CHILDREN, INVESTMENTS IN EDUCATION AND POVERTY IN MALAWI

RICHARD MUSSA

Thesis Presented for the Degree of
DOCTOR OF PHILOSOPHY
in the School of Economics
UNIVERSITY OF CAPE TOWN

Supervisor: Associate Professor Martin Wittenberg

January 2009

Dedication

To my dearest parents, brothers and sisters.

Abstract

This thesis presents an examination of three interrelated themes concerning household investments in the education of children, and how the number of children in a household affects a household's poverty situation in Malawi. These themes cover three chapters. Chapter 2, investigates two issues regarding household expenditure on primary education of own children using the Second Malawi Integrated Household Survey (IHS2) data. Firstly, we look at factors which influence a household's decision to spend or not (the participation decision), and by how much (the expenditure decision). This is done for urban and rural households. We find that there are differences in the factors which influence both decision levels for the two groups of households. Secondly, to get a deeper understanding of these rural-urban spending differences, the study develops the Blinder-Oaxaca decomposition technique for the independent Double Hurdle model. The proposed decomposition is done at the aggregate and disaggregated levels. The aggregated decomposition allows us to isolate the expenditure differences into a part attributable to differences in characteristics and a part which is due to differences in coefficients. The detailed (disaggregated) decomposition enables us to pinpoint the major factors behind the spending gap. At the aggregate decomposition level, our results show that at least 66% of the expenditure differential is explained by differences in characteristics between rural and urban households, implying that an equalization of household characteristics would lead to about 66% of the spending gap disappearing. At the disaggregated decomposition level, the rural-urban difference in household income is found to be the largest contributor to the spending gap, followed by quality of access of primary schools. Besides, rural-urban differences in mothers education and employment are found to contribute more to the spending differential relative to the same for fathers.

Recognizing that in many African countries parents are not just responsible for the education of their own children, Chapter 3 examines the relationship between household income and schooling costs in the presence of intrahousehold schooling bias against non-biological children. To this end, we construct a two-period model of intrahousehold schooling bias. The model predicts that there is an asymmetry in the impact of changes in costs and income on schooling in the sense that the impact is larger for the non-biological child. It predicts that the asymmetry increases as the relationship distance between the non-biological child and the parents gets wider. It also shows that an increase in cost of schooling leads to a bigger reduction in schooling for poor households, and that the difference in the impact of cost changes between the biological and the non-biological child declines as household income increases i.e. there is convergence. The convergence is faster the more distantly related to the parents the non-biological child is. An empirical investigation of these predictions using IHS2 data, shows that when current enrolment and grade attainment are used to measure schooling, the price and income elasticities of schooling are larger for non-biological children. The results also indicate that households in the lowest income quintile (the poorest) have the largest price elasticities, and households in the highest income quintile (the wealthiest) have the smallest price elasticities. We also find that the price elasticities for biological and non-biological children converge as we move from the lowest income quintile to the highest income quintile, and that the convergence is faster for non-biological children who are non-relatives.

Having looked at among other things the role of household economic status on the education investments of children in Malawi in Chapters 2 and 3, in Chapter 4 we turn the question on its head and investigate the impact of fertility (number of children) on poverty in rural Malawi. We use two measures of poverty; the objective and the subjective. After accounting for endogeneity of fertility by using son preference as an instrumental variable, we find that fertility increases the probability of being objectively poor. This effect is robust for all poverty lines used. It is also robust to accounting for economies of scale and household composition as well as assuming that poverty is continuous. We also find that when fertility is treated as an exogenous variable its impact is underestimated. When poverty is measured subjectively, the results are opposite to those of objective poverty. We find that fertility lowers the likelihood of feeling poor, and that fertility is exogenous with respect to subjective poverty.

Acknowledgements

I would like to express my very sincere gratitude to my supervisor Associate Professor Martin Wittenberg for his invaluable guidance, critical comments, and suggestions. Working with him has without a doubt deepened and widened my knowledge of economic theory and econometrics.

I am deeply indebted to the African Economic Research Consortium (AERC) for providing me with a scholarship under the Collaborative PhD Programme (CPP). Am also grateful to the AERC for affording me the opportunity to visit the IZA (Institute for the Study of Labour) in Germany. My thesis greatly benefitted from the comments received during the visit. It also benefitted immersely from the various comments and suggestions received at the biannual conferences organized by the AERC.

I would also like to thank staff members in the School of Economics, University of Cape Town, for the wonderful help given to me during the course of writing the thesis. To my classmates and all friends I say thank you for your understanding and encouragement.

Contents

Chapter 1

Introduction

1.1 Background and motivation

Malawi is one of the poorest countries in the world. In 2005, 52.4 % of Malawians were living below the national poverty line of 44.30 Malawi Kwacha (US$0.50) per day. 22.4% of the population was classed as ultra poor, living on less than 27.50 Malawi Kwacha (US$0.31) per day. About 25 percent of the population in urban areas was living in poverty, compared to 56 percent of the rural population. That is, a person in a rural area was more than twice as likely to be poor (NSO 2005). For human development, Malawi was ranked by the United Nations at 164 out of 177 countries in 2007 (UNDP 2007)[1]. Malawi's poverty situation like in many Sub-Saharan African countries is further compounded by the HIV/AIDS epidemic. In 2005, the estimated adult (age 15-49) HIV prevalence rate for Malawi was 14.1%. With this high prevalence rate, Malawi was ranked number eight in the world (UNAIDS 2006). Thus, the twin challenges of high poverty and high HIV prevalence continue to occupy the attention of donors, government development planners as well as non-governmental organizations (NGOs). Over the years, economists and other researchers have proposed policy prescriptions to eradicate or alleviate poverty.

One of the channels out of the dire poverty situation is investment in human capital formation through the provision of health and education. Low levels of education are widely considered to be a major impediment to economic growth and the eradication of poverty in sub-Saharan Africa (Glick and Sahn 2000). It has been argued (e.g. Appleton *et al.* 1996), that countries with low literacy and numeracy levels would find it difficult to attain high and sustainable levels of economic growth necessary to have a significant impact on poverty. Low levels of education in a country can make it more difficult for it to take advantage of the opportunities offered by a globalized economy where liberalization

[1] Malawi's GDP per capita (PPP US$) was 667 in 2007. This gives the country a rank of 174^{th} out of 177 countries with data (UNDP 2007).

and educational expansion can either, in a virtuous cycle, reinforce each other or lead to a low level equilibrium (Kim and Kim 2000). An educated workforce makes it easier for a country to adopt the requisite new technologies which would enable it compete effectively in an increasingly globalized world. Investments in education have other more subtle and indirect benefits. For example, societies with high illiteracy levels have been shown to be more susceptible to political manipulation, corruption and bad governance, as well as civil strife and violence, phenomena that undermine human as well as economic development (Gupta *et al.* 1998; Ranis *et al.* 2000).

The accumulation of human capital through investments in education is mostly done by families and governments. Parents play a crucial role in the education investment process through their decisions to enrol their children in school, making sure that once enrolled they stay in school, as well as spending on their education. Economists have long been concerned with modelling decisions that parents make regarding investments in the education of their children (see Haveman and Wolfe 1995 for a review). They have looked at the factors which influence the direct education expenditures that parents make on their children. Here there are two strands of literature; those that use aggregated expenditure where expenditure on education is combined with other items (e.g. Lazear and Michael 1988), and another strand which uses education expenditure as a stand alone item (e.g. Mauldin *et al.* 2001; Yueh 2006; Beneito *et al.* 2001; Song *et al.* 2006; Kingdon 2005). While focussing on household expenditure on primary education as a stand alone item, we advance the understanding of the direct expenditures that parents make on their children on two fronts;

- we make a distinction between households by whether they reside in rural or urban areas. Most studies looking at spending on education of children either pool the rural and urban samples or just look at one sample (e.g. Mauldin *et al.* (2001) focus on a pooled sample while Yueh (2006), Song *et al.* (2006), and Kingdon (2005) look at rural households only). The rural-urban distinction is important as the perceived expected rate of return to education may not be the same between rural and urban areas, due to differences in returns between the formal sector (mostly urban) and the agricultural sector (mostly rural) (Al-Samarrai and Reilly 2000). The implication of this is that a household's expected return to investing in education may be different between the two areas, and hence the spending would also be different. The characteristics between the two areas may be dissimilar in the sense that for example access to schools in terms of distance would be poorer in rural areas, reflecting an urban bias in terms of developmental projects.The rural-urban divide is also interesting aside from the education spending argument in that most poor people live in rural areas. One can therefore loosely look at the spatial demarcation as poor-rich one. With this distinction in mind, we look at factors which influence a family's decision to spend on own children's primary education in rural and urban Malawi.

- in the light of the rural-urban differences in expenditure, we go a step further to explain these differences. We propose an extension of the decomposition technique developed by Blinder (1973) and Oaxaca (1973) to the independent double hurdle model. We then use the proposed Blinder-Oaxaca decomposition of the independent double hurdle model to conduct a decomposition of the gap in household expenditure on education between the two areas. The decomposition isolates how much of the differential in expenditure can be attributed to characteristics (*characteristic effect*) and how much is due to differences in returns to those characteristics (*coefficient effect*). The two effects give us an aggregated picture of the reasons for the expenditure gap, and to move on from this black box explanation of the expenditure gap, we further propose a disaggregated decomposition of the *characteristic effect* of the independent double hurdle model. This detailed decomposition allows us to pinpoint the major factors behind the spending gap. For example, a detailed decomposition of the *characteristic effect* provides an understanding of the role of household income in the rural-urban spending differential.

In many African countries, Malawi inclusive, parents are not only responsible for the education of their own children, but they are also responsible for the investments in education of non-biological children who are put in their care. With the high prevalence rate of HIV as alluded to earlier, the number of HIV/Aids orphans in Malawi is bound to grow. Thus, the role that parents play in educating non-biological children in addition to their own takes on an added significance. This blending of biological and non-biological children in families may give rise to schooling discrimination, in the sense that non-biological children may receive less schooling relative to biological ones. Given the benefits of education, schooling discrimination may have negative implications on future economic growth and the fight against poverty. The literature is few and far between on the possible sources of this discrimination. For example, Case *et al.* (2004) show that the probability of school enrollment is inversely proportional to the degree of relatedness of the child to the household head, regardless of whether the child is an orphan or not. In addition to the dearth of economic studies on sources of schooling bias against non-biological children, to the best of our knowledge there is no study which addresses the issue of what happens to schooling bias following household income and school cost changes. We therefore make the following contributions;

- we propose a theoretical model which offers possible sources of schooling bias against non-biological children.

- we demonstrate in the presence of discrimination, how households respond to changes in household income and school costs, and how the household's response to cost changes varies with income.

- then, we empirically investigate the theoretical predictions. Specifically, the empirical analysis seeks to examine using Malawian data, how households respond to changes in household income and school costs, and how the household's response to cost changes varies with income.

Parents do not just make decisions on the education investments of children, biological or otherwise, they also decide on how many children to have. The choices that parents make regarding their family size may have consequences on their household's poverty situation. Studies looking at the impact of fertility (number of children) in Africa are few, despite the fact that Africa has some of the highest levels of both fertility and poverty in the world. The few studies (e.g. Chernichovsky 1984; Langani 1997; Cohen and House 1994; Noumbissi and Sanderson 1998) have treated fertility as exogenous, thus ignoring the potential simultaneity that exists between the two variables as well as the fact that there are unobserved factors which influence both variables i.e. unobserved heterogeneity. Besides, they only focussed on poverty defined in the objective monetary sense which is a narrower definition of household welfare. Subjective measures of welfare better capture the multidimensional nature of poverty. In the light of these shortfalls, we make the following contributions;

- while accounting for the simultaneity and unobserved effects, we investigate how fertility impacts on objective poverty in Malawi

- we examine how fertility impacts on subjective poverty in Malawi[2]

The aforementioned contributions are presented in three chapters of the thesis, structured as follows. In Chapter 2 we look at the rural-urban differences in parental spending on children's primary education. Schooling bias against non-biological children is the focus of Chapter 3. In Chapter 4 we flip the question on its head and look at how fertility impacts on objective and subjective poverty. We finally present the major conclusions of the thesis in Chapter 5.

[2] As is shown later, fertility is found to be exogenous to subjective poverty.

Chapter 2

Rural-urban differences in parental spending on children's primary education in Malawi

2.1 Introduction

One of the costs of raising children that must be incurred by parents is investing in their education. There are two major players in investments in human capital of children namely; the household and the government. Household and government expenditure on education is both an end in itself and a means for achieving other goals of development, such as economic growth, poverty reduction, improved health status, greater equity and reduced fertility (Glewwe and Ilias 1996). The low level of human capital development in most African countries is considered an obstacle for economic growth as well as the alleviation of poverty (Glick and Sahn 2000).

The Malawi government in recognition of the crucial role that human capital accumulation and development plays in fostering economic growth among other benefits introduced free primary education (FPE) in 1994. Under FPE, parents no longer pay tuition fees, however they still have to pay for other educational expenses including books, uniforms, transport, contribution for school building and maintenance among other expenses. This means that households still have to play a role in investing in the primary education of their children. Besides, they also have to pay for the education of their children when they go to secondary school.

In this study, we focus on investment in education by families and not government. Economists have long been concerned with modelling decisions that parents make regarding investments in the education of their children (see Haveman and Wolfe 1995 for a review). They have investigated the time parents allocate to their children (e.g. Lazear and Michael

1988; Leibowitz 1974, 1977; van der Gaag 1982). They have focused on the factors which influence enrolment in primary and secondary schools (e.g. Kabubo-Mariara and Mwabu 2007; Glewwe and Ilias 1996). Others have looked at household willingness to pay for the education of children (e.g. Gertler and Glewwe 1989). Other studies have looked at the factors which influence direct education expenditures that parents make on their children. Here, there are two strands of literature; those that use aggregated expenditure where expenditure on education is combined with other items (e.g. Lazear and Michael 1988), and another strand which uses education expenditure as a stand alone item (e.g. Mauldin *et al.* 2001; Yueh 2006; Beneito *et al.* 2001; Song *et al.* 2006; Kingdon 2005). In this study, we look at education expenditure as a stand alone item.

While focussing on household expenditure on primary education as a separate item, the study advances the understanding of the direct expenditures that parents make on their children in two ways. First, we make a distinction between households by whether they reside in rural or urban areas[1]. Most studies looking at spending on education of children either pool the rural and urban samples or just look at one sample (e.g. Mauldin *et al.* (2001) focus on a pooled sample while Yueh (2006), Song *et al.* (2006), and Kingdon (2005) look at rural households only). The rural-urban distinction is important as it is shown in Section 2.4 that there are differences in average expenditure between households in rural and urban areas. The rural-urban divide is also interesting aside from the education spending argument in that most poor people live in rural areas. One can therefore loosely look at the spatial demarcation as poor-rich one. Further to that, Al-Samarrai and Reilly (2000) contend that the perceived expected rate of return to education may not be the same between rural and urban areas, due to differences in returns between the formal sector (mostly urban) and the agricultural sector (mostly rural). The implication of this is that a household's expected return to investing in education may be different between the two areas, and hence the spending would also reflect this. The characteristics between the two areas can be dissimilar in the sense that for example access to schools in terms of distance would be poorer in rural areas, reflecting an urban bias in terms of developmental projects. A more detailed discussion of the reasons why we would expect rural-urban differences in investment in education are given in the theoretical section. With this distinction in mind, the study looks at factors which influence a family's decision to spend on own children's primary education in rural and urban Malawi. Specifically, here we seek to answer two interrelated questions; a) what factors influence the probability that a household spends or does not spend on own children's education? This is the participation decision. and b) what factors affect educational expenditure if a household decides to spend? This is the expenditure decision.

Second, in the light of these rural-urban differences in expenditure, we go a step further

[1] Al-Samarrai and Reilly (2000) make the rural-urban distinction with respect to school enrolment.

to explain these differences. To this end, we propose an extension of the decomