V	N	△mean [2007-2000] (%)	△CV [2007-2000]
1	442	58	13.12	712	460	62	13.48	1024	4.07	0.36
2	446	57	12.78	724	462	61	13.20	807	3.59	0.42
3	444	59	13.29	668	471	64	13.59	771	6.08	0.30
4	452	64	14.16	603	486	67	13.79	761	7.52	-0.37
5	453	64	14.12	437	501	71	14.17	876	10.60	0.05
Overall	447	60	13.50	3155	477	67	14.11	4239	6.71	0.61

Source: Own calculations based on SACMEQ II and SCMEQ III. *Notes*: C.V stands for Coefficient of Variation.

has been an increase in educational achievement across SES quintiles. But it seems that wealthy students are the ones driving much of the changes in reading and maths test scores. From both tables, the percentage change in achievement is increasing with student's wealth position. For example, while the reading performance of students in the first wealth quintile (the poorest) increased by 1.59%, that of students in the fifth quintile (the richest) increased by 6.38%.

Table 3.3a further shows that reading achievement is more unequal in 2007 compared to 2000, as revealed by the increased coefficient of variation over the period (see column 9 of the table)[9]. Much of this increase in reading scores' inequality is concentrated in the middle of the wealth distribution. For example, the change in the coefficient of variation at the third and fourth wealth quintiles is 1.70 and 2.09, respectively, compared to 1.62 at the first quintile and 0.68 at the fifth quintile.

Table 3.3b shows that maths performance of students at first quintile increased by a smaller percentage, 4.07%, compared to 10.60% increase in that of students at the fifth quintile. However, the increase in educational inequality in maths is largely concentrated at the second quintile, and it declined among students at the fourth quintile between 2000 and 2007. The gap in educational achievement between the first and the fifth quintile is highest in reading performance, but it increased the most in maths performance. The reading test score gap between the rich and the poor is 30 points and 53 points, in favour of the rich, in 2000 and 2007, respectively. On the other hand, the maths score gap between the rich and poor is about 11 points and 41 points in 2000 and 2007, respectively. Thus the wealth gap in performance increased by 30 points in maths, it increased by 23 points in reading.

These changes in performance by wealth suggest that pupils' wealth is associated with the change in performance over time. But what else could explain these changes in performance over time? In the next section, I explore changes in the distribution of policy variables that could potentially be associated with the observed changes in educational quality and educational inequality[10].

3.2.3 Changes in the distribution of educational policy variables

Table 3.4 presents the distributions of teacher test scores, student textbook ownership, and pupil-teacher ratio (PTR) by wealth quintiles, school proprietor, and location, in 2000 and 2007.

Starting with teacher-subject scores by students' wealth, we can see that there is a positive

[9]We use the coefficient of variation as a measure of educational inequality in comparing test scores variances for distributions with different means and standard deviations.

[10]We consider pupil teacher ratio, teacher education and textbook access to be educational policy variables because they are widely employed supply-side policies, and are often used in the literature and popular press as indicators of school quality.

Table 3.4: Distribution of educational policy variables

		Teacher test scores				*Textbook ownership*				*Pupil-Teacher Ratio (PTR)*	
		2000		2007		2000		2007		2000	2007
		Reading	Math	Reading	Math	Reading	Math	Reading	Math		
SES index	Quintile 1	722.8 (2.2)	740.4 (2.8)	718.3 (1.8)	735.9 (2.1)	0.51 (0.02)	0.42 (0.02)	0.59 (0.02)	0.60 (0.02)	56.6 (0.7)	45.7 (0.6)
	Quintile 2	717.8 (2.1)	739.0 (2.6)	717.3 (2.0)	735.9 (2.5)	0.58 (0.02)	0.46 (0.02)	0.59 (0.02)	0.60 (0.02)	54.3 (0.7)	41.5 (0.7)
	Quintile 3	720.4 (2.2)	740.2 (2.5)	721.5 (2.2)	739.3 (2.6)	0.55 (0.02)	0.44 (0.02)	0.53 (0.02)	0.58 (0.02)	53.0 (0.7)	39.7 (0.5)
	Quintile 4	721.0 (2.4)	735.1 (2.7)	722.8 (2.2)	740.6 (2.4)	0.60 (0.02)	0.48 (0.02)	0.56 (0.02)	0.55 (0.02)	53.2 (0.7)	41.5 (0.6)
	Quintile 5	730.8 (3.4)	743.1 (3.8)	726.1 (2.2)	741.8 (2.6)	0.52 (0.02)	0.47 (0.02)	0.52 (0.07)	0.49 (0.02)	51.7 (0.9)	40.6 (0.4)
School Proprietor	Government	882.6 (0.0)	986.3 (0.0)	738.8 (3.0)	746.2 (4.1)	0.58 (0.12)	0.79 (0.10)	0.76 (0.02)	0.71 (0.02)	21.1 (0.0)	37.8 (0.9)
	Community	721.7 (7.0)	722.2 (7.1)	726.6 (2.3)	782.7 (2.8)	0.43 (0.08)	0.35 (0.07)	0.61 (0.03)	0.71 (0.03)	47.1 (0.7)	40.9 (0.6)
	RCM	719.9 (1.6)	740.7 (1.7)	717.3 (1.6)	736.9 (1.7)	0.50 (0.01)	0.39 (0.01)	0.51 (0.01)	0.52 (0.01)	53.2 (0.5)	41.4 (0.3)
	LEC	728.9 (1.5)	742.1 (2.0)	721.8 (1.5)	741.7 (1.9)	0.60 (0.01)	0.54 (0.01)	0.56 (0.01)	0.56 (0.01)	54.3 (0.4)	45.4 (0.5)
	ACL	701.6 (3.2)	723.2 (3.6)	708.7 (2.5)	727.4 (3.4)	0.58 (0.02)	0.42 (0.04)	0.60 (0.03)	0.63 (0.03)	59.6 (1.3)	37.5 (0.5)
	AME	707.7 (6.9)	713.3 (6.3)	753.9 (0.0)	748.6 (0.0)	0.88 (0.04)	0.83 (0.05)	1 (0.00)	1 (0.00)	34.6 (1.3)	27.7 (0.0)
	Private	—	—	679.8 (0.0)	629.2 (0.0)	—	—	0.28 (0.01)	0.04 (0.04)	—	45.1 (0.0)
	Other churches	766.3 (11.6)	775.3 (14.7)	725.4 (5.7)	717.1 (4.3)	0.45 (0.05)	0.47 (0.05)	0.40 (0.03)	0.38 (0.03)	45.1 (1.7)	38.5 (0.7)
Location	Rural	720.3 (1.3)	737.4 (1.6)	717.3 (1.1)	737.1 (1.3)	0.57 (0.01)	0.48 (0.01)	0.55 (0.01)	0.58 (0.01)	53.1 (0.4)	41.2 (0.3)
	Urban	725.2 (2.0)	743.1 (2.1)	730.1 (1.7)	742.3 (2.1)	0.53 (0.02)	0.41 (0.01)	0.57 (0.01)	0.52 (0.01)	55.1 (0.5)	42.9 (0.5)
Overall		722.0	739.4	721.3	738.8	0.553	0.456	0.559	0.563	53.85	41.80

Source: Own calculations based on SACMEQ II and SACMEQ III. We report the means with standard errors in parentheses.

relationship between wealth and teacher test scores. For instance, in 2000, the average teacher reading score in the first wealth quintile is 723 points. It dips to 718 points in the second quintile, and then steadily rises up to 731 points in the fifth quintile. We observe a similar pattern in 2007. Teacher maths scores, on the other hand, very much oscillate across wealth quintiles in 2000. In 2007, however, teacher maths scores steadily rise from 736 points in the first quintile to 742 points in the fifth quintile. This positive relationship between teacher test scores and wealth may explain the relatively high maths and reading achievement inequality in 2007 seen in Table 3.3a.

Furthermore, the last two rows of Table 3.4 indicate that teacher-subject knowledge is generally skewed towards urban areas. In 2000, an average urban teacher scored 5 points (in reading) and 6 points (in maths) above his/her rural counterpart. In 2007, the gap widens up to 13 points in reading performance, but remains almost constant at 5 points in maths performance. This indicates that low quality teachers are largely serving students from rural areas (with many poor households), while high quality teachers are teaching students from urban areas (with many rich households).

The table further reveals that the proportion of students from poor backgrounds with textbooks is higher than that of their affluent counterparts in 2007, which is the reverse of the situation in 2000. For example, in 2007, about 60% of students in the first quintile have textbooks, while only 52% and 49% in the fifth quintile have access to reading and maths textbooks, respectively. Across wealth quintiles, there is an increase in the proportion of students with maths textbooks over the period. Moreover, rural students generally have better access to maths textbooks than their urban counterparts. For example, 58% of rural students have maths textbooks in 2007 compared to 52% of urban students.

Even though pupil-teacher ratio has declined across socio-economic groups and locations between 2000 and 2007, students from wealthy families generally attend less crowded schools compared to those from disadvantaged families. In 2000, the pupil-teacher ratio in the first quintile was 57 versus 52 in the fifth quintile. These ratios dropped, respectively, to 45 and 41 students per teacher in 2007. However, this was to be expected because the FPE policy involved construction of more primary schools in rural (and largely hard-to-reach mountainous) areas.

Another particularly interesting finding in this table is that pupil-teacher ratio fell in all schools except government schools, where it increased by about 16 students per teacher. This could be explained by the location of newly built government schools and where new teachers are likely to work. As mentioned above, under the government school building programme, more schools were built in previously underserved and hard-to-reach mountainous areas, in addition to building additional classrooms to existing schools. In fact this is what Table 3.2 shows: the share of urban schools decreased from 35 percent to 33.6 percent. This, coupled with the fact that many children are from rural areas, imply that government schools attracted more children. It is also possible that children changed to government schools because of them being nearer to their homes. Attracting teachers to rural and hard-to-reach schools is always a problem in Lesotho. Thus, the increase in the number of students in rural (and mainly government schools) was not accompanied by an increase in the number of teachers.

All in all, these average statistics appear to show some correlation between the changes in the distribution of policy variables and changes in educational achievement. For instance, the relatively high pupil teacher ratio and low teacher subject scores in the first wealth quintile are possibly associated with the low student performance in the first wealth quintile. But, as pointed out earlier, changes in averages obscure much of educational quality distributional dynamics. Therefore, there is need to go beyond these simple, but informative, descriptives to uncover the distributional changes. I do this in the next two sections.

3.3 Relative Distribution Analysis

3.3.1 The method

To further explore the distributional changes in educational achievement over the period, I use the relative distribution method developed by Handcock and Morris (1998, 1999). This method is a fully non-parametric statistical framework for analysing changes in the distribution of an attribute (e.g. test scores) between groups (rural-urban) or periods (2000-2007)[11]. Rodrigues et al. (2013) are the first to employ this method in education economics to investigate distributional changes in educational achievement in Brazil between 1997 and 2005.

Let the continuous cumulative distribution function (CDF) of test scores (Y) for the reference year, 2000, be $F_0(y_0)$, and that of the comparison year, 2007, be $F_1(y_1)$. Again, assume that both of these CDFs have a common support with their respective density functions given as $f_0(y_0)$ and $f_1(y_1)$. The relative distribution of Y_1 to Y_0 is then defined as

$$R = F_0(Y_1) \tag{3.1}$$

This produces a relative data r (i.e. the realization of R) interpreted as the percentile rank that the 2007 test score y_1 would have in the 2000 test scores distribution, and is continuous in the interval [0, 1]. R is a random variable with a CDF given as

$$G(r) = F_1(F_0^{-1}(r)) \equiv F_1(y_r) \;\; r \in [0, 1] \tag{3.2}$$

[11]This presentation draws heavily from Handcock and Morris (1998, 1999) who pioneered the method.

and a probability density function (PDF), the *relative density*, defined as

$$g(r) = \frac{f_1(F_0^{-1}(r))}{f_0(F_0^{-1}(r))} = \frac{f_1(y_r)}{f_0(y_r)} \ \ r \in [0,1] \tag{3.3}$$

where y_r is the τ^{th} quantile test score in the 2000 test score distribution.

Thus, for each quantile of the 2000 test scores' distribution, there are three possible results: if $g(r) > 1$ (or $g(r) < 1$), there is an over-representation (or under-representation) of year 2007 students relative to 2000 students; and if $g(r) = 1$, we have distributional equivalence such that the performance levels of students in both years is the same across quantiles.

If $g(r) > 1$ or $g(r) < 1$, it implies that the two PDFs are different. To decompose these differences between PDFs into differences that are due to location (mean or median) shift and shape (scale or skewness) effects, I further make the following assumptions. Let Y_A denote a hypothetical random variable indicating the reference year test scores location-adjusted to have the same median as the comparison year test scores. That is, Y_A has the same median of the 2007 distribution, but the shape of the 2000 distribution. For an additive shift in location, Y_A is the random variable defined as $Y_0 + \rho$, where $\rho = \text{median}(Y_1) - \text{median}(Y_0)$. Thus, Y_A is equal to the educational achievement in 2000 plus the 2007 median score less the 2000 median score. The CDF of Y_A is given as $F_A(Y_1) = F_0(Y_1 - \rho)$.

From this transformation, I now have three distributions, Y_0, Y_A, and Y_1, with which I can construct two relative distributions that isolate the effects of changes in location and shape (or structure) of the distribution. Using the notation of equation 3.1, let $R \equiv R_0^1 = F_0(Y_1)$ be the relative distribution of Y_1 to Y_0, $R_0^A = F_0(Y_A) = F_0(Y_0 + \rho)$ be the relative distribution of Y_A to Y_0, which isolates the location effect, and $R_A^1 = F_A(Y_1) = F_0(Y_1 - \rho)$ be the the relative distribution of Y_1 to Y_A isolating the shape effect. This is represented in terms of density ratios from equation 3.3 as:

$$\frac{f_1(y_r)}{f_0(y_r)} \quad = \quad \frac{f_A(y_r)}{f_0(y_r)} \times \frac{f_1(y_r)}{f_A(y_r)} \tag{3.4}$$

The results of this decomposition are displayed graphically, and quantified using summary measures such as the median relative polarization (MRP) index.

The MRP index helps to tease out whether changes in educational achievement over time are associated with changes in educational inequality by measuring the relative density in the center and/or tails of the distribution. It corresponds to the median absolute deviation of the relative distribution. Because I am only concerned with shape effects, I use the median-matched relative distribution of Y_1 to Y_0, which is defined as $R_0^A = F_0(Y_1 - \rho)$. Therefore, the median of $R_0^A = 1/2$ and the MRP of Y_1 to Y_0 is defined as

$$MRP(F_1;F_0) = 4\int_0^1 \mid r - \frac{1}{2} \mid g_0^A(r)dr - 1 \qquad (3.5)$$

where $g_0^A(r)$ is the relative distribution of the 2007 distribution, location-adjusted to have the same median as the 2000 distribution. The MRP produces values between -1 and 1, where zero implies no shape differences, positive values mean more polarization (i.e. increases in the tails of the distribution), and negative values show convergence towards the center of the distribution (i.e. decreases in the tails of the distribution). The MRP is further decomposed into lower tail polarization index (LRP), and upper tail polarization index (URP), respectively, defined as

$$LRP(F_1;F_0) = 8\int_0^{\frac{1}{2}} \mid r - \frac{1}{2} \mid g_0^A(r)dr - 1 \qquad (3.6)$$

$$URP(F_1;F_0) = 8\int_{\frac{1}{2}}^1 \mid r - \frac{1}{2} \mid g_0^A(r)dr - 1 \qquad (3.7)$$

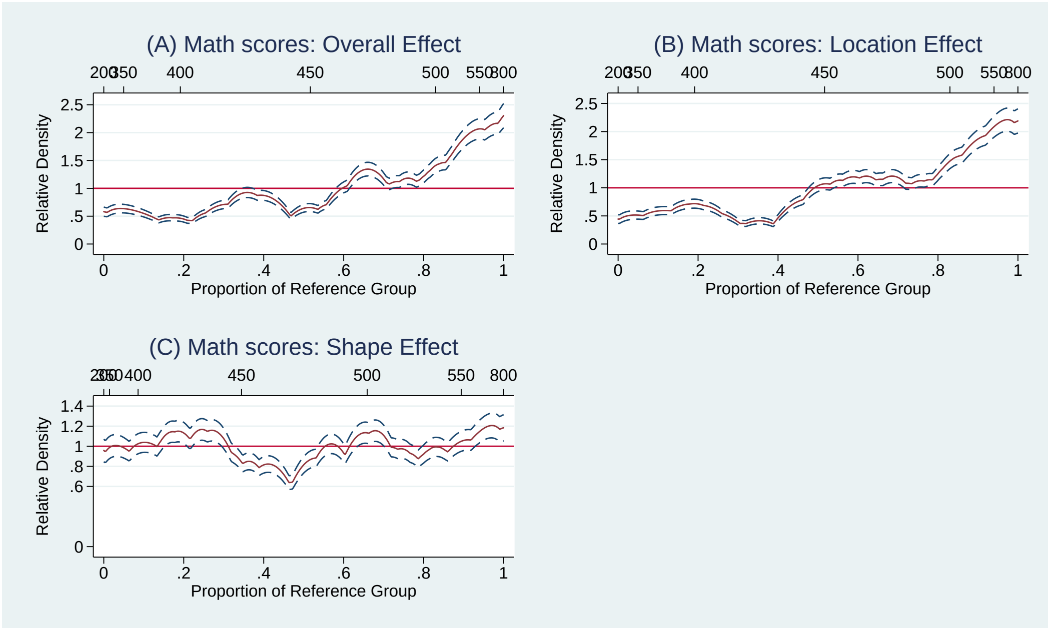
3.3.2 Changes in Educational Achievement

Figures 3.2 and 3.3, and Table 3.5 present the relative distribution analysis results[12]. Figure 3.2 shows relative distributions of maths scores between 2000 and 2007, and Figure 3.3 shows the relative distributions of reading scores. In each figure, panel A shows the overall distributional change, panel B shows the location effect, and panel C shows the shape effect. From panel A of each of these figures, we can see that there is an increase in the relative densities of maths and reading scores along the quantiles of the distributions. This implies a decline in the density of students with lower performance levels on both proficiency tests (maths and reading), concurrent with an increase in the density of students with higher performance levels above the 80^{th} percentile.

Focusing on panel A of Figure 3.2, for example, we can see that the density of 2007 maths scores relative to 2000 maths scores at the 20^{th} percentile is about 0.5 (i.e. $g(0.2) = 0.5$). This implies that the 2007 students are 50% less likely to be at the 20^{th} percentile of the 2000 maths scores' distribution. Similarly, panel A of Figure 3.3 shows that $g(0.2) = 0.8$ for reading scores, implying that the 2007 students are 20% less likely to be at the 20^{th} percentile of the 2000 reading scores' distribution. At the other extreme, we can see that the relative density is approximately 2.3, i.e. $g(1.0) = 2.3$, for maths, and $g(1.0) = 1.8$ for

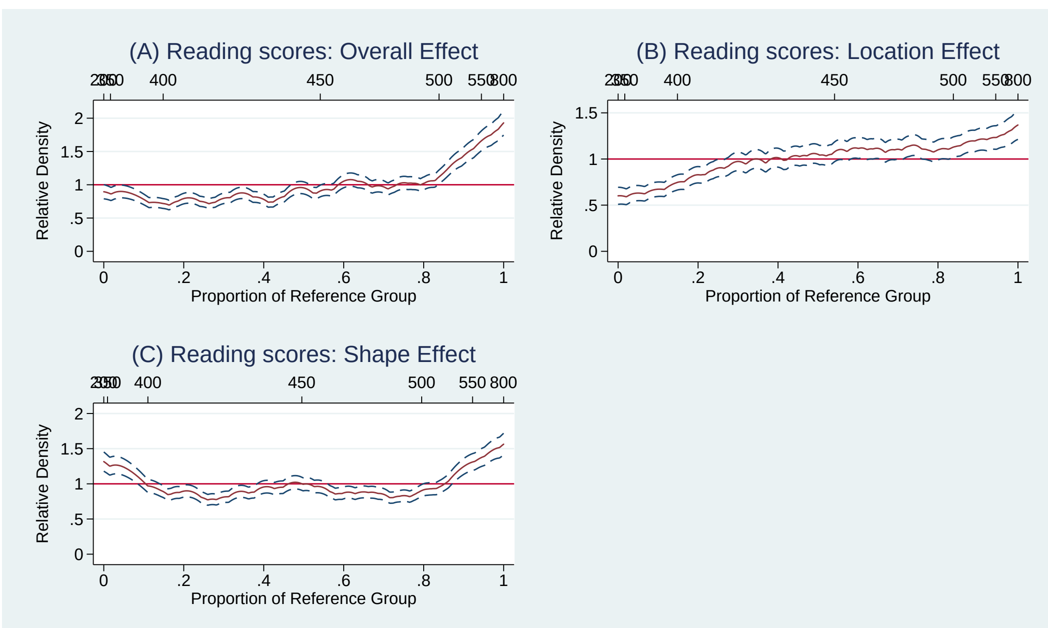
[12]Ben Jann's *reldist* Stata code is used for this analysis (see Jann, 2008).

Figure 3.2: Relative distribution of maths test scores: 2000–2007



Source: Own calculations based on SACMEQ II and SACMEQ III data. *Notes*: The x-axis shows the quantiles of the 2000 test scores while the top horizontal line shows actual scores corresponding to each quantile. The solid horizontal line is the line of distributional equivalence. The dashed lines are the 95% confidence intervals for the relative density.

Figure 3.3: Relative distribution of reading test scores: 2000–2007



Source: Own calculations based on SACMEQ II and SACMEQ III data. *Notes*: The x-axis shows the quantiles of the 2000 test scores while the top horizontal line shows actual scores corresponding to each quantile. The solid horizontal line is the line of distributional equivalence. The dashed lines are the 95% confidence intervals for the relative density.

reading. These results suggest that the proportion of students who, in 2007, scored between 550 and 800 points in maths and reading, respectively, is 130% and 80% more than that of the 2000 student population that reached the same proficiency levels. All in all, these results show a huge increase in maths performance relative to reading performance.

Turning to the location effects in panel B of Figures 3.2 and 3.3, I find that the results are much in line with those seen in Figure 3.1 and Table 3.2. That is, there is a strong (less strong) location shift in the maths (reading) scores between 2000 and 2007, indicating that the 2007 students' test scores are far superior than the 2000 scores. This is largely driven by students below the median and those above the 80^{th} quantile for maths, and those below the 20^{th} quantile and those above the 80^{th} quantile for reading.

Figure 3.1 and Table 3.2 have clearly shown that the 2007 test scores' distributions are more unequal than the 2000 ones. Panel C of Figures 3.2 and 3.3 show us who is driving the inequality between these distributions. These figures plot the relative density of Y_1 to Y_A. Focusing on panel C of Figure 3.2 for maths scores, we can see that the relative density is oscillating around the line of distributional equivalence. Therefore, there is no clear evidence of polarization in maths performance. Looking at the reading scores shape effect in panel C of Figure 3.3, however, we can see clear evidence of an increase in polarization (i.e. an increase in the tails of the relative distributions) between 2000 and 2007. The U shape in panel C of Figure 3.3 shows an increase in reading scores polarization (or inequality).

Table 3.5 presents education quality polarization indices; the MRP, LRP, and URP indices. The MRP index for reading scores indicates that, over the period, about 9% of the student population has shifted away from the centre of the reading scores' distribution, a percentage point higher than the 8% shift in the maths scores' distribution, and these changes are statistically significant at 5% level. This implies that there is relatively higher increase in educational inequality in reading performance compared to that in maths performance. The results further show that lower tail polarization (LRP) has contributed more to the overall polarization index, MRP, than the upper tail polarization (URP). For example, the LRP index for reading scores is 10.6%, almost 4 percentage points higher than the URP index, and is statistically significant at 5% level. Thus, there is a high proportional representation of the 2007 cohort at the lower and upper tails of the reading distribution, but largely at the lower tail. As for the maths relative distribution, neither LRP index nor URP index is statistically significant, even though LRP seems to be relatively higher in magnitude.

It is possible that the observed changes in educational achievement and educational inequality are a mix of "composition effects" from rising enrolment and "treatment effects" from changes in teacher and school characteristics. In the next two sections, I attempt to link these distributional changes in test scores to changes in the covariates (e.g. the policy variables) seen in Section 3.2.

Table 3.5: Grade 6 test scores' distribution polarization between 2000 and 2007

Polarization index	Reading scores		Maths scores	
	Coefficient	Std. error	Coefficient	Std. error
Median (MRP)	0.088**	0.0162	0.079**	0.0160
Lower tail (LRP)	0.106**	0.0349	0.126	0.0942
Upper tail (URP)	0.069**	0.0288	0.032	0.0825

Source: Own calculations based on SACMEQ II and SACMEQ III data. *Notes*: Bootstrapped standard errors are reported. * significant at 10% level, ** significant at 5% level, *** significant at 1% level.

3.4 Decomposing changes in educational achievement and educational inequality

When one is interested in comparing average test scores between two periods (or groups), it is possible to apply the workhorse decomposition method proposed by Blinder (1973) and Oaxaca (1973), the Oaxaca-Blinder (O-B) decomposition method. The O-B method enables one to decompose the mean change in an outcome variable into covariates-explained change and 'unexplained' (or coefficients-explained) change, and further divide these components into the contribution of each covariate. Suppose that student i's test score in period t, y_{it}, is linearly dependent on a vector of characteristics, $\mathbf{X}_{it}$, through the following education production function

$$y_{it} = \mathbf{X}_{it}\beta_t + \varepsilon_{it}, \text{ for } t = 0, 1 \tag{3.8}$$

where β_t is a time-specific parameter vector, and ε_{it} is a mean-zero random error. The mean difference in test scores between 2000 ($t = 0$) and 2007 ($t = 1$) is therefore given as:

$$\begin{aligned} \triangle_T^\mu &= \overline{y}_1 - \overline{y}_0 \\ &= \overline{\mathbf{X}}_1\beta_1 - \overline{\mathbf{X}}_0\beta_0 \\ &= \overline{\mathbf{X}}_1\beta_1 - \overline{\mathbf{X}}_1\beta_0 + \overline{\mathbf{X}}_1\beta_0 - \overline{\mathbf{X}}_0\beta_0 \\ \triangle_T^\mu &= \overline{\mathbf{X}}_1\left(\beta_1 - \beta_0\right) + \left(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_0\right)\beta_0 \\ \triangle_T^\mu &= \triangle_r^\mu + \triangle_X^\mu \end{aligned} \tag{3.9}$$

where $\triangle_r^\mu$ is the unexplained (or coefficients) effect, and $\triangle_X^\mu$ is the explained (or characteristics) effect.[13] The detailed decompositions of these effects are written in terms of the sums of the covariates' contributions as:

[13] Note that the use of "effects" here does not refer to causal effects but rather the contribution of coefficients and characteristics to the total change. This is in line with the literature on decomposition.

$$\begin{aligned}\triangle_r^{\mu} &= \sum_{k=1}^{K} \overline{x}_1^k \left(\beta_{1,k} - \beta_{0,k}\right) &(3.10)\\ \triangle_X^{\mu} &= \sum_{k=1}^{K} \left(\overline{x}_1^k - \overline{x}_0^k\right) \beta_{0,k} &(3.11)\end{aligned}$$

where $\overline{x}^k$ is the mean of the k^{th} variable, and $\beta_{t,k}$ is the coefficient of the k^{th} variable.

One of the limitations of the O-B decomposition is that it cannot be used to decompose other distributional statistics, such as the median and variance, other than the mean (Firpo et al., 2007; Fortin et al., 2011). Since the primary interest is in learning about how pupils' characteristics affect not only the mean test scores but also other distributional statistics, I use the recentered influence function (RIF) regression decomposition method by Firpo et al. (2007, 2009). The RIF-regression decomposition method is a generalization of the O-B that allows for the decomposition of changes in any distributional statistic.

To see this, assume that the education production function in equation 3.8 is some unknown flexible function,

$$y_{it} = f\left(\mathbf{X}_{it}, \varepsilon_{it}\right), \text{ for } t = 0, 1 \tag{3.12}$$

where y, t, and $\mathbf{X}$ are jointly distributed. Let F_1 be the distribution of test scores (y) at time $t = 1$ (i.e. $y_{t=1}|t = 1 \overset{d}{\sim} F_1$) and similarly, $y_{t=0}|t = 0 \overset{d}{\sim} F_0$. Let F_C be the distribution of test scores that 2007 (i.e. $t = 1$) students would have got had their characteristics been rewarded as in 2000 ($t = 0$). We denote such counterfactual scores as $y_{t=0}|t = 1 \overset{d}{\sim} F_C$. The corresponding distributional statistics (e.g. the median) for our three respective distributions are denoted as $\varphi(F_1)$, $\varphi(F_0)$ and $\varphi(F_C)$. Therefore, the change in this distributional statistic between $t = 0$ and $t = 1$ is given as:

$$\begin{aligned}\triangle_T^{\varphi} &= \varphi(F_1) - \varphi(F_0)\\ \triangle_T^{\varphi} &= \varphi(F_1) - \varphi(F_C) + \varphi(F_C) - \varphi(F_0)\\ \triangle_T^{\varphi} &= \triangle_r^{\varphi} + \triangle_X^{\varphi} &(3.13)\end{aligned}$$

In order to get the detailed decompositions of $\triangle_r^{\varphi}$ and $\triangle_X^{\varphi}$, I estimate the τ^{th} quantile (q_τ) RIF regressions. First, I calculate the RIF for each observation as

$$
\begin{aligned}
RIF(y; q_\tau) &= q_\tau + IF(y; q_\tau) \\
&= q_\tau + \frac{\tau - \mathbf{1}\{y \leq q_\tau\}}{f_Y(q_\tau)}
\end{aligned}
$$

where q_τ is the sample quantile, $f_Y(q_\tau)$ is the density of y at point q_τ (estimated using kernel methods) and $\mathbf{1}\{y \leq q_\tau\}$ is an indicator function taking the value of 1 when the value of the test score (y) is less or equal to q_τ. Second, I run the OLS regression of this new dependent variable, $RIF(y; q_\tau)$, on the covariates, ($\mathbf{X}$) as follows:[14]

$$
RIF(y; q_\tau) = \mathbf{X}_{it}\beta_t + \varepsilon_{it} \tag{3.14}
$$

and then apply the O-B decomposition to get $\triangle_r^\varphi$ and $\triangle_X^\varphi$ by quantile, and their detailed decompositions as in equations 3.10 and 3.11. This decomposition gives the influence of changes in covariates and coefficients over time on changes in achievement (Firpo et al., 2007, 2009). Since the expected value of $RIF(y; \varphi, F_t)$ equals φ, these are unconditional quantile decompositions. Moreover, if $\varphi = \mu$, then $RIF(y; \varphi, F_t)$ equals y, so that the decomposition of equation 3.14 becomes the traditional O-B decomposition.[15]

3.5 RIF regression results

The object of this section is to see whether changes in any of the education policy variables (for example, PTR, access to textbooks and teacher education) and any other covariates are associated with the observed changes in educational achievement and educational inequality between 2000 and 2007. I first report the results of the RIF-regression decomposition of changes in educational achievement, and then the results for changes in educational inequality (i.e. changes in test scores' variances).

Table 3.6 presents the RIF regression quantile decomposition of changes in reading performance. It shows the total change in reading achievement by quantiles, and the contribution of each covariate to the changes. Table 3.7, on the other hand, presents the RIF regression quantile decomposition of changes in maths performance. Just like Table 3.6, it shows the total change in maths scores by quantiles and the contribution of each covariate to these changes.

The overall message from these tables is that educational achievement increased across all quantiles between 2000 and 2007, with much of the increase being driven by the performance

[14] Equation 3.14 is called the *Uconditional Quantile Regression* (Firpo et al., 2007, 2009).

[15] To implement this decomposition method, I use the `rifreg` Stata code by Nicole Fortin, dowloadable at `http://faculty.arts.ubc.ca/nfortin/datahead.html`.

of students at the lower quantiles. These results are much in line with earlier results presented in Figures 3.2 and 3.3. More importantly, we can see from both Tables 3.6 and 3.7 that the coefficients effect (or unexplained effect) is statistically significant, and it is almost as large as the total differential across quantiles, while the explained effect is relatively smaller and statistically insignificant. This signifies that changes in observed students' characteristics have had very little influence on the observed changes in educational achievement.

Looking at individual tables, and starting with Table 3.6, I find that changes in pupil teacher ratio (PTR), speaking English at home, grade repetition, and teacher effort (i.e. teaching hours per week and testing frequency), all have a statistically significant relationship with changes in reading achievement. Interestingly, all these covariates, except grade repetition, only have a significant influence on test score changes at the lower quantiles. For example, pupil-teacher ratio is negatively associated with changes in reading scores at the 10^{th} and 50^{th} quantiles. This implies that the fall in pupil teacher ratio between 2000 and 2007 is related to the increase in pupils' reading achievement at these quantiles. Speaking English at home is positively related to the increase in reading performance of pupils at the 10^{th} quantile only. An increase in teacher effort, either by giving more tests and/or teaching more hours per week, is positively related with the increased reading performance of pupils at the 10^{th} and 50^{th} quantiles. Lastly, grade repetition has a strong positive influence on maths scores at the top quantiles. Although it is not a policy variable, it is also interesting to note that the socio-economic status (SES) index appears to have strong positive influence on students' reading performance at the upper tails of the distribution.

Turning to Table 3.7, I find that education policy variables have no significant relationship with changes in maths performance. However, pupil maths performance is statistically associated with the socio-economic status (SES) index, school wealth (i.e. school total assets and school building condition), pupil age, and grade repetition. Specifically, I find that an increase in pupil socio-economic status is significantly associated with an increase in maths performance across all quantiles. But the influence of the SES index seems to be stronger at all quantiles of the maths distribution.

Table 3.6: Changes in reading scores, 2000-2007: RIF-Reg decomposition results

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Total Differential	156.8*** (7.957)		91.35*** (2.615)		86.08*** (6.014)	
Pupil Teacher Ratio	-3.306*	-38.96*	-3.167*	9.070	-3.948	17.49
	(1.715)	(22.56)	(1.654)	(7.332)	(4.025)	(17.29)
Teacher quality	-1.174	-134.4	0.833	-56.50*	4.220	-64.26
	(2.325)	(99.48)	(2.021)	(29.27)	(4.680)	(63.92)
School Ownership	-1.738	193.6***	-1.944	75.93***	-1.591	167.4***
	(2.065)	(71.00)	(1.705)	(10.13)	(2.991)	(20.27)
Teacher sex	0.172	16.59*	-0.0216	-0.320	0.236	-5.596
	(0.410)	(9.330)	(0.114)	(3.682)	(0.509)	(8.121)
Pupil sex	0.194	11.09*	0.232	4.862***	0.0232	-5.593
	(0.418)	(6.351)	(0.194)	(1.418)	(0.126)	(4.088)
Textbook access	0.0524	13.44*	0.134	5.446***	0.320	10.00
	(0.301)	(7.282)	(0.327)	(2.072)	(0.747)	(6.087)
Teacher age	0.199	122.8**	0.188	2.037	-0.510	15.79
	(0.622)	(48.93)	(0.517)	(13.59)	(1.569)	(33.42)
Pupil repeated	0.853	6.477	1.690***	-5.936***	6.423***	-34.47***
	(0.554)	(10.91)	(0.509)	(2.268)	(1.536)	(5.394)
Speak English home	1.694*	2.339	0.632	4.201*	0.999	13.05**
	(0.878)	(7.609)	(0.390)	(2.230)	(0.708)	(5.407)
Teacher effort	3.307**	-16.30	3.570**	-24.22***	-0.134	4.060
	(1.412)	(23.77)	(1.426)	(7.706)	(2.444)	(14.42)

Continued on next page

Table 3.6: *Continued from previous page*

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
SES index	1.019	-3.059	1.473**	-1.541***	4.028**	-3.751***
	(0.658)	(2.281)	(0.593)	(0.505)	(1.977)	(1.434)
School location	1.323	0.235	0.924	4.533***	-0.642	7.912*
	(1.011)	(3.919)	(1.033)	(1.566)	(1.331)	(4.114)
Pupil age	0.575	98.73**	1.699***	-57.14***	1.938**	-125.7***
	(0.454)	(48.69)	(0.507)	(11.85)	(0.836)	(31.10)
Pupil Problem index	-0.283	-0.142	-0.0265	-0.0218	-0.714	1.035
	(0.493)	(0.992)	(0.138)	(0.191)	(1.430)	(0.836)
Teacher Problem index	-0.191	0.168	0.0273	-0.0368	-0.287	0.130
	(0.326)	(0.518)	(0.176)	(0.109)	(0.817)	(0.460)
Lost School days	3.552	-2.830	-0.474	0.844	-1.417	1.778
	(3.384)	(6.818)	(2.666)	(3.080)	(9.288)	(10.65)
School wealth	0.727	10.84	-1.116	28.30***	-2.742	38.48*
	(1.323)	(23.34)	(1.617)	(7.873)	(3.321)	(21.06)
School distance from public services	1.746	-9.041	0.507	-2.692	0.880	-6.534
	(1.720)	(6.717)	(0.641)	(2.495)	(1.245)	(5.244)
Social Capital	-1.231	21.71	0.621	6.976	-30.99**	34.69*
	(4.155)	(15.97)	(3.810)	(7.459)	(11.98)	(17.63)
Explained Effect \ Unexplained Effect	7.492	149.3***	5.782	60.77***	-23.90	110.0***
	(7.604)	(9.253)	(6.017)	(5.547)	(18.46)	(19.12)

Continued on next page

Table 3.6: *Continued from previous page*

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Constant	-143.9	(155.8)	66.98*	(38.47)	44.09	(79.28)
Observations	7,370	7,370	7,370	7,370	7,370	7,370

Source: Own calculations based on SACMEQ II and SCMEQ III. *Notes*: Teacher quality represents the combined coefficient of teacher subject test score, qualification, years of professional training, and years of experience. School ownership represents the combined coefficient of different school ownership dummies where the government ownership is a reference category. Teacher effort represents the combined influence of teaching hours per week, and the frequency of subject tests. Pupil-Problem and Teacher-Problem indices are composite indices made up of behaviours such as pupil/teacher drug abuse, alcohol abuse, late arrivals, unjustified absenteeism, bullying of pupils/teachers. School wealth represents the combined influence of school asset index and school condition. School capital is a dummy that equals 1 if a community contributes to school activities such as school building, maintenance, etc. Combined influence/coeficient is the total across multiple indicator variables for teacher quality and teacher effort, for example.

Bootstrapped standard errors (100 replications) are in parentheses. Significance level: *** p<0.001, ** p<0.05, * p<0.10.

Table 3.7: Changes in maths scores, 2000-2007: RIF-Reg decomposition results

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Total Differential	214.9*** (9.206)		91.35*** (2.615)		71.99*** (3.932)	
Pupil Teacher Ratio	-3.088	5.834	-0.911	0.330	-1.347	4.654
	(2.368)	(34.17)	(1.067)	(8.166)	(1.929)	(9.263)
Teacher quality	2.066	-551.8***	0.861	-100.2***	0.619	-147.8***
	(3.184)	(118.9)	(2.087)	(27.48)	(4.199)	(39.71)
School Ownership	-1.086	61.31	0.739	124.7***	1.494	77.47***
	(2.995)	(91.15)	(1.376)	(13.49)	(1.608)	(18.02)
Teacher sex	0.416	-4.700	0.172	-3.988	-0.404	3.567
	(0.765)	(16.02)	(0.303)	(3.724)	(0.564)	(5.049)
Pupil sex	-0.00253	-0.800	0.0290	-1.048	0.0562	-3.683
	(0.0809)	(6.088)	(0.0536)	(1.654)	(0.0945)	(2.864)
Textbook access	0.522	4.782	0.426	1.997	0.696	2.831
	(0.673)	(7.822)	(0.339)	(1.712)	(0.647)	(2.616)
Teacher age	0.0249	51.62	0.188	-4.922	1.914	-35.75
	(1.397)	(70.90)	(1.153)	(22.04)	(1.979)	(29.17)
Pupil repeated	1.242	6.670	1.049***	-3.851*	3.164***	-17.83***
	(0.911)	(9.157)	(0.397)	(2.125)	(0.883)	(3.795)
Speak English home	1.709	-5.729	0.770	4.080	0.494	8.103*
	(1.219)	(8.990)	(0.500)	(2.698)	(0.445)	(4.714)
Teacher effort	0.833	11.48	0.358	-5.934	-1.968	14.02
	(1.554)	(38.20)	(0.936)	(8.485)	(1.344)	(10.57)

Continued on next page

Table 3.7: *Continued from previous page*

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
SES index	2.033*	-3.738*	0.827**	-0.915**	2.381**	-1.914**
	(1.098)	(2.200)	(0.410)	(0.430)	(1.113)	(0.858)
School location	1.448	-3.771	0.562	2.226	-0.0604	0.400
	(1.310)	(5.024)	(0.676)	(1.589)	(0.237)	(2.277)
Pupil age	2.803**	19.59	1.420***	-45.66***	0.596	-48.00**
	(1.158)	(55.98)	(0.428)	(13.00)	(0.370)	(21.56)
Pupil Problem index	0.129	2.201	-0.0124	0.399	-0.428	1.037
	(0.516)	(2.292)	(0.166)	(0.425)	(1.004)	(0.649)
Teacher Problem index	0.505	0.0235	-0.0955	0.0348	-0.249	0.188
	(0.792)	(0.379)	(0.252)	(0.153)	(0.577)	(0.322)
Lost School days	0.616	-20.97**	-1.462	0.687	-7.739	8.086
	(4.898)	(9.866)	(3.306)	(3.880)	(9.474)	(10.89)
School wealth	-1.431	11.81	-2.192	28.12***	-3.920*	36.92***
	(2.265)	(32.93)	(1.680)	(7.779)	(2.107)	(10.92)
School distance from public services	2.063	-22.89**	0.221	-3.471	0.669	-5.069*
	(1.931)	(9.811)	(0.470)	(2.925)	(0.801)	(2.734)
Social Capital	1.581	374.6***	-0.425	159.0***	-15.25*	26.89**
	(8.105)	(38.63)	(5.360)	(9.403)	(8.412)	(11.81)
Explained Effect \ Unexplained Effect	12.38	202.5***	2.524	88.82***	-19.28	91.27***
	(13.70)	(15.90)	(7.404)	(7.221)	(14.69)	(14.98)

Continued on next page

Table 3.7: *Continued from previous page*

VARIABLES	Quantile_10		Quantile_50		Quantile_90	
	Explained	Unexplained	Explained	Unexplained	Explained	Unexplained
Constant	267.0 (180.6)		-62.78 (41.24)		167.1*** (53.94)	
Observations	7,370	7,370	7,370	7,370	7,370	7,370

Source: Own calculations based on SACMEQ II and SCMEQ III. *Notes*: Teacher quality represents the combined coefficient of teacher subject test score, qualification, years of professional training, and years of experience. School ownership represents the combined coefficient of different school ownership dummies where the government ownership is a reference category. Teacher effort represents the combined influence of teaching hours per week, and the frequency of subject tests. Pupil-Problem and Teacher-Problem indices are composite indices made up of behaviours such as pupil/teacher drug abuse, alcohol abuse, late arrivals, unjustified absenteeism, bullying of pupils/teachers. School wealth represents the combined influence of school asset index and school condition. School capital is a dummy that equals 1 if a community contributes to school activities such as school building, maintenance, etc. Combined influence/coeficient is the total across multiple indicator variables for teacher quality and teacher effort, for example.

Bootstrapped standard errors(100 replications) are in parentheses. Significance level: *** p<0.001, ** p<0.05, * p<0.10.

Furthermore, I find that grade repetition is positively associated with changes in maths performance at the 90th quantile. This implies that increased repetition leads to increased performance. Thus, this result seems to support the evidence that automatic grade promotion (i.e. relaxing the grade retention policy)[16] has a negative effect on student performance (see Foureaux Koppensteiner, 2014). Lastly, school wealth and school social capital (i.e. whether the surrounding community contributes to school activities such as school building and maintenance) are negatively associated with pupil maths performance at the 90th quantile.

These pupil-teacher ratio results are in line with the literature on class size and/or school resources which shows that school resources do matter for students' performance. Moreover, that grade repetition leads to increased performance implies that by repeating grades, students are able to acquire the grade-appropriate material which also enables them to perform better in subsequent grades.

Given that a large portion of the increase in students' performance is largely unexplained by the covariates, this suggests that other family background factors and intangible school quality factors, such as school management, autonomy, perceptions of staff and students, and accountability, have possibly contributed to the observed changes in achievement (Lounkaew, 2013).

Finally, I turn to Table 3.8 which presents the decomposition of reading and maths test scores' variances. The first two columns present reading scores' variance decomposition results, while the last two columns present results for maths scores' variance decomposition.

We can see from the table that the total differential (i.e. the increase) in reading scores' variance is statistically significant and is much larger than the increase in maths scores' variance. These results are much in line with those presented in Figures 3.2 and 3.3 and Table 3.5. The main contribution of this table is, however, that it shows the detailed decomposition of the variance differential. Specifically, results from Table 3.8 show that the covariates (teacher effort and school social capital) have negatively contributed to test scores inequality in both subjects. This, however, is cancelled by the huge positive contribution of grade repetition and the coefficients effects, mainly school ownership coefficients effects.

[16] Notice, however, that the drop in pupil repetition rate by 10 percentage points between 2000 and 2007 was most likely not due a deliberate policy, but rather forced on schools by the huge influx of students after FPE, to such an extent that they had to relax their retention policies.

Table 3.8: Variance RIF-Regression decomposition for Reading and Maths test scores

VARIABLES	Reading Variance Decomposition		Maths Variance Decomposition	
	Explained	Unexplained	Explained	Unexplained
Total Differential	1,339*** (366.2)		849.7*** (326.0)	
Pupil Teacher Ratio	-69.71	1,143	-11.66	-1,573
	(140.6)	(923.3)	(123.0)	(1,360)
Teacher quality	129.2	-3,384	-107.0	-6,889*
	(194.0)	(4,244)	(266.4)	(3,565)
School Ownership	3.271	25,529***	136.5	22,804***
	(111.2)	(5,506)	(117.1)	(2,874)
Teacher sex	19.45	-679.6	-63.74	222.2
	(31.19)	(468.6)	(50.43)	(432.5)
Pupil sex	8.803	-273.3	2.813	42.74
	(9.618)	(191.0)	(4.547)	(206.5)
Textbook access	3.026	216.8	38.91	184.8
	(12.41)	(296.7)	(40.79)	(203.9)
Teacher age	-47.58	4,358*	128.2	-1,314
	(69.67)	(2,564)	(107.5)	(2,102)
Pupil repeated	161.2***	-246.6	123.4***	-691.1**
	(47.49)	(274.1)	(43.38)	(276.4)
Speak English home	-1.790	-86.78	25.33	1,191***
	(12.29)	(307.6)	(23.62)	(396.7)
Teacher effort	-176.9*	2,166*	-162.3*	1,646
	(99.34)	(1,150)	(84.25)	(1,024)

Continued on next page

Table 3.8: *Continued from previous page*

VARIABLES	Reading Variance Decomposition		Maths Variance Decomposition	
	Explained	Unexplained	Explained	Unexplained
SES index	100.3	-8.349	63.57	-17.50
	(67.99)	(28.07)	(41.42)	(25.93)
School location	2.626	-85.87	6.979	-204.3
	(21.35)	(248.2)	(21.03)	(224.6)
Pupil age	33.61	-1,352	-13.28	2,401
	(24.67)	(1,651)	(13.50)	(1,571)
Pupil Problem index	-20.56	26.37	-8.541	59.62
	(43.77)	(34.75)	(39.12)	(51.47)
Teacher Problem index	-19.63	-22.59	-15.94	8.124
	(31.57)	(35.16)	(36.10)	(22.51)
Lost School days	-162.0	-8.081	-509.3	519.6
	(330.0)	(428.7)	(600.4)	(701.8)
School wealth	-115.2	-749.7	-155.9	1,471
	(123.0)	(1,402)	(119.2)	(994.0)
School distance from public services	-0.869	155.7	23.52	-332.3
	(22.29)	(258.1)	(35.68)	(202.9)
Social Capital	-1,624***	1,752	-1,245**	-499.0
	(570.9)	(1,836)	(611.1)	(1,371)
Explained Effect \ Unexplained Effect	-1,777**	3,115***	-1,744*	2,593**
	(742.2)	(774.2)	(1,019)	(1,088)

Continued on next page

Table 3.8: *Continued from previous page*

VARIABLES	Reading Variance Decomposition		Maths Variance Decomposition	
	Explained	Unexplained	Explained	Unexplained
Constant		-25,335***		-16,436***
		(5,944)		(5,009)
Observations	7,370	7,370	7,370	7,370

Source: Own calculations based on SACMEQ II and SCMEQ III. *Notes*: Teacher quality represents the combined coefficient of teacher subject test score, qualification, years of professional training, and years of experience. School ownership represents the combined coefficient of different school ownership dummies where the government ownership is a reference category. Teacher effort represents the combined influence of teaching hours per week, and the frequency of subject tests. Pupil-Problem and Teacher-Problem indices are composite indices made up of behaviours such as pupil/teacher drug abuse, alcohol abuse, late arrivals, unjustified absenteeism, bullying of pupils/teachers. School wealth represents the combined influence of school asset index and school condition. School capital is a dummy that equals 1 if a community contributes to school activities such as school building, maintenance, etc. Combined influence/coeficient is the total across multiple indicator variables for teacher quality and teacher effort, for example.

Bootstrapped standard errors(100 replications) are in parentheses. Significance level: *** p<0.001, ** p<0.05, * p<0.10.

Changes in *teacher effort* (which is the combined effect of teacher's frequency of giving tests and teaching hours per week) are negatively correlated with the increase in reading achievement inequality. Nevertheless, changes in grade retention appears to be the main contributor to the increase in maths and reading achievement inequality, as indicated by the statistically significant positive contribution of the students repetition variable to the increase in performance and inequality. More importantly, I find that neither changes in school quality (for example, PTR) nor changes in textbook access have significantly influenced the observed changes in the variance. As such, the increase in educational inequality is largely unexplained by the covariates.

3.6 Conclusion

Understanding changes in educational quality and educational inequality is essential for education and social policies. This is particularly so for most developing countries which have recently had a rapid expansion in access to primary education following the introduction of FPE policies.

In this chapter, I use SACMEQ grade six standardised test scores to analyse the changes in educational achievement and educational inequality, and their determinants. Using both the relative distribution method and the RIF-regression decomposition method, several results stand out from the analysis. I find that the increase in educational quality between 2000 and 2007 is largely driven by improved students' performance across quantiles (or abilities), but mainly due to the increase in the performance of those at the lower tails of the achievement distribution. Some of the increase in reading performance is explained by changes in pupil teacher ratio, speaking English at home, grade repetition, and teacher effort (that is, teaching hours per week and testing frequency). The increase in maths performance is not explained by any of the education policy variables, but partly by grade repetition, age, school wealth and school social capital (that is, whether the surrounding community contributes to school activities such as school building and maintenance). By and large, much of the increase in educational performance remains unexplained.

With regard to the significant influence of grade repetition, the results seem to support Rodrigues et al. (2013)'s conjecture, and later confirmed by Foureaux Koppensteiner (2014), that relaxing the retention policy reduces students' performance. These results thus suggest that the automatic grade promotion policy that is now being implemented in Lesotho could hurt educational quality.

I also find that educational quality polarization (or inequality) has increased during this period, more so in reading performance, and this is largely driven by the increase in spread in test scores below the median. This increase in educational inequality is strongly associated with changes in grade repetition, teacher effort and school social capital. Grade repetition has a positive influence on educational inequality, while high teacher effort and school social capital negatively influence educational inequality.

Finally, the results reveal that much of the observed changes in maths and reading scores changes remain unexplained, as shown by the large and statistically significant unexplained (coefficients) effects component. This suggests that there are potentially other intangible school quality variables (for example, network effects) and student family background factors (for example, parental child's sex preferences, birth order effects, etc.) that could have significantly influenced educational performance after the FPE programme. In order to draw policy lessons from Lesotho's successful FPE policy implementation, it is important to try to uncover the factors that could have influenced educational achievement and inequality during this period. The task is to unpack the influence of those factors that are currently unobserved in our data. In the next chapter, I shift focus from FPE policy effects and educational quality changes over time, and look at the impact of family factors on human capital accumulation. Specifically, I examine the effect of a child's order of birth on her development outcomes. This is a topic which has recently received much attention in the economics and psychology literatures.

Appendix

Figure A.1: SES index distribution with and without imputed parental education

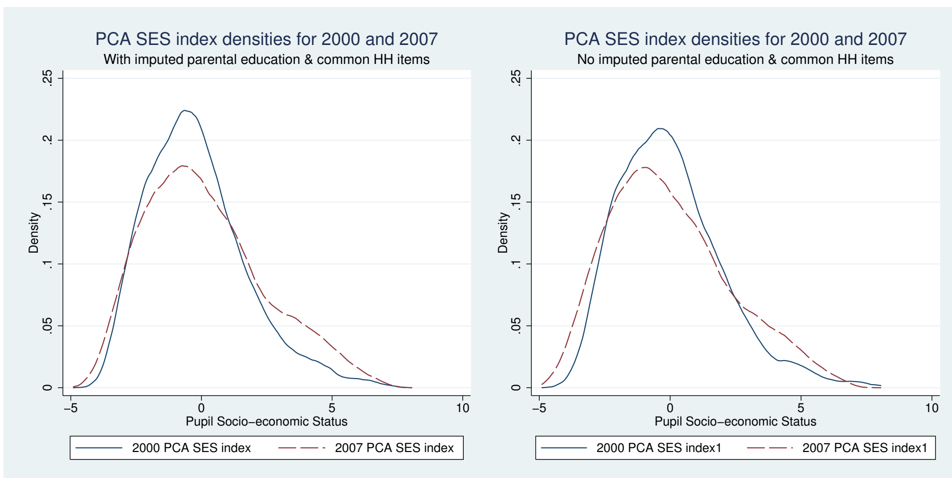


Figure A.2: Cumulative Distribution Functions (CDFs) of grade pupil test scores: 2000–2007

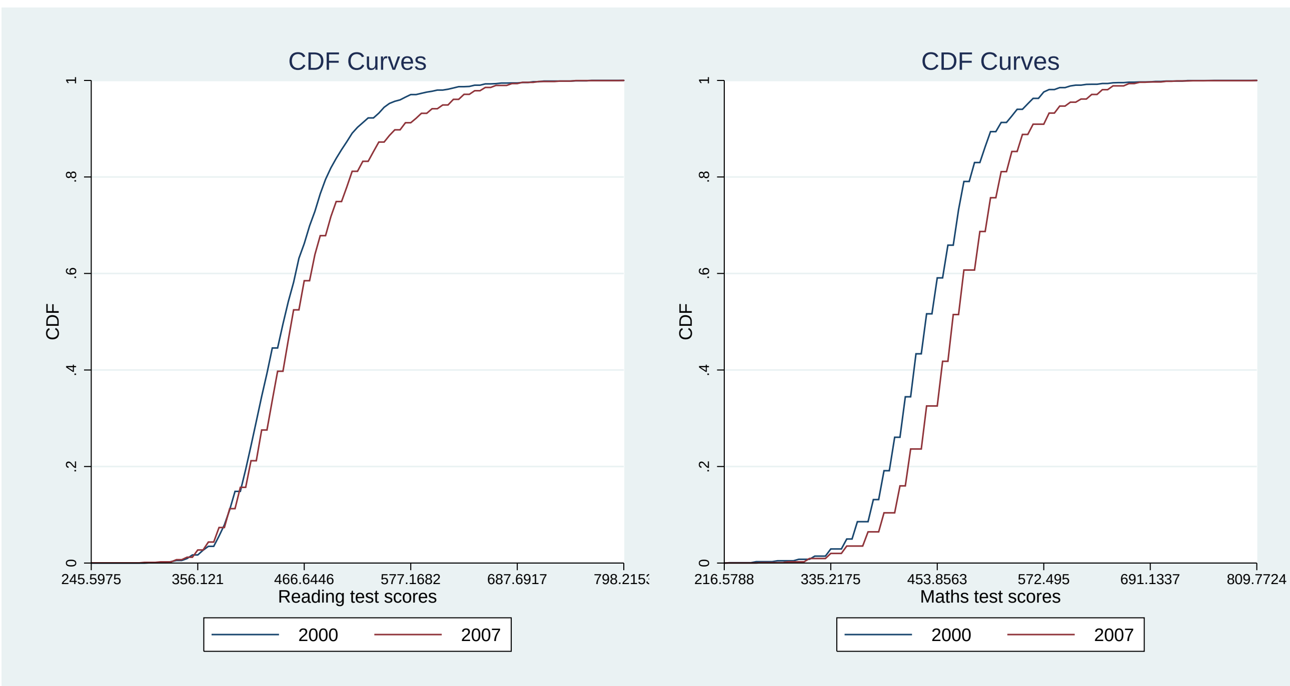


Table A.1: PCA Socio-economic Index Components

Variable	2000			2007		
	mean	sd	N	mean	sd	N
Home items						
mother years of education	7.993	3.714	3155	7.986	3.617	4240
father years of education	7.067	4.430	3155	7.284	4.215	4240
Daily newspaper	0.341	0.474	3155	0.227	0.419	4240
Magazine	0.282	0.450	3155	0.252	0.434	4240
Radio	0.932	0.251	3155	0.911	0.284	4240
TV	0.338	0.473	3155	0.367	0.482	4240
Video Tape Player (VCR)	0.137	0.344	3155	0.119	0.323	4240
Audio cassette player	0.413	0.492	3155	0.355	0.479	4240
Telephone / cellphone	0.132	0.339	3155	0.694	0.461	4240
Refrigerator/freezer	0.176	0.381	3155	0.204	0.403	4240
Car	0.179	0.383	3155	0.152	0.359	4240
Motorcycle	0.0672	0.250	3155	0.0480	0.214	4240
Bicycle	0.278	0.448	3155	0.221	0.415	4240
Piped water	0.267	0.442	3155	0.342	0.474	4240
Electricity	0.138	0.345	3155	0.232	0.422	4240
Table to write on	0.686	0.464	3155	0.744	0.436	4240
Number of books at home	16.35	41.95	3155	10.31	37.56	4240
Source of lighting						
Fire	0.0193	0.138	3155	0.0126	0.112	4240
Candle	0.588	0.492	3155	0.520	0.500	4240
Paraffin or oil lamp	0.294	0.456	3155	0.297	0.457	4240
Gas lamp	0.0261	0.160	3155	0.0325	0.177	4240
Electric lighting	0.0700	0.255	3155	0.136	0.343	4240
No lighting	0.00307	0.0553	3155	0.00157	0.0395	4240
House Floor						
Clay	0.219	0.413	3155	0.226	0.418	4240
Canvas	0.0745	0.263	3155	0.0240	0.153	4240
Wooden planks	0.0358	0.186	3155	0.0333	0.180	4240
Cement	0.442	0.497	3155	0.434	0.496	4240
Carpet/tiles (plastic, ceramic)	0.230	0.421	3155	0.283	0.451	4240
Wall material						
Cardboard/plastic sheeting	0.0314	0.174	3155	0.0241	0.154	4240

Continued on next page

Table A.1: *Continued from previous page*

Variable	2000			2007		
	mean	sd	N	mean	sd	N
Reeds/Sticks/Grass thatch	0.0627	0.242	3155	0.0454	0.208	4240
Stones / Mud bricks	0.447	0.497	3155	0.484	0.500	4240
Metal sheets / Asbestos sheets	0.0654	0.247	3155	0.0498	0.217	4240
Wood (planks/timber)	0.0264	0.160	3155	0.0280	0.165	4240
Cut stone /concrete blocks	0.368	0.482	3155	0.369	0.482	4240
Roofing material						
Cardboard/plastic sheeting	0.0319	0.176	3155	0.0274	0.163	4240
Grass thatch and mud	0.264	0.441	3155	0.273	0.445	4240
Metal sheets / Asbestos sheets	0.519	0.500	3155	0.483	0.500	4240
Cement	0.0771	0.267	3155	0.0624	0.242	4240
Tiles	0.108	0.310	3155	0.154	0.361	4240
Other						
Pupil has regular meal	0.589	0.492	3155	0.716	0.451	4240
Family configuration	0.775	0.417	3155	0.506	0.500	4240
Pupil pays for extra classes	0.493	0.500	3155	0.0938	0.292	4240

Source: SACMEQ II and SACMEQ III data. *Notes*: Family configuration is a dummy equal to one if a child lives with both parents.

Chapter 4

Birth Order Effects on Educational Attainment: Evidence from Lesotho

4.1 Introduction

Over the past two decades, increasing education levels in the developing world, particularly sub-Saharan Africa, has been of paramount importance. Morduch (2000) argues that an effective policy for achieving this goal depends on a better understanding of the nature of schooling decisions in these countries. According to Leibowitz (1974), Cunha et al. (2006), and Cunha and Heckman (2009), differences in child outcomes emerge from an early age, due in part to the amount of human capital (cognitive and non-cognitive skills) acquired before age six. The different environments chosen and created by families help create these differences in child outcomes. Moreover, much of schooling decisions, which affect later life outcomes, take place within the family. Hence, it is important to understand specific family factors responsible for these differences in child outcomes.

One of the family environment factors, birth order (that is, a child's order of birth), is a recurrent theme in the economics and psychology literatures. Even though any particular child's order of birth is biologically determined, parents can actively or passively choose to create different home environments for children of different birth orders, which will then affect their cognitive and non-cognitive skills' development. The available evidence on birth order effects on educational attainment shows a consistent divide between the developing world and the developed world: there tend to be negative birth order effects in developed countries, while there is evidence of positive birth order effects from developing countries (De Haan et al., 2014). Therefore, birth order effects on child outcomes are context-specific, as it relates to countries' levels of development.

However, within less developed countries there are heterogeneities in terms of social norms that shape parental preferences towards children of different birth orders, which may lead to different birth order effects (see Jayachandran and Kuziemko, 2011). Thus, the evidence of positive birth order effects found in some developing countries may not be generalised to other developing countries with different contexts. More importantly, the underlying causal mechanisms through which birth order affects educational attainment are still unsettled, and also appear to be context-specific.

In this chapter, I examine the effect of birth order on children's educational attainment in Lesotho. I use the 2006 census data for children aged six to twelve years to estimate birth order effects on educational attainment. To deal with potential endogeneity of birth order resulting from its correlation with family size and other unobserved family factors, I use family fixed effects models.

Apart from this paper being the first to estimate birth order effects on human capital development in Lesotho, it adds to the literature in three other ways. Firstly, I use different measures of educational attainment, instead of one as is common in the literature. I estimate birth order effects on a short-run measure of education (that is, enrolment), and two long-run measures of education (that is, completed years of education, and *schooling progression* or age-adjusted schooling). It is interesting to not only know the birth order effects on short-term outcomes, but also its effects on long-term educational outcomes. Moreover, the fact that these measures have different strengths and weaknesses, as detailed below, means that estimating birth order effects on all of them will also allow one to check the robustness of the birth order results. Secondly, I examine two possible sources of heterogeneity in birth order effects: differences due to families of different sizes, and gender bias (that is, gender of the first-born). Lastly, I investigate whether family wealth and child-spacing can explain the observed birth order effects.

Different from the available evidence based on many developing countries data, I find large and significant negative birth order effect on child educational attainment. I also find strong evidence that this negative birth order effect is transmitted through birth-spacing, and not family wealth, contrary to earlier evidence from other developing countries (for example, Ecuador and some sub-Saharan countries) which shows that wealth is the underlying causal mechanism behind the positive, not negative, birth order effect on attainment (Tenikue and Verheyden, 2010; De Haan et al., 2014). Surprisingly, my results affirm the negative birth order effects found in developed countries (for example, Norway and the United States) (Black et al., 2005; De Haan, 2010). The difference with the latter evidence is that, in the United States, for example, De Haan (2010) finds that family wealth, and not birth-spacing, explains the negative birth order effect. Therefore, I tentatively conclude that these findings are consistent with the *confluence* model's predictions, even though I cannot rule out the hypothesis that first-borns do better because of being brought up under tougher parental disciplinary rules than their younger siblings.

The chapter proceeds as follows. Section 2 gives the context of Lesotho. Section 3 details the related literature. Section 4 describes the data and gives descriptive analysis. Section 5 discusses the empirical strategy: the fixed effects model of birth order effects. Section 6 presents the main birth order effect results, and the heterogeneities of birth order effect by family size and family gender preferences. Section 7 examines the pathways through which this birth order effect is transmitted. Section 8 concludes the chapter.

4.2 The context

Lesotho is a small landlocked, lower middle income sub-Saharan African country with an estimated population of 2 million. Like many other developing countries, unemployment, poverty and income inequality are major concerns. The proportion of people living below the national poverty line is 56.6 percent, and there is a great divide between the rich and the poor with an estimated Gini coefficient of 0.52 (Bureau of Statistics, 2006).

During the South African apartheid era, Lesotho served as a labour reserve for the South African mining sector. This situation helped create a culture of labour migration, particularly among prime-age men. Although the number of Basotho men currently working in the South African mines has drastically declined over the last twenty years, being a migrant miner is still regarded as the best employment avenue by many prime-age men. For example, according to the 2008 Labour Force Survey (2008 LFS), about 14 percent of individuals in the prime-age (25-54 years old) group were living outside the country (Bureau of Statistics, 2008), compared to about 0.7 percent, 1.1 percent, and 3.5 percent for those aged 5-9 years, 10-14 years, and 15-19 years, in that order. In the absence of adult males, young boys (mainly in rural areas) have to look after their families' livestock. According to the 2008 LFS, of all children aged 6-14 years, 2.9 percent are working, with girls making up only 0.3 percentage points of child labour force participation. Many of these children work as herd boys and do not attend school.

The FPE programme has been introduced to encourage these children to enrol and acquire basic education, which comprises seven years of primary education and three years of secondary education. Even though secondary education is still not free, there are programmes, which include scholarships for orphaned children and book rental schemes to reduce costs of textbooks, that are meant to assist disadvantaged children. Nonetheless, enrolment drops drastically when children transition from primary to secondary school.

4.3 Birth Order: Theory and Evidence

In this section, I review the literature on birth order effects on child outcomes. There are several theories that attempt to explain the birth order effects on human capital development of children: some theories predict negative birth order effects, while others predict positive birth order effects. I first review the theories that predict the negative birth order effects and the evidence consistent with this prediction. I then review the theories that predict positive birth order effects, and discuss the evidence consistent with this prediction.

Birth order has been an active research theme in the psychology literature where researchers are interested in its effects on individuals' intelligence. The confluence model, which predicts negative birth order effects on a child's intelligence level, is the operating theory in the psychology literature (Zajonc, 1976). According to this model, the intellectual performance of a child depends on his/her intellectual environment, which is a function of the average of the absolute intellectual levels (or age levels) of the child and his/her family members. For example, the first-born child enters a high intellectual environment with two adult parents. The second-born child enters a relatively lower intellectual environment because of the presence of her first-born sibling in the family, and the intellectual environment for the third-born is much lower. The model, therefore, predicts a negative relation between birth order and educational attainment.

Child-spacing (or birth spacing) can either perpetuate or attenuate the negative birth order effects. Parents can stimulate their offspring's intellectual ability through talking and playing with them. Thus, a large gap between siblings increases parent-child interaction, all else equal, and this can translate into better outcomes for the earlier-born sibling. Using American data, Baydar et al. (1997) find that the birth of a sibling results in changes in the family environment (e.g. changes in maternal labour participation and family income), a decline in maternal interactions with the older child, especially when the birth-space is short, and that mothers adopt controlling parenting styles toward the older sibling. They discover that these changes result in negative verbal development of the older child. So, a short birth interval between siblings is harmful to cognitive and non-cognitive development of the earlier-born child.

The longer birth interval can further perpetuate the negative birth order effects through the tutoring effect (Sulloway, 2007). According to this tutoring hypothesis, first-borns develop more intellectual abilities, through organisation and expression of thoughts, as they teach their younger siblings, while the last-borns have no one to teach. However, another implication of the confluence model is that a larger birth space between siblings increases the younger sibling's intellectual environment. This can attenuate the negative birth order effects. Therefore, the mediating effects of birth-spacing are ambiguous *a priori*. That is, the larger gap is better for both first- and second-borns, leading to ambiguous effects on birth order effects.

There is also the hypothesis that birth order effects have biological or prenatal origins. According to this hypothesis, later-borns face higher prenatal environmental risks because levels of maternal antibody increase with birth order (Gualtieri and Hicks, 1985). It, therefore, predicts a negative relation between birth order and child development. The evidence, however, seems to refute this hypothesis. Kristensen and Bjerkedal (2007) use Norwegian data to show that the negative birth order effect on intelligence depends on the child's social rank within the family, and not her birth order *per se*. For instance, second-born children who have deceased first-born siblings, and hence brought up as 'first-borns', have equally high levels of intelligence as biological first-borns.

Most economic theories of intra-household resource allocation emphasise the resource dilution hypothesis as the mechanism behind negative birth order effects on human capital development. The first-born child becomes the only-child of the family, hence she enjoys a higher stock of parental resources (including time and financial resources) than the later-born siblings who have to share parental resources with all other earlier born siblings. According to Cunha et al. (2006), early child investments are the most productive, and they increase the productivity of later investments. Therefore, the theory predicts that first-borns will have higher intellectual ability than their later-born siblings because of high investment enjoyed during the sensitive formative years. This first-born advantage gets larger the longer the birth space.

Hao et al. (2008), however, posit that the negative birth order effects arise endogenously due to the strategic interaction between parents and their offspring. Parents impose more stringent disciplinary measures on their first-born children in response to their bad behaviour and/or poor school performance in order to establish a reputation of toughness and deter similar behaviour amongst their later-born children. This increased attention on the first-born leads to better outcomes for the first-born relative to the later-borns. Using data from the National Longitudinal Survey of Youth, Hotz and Pantano (2013) find robust evidence that children's school performance declines with birth order, as does the toughness of their parents' disciplinary actions.

Black et al. (2005) use family fixed-effects models to tease out birth order effects on child outcomes from a large administrative dataset of Norwegians aged 16-74. They find large and statistically significant negative birth order effects on children's education and their later life outcomes, such as earnings and teenage childbearing. They then posit that their findings are consistent with the optimal fertility stopping model where parents stop having children if the last one has low endowments. It is, however, hard to divorce these findings from the predictions of models by Zajonc (1976), Cunha et al. (2006), and Hao et al. (2008).

De Haan (2010) uses the Wisconsin Longitudinal Study data to estimate birth order effects. To purge any potential endogeneity of birth order due to its correlation with family size and/or any other family size related factors, she estimates birth order effects separately for families with different number of children. Like Black et al. (2005), she finds strong negative birth order effects on completed years of education, evidently due to high parental financial transfers to earlier-borns, and not due to the average age gap between successive children posited by the confluence model. Härkönen (2014) also finds similar evidence in West Germany using family fixed-effects models, and propounds that the dilution of parental time is a plausible mechanism behind these effects. Likewise, Price (2008) finds evidence supporting the parental time dilution hypothesis using the American Time Use Survey data. He finds that parents spent significantly more quality time with their first-born children than with later-born children.

The evidence on birth order effects discussed thus far is consistent with the negative birth order effects on human capital development predicted by the confluence model, the resource-dilution, and strategic parenting hypotheses. This evidence is exclusively from the developed world. The evidence from developing countries, in contrast, reveals positive birth order effects.

How could birth order positively affect human capital development? The resource dilution hypothesis provides a different mechanism through which this could happen. Horowitz and Wang (2004) develop a model of intra-household allocation of resources (that is, time of children) across labour market and education activities when children are different in their human capital accumulation abilities.[1] As the sibship size increases, per capita familial resources decline. Therefore, according to this model, poor families supply too much labour of (that is, provide too little education to) the child with human capital accumulation comparative advantage (that is, the first-born) resulting in 'reverse specialisation'. That is, increased pressure on familial resources may force a poor family to send the first-born child out to work to compensate familial resources and finance education of her younger siblings. This will have a negative impact on human capital accumulation of the first-born child and a positive effect on attainment of the younger siblings. Tenikue and Verheyden (2010)'s model, which explicitly models child heterogeneity in terms of birth order, gives similar predictions to those of Horowitz and Wang's model.

Relatedly, Lafortune and Lee (2014) examine birth order and gender bias effects on human capital accumulation within a model that combines convex returns to education and credit constraints. In this model, higher birth order children are favoured for schooling in credit-constrained families. Additionally, the model predicts that if there are higher opportunity costs of educating boys compared to girls, for instance, having male siblings will lead to higher education for girls. An implicit prediction of this model is that a first-born male will

[1]Children are different either due to their innate abilities, gender, birth order or environmental factors.

have lower schooling relative to his later-born siblings, and much more so if he is the only son in the family. Lafortune and Lee (2014) use data from the United States, South Korea and Mexico to test the theory's predictions. They find that birth order is positively related with a child's schooling in low income families, but in high income families, the first-born gets more education. Moreover, they find that in South Korea, where boys are preferred over girls, having more younger female siblings benefits the boys.

Much of the evidence from developing countries consistently supports the theoretical prediction that birth order positively affects schooling. Applying the family fixed-effects estimation strategy on data from the Philippines, Ejrnæs and Pörtner (2004) find positive birth order effects on completed years of schooling, and that these effects are more pronounced in low-educated (or low-income) families. Tenikue and Verheyden (2010) also use family fixed effects strategy on data from 12 sub-Saharan African countries[2]. They find positive (negative) birth order effects on educational attainment, measured by completed years of education during the survey period, of children aged six to 18 years in poor (rich) families.

Tenikue and Verheyden (2010)'s results, however, are potentially confounded by measurement error in birth order. As the authors rightly acknowledged, there is a high likelihood that the observed 18 year-old child in the household is not the first-born but actually a second-born, if the first-born has moved out of the household and hence not observed. I show in the data analysis section below that this is a real threat in this study. In addition, Tenikue and Verheyden (2010) estimate family fixed effects models, but do not deal with the problem of increased correlation between child age and birth order within the family, which may also bias their results. I later on deal with this problem by including age fixed effects following Jayachandran and Kuziemko (2011) and De Haan et al. (2014). Finally, as pointed out by De Haan et al. (2014), Tenikue and Verheyden's finding that birth order effects are transmitted by wealth may be confounded with age effects, for example, because they do not estimate a fully interacted model.

De Haan et al. (2014) use Ecuadorian survey data on infants (the less than six year old children) and adolescents (the 12-18 year olds) to estimate long term effects of birth order on human capital development. They find positive and persistent birth order effects on achievement; that is, first-born children lag behind in educational achievement from infancy to adolescence, evidently due to mothers spending less quality time with first-borns, and breastfeeding them for a shorter period than later-born children. They also find that first-born adolescents are more likely be involved in child labour than their younger siblings. Furthermore, they find that the positive birth order effects on human capital development are larger in poor and low-educated families, but are negative in rich and high-educated families. This is consistent with Tenikue and Verheyden (2010)'s and Lafortune and Lee

[2] These countries are Benin, Burkina Faso, Cameroon, Ghana, Kenya, Mali, Niger, Senegal, Tanzania, Uganda, Zambia, and Zimbabwe.

(2014)'s findings.

To sum up, the confluence model, the resource dilution hypothesis, and the strategic-parenting hypothesis predict negative birth order effects on child outcomes. According to the confluence model, these birth order effects are wholly mediated by birth spacing, which may intensify or reverse them owing to the tutoring effect that kicks-in as sibship size increases. Empirical evidence from developed countries, namely the United States, Norway, and Germany, is largely consistent with the negative birth order effects predicted by these theories. On the contrary, available evidence from developing countries, namely the Philippines and Ecuador, reveals positive birth order effects on human capital development. This latter evidence is largely consistent with a different resource dilution hypothesis mechanism; in resource-poor families, an increase in sibship size reduces per capita resources, and may force the family to sacrifice schooling of the first-born child by sending the child to work to compensate familial resources.

Even though the evidence from developing countries largely supports the positive birth order effects and the family resource dilution hypothesis, it may not be generalised to other developing countries with different contexts. For instance, the context in all the countries included in Tenikue and Verheyden's study is different from that in Lesotho in two important ways. First, anecdotally, Lesotho has female (not male) bias in education. Second, the fertility rate (that is, average family size as used here) in Lesotho is about 3.3 births per woman, much lower than any of the countries in Tenikue and Verheyden (2010). Therefore, to the extent that birth order effects differ by family size, and that birth-spacing transmits birth order effects, we might expect to find different birth order effects in Lesotho.

4.4 Data and Descriptive Statistics

I use data from the Lesotho Population and Housing Census (the Census, hereafter) which was collected on the 19$^{\text{th}}$ April 2006. This data contains information on household (or family)[3] socio-economic background, employment status, school enrolment, highest level of education completed, family relations, and demographic characteristics of all family members. For each family, information is collected about all members, those present during the census and those temporarily absent, including their relation to the household head. As per the Census, the population of Lesotho is 1,868,526.

For purposes of this chapter, I assume that all children identified as the household head's children are his biological children. The majority of household heads are men, about 76 percent. I use information on family members' relations with the household head, irrespective of the gender of the household head, and age (in years at last birthday) to construct the

[3] In this chapter, I use family and household interchangeably.

absolute birth order measure: Birth order equals 1 for the first born, 2 for the second-born, et cetera. If there are polygamous families, then there will be measurement error in the birth order measure. However, the extent of polygyny in the 2006 Census data is only 0.74 percent. This is unlikely to cause any significant bias on the estimated birth order effects.

To get to the working sample, I sequentially impose the following restrictions. Firstly, I restrict the sample to all children of the household head aged between 6 and 18 years. This is to ensure that all primary and secondary school-going children are included. Following Tenikue and Verheyden (2010) and De Haan (2010), I further reduce the sample to households where the eldest child is at most 18 years old, and where there are a minimum of two and a maximum of five children. The second restriction, that families should have at least two children, is a technical requirement for studying birth order effects. Because the total fertility rate in Lesotho is 3.3 children per woman, limiting the sample size to households with a maximum of five children increases the chances that all children observed within the household are the biological children of the household head. In addition, by including families with up to five children increases chances that households with completed family sizes are included[4]. It also ensures that we have enough observations in each family size cell. This restriction, however, does not have implications for the estimations later on given that the fertility rate is exactly 3.3 in the final sample. Finally, I drop all families with multiple births[5] (for instance, twins) because of the ambiguities of assigning birth order in such instances. This restriction shrinks the sample size by 6.6 percent (and 5.5 percent decline in the number of families).

If there are some household head's biological children who have moved out of the household, this will introduce measurement error in the birth order measure. The incidence of this happening for those aged below 18 appears to be low. For instance, I find that only about 0.08 percent and 0.15 percent of children aged 6-18 years old are listed as household heads or spouses and household head's sons/daughters-in-law, respectively. These are children that must have moved out of their biological families. However, there is a real concern that many of those aged above 18 years may have moved out of their biological families. Figure A.1 shows that the proportion of those living outside their biological families (that is, young household heads and young sons- or daughters-in-law) increases just after age 18. Therefore, to deal with this problem, I restrict the analytical sample to include only children of primary-school-going age; that is, all 6-12 year olds. The advantages of this sample restriction are that (1) it is less likely that a 12-year old has an elder sibling who is outside the family, hence making the 12-year old a 'social first-born', and (2) all 6-12 year-old children are potential

[4] According to the 2009 Demographic Health Survey data, 95 percent of all women of productive age, 15-49, with five living children do not want to have any more children (SIF, 2011).

[5] The Census data does not come with birth date information, but only age in years at last birthday, even though the questionnaire also has a birth date question. Therefore, within each household, all household head's children with the same reported age are considered to be twins, triplets, etc.

beneficiaries of the FPE programme, making it easy to isolate birth order effects from FPE effects. The down-side, however, is that one cannot look at birth order effects on child labour participation on human capital development. These restrictions together reduce the working sample by 73 percent to 77,730 children living in 46,973 (or 9.6 percent fewer) families.

In sum, the working sample includes 6 to 12 year olds, coming from households with between 2 and five children, where the eldest child is aged at most 18 years.

Educational attainment is measured in three ways. The first measure is enrolment where a child aged between 6 and 12 years is considered enrolled if her parent reports that the child is enrolled. This is a short-run measure of education, but it has the advantage of being easy to calculate and interpret. The second measure is the number of completed years of education at the time of the census. The strength of this measure is that it is a long-term measure of education that is easy to read and interpret. Its limitation, however, lies in the fact that it is right-censored because many children are still in school and their final attainment will most likely differ from the currently reported. The third measure is schooling progression (or relative grade attainment or age-adjusted schooling), which is a long-term measure of education that is good in an environment where there is high grade repetition and high school entry delays. Schooling progression is defined as completed years of schooling divided by potential years of schooling, which are the total number of schooling years a child would have accumulated had she completed a year of schooling by age 7 and continued to add one more in each subsequent year (Mani et al., 2013).[6] The downside of schooling progression is that for any two children of different ages (say, 7 and 10 years old) with zero completed years of education, one cannot tell which one of the two is more disadvantaged than the other as they will both have zero schooling progression. Additionally, it is only defined for children aged 7 and older.

To get the completed years of education variable from the data, I use the highest level of education completed as is reported for the first seven grades, and then add one year for each of the three years of secondary, and two years of high school grades completed. Therefore, individuals who have completed high school must have 12 years of education, while those with graduate and post-graduate degrees have 16 and 18 years of education, respectively.[7] I then convert other qualifications into completed years of schooling as follows. *Non formal education* is converted to two years of schooling, *Diploma/Vocational training after primary* to 8 years, *Diploma/Vocational training after secondary* to 11 years of schooling, and *Diploma/Vocational training after high school* to 14 years of schooling.

One of the key control variables of interest is household wealth. The Census does not have household income or expenditure but does have information on household ownership of

[6]In this study, schooling progression equals $education/(age - 6)$, where education is completed years of education.

[7]This information is only relevant for the older siblings in the household.

Table 4.1: Summary Statistics

Variables	Number of children	Mean	Overall Std.Dev.	Between Family Std.Dev	Within Family Std.Dev
Outcome variables					
Enrolment	77730	0.911	0.285	0.231	0.172
Education	77471	2.512	1.720	1.403	1.144
Schooling progression	67124	0.807	0.406	0.357	0.206
Child characteristics					
First-Born	77730	0.204	0.403	0.228	0.326
Second-Born	77730	0.481	0.500	0.350	0.403
Third-Born	77730	0.249	0.433	0.350	0.306
Fourth-Born	77730	0.061	0.239	0.165	0.173
Fifth-Born	77730	0.005	0.073	0.041	0.056
Male	77730	0.506	0.500	0.418	0.314
Age	77730	9.110	1.992	1.443	1.515
Family characteristics					
Household head education	73996	5.168	3.920	3.969	0
Rural	77730	0.764	0.425	0.434	0
Children	77730	3.287	0.973	0.951	0
Wealth Index	77730	0.964	29.977	30.686	0
Number of families	46973				

Source: Own calculations from 2006 Census. *Notes*: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. Education is completed years of education during the census period, Schooling progression equals $education/(age-6)$, and Enrolment equals 1 if a child is reported to be enrolled by the parent. The sample size is smaller for schooling progression because, for children aged 6, the measure is unidentified. I use the uncentered PCA (see Banerjee, 2010 and Wittenberg and Leibbrandt, 2015) to calculate the wealth index from household ownership of durable goods (e.g. TV, radio, etc.), land (i.e. number of fields), and livestock (e.g. number of cattle, sheep, horses, etc.). I use Martin Wittenberg's Stata code to calculate this index.

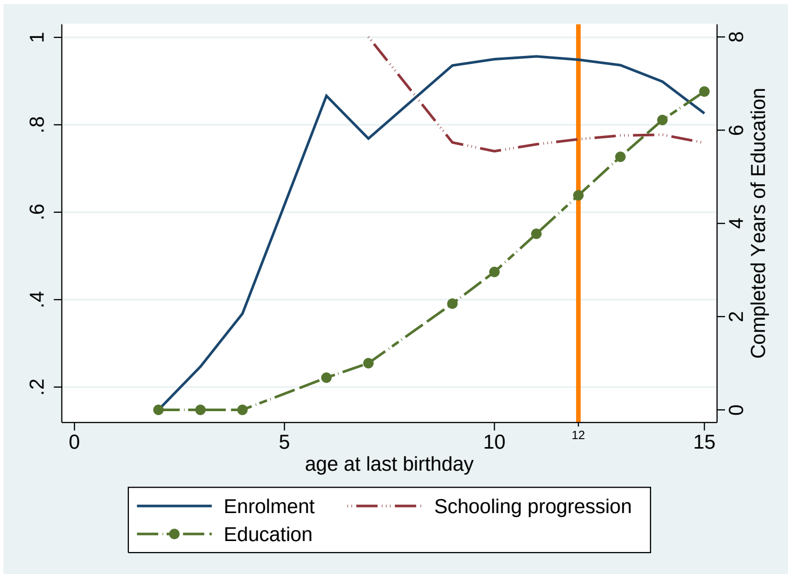
durable goods[8], land (that is, the number of fields), and livestock (for instance, number of cattle, sheep, horses, chicken, et cetera.). In such situations, the best available option widely applied in the literature is the use of wealth or asset indices constructed using the principal components analysis (PCA). Under the PCA, the variables are first standardised (that is, they are demeaned and divided by their standard deviations), and then the asset/wealth index is constructed as the first principal component of the correlation matrix. The main problem with this *centered* PCA index is that it tends to give negative scores to assets that are owned only in rural areas, such as having cattle, and hence exaggerates the rural-urban divide (Wittenberg and Leibbrandt, 2015). As Wittenberg and Leibbrandt (2015) argue, it gives us information about wealth and the degree of 'urbanness' in the data.

Given this, Wittenberg and Leibbrandt (2015) advocate for the use of the procedure proposed by Banerjee (2010) to construct an *uncentered* PCA asset index. Under this pro-

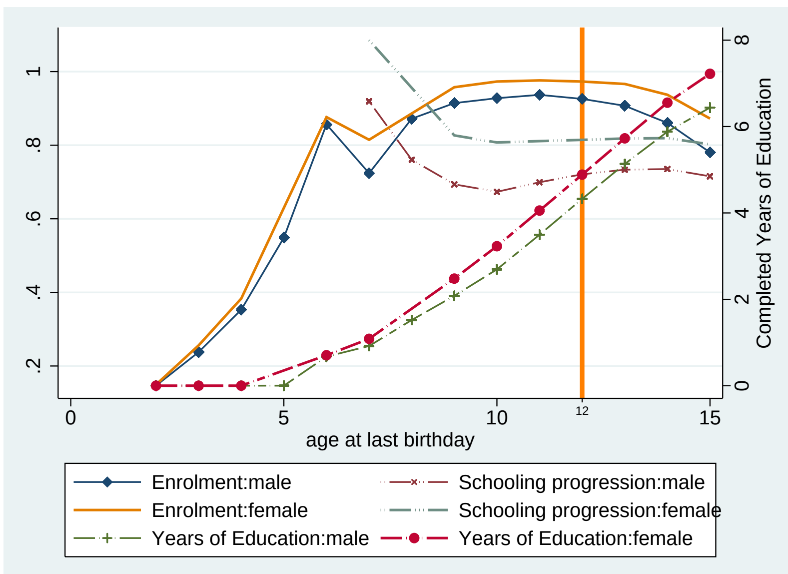
[8]These are radio, TV, telephone, cellular phone, refrigerator, bed/mattress, car, scotch cart, computer, and internet.

Figure 4.1: Educational Attainment, Age and Gender

(a) Educational Attainment by Age



(b) Educational Attainment by Age and Gender



Source: Own calculations from 2006 Census. *Notes*: The sample is all children aged 6-15 from families with 2-5 children, and where the oldest observed child is at most 18 years.

cedure, we first divide the variables (assets) by their respective means to form a matrix A, and then create the asset index as the first principal component of the non-negative square matrix $A'A$ (see Banerjee, 2010 and Wittenberg and Leibbrandt, 2015). Unlike the wealth index constructed using the ordinary PCA procedure, the uncentered PCA index is not only externally consistent in that it is a good proxy for household income and expenditure, but it is also internally consistent (Wittenberg and Leibbrandt, 2015)[9]. It, however, has the limitation that it pays more attention (that is, it gives large scores) to rare assets in the binary variable cases. Despite the differences, these indices (*the centered* PCA and *uncentered* PCA indices) are highly correlated in practice. I, therefore, construct the wealth index using the uncentered PCA procedure, and also create wealth quintiles from this index.

Table 4.1 presents the descriptive statistics of the outcome and control variables. We can see that about 91 percent of 6-12 year old children are enrolled, and their average completed years of education is 2.5. The average schooling progression (or relative grade attainment) is 0.81, which implies that children accumulate an average of 0.81 grades per year of schooling. The table also shows that there is more between-family variation than within-family variation in educational attainment. For example, the between- and within-family standard deviations in enrolment are 0.231 and 0.172, respectively.

Consistent with the 3.3 average children per family, there are more second-born children, about 48 percent, and relatively fewer first- and third-borns who, respectively, make up 20 percent and 25 percent. The share of fourth-borns is 6 percent, and the fifth-borns are least represented, making up only about 0.5 percent. The sample is equally split between males and females, and the average age is 9.1. In all these variables, unlike in educational attainment outcomes, there is almost as much within-family variation as there is between-family variation. The average household head's education is 5.3, and about 76 percent of children live in rural areas.

Figure 4.1 shows the relationship between educational outcomes and age. Figure 4.1a presents the overall picture, while Figure 4.1b shows the relationship by gender. We can see from Figure 4.1a that enrolment has an inverted U-shape relationship with age: it first increases with age up to 90 percent at age 10, and then starts to decline. Schooling progression first drops from about 1 grade per year of schooling at age 7 to a low of 0.75 at age 10. It then steadies just below 0.8 grades per year of schooling. On the other hand, completed years of education have a strong positive relationship with age; they strongly increase with age from age 6. Figure 4.1b reveals that, across all three measures of educational attainment, girls outperform boys. Therefore, in order to tease out the association between birth order and educational outcomes, one needs to control for age, among other potential confounding

[9] An internally consistent asset index ranks individuals with more of anything good (in this case assets) higher than individuals with less. Therefore, by satisfying these criteria, the uncentered PCA wealth index ranges from zero upwards, i.e. it is never negative, and can be used to calculate asset inequality.

factors. I turn to this next.

4.5 Empirical Model

To estimate the effect of birth order on educational attainment, I follow De Haan et al. (2014) and estimate a family fixed effects model to address the endogeneity between birth order and observed and unobserved family specific fixed effects, including family size. The model is specified as

$$y_{if} = \alpha + \beta_2 \cdot Second_{if} + \beta_3 \cdot Third_{if} + \beta_4 \cdot Fourth_{if} + \beta_5 \cdot Fifth_{if} + X_{if}\theta + \lambda_f + \varepsilon_{if} \quad (4.1)$$

where y_{if} is the outcome of child i in family f; $Second$, $Third$, $Fourth$, and $Fifth$ are dummies for second, third, fourth, and fifth-born children, respectively; X_{if} is a vector of controls, including dummies for age and gender of the child, family wealth and birth order interaction dummies, and λ_f are family fixed effects. I add age dummies (that is, age fixed effects or birth cohort effects) to address the fact that the correlation between age and birth order within a family is high, which could bias the results. Controlling for birth cohort effects also addresses the fact that later-borns may face different educational opportunities compared to first-borns, either due to changes in parental tastes for education or otherwise. In the OLS specification, X_{if} also includes location, household head's education, number of children in the family (or family size) as dummies. The estimation results allow for any arbitrary correlation within the family.

4.6 Results

In this section, I first present the chapter's main results on birth order effects on educational attainment. Thereafter, I discuss the heterogeneities in birth order effects by family size and gender, and then discuss some sensitivity checks to the main results.

4.6.1 Birth order effects on educational attainment

Table 4.2 presents the estimated effects of birth order on educational attainment (that is, enrolment, completed years of education, and schooling progression) of 6-12 year old children. For each outcome, the table reports estimates for two different estimation strategies: the OLS estimates are reported in columns 1, 3, and 5; and the family fixed effects model estimates are reported in columns 2, 4, and 6. All specifications include dummies for child age, to control for cohort effects and sample selection bias[10], child gender, location, and location interacted with gender. In the OLS regressions, I also include household head's years of education, and a full set of dummies for the number of children in the family (four in all), which control for the correlation between birth order and family size.

Looking at the OLS estimates in columns 1, 3, and 5 of the table, they reveal a significant positive relation between birth order and educational attainment. For example, relative to the first-born, an average second-born child has 0.05 more years of education, and accumulates 0.01 more years of education per year of schooling. We also see that educational attainment is negatively associated with family size, and this is consistent with previous findings by Black et al. (2005), De Haan (2010), De Haan et al. (2014), and Ponczek and Souza (2012).

However, these results cannot be given any causal interpretation because of the possible bias due to omitted unobserved family factors. In fact, we have seen in Table 4.1 that there is more between-family variation than within-family variation in all educational attainment measures. Therefore, the positive OLS birth order coefficients could be due to differences in educational attainment across families, not necessarily due to differential investments on children of different orders of birth within families. This is particularly so because households with only one child aged between 6 and 12 years old are only included in OLS regressions, and not in the family fixed effects estimations. So much of the variation comes from between rather than within families. Table A.1 reports the OLS results from a restricted sample of households with at least 2 children between the ages of 6 and 12. These results turn negative, although not always significant, clearly indicating this potential bias. Nonetheless, I follow the literature (see for example De Haan et al., 2014) and report the OLS results based on the working sample described above.

In order to isolate the effect of birth order on educational attainment, I turn to family fixed effects specifications presented in columns 2, 4, and 6, respectively, for enrolment, completed years of education, and schooling progression. We immediately see that the estimates for educational attainment show a reverse birth order pattern from that produced by the OLS regression estimates, and that the fixed effects estimates are much larger, in absolute values, than the OLS estimates. This suggests that OLS estimates are potentially upward biased.

Looking at enrolment results in column 2 of the table, I find that higher birth order children are less likely to be enrolled compared to their first-born sibling, and this effect increases, in absolute values, with birth order. For example, relative to the first-born, a second-born has 6 percentage points less probability of enrolment, while the fifth-born is

[10]Since educational attainment is only observed for children from families with at least two children aged between 6 and 12 years old, sample selection bias could arise. Controlling for age dummies makes the estimates consistent because, conditional on family fixed effects, whether or not we observe the child's outcome is wholly determined by her age at the time of the census (see De Haan et al., 2014).

Table 4.2: The Effect of Birth Order on Educational Attainment

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Enrolment		Education		Schooling progression	
	OLS	FE	OLS	FE	OLS	FE
Second-Born	0.00304	-0.0580***	0.0458***	-0.114***	0.0122***	-0.130***
	(0.00218)	(0.00620)	(0.0111)	(0.0207)	(0.00217)	(0.0105)
Third-Born	0.00402	-0.131***	0.0954***	-0.182***	0.0209***	-0.264***
	(0.00325)	(0.0121)	(0.0140)	(0.0398)	(0.00328)	(0.0205)
Fourth-Born	-0.0115*	-0.221***	0.133***	-0.255***	0.00826	-0.424***
	(0.00638)	(0.0184)	(0.0197)	(0.0595)	(0.00636)	(0.0312)
Fifth-Born	-0.0248	-0.303***	0.192***	-0.332***	-0.00488	-0.626***
	(0.0204)	(0.0311)	(0.0429)	(0.0919)	(0.0210)	(0.0554)
Three children	-0.00430		-0.0452***		-0.0113***	
	(0.00269)		(0.0111)		(0.00298)	
Four children	-0.0140***		-0.123***		-0.0270***	
	(0.00333)		(0.0131)		(0.00345)	
Five children	-0.0326***		-0.245***		-0.0552***	
	(0.00464)		(0.0175)		(0.00453)	
Constant	0.859***	0.997***	0.647***	1.103***	0.682***	1.292***
	(0.00531)	(0.0216)	(0.0192)	(0.0828)	(0.00657)	(0.0373)
Observations	73,996	73,996	73,758	73,758	73,758	63,917
R-squared	0.075	0.099	0.656	0.775	0.116	0.118
Number of Families		44,732		44,672		43,316

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

30 percentage points less likely to be enrolled. This effect is sizeable. Relative to the average enrolment rate of 91 percent in the sample, this means that being a fifth-born reduces the enrolment probability by 33 percent (= 30/90) or 0.77 standard deviations ($= 0.30 \times (0.073/0.285)$).

Column 4 of Table 4.2 shows that a second-born completes 0.11 less years of education compared to her elder sibling. The later-born's disadvantage increases to 0.33 fewer years of education for the fifth-born. Similarly, there is a strong negative birth order effect on schooling progression (or relative grade attainment) in column 6, and the absolute effect is increasing with birth order. For example, relative to the first-born, a second-born child accumulates 0.13 less years of education (or grades) per year of schooling, while the fifth-born accumulates 0.63 fewer years of education for each year of schooling. This implies that first-borns progress much faster in school than their younger siblings, either due to early school entry or less grade repetition.

Overall, these results are in stark contrast with the previous evidence from other developing countries, including sub-Saharan Africa, showing strong positive birth order effects (Ejrnæs and Pörtner, 2004; Tenikue and Verheyden, 2010; De Haan et al., 2014). The positive birth order effects have been interpreted as partial evidence for the financial resource dilution hypothesis. In other words, it implies that liquidity constraints force families to invest less on earlier born children in spite of their comparative advantage in human capital accumulation (Horowitz and Wang, 2004; Tenikue and Verheyden, 2010; De Haan et al., 2014). These results surprisingly affirm evidence from the developed world, and this is the first indication that they are less likely to be explained by the liquidity effect hypothesis (Black et al., 2005; De Haan, 2010).

One difference between these results and those of Tenikue and Verheyden (2010) and De Haan et al. (2014), for instance, is that they attach greater weight to households with more young children and fewer teenage children. I later on show that these results are robust to the 6-12 year olds sample restriction.

The evidence presented here is consistent with the predictions of the confluence and 'strategic' parenting models of Zajonc (1976) and Hao et al. (2008). The first-borns not only have higher enrolment rates, but they also progress faster in school than their younger siblings, potentially due to the high quality investments they must have enjoyed while young (Cunha et al., 2006; Heckman and Mosso, 2014). These childhood investments could have come in the form of high family intellectual environment, teaching younger siblings, breastfeeding longer than the younger siblings, and/or stricter parenting that instils much better discipline. It is important, therefore, to investigate some of these possible mechanisms through which these birth order effects are being propagated. But, before doing that, I first look at the differential effects of birth order according to family size and child gender.

4.6.2 Heterogeneities in birth order effects: family size and gender

Birth order effects may be different due to different family environments, for instance, different family sizes and differences in family gender preferences (Zajonc, 1976; Jayachandran and Kuziemko, 2011; Lafortune and Lee, 2014). Here I specifically test two hypotheses. First, if the confluence model is a good candidate model for explaining the observed birth order effects, then we should expect to see more pronounced birth order effects in large families because, all else equal, younger children are born in a low intellectual environment, and have to compete for limited parental time with their elder siblings. Second, the first-borns develop more skills as they teach their younger siblings. However, as families grow larger, birth order effects might get smaller (in absolute terms) or even dissipate for the middle-born children as the tutoring effect kicks in. That is, middle-born children may develop skills through teaching their younger siblings, thereby reducing the knowledge gap between them and the first-born.

According to Jayachandran and Kuziemko (2011), if parents prefer sons to daughters, girls will be weaned faster than boys so that parents can try again for a son, and once a boy is born, they nurse him for a longer time to reduce their fecundity (see also Zajonc, 1976). To wit, parents may passively invest differently in children of different birth orders based on their gender preferences. It follows then that the second hypothesis I test is that, if parents prefer girls' education over that of boys, then the negative birth order effects will be stronger for first-born girls than first-born boys. That is, these negative effects will be attenuated (or even reverse signs) in first-born boy families. I follow De Haan et al. (2014) in studying these two possible sources of heterogeneity in birth order effects.

Birth order effects by family size

I first explore the birth order effects by family size. Table 4.3 presents family fixed effects estimation results for families with two, three, four, and five children, in that order. The three columns present results for educational attainment outcomes: enrolment, years of education, and schooling progression, in that order. By and large, these results are consistent with the main findings: birth order negatively affects educational attainment, and the absolute effect intensifies with birth order.

In line with De Haan et al. (2014), I find that birth order effects are different for different family sizes. Looking at schooling progression results in column 3, for example, we can see that a second-born child in a two-child family accumulates 0.09 fewer years of education for each year of schooling relative to her first-born sibling, while a second-born in a four-child family accumulates twice as much less years of education per year of schooling as her counterpart in a two-child family: she accumulates 0.17 fewer years of education for each year of schooling than her first-born sibling. The second-born disadvantage relative the first-born

Table 4.3: Birth Order Effect on Education Attainment and Child labour, Fixed Effects by Family Size

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
	Two-child families		
Second-Born	-0.0411***	-0.102**	-0.0912***
	(0.0125)	(0.0446)	(0.0244)
Observations	17,812	17,771	15,507
	Three-child families		
Second-Born	-0.0467***	-0.111***	-0.109***
	(0.00985)	(0.0340)	(0.0175)
Third-Born	-0.101***	-0.312***	-0.220***
	(0.0198)	(0.0671)	(0.0349)
Observations	26,882	26,803	23,237
	Four-child families		
Second-Born	-0.0869***	-0.0900**	-0.170***
	(0.0118)	(0.0378)	(0.0191)
Third-Born	-0.162***	-0.285***	-0.326***
	(0.0221)	(0.0708)	(0.0367)
Fourth-Born	-0.243***	-0.577***	-0.516***
	(0.0332)	(0.106)	(0.0551)
Observations	19,570	19,494	16,799
	Five-child families		
Second-Born	-0.0493**	-0.0733	-0.142***
	(0.0194)	(0.0641)	(0.0267)
Third-Born	-0.132***	-0.221**	-0.293***
	(0.0329)	(0.102)	(0.0485)
Fourth-Born	-0.224***	-0.461***	-0.445***
	(0.0478)	(0.147)	(0.0713)
Fifth-Born	-0.286***	-0.734***	-0.633***
	(0.0643)	(0.194)	(0.0992)
Observations	9,732	9,690	8,374

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

drops to 0.14 grades per year of schooling in a five-child family. A similar pattern is observed for enrolment. The second-born in a two-child family is 4 percentage points less likely to enrol compared to the first-born, while her counterparts in four- and five-child families are, respectively, 9 and 5 percentage points less likely to enrol relative to the first-born. There is no such clear birth order effects pattern in the case of completed years of education.

To formally test for the equality of the coefficients, I estimate three fully interacted fixed effects models where every variable, not just birth order, is interacted with family size. This is equivalent to jointly estimating panels 1-4 of Table 4.3 for each outcome variable. The results are reported in Table A.2. The interactions between birth order and family size are statistically significant for enrolment and schooling progression, which implies that birth order effects on enrolment and schooling progression get larger (in absolute values) as the sibship size increases, but not on completed years of education. The pattern is more robust for schooling progression where all birth order interactions with family size are negative and statistically significant at 5 percent level. Furthermore, these effects seem not to dissipate as the family size increases.

Birth order effects by gender

I now turn to the heterogeneity in birth order effects by gender of the child. To disentangle birth order effects purely driven by gender biases from those due to sibling sex composition, I compare birth order effects for families of first-born sons versus those for families of first-born daughters (De Haan et al., 2014). This is a relevant margin to examine because, if parents do have strong preferences for girls, a favourable treatment for girls should give us stronger birth order effects among families with first-born girls. Put differently, the first-born boys will have less or no advantage over their younger siblings.

Table 4.4 presents family fixed effects birth order estimates for two separate samples of, respectively, families with first-born boys (in columns 1 to 3), and families with first-born girls (in columns 4 to 6).[11] There are huge differences in birth order effects across these families, and these results are largely consistent with the girl-bias hypothesis. While birth order effects on completed years of education for families with first-born girls are strongly negative, and consistent with the main results, there are no significant birth order effects on completed years of education in families with first-born boys. We can further see that birth order effects on enrolment and schooling progression in first-born-girl families are slightly larger, in absolute values, than in first-born-boy families, especially for fifth-borns relative to the first-born. Taking the fifth-borns in first-born-boy families, for example, we can see that, compared to the first-born, they are 28 percentage points less likely to be enrolled, and they

[11] These families (that is, first-born boy and those of first-girl families) are identical in terms of the average years of education. The mean years of education for first-born boy families is 9.59 versus 9.57 for first-born girl families, with a p-value of 0.5878.

Table 4.4: The Effect of Birth Order on Educational Attainment and Child labour, Fixed Effects by Gender of the First-Born

VARIABLES	First-Born is a Boy			First-Born is a Girl		
	(1)	(2)	(3)	(4)	(5)	(6)
	Enrolment	Education	Schooling progression	Enrolment	Education	Schooling progression
Second-Born	-0.0612***	-0.0220	-0.138***	-0.0616***	-0.196***	-0.119***
	(0.00948)	(0.0315)	(0.0157)	(0.00878)	(0.0287)	(0.0150)
Third-Born	-0.123***	-0.0655	-0.257***	-0.143***	-0.293***	-0.266***
	(0.0181)	(0.0596)	(0.0302)	(0.0167)	(0.0534)	(0.0285)
Fourth-Born	-0.207***	-0.117	-0.400***	-0.237***	-0.390***	-0.440***
	(0.0275)	(0.0889)	(0.0461)	(0.0250)	(0.0796)	(0.0426)
Fifth-Born	-0.278***	-0.164	-0.575***	-0.331***	-0.507***	-0.667***
	(0.0457)	(0.138)	(0.0834)	(0.0428)	(0.122)	(0.0741)
Observations	36,605	36,475	31,675	37,391	37,283	32,242
R^2	0.089	0.753	0.135	0.110	0.795	0.104

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

accumulate 0.6 fewer grades per year of schooling, while their counterparts in first-born-girl families are 33 percentage less likely to enrol, and accumulate 0.7 fewer grades for each year of schooling.

I formally test for the equality of birth order effects between first-born-boy and first-born-girl families by estimating fully interacted models, where every variable is interacted with a dummy for first-born-girl families. The results are presented in Table A.3. The results show that there are statistically significant differences in birth order effects on completed years of education between first-born-boy and first-born-girl families, but not on enrolment and schooling progression. For example, a second-born in a first-born-girl family has 0.193 less years of education than his/her elder sister, which is 0.18 years lower than what his/her counterpart in a first-born-boy family gets. This implies that later-borns are more disadvantaged, in terms of completed years of education, if they have a first-born sister than when they have a first-born brother.

To sum up, the evidence presented in this subsection shows that birth order effects on educational attainment are larger, in absolute terms, in larger families, and they appear not to evaporate with the increase in family size. Moreover, birth order effects are larger in first-born-girl families, and this is consistent with the hypothesis that parents prefer to educate girls more than boys. The fact that first-borns' advantage increases with sibship size could possibly be due to: (1) the dilution of familial resources as the family size increases, (2) the lower intellectual environment, relative to the first-born, that later-borns are born into, or (3) the strict disciplinary environment that the first-born are brought up with. In the following section, I narrow down the potential transmission mechanisms of these birth order effects. I begin with checking the sensitivity of the main results.

4.6.3 Sensitivity checks

In this subsection, I present sensitivity checks of the main birth order results to different sample restrictions made to reduce measurement error in the birth order measure. The sample restrictions imposed in Section 4.4 are: (1) that families should have at least two children, with the eldest aged at most 18 years old, and (2) that families should have a maximum of five children. I further restricted the analytical sample to children aged 6-12 years, which could introduce sample selection bias because educational attainment is only observed for some but not all children in families with 6-18 year old children. I first relax these restrictions and compare the results with the main results which are based on stricter sample selection criteria. Relaxing these restrictions increases the sample size as well as the measurement error.

Table 4.5 presents fixed-effects model results based on a much bigger sample of 6-18 year-old children from families where the eldest child is at most 18 years old. Since the oldest child

Table 4.5: The Effect of Birth Order on Educational Attainment: 6-18 year-old children sample

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
	Enrolment		Education		Schooling progression	
	OLS	FE	OLS	FE	OLS	FE
Second-Born	0.00420**	-0.0106***	0.0275***	-0.0408**	0.0111***	-0.0464***
	(0.00194)	(0.00365)	(0.0103)	(0.0173)	(0.00182)	(0.00390)
Third-Born	0.00883***	-0.0219***	0.107***	-0.0184	0.0183***	-0.0939***
	(0.00303)	(0.00687)	(0.0138)	(0.0323)	(0.00377)	(0.00719)
Fourth-Born	-0.00131	-0.0455***	0.199***	0.0579	-0.00705	-0.163***
	(0.00609)	(0.0107)	(0.0203)	(0.0484)	(0.00943)	(0.0122)
Fifth-Born	-0.00983	-0.0749***	0.301***	0.118	-0.0578	-0.245***
	(0.0201)	(0.0234)	(0.0468)	(0.0889)	(0.0404)	(0.0394)
Three children	-0.00516**		-0.0347***		-0.00895***	
	(0.00247)		(0.0129)		(0.00262)	
Four children	-0.0146***		-0.119***		-0.0289***	
	(0.00298)		(0.0152)		(0.00301)	
Five children	-0.0413***		-0.311***		-0.0564***	
	(0.00414)		(0.0212)		(0.00389)	
Constant	0.858***	0.907***	0.623***	1.007***	0.992***	1.148***
	(0.00506)	(0.0149)	(0.0208)	(0.0737)	(0.00799)	(0.0166)
Observations	123,603	123,603	123,304	123,304	113,463	113,463
R^2	0.153	0.170	0.740	0.768	0.147	0.113
Number of Families		49,484		49,460		49,458

Notes: The sample is all children aged 6 to 18 years old from families with 2-5 children, and where the oldest observed is at most 18 years. All regressions include household head's education, dummies for age (in years), correcting for cohort effects, and gender of the child, and dummies for location, and location interacted with child's gender. Enrolment is a dummy which equals 1 if a child reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

observed in the family may not necessarily be the eldest child, there is high measurement error in the birth order measure. As we can see from the table, the birth order effects on educational attainment are still negative, and are very similar to the results reported in Table 4.2, in spite of the increased measurement error in the birth order measure. However, the birth order effects in this sample appear to be underestimated (that is, biased towards zero).

The possibility that the observed oldest child in the household might not be the eldest living child of the household forced one to restrict the analytical sample to 6-12 year old children. Thus, the results may suffer from sample selection bias, even though controlling for age fixed effects is an attempt to address this problem, as mentioned in footnote 10.

To formally check the robustness of the results against this potential sample selection bias, I estimate birth order effects using a sample of families of 'young mothers'. I define 'young mothers' as all women aged 35 years and below, and are either household heads or household head spouses. These are mothers who are less likely to have children older than 18 years, which allows one to estimate birth order effects on a sample of 6-18 year olds with little measurement error.

According to the 2009 Lesotho Demographic Health Survey report (SIF, 2011), Lesotho's

Table 4.6: Birth Order Effects on Educational Achievement: Young Mothers Families

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	-0.0241***	-0.0475**	-0.0667***
	(0.00524)	(0.0239)	(0.00652)
Third-Born	-0.0599***	-0.0444	-0.141***
	(0.0102)	(0.0451)	(0.0125)
Fourth-Born	-0.112***	0.0351	-0.225***
	(0.0173)	(0.0702)	(0.0241)
Fifth-Born	-0.118***	0.167	-0.265***
	(0.0443)	(0.156)	(0.0781)
Constant	0.922***	1.001***	1.158***
	(0.0177)	(0.0843)	(0.0219)
Observations	56,750	56,586	50,311
R^2	0.113	0.789	0.113

Notes: Fixed-effects regression results. The sample is all children aged 6 to 18 years old from families of young mothers (i.e. women aged 35 years or less and are household heads or household head spouses) with 2-5 children, and where the oldest observed is at most 18 years old. All regressions include household head's education, dummies for age (in years), correcting for cohort effects, and gender of the child, and dummies for location, and location interacted with child's gender. Enrolment is a dummy which equals 1 if a child reported as enrolled. Estimates are from fully interacted fixed effects models where all birth order dummies and controls variables are interacted with the wealth index. Education is completed years of schooling. Schooling progression equals $education/(age - 6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

overall fertility rate has been on the decline since the 1970s, but that of women aged 15-19 years has been on the rise over the same period. Between 2004 and 2009, the proportion of 15-19 year old women who have ever given birth to more than one child increased from 0.8 percent to 1.5 percent. This implies that in the 1980s, when most of our young mothers were teenagers, giving birth to at least two children as a teenager was a rarity. If we assume that a woman gave birth for the first time at age 17, her eldest child, if alive, must be 18 years old in 2006. Therefore, restricting the sample to families of mothers aged at most 35 years essentially ensures that there are no over-eighteen year old siblings omitted in the sample, and hence provides the most accurate birth order measure among the 6-18 year olds, although at a cost of a relatively small sample size.

Table 4.6 presents fixed effects model results using a sample of young mothers' families with at least two children aged between 6 and 18 years old, where the eldest is at most aged 18. We can see from the table that birth order effects on educational attainment and child labour participation are negative. However, it appears that the birth order effects on educational attainment in this table are almost half, in absolute values, the main effects reported in Table 4.2, but larger than those reported in Table 4.5. Notwithstanding these differences, it appears that, by and large, the results presented in Tables 4.5 and 4.6 are qualitatively the same as the main birth order effects findings. Therefore, I conclude that the main results are not very sensitive to the imposed sample selection criteria.

4.7 Potential pathways of negative birth order effects

In this section I investigate two mechanisms through which birth order effects can be propagated: the liquidity effect (or wealth) channel, and the birth-spacing (or child-spacing) channel. The decision to look at these two potential transmission mechanisms is purely based on data availability. I begin with the wealth channel and then turn to the child-spacing channel.

4.7.1 Family wealth

According to the liquidity effect hypothesis, as the household gets larger, per capita household resources get depleted, and this may force the household to send the first born to the labour market to relax the household's budget constraint (Tenikue and Verheyden, 2010). This hypothesis predicts positive birth order results, contrary to the evidence presented in this chapter thus far.

In order to test whether family wealth drives the observed birth order effects, I use the family wealth index discussed earlier. The wealth index reflects the accumulated pattern of household income and expenditure to date, including education investments. For example, there might be a positive effect of unobserved permanent income on observed assets and observed education investment. However, households may invest more (less) in children's education and less (more) in asset accumulation, conditional on permanent income. This will attenuate any relationship between education and "wealth" when using the wealth index. In spite of this potential attenuation bias, Tenikue and Verheyden (2010) use a similar wealth index from the Demographic Health Surveys and find significant negative wealth effects in a number of sub-Saharan Africa. The consistency of their results across different countries, therefore, allays fears that the results reported in this chapter are an artefact of the wealth index used.

Table 4.7 discloses the fixed effects birth order results, with wealth index interactions. Following De Haan et al. (2014), I estimate fully interacted fixed effects models where birth order dummies and included control variables are interacted with the wealth index. We can see from the table that birth order effects are largely consistent with the earlier findings: birth order negatively affects educational attainment. More importantly, these results do not support the hypothesis that the observed birth order effects are transmitted through wealth. All the wealth-index–birth-order interactions are insignificant, and the coefficients are approximately equal to zero. Moreover, comparing the birth order results from this table with the main birth order results reported in Table 4.2, we can see that the magnitudes

Table 4.7: Fixed Effects Estimates of Birth Order, Interacted with Wealth Index

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	-0.0578***	-0.114***	-0.130***
	(0.00620)	(0.0207)	(0.0105)
Third-Born	-0.130***	-0.181***	-0.263***
	(0.0121)	(0.0398)	(0.0205)
Fourth-Born	-0.220***	-0.255***	-0.423***
	(0.0184)	(0.0595)	(0.0312)
Fifth-Born	-0.304***	-0.334***	-0.628***
	(0.0312)	(0.0921)	(0.0559)
Second-Born×wealth_index	4.57e-05	-0.000282	0.000638
	(0.000411)	(0.00119)	(0.000886)
Third-Born×wealth_index	0.000287	-0.000774	0.00109
	(0.000874)	(0.00227)	(0.00207)
Fourth-Born×wealth_index	-9.11e-05	-0.000671	0.00148
	(0.00129)	(0.00330)	(0.00286)
Fifth-Born×wealth_index	0.00109	-6.28e-05	0.00284
	(0.00180)	(0.00470)	(0.00440)
Constant	0.962***	1.121***	1.237***
	(0.0273)	(0.114)	(0.0449)
Observations	73,996	73,758	63,917
R-squared	0.099	0.776	0.118

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

are almost the same. Therefore, based on these results, there is no evidence to support the hypothesis that the observed birth order effects are propagated by wealth.

4.7.2 Child-spacing

I now turn my attention to the birth- or child-spacing channel. According to the confluence model, birth order effects on human capital development are wholly transmitted by the birth interval between successive siblings (Zajonc, 1976).

For the analysis that follows, I define birth-spacing as the actual age gap between two successive siblings within a family. For example, the age between the first- and second-born is given by the difference between their respective ages. Because the first-born follows no one, the age gap for this child is undefined.

The regression results for birth order, child-spacing, and their interaction (with family fixed effects) are presented in Table 4.8. As in the case of wealth effects above, these results are from fully interacted models where all variables, not just birth order, are interacted with birth-space. Looking at these results, we can see that birth order effects on enrolment, completed years of education, and schooling progression are still negative, as in the main results. In addition, the magnitudes of the coefficients are slightly larger than those reported in the main results, Table 4.2, particularly the fifth-borns' disadvantage in completed years of education relative to first-born where the difference is largest, about 0.1 years.

More interestingly, birth order interactions with child-spacing are negative and statistically significant in the completed years of education regression. This indicates that child-spacing amplifies the negative birth order effects on completed years of education. For example, increasing the age gap between the third- and fourth-born by one year reduces the fourth-born's completed years of education by 0.24 (that is, $-0.230 - 0.00848$) years relative to the first born.

These results could be an artefact of infant and child mortality rates that change the actual birth spacing pattern. Nonetheless, according to the Ministry of Health et al. (2005), the difference between the mean number of children ever-born (i.e. 1.80) and the mean number of living children (i.e. 1.62) of women aged 15 to 49 years is 0.18. This is a numerically small difference. Therefore, infant and child mortality rates are likely to have a minimal influence, if any, on the observed birth spacing effects.

These results are consistent with models that emphasise the importance of early quality investments (in the form of high intellectual enrolment and/or strict parenting which instils discipline) enjoyed by the first-born children as the driving force behind birth order effects. Without better data, however, it is difficult to pin down which mechanism, whether the intellectual environment or the strict parenting, birth-spacing actually captures in this context.

Table 4.8: Effect of the child-spacing, birth order and their interaction on educational development and child labour, with Fixed Effects

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	-0.0596***	-0.116***	-0.132***
	(0.00623)	(0.0209)	(0.0105)
Third-Born	-0.126***	-0.177***	-0.263***
	(0.0121)	(0.0397)	(0.0204)
Fourth-Born	-0.214***	-0.230***	-0.423***
	(0.0187)	(0.0601)	(0.0314)
Fifth-Born	-0.311***	-0.405***	-0.654***
	(0.0350)	(0.104)	(0.0619)
Second-Born×Birth-Space	-0.000792	-0.00649	-0.00285
	(0.00144)	(0.00412)	(0.00308)
Third-Born×Birth-Space	-0.00118	-0.00716*	-0.00305
	(0.00145)	(0.00412)	(0.00308)
Fourth-Born×Birth-Space	-0.00104	-0.00848**	-0.00299
	(0.00150)	(0.00421)	(0.00314)
Fifth-Born–Birth-Space (omitted)	0	0	0
	(0)	(0)	(0)
Constant	0.994***	1.183***	1.298***
	(0.0220)	(0.0842)	(0.0380)
Observations	77,730	77,471	67,124
R-squared	0.098	0.776	0.121

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age - 6)$. Birth-space is the actual age gap between successive siblings. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

4.8 Conclusion

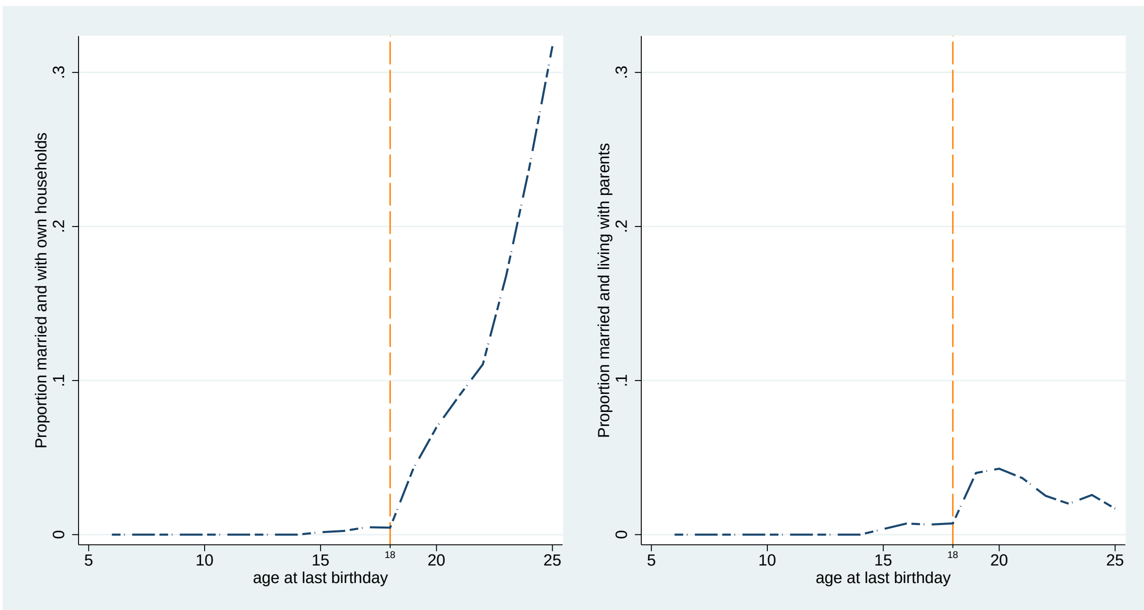
In this chapter, I examine the effect of birth order on educational attainment in Lesotho using the 2006 Census data for children aged 6 to 12 years. Applying family fixed effects models, I find robust negative birth order effects on all measures of educational attainment (enrolment, completed years of education, and schooling progression). These results are in stark contrast with previous evidence from developing countries, but much in line with that from the developed world, and hence strongly challenge the idea that birth order effects in developing countries are largely due to wealth constraints.

I also investigate heterogeneities in birth order effects due to family size and family gender bias, and find that birth order effects are more pronounced in large families. I take this as part evidence for the confluence model predictions. Furthermore, there is strong evidence that girls' education is favoured over that of boys.

Turning to the potential pathways through which these birth order effects operate, there is evidence that they are mainly transmitted through the average age gap between two close siblings. I find no support for the hypothesis that birth order effects are propagated through family wealth as previously found in other developing countries. Taken together, the evidence presented here largely supports the notion that it is early quality investments enjoyed by earlier-born children that gives them a comparative advantage over their siblings in human capital accumulation. More research is, however, needed to test many other potential pathways that could strengthen this explanation. For example, do pre- and post-natal parental investments on children vary with birth order?, If so, how? Answering these questions could help narrow down the main transmission mechanisms of these birth order effects.

Appendix

Figure A.1: Proportion of children out of their biological families



Source: Own calculations from 2006 Census. *Notes*: The left panel of the figure shows the proportion of 6-25 year olds who are household heads or spouses by age, while the right panel shows the proportion of 6-25 year olds who are sons/daughters-in-law in a household by age.

Table A.1: The Effect of Birth Order on Educational Attainment

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	0.00198	-0.0354***	-0.000591
	(0.00251)	(0.0122)	(0.00243)
Third-Born	-0.00371	-0.0247	-0.00106
	(0.00368)	(0.0151)	(0.00370)
Fourth-Born	-0.0250***	-0.0494***	-0.0268***
	(0.00655)	(0.0191)	(0.00662)
Fifth-Born	-0.0378*	-0.0543	-0.0502**
	(0.0204)	(0.0402)	(0.0212)
Three children	-0.0196***	-0.0855***	-0.0251***
	(0.00396)	(0.0138)	(0.00395)
Four children	-0.0358*	-0.101	-0.0359*
	(0.0201)	(0.0645)	(0.0184)
Five children	-0.0696	-0.0280	0.0590
	(0.0758)	(0.526)	(0.102)
Constant	0.856***	0.707***	0.689***
	(0.00601)	(0.0211)	(0.00729)
Observations	55,137	54,947	54,947
R^2	0.077	0.661	0.111

Notes: The sample is all children from families with 2-5 children aged 6-12 , and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

Table A.2: Birth Order Effect on Education Attainment, Fixed Effects by Family Size: Fully interacted models

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	-0.0296	-0.157**	-0.0464
	(0.0219)	(0.0743)	(0.0377)
Third-Born	-0.0505	-0.355**	-0.0654
	(0.0442)	(0.147)	(0.0768)
Fourth-Born	-0.00475	-0.639**	-0.149
	(0.0824)	(0.260)	(0.142)
Fifth-Born (omitted)	-	-	-
Second-Born×Family Size	-0.00839	0.0194	-0.0238**
	(0.00650)	(0.0216)	(0.0106)
Third-Born×Family Size	-0.0202	0.0249	-0.0538**
	(0.0126)	(0.0408)	(0.0210)
Fourth-Born×Family Size	-0.0482**	0.0291	-0.0731**
	(0.0214)	(0.0673)	(0.0360)
Fifth-Born×Family Size	-0.0628***	-0.159***	-0.145***
	(0.00963)	(0.0293)	(0.0160)
Constant	0.996***	1.191***	1.301***
	(0.0214)	(0.0850)	(0.0383)
Observations	77,730	77,471	67,124
R^2	0.099	0.778	0.121
Number of Families	46,973	46,905	45,500

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

Table A.3: The Effect of Birth Order on Educational Attainment, Fixed Effects by Gender of the First-Born: Fully Interacted model

VARIABLES	(1) Enrolment	(2) Education	(3) Schooling progression
Second-Born	-0.0602***	-0.0130	-0.138***
	(0.00923)	(0.0306)	(0.0153)
Third-Born	-0.122***	-0.0522	-0.257***
	(0.0176)	(0.0580)	(0.0294)
Fourth-Born	-0.206***	-0.0995	-0.403***
	(0.0267)	(0.0864)	(0.0451)
Fifth-Born	-0.285***	-0.137	-0.584***
	(0.0445)	(0.134)	(0.0815)
Second-Born ×FirstBorn_Girl	-0.00156	-0.180***	0.0179
	(0.0126)	(0.0415)	(0.0211)
Third-Born×FirstBorn_Girl	-0.0193	-0.239***	-0.00575
	(0.0239)	(0.0780)	(0.0405)
Fourth-Born×FirstBorn_Girl	-0.0248	-0.283**	-0.0303
	(0.0361)	(0.116)	(0.0614)
Fifth-Born×FirstBorn_Girl	-0.0326	-0.367**	-0.0657
	(0.0606)	(0.179)	(0.109)
Constant	1.001***	1.056***	1.295***
	(0.0216)	(0.0819)	(0.0373)
Observations	77,730	77,471	67,124
R-squared	0.098	0.776	0.122
Number of num_id	46,973	46,905	45,500

Notes: The sample is all children aged 6-12 from families with 2-5 children, and where the oldest observed child is at most 18 years. All regressions include household head's education, dummies for age (in years), and gender of the child, and location, location interacted with gender, and household head's education. Enrolment is a dummy which equals 1 if a child is reported as enrolled. Education is completed years of schooling. Schooling progression equals $education/(age-6)$. Estimates are from fully interacted fixed effects models where all birth order dummies and controls variables are interacted with First-Born-Girl dummy. Standard errors (in parentheses) are clustered at the family level. ***significant at 1 percent, **significant at 5 percent, *significant at 10 percent.

Chapter 5

Summary and Conclusions

This thesis investigates how a large public education programme, the Free Primary Education (FPE) programme, promotes human capital accumulation of primary school-age children, and how a child's specific characteristic, the order of birth, might influence intra-household allocation of resources and hence differences in education outcomes of siblings. These questions are addressed in three self-contained chapters, each making a contribution towards our understanding of how public policy and family environment matter for human capital development. The data used in this thesis is from Lesotho, a less developed country characterised by high levels of unemployment, poverty, and inequality.

In 2000, Lesotho instituted the FPE programme, with the goal of improving the level of education by encouraging primary school-age children to enrol and stay in school until they have at least completed the seven years of primary schooling. The FPE programme, therefore, involved abolition of school fees, provision of free textbooks and stationery, and building and/or refurbishment of new primary schools. It was implemented sequentially, grade-by-grade, starting with the abolition of school fees in grade one, such that, by 2006, all primary school grades were covered by the programme. Chapter 2 of the thesis examines the effect of this programme on human capital development; specifically, its impact on school enrolment and gender– and wealth–related schooling inequalities.

Owing to the implementation strategy of the FPE programme, there were variations in enrolment rates over time and across grade-groups (that is, grades covered versus those not-yet covered), and age-groups (that is, age-appropriate cohorts for covered grades versus age-appropriate cohorts for the not-yet-covered grades). Thus, the control and treatment groups are defined in two ways; by grade-group and by age-group. I employ a difference-in-differences strategy that exploits these variations to tease out the treatment effect of the policy on enrolment. I use data from the 1995 Household Budget Survey (HBS) and the 2002 Core Welfare Indicators Questionnaire (CWIQ) survey.

The results of this chapter show that the FPE programme had a significant positive effect on school enrolment of primary school-age children between 1995 and 2002. Most importantly, the results indicate that the programme raised enrolment of boys (the historically disadvantaged group in Lesotho) and children from poor backgrounds the most, thereby enhancing equity in access. When the treatment is defined by age-group I find a lower-bound FPE effect of 9.3 percentage points (or 13.2 percent) increase in enrolment. The FPE effect rises to 28.8 percentage points (or 40.7 percent) increase in enrolment when the treatment group is defined by grade-group.

On the whole, the results presented in chapter 2 are consistent with the evidence found in Uganda (Deininger, 2003), and in Kenya (Lucas and Mbiti, 2012). More importantly, they show that in an environment where, unlike in Kenya, there is little scope for schooling demand shifts from public schools (that is, those covered by the FPE policy) to private schools, FPE does indeed significantly raise access to schooling. From this perspective, FPE has value for money in that the resources spent have resulted in the desired goals of increasing enrolment and bridging gender-related inequalities in schooling.

However, the results in chapter 2 tell us nothing about the possible influence of the FPE programme on changes in educational achievement and educational inequality. Chapter 3 of the thesis attempts to bridge this knowledge gap. Specifically, the chapter uses the Southern and Eastern African Consortium for Monitoring Educational Quality (SACMEQ) grade six pupils' standardised maths and reading test scores to analyse changes in educational achievement and educational inequality, and the determinants of such changes. Thus far, SACMEQ has successfully carried out three survey projects: SACMEQ I conducted in 1995-1998; SACMEQ II in 1998-2000; and SACMEQ III in 2005-2007. Lesotho only has data for the last two surveys (SACMEQ II collected in 2000, and SACMEQ III collected in 2007). Thus, the educational achievement (or educational quality) data used in chapter 3 is for a cohort that did not go through the FPE programme (the 2000 cohort) and one that went through the FPE programme (the 2007 cohort).

I use the relative distribution method to analyse the changes in educational achievement and their implications for educational inequality since the introduction of the FPE programme. In addition, I employ the Recentered Influence Function (RIF) regression decomposition method to analyse the determinants of changes in educational achievement and changes in educational inequality. Several results stand out from the analysis. I find that educational quality increased between 2000 and 2007, and this increase was driven by an improved performance of both low- and high-ability students, but mainly that of low-ability students. However, during the same period, the results reveal that educational quality polarization (or inequality) increased, particularly so in reading performance, and this was largely driven by the increase in the spread of test scores below the median.

Turning to results on determinants of changes in educational achievement, they show that some of the increase in reading performance is explained by changes in pupil teacher ratio (PTR), the proportion of children speaking English at home, grade repetition, and teacher effort (that is, teaching hours per week and testing frequency). The increase in maths performance is not explained by any of the education policy variables (for example, PTR), but partly by grade repetition, age, school wealth and school social capital (that is, whether the surrounding community contributes to school activities such as school building and maintenance). Nonetheless, much of the increase in educational performance remains unexplained.

With regard to the determinants of educational inequality, I find that the increase in educational inequality is strongly associated with changes in grade repetition, teacher effort and school social capital. Grade repetition has a positive influence on educational inequality, while high teacher effort and school social capital negatively influence educational inequality. Furthermore, the results show that much of the observed changes in maths and reading scores' inequality remain unexplained.

In chapter 4 , I shift focus from the FPE programme effects to analysing how a child's specific characteristic, the order of birth, impacts on human capital accumulation through its influence on parents' decisions to invest in their children's education. Specifically, the chapter evaluates the effect of birth order on school enrolment, completed years of education, and schooling progression (or relative grade attainment) of primary school-age children (that is, the 6 to 12 year olds). Schooling progression is defined as completed years of education (or completed school grades) divided by potential years of education, which is the total number of years of education that an individual would have completed had she started schooling at age 6 and successfully completed a grade each year. The empirical analysis uses data from the 2006 Lesotho Population and Housing Census (the Census). In order to isolate the effect of birth order on human capital accumulation, I use the within family fixed effects model. This model enables one to purge any potential endogeneity of birth order that is due to its correlation with family size and/or any other omitted factors that are common among siblings. In addition, I control for age fixed effects (or birth cohort effects).

In contrast to the available evidence from many other developing countries data, I find large and significant negative birth order effects on a child's educational attainment. On potential transmission mechanisms, the evidence shows that this negative birth order effect is propagated through birth-spacing, and not family wealth. These results are contrary to earlier evidence from other developing countries (for example, Ecuador, Ghana, and Kenya) which shows that wealth is the underlying causal mechanism behind the positive, not negative, birth order effect on attainment (Tenikue and Verheyden, 2010; De Haan et al., 2014). Surprisingly, chapter 4's results affirm the negative birth order effects found in developed countries (for example, Norway and the United States) (Black et al., 2005; De Haan, 2010).

The difference with the latter evidence is that, in the United States, for example, De Haan (2010) finds that family wealth, and not birth-spacing, explains the negative birth order effect. The findings are consistent with the *confluence* model's hypothesis that the high family intellectual environment that first-borns are born in gives them an advantage over their later born siblings in human capital accumulation. However, I cannot rule out the hypothesis that first-borns do better because of being brought up under tougher parental disciplinary rules than their younger siblings.

Implications of the findings: The results of this thesis have a number of implications for policy and the literature. The first and most important implication of the results in this thesis is that it is possible that, if the FPE policy can be instituted in those sub-Saharan African countries still lagging behind in education indicators, the 'Education for All' goal can be achieved.

Second, Lesotho is one of the first countries to have sequentially implemented the FPE programme in sub-Saharan Africa. The fact that educational achievement increased since the policy was introduced, and that this increase is positively associated with educational policy variables (for instance, PTR and instructional time) suggests that the sequential implementation strategy potentially helps to minimise the disruptions in the system and reduce potential negative quality effects of the programme on quality. This is possibly a good policy lesson to those countries that are yet to institute FPE programmes.

Third, the increase in educational achievement inequality will have negative implications on health and earnings inequalities (Becker and Tomes, 1986a; Alves, 2012; Ferreira and Gignoux, 2014). The Lesotho government thus should implement policies to address educational inequality in order to reduce the high level of income inequality. More generally, the evidence suggests that for countries to effectively address the income inequality problem, distributional changes in educational quality have to be closely monitored, such that any emerging inequalities can be bridged in time.

Fourth, the result that grade repetition has a significant influence on achievement seems to support the evidence found by Foureaux Koppensteiner (2014) in Brazil (see also Rodrigues et al., 2013) which indicates that relaxing the retention policy reduces students' performance. Therefore, the finding suggests that the automatic grade promotion policy that is now being implemented in Lesotho could hurt educational quality.

Fifth, the fact that much of the changes in both educational achievement and inequality are unexplained by the observed characteristics indicates that there could be other intangible school quality variables and student family background factors that potentially have a significant influence on educational performance.

The implications of the negative birth order results are twofold. One is that, the theories that try to explain birth order effects in developing country contexts cannot be generalised as yet. We still have to identify those conditions or contexts that fit the negative birth order effects and those that fit the positive birth order effects because it is evident that the level of development cannot be the only factor. Lastly, for policy direction, the results suggest that, in order to attenuate these birth order effects and increase educational attainment of later-borns, it is essential that the government designs policies that will improve school participation of later-borns, especially for boys and those from larger families. Without speculating on the general equilibrium effects, one such policy intervention could be the introduction of conditional cash transfers to larger families, particularly those with first-born boys.

Directions for future research: This thesis has shown that FPE does increase school enrolment of primary school-aged children. However, schooling is not education. It is just an input into education production. The critical question of whether the programme increased educational achievement still remains open. In addition, we have seen that most of the increase in educational achievement and educational inequality is not related to changes in any of the observed variables. That is, the increase in educational achievement is potentially explained by unobserved school and family factors. This then calls for more research work focusing on the effects of intangible family and school factors on educational achievement. The results in chapter 3 seem to support the *confluence* model's hypothesis (Zajonc, 1976). However, the fact that the results cannot be divorced from the predictions of the *strategic parenting* hypothesis implies that more research is needed to test many other potential pathways through which birth order affects educational achievement to either support or discount the explanation given in this thesis.

Bibliography

(2011). Lesotho 2009: Results from the Demographic and Health Survey. *Studies in Family Planning*, 42(4):305–310.

Al-Samarrai, S. and Zaman, H. (2007). Abolishing School Fees in Malawi: The Impact on Education Access and Equity. *Education Economics*, 15(3):359–375.

Alves, N. (2012). The impact of education on household income and expenditure inequality. *Applied Economics Letters*, 19(10):915–919.

Ambrose, D. (2007). The History of Education in Lesotho: Six Brief Subsectoral Studies. *Lesotho Miscelaneous Documents*, No. 23.

Angrist, J. D. and Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press, New Jersey.

Banerjee, A. K. (2010). A multidimensional Gini index. *Mathematical Social Sciences*, 60(2):87–93.

Baydar, N., Greek, A., and Brooks-Gunn, J. (1997). A Longitudinal Study of the Effects of the Birth of a Sibling During the First 6 Years of Life. *Journal of Marriage and the Family*, 59(4):939–956.

Beatty, A. and Pritchett, L. (2012). From Schooling Goals to Learning Goals: How Fast can Student Learning Improve? *Center for Global Development Policy Paper*.

Becker, G. and Tomes, N. (1986a). Human capital and the rise and fall of families. *Journal of labor economics*, 4(3 Pt. 2):1–47.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. National Bureau of Economic Research, New York. Distributed by Columbia University Press.

Becker, G. S. (1965). A theory of the allocation of time. *Economic Journal*, 75(299):493–517.

Becker, G. S. and Tomes, N. (1986b). Human Capital and the Rise and Fall of Families. *Journal of labor economics*, pages S1–S39.

Behrman, J. R. (2010). Investment in education – inputs and incentives. In Rodrik, D. and Rosenzweig, M., editors, *Handbook of Development Economics*, volume 5, pages 4883 – 4975. Elsevier.

Bettendorf, L. J., Jongen, E. L., and Muller, P. (2012). Childcare Subsidies and Labor Supply: Evidence from a large Dutch Reform. *Tinbergen Institute Discussion Paper, No.12-093/I.*

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*, 120(2):669–700.

Blinder, A. S. (1973). Wage discrimination: reduced form and structural estimates. *Journal of Human resources*, 8(4):436–455.

Blundell, R. and Dias, M. C. (2009). Alternative Approaches to Evaluation in Empirical Microeconomics. *Journal of Human Resources*, 44(3):565–640.

Bold, T., Kimenyi, M., Mwabu, G., and Sandefur, J. (2015). Can free provision reduce demand for public services? Evidence from Kenyan education. *The World Bank Economic Review*, 29(2):293–326.

Bureau of Statistics (2002). Lesotho Core Welfare Indicators Questionnaire CWIQ Survey, 2002. Technical report, Bureau of Statistics.

Bureau of Statistics (2006). 2002/03 and 1994/95 Household Budget Survey Analytical Report: Volume I. Technical report, Bureau of Statistics.

Bureau of Statistics (2008). 2008 Lesotho Integrated Labour Force Survey. Technical report, Bureau Of Statistics.

Carnoy, M. and Rothstein, R. (2013). What do international tests really show about US student performance. Technical report.

Case, A. and Deaton, A. (1999). School inputs and educational outcomes in South Africa. *The Quarterly Journal of Economics*, 114(3):1047–1084.

Chyi, H. and Zhou, B. (2014). The effects of tuition reforms on school enrollment in rural china. *Economics of Education Review*, 38:104–123.

Colclough, C., Kingdon, G., and Patrinos, H. A. (2008). The pattern of returns to education and its implications. *Research Consortium on Educational Outcomes and Poverty Policy Brief*, No. 4.

Cunha, F. and Heckman, J. J. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, 7(2-3):320–364.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the Evidence on Life Cycle Skill Formation. In Hanushek, E. and Welch, F., editors, *Handbook of the Economics of Education*, volume 1, pages 697 – 812. Elsevier.

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.

Deaton, A. (2003). Health, income, and inequality. *NBER Reporter Online*, (Spring 2003):9–12.

Deininger, K. (2003). Does cost of schooling affect enrollment by the poor? Universal primary education in Uganda. *Economics of Education Review*, 22(3):291–305.

Eide, E. and Showalter, M. H. (1998). The effect of school quality on student performance: A quantile regression approach. *Economics Letters*, 58(3):345–350.

Ejrnæs, M. and Pörtner, C. C. (2004). Birth Order and the Intrahousehold Allocation of Time and Education. *Review of Economics and Statistics*, 86(4):1008–1019.

Ferreira, F. H. and Gignoux, J. (2014). The Measurement of Educational Inequality: Achievement and Opportunity. *The World Bank Economic Review*, 28(2):210–246.

Firpo, S., Fortin, N., and Lemieux, T. (2007). Decomposing wage distributions using recentered influence function regressions. *Working Paper. Department of Economics, University of British Colombia*.

Firpo, S., Fortin, N. M., and Lemieux, T. (2009). Unconditional Quantile Regressions. *Econometrica*, 77(3):953–973.

Fortin, N., Lemieux, T., and Firpo, S. (2011). Decomposition methods in economics. *Handbook of labor economics*, 4:1–102.

Foureaux Koppensteiner, M. (2014). Automatic grade promotion and student performance: Evidence from Brazil. *Journal of Development Economics*, 107:277–290.

Glewwe, P. W., Hanushek, E. A., Humpage, S. D., and Ravina, R. (2011). School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010. *NBER Working Paper*, No. 17554.

Grogan, L. (2009). Universal primary education and school entry in Uganda. *Journal of African Economies*, 18(2):183–211.

Gualtieri, T. and Hicks, R. E. (1985). An Immunoreactive Theory of Selective Male Affliction. *Behavioral and Brain Sciences*, 8(03):427–441.

Handcock, M. S. and Morris, M. (1998). Relative distribution methods. *Sociological Methodology*, 28(1):53–97.

Handcock, M. S. and Morris, M. (1999). *Relative Distribution Methods in the Social Sciences (Statistics for Social Science and Behavorial Sciences)*. Springer.

Hanushek, E. A. and Woessmann, L. (2012). Schooling, educational achievement, and the Latin American growth puzzle. *Journal of Development Economics*, 99(2):497–512.

Hao, L., Hotz, V. J., and Jin, G. Z. (2008). Games Parents and Adolescents Play: Risky Behaviour, Parental Reputation and Strategic Transfers. *The Economic Journal*, 118(528):515–555.

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

Heckman, J. J. and Mosso, S. (2014). The Economics of Human Development and Social mobility. *NBER Working Paper 19925*.

Hoogeveen, J. and Rossi, M. (2013). Enrolment and Grade Attainment following the Introduction of Free Primary Education in Tanzania. *Journal of African Economies*, 22(3):375–393.

Horowitz, A. W. and Wang, J. (2004). Favorite son? Specialized child laborers and students in poor LDC households. *Journal of Development Economics*, 73(2):631–642.

Hotz, V. J. and Pantano, J. (2013). Strategic parenting, birth order and school performance. *NBER Working Paper No. 19542*.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jann, B. (2008). Relative Distribution Methods in Stata. In *6th German Stata Users Group Meeting, Berlin*.

Jayachandran, S. and Kuziemko, I. (2011). Why do mothers breastfeed girls less than boys? Evidence and implications for child health in India. *The Quarterly Journal of Economics*, 126:1485–1538.

Jopo, H., Maema, M., and Ramokoena, M. (2011a). Quality of primary school inputs in Lesotho. *SACMEQ Policy Brief*, No. 2.

Jopo, H., Maema, M., and Ramokoena, M. (2011b). Trends in achievement levels of grade 6 learners in lesotho. *SACMEQ Policy Brief*, No. 1.

Kristensen, P. and Bjerkedal, T. (2007). Explaining the Relation Between Birth Order and Intelligence. *Science*, 316(5832):1717–1717.

Lafortune, J. and Lee, S. (2014). All for One? Family Size and Children's Educational Distribution under Credit Constraints. *The American Economic Review*, 104(5):365–369.

Lee, V. E. and Zuze, T. L. (2011). School resources and academic performance in Sub-Saharan Africa. *Comparative Education Review*, 55(3):369–397.

Leibowitz, A. (1974). Home investments in children. *The Journal of Political Economy*, 82(2):S111–S131.

Lekhetho, M. (2013). The Impact of Free Primary Education on Access and Quality of Primary Education in Lesotho. *International Journal of Educational Science*, 5(4):397–405.

Leuven, E., Oosterbeek, H., and Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics*, 110(4):663–693.

Lounkaew, K. (2013). Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy data. *Economics of Education Review*, 37:213–225.

Lucas, A. M. and Mbiti, I. M. (2012). Access, Sorting, and Achievement: The Short-Run Effects of Free Primary Education in Kenya. *American Economic Journal: Applied Economics*, 4(4):226–253.

Mani, S., Hoddinott, J., and Strauss, J. (2013). Determinants of Schooling: Empirical Evidence from Rural Ethiopia. *Journal of African Economies*, 22(5):693–731.

Ministry of Health, Social Welfare (MOHSW) [Lesotho], Bureau of Statistics (BOS) [Lesotho], and Macro, O. (2005). *Lesotho Demographic and Health Survey 2004*. Calverton, Maryland: MOH, BOS, and ORC Macro.

MOET (2011). Presentation on Education Management Information System (EMIS): 1999-2011. Technical report.

Morduch, J. (2000). Sibling Rivalry in Africa. *American Economic Review*, 90(2):405–409.

Murimba, S. (2005). The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ): Mission, approach and projects. *Prospects*, 35(1):75–89.

Nishimura, M., Yamano, T., and Sasaoka, Y. (2008). Impacts of the universal primary education policy on educational attainment and private costs in rural Uganda. *International Journal of Educational Development*, 28(2):161–175.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3):693–709.

Ponczek, V. and Souza, A. P. (2012). New Evidence of the Causal Effect of Family Size on Child Quality in a Developing Country. *Journal of Human Resources*, 47(1):64–106.

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

Pritchett, L. (2004). Towards a new consensus for addressing the global challenge of the lack of education. *Center for Global Development Working Paper*, (43).

Rodrigues, C. G., Rios-Neto, E. L. G., and de Xavier Pinto, C. C. (2013). Changes in test scores distribution for students of the fourth grade in Brazil: A relative distribution analysis for the years 1997–2005. *Economics of Education Review*, 34:227–242.

Ross, K., Saito, M., Dolata, S., Ikeda, M., Zuze, L., Murimba, S., Postlethwaite, N., and Griffin, P. (2005). The Conduct of the SACMEQ II Project. In *The SACMEQ II Project in Kenya: A Study of the Conditions of Schooling and the Quality of Education*. SACMEQ.

Schultz, T. P. (2004). School subsidies for the poor: Evaluating the Mexican Progresa poverty program. *Journal of Development Economics*, 74(1):199–250.

Shanks, T. R. W. and Robinson, C. (2013). Assets, economic opportunity and toxic stress: A framework for understanding child and educational outcomes. *Economics of Education Review*, 33:154–170.

Sulloway, F. J. (2007). Birth Order and Intelligence. *Science*, 316(5832):1711–1712.

Taylor, S. and Spaull, N. (2013). The effects of rapidly expanding primary school access on effective learning: The case of Southern and Eastern Africa since 2000. *Stellenbosch Economic Working Papers*, No. 01/13.

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.

Thomas, D., Strauss, J., and Henriques, M.-H. (1990). Child survival, height for age and household characteristics in Brazil. *Journal of Development Economics*, 33(2):197–234.

UNESCO (2000). The Dakar Framework for Action: Education for All, Meeting our Collective Commitments. Technical report, UNESCO.

UNESCO (2014). Education for All Global Monitoring Report 2013/4: Teaching and Learning, Achieving quality for all. Technical report, UNESCO.

Urwick, J. (2011). "Free Primary Education in Lesotho" and the disadvantages of the highlands. *International Journal of Educational Development*, 31(3):234–243.

Wittenberg, M. and Leibbrandt, M. (2015). Measuring Inequality by Asset Indices: A general approach with application to South Africa. *A Southern Africa Labour and Development Research Unit Working Paper*, No. 141.

Zajonc, R. (1976). Family configuration and intelligence. *Science*, 192:227–236.

Zuze, T. L. and Leibbrandt, M. (2011). Free education and social inequality in Ugandan primary schools: A step backward or a step in the right direction? *International Journal of Educational Development*, 31(2):169–178.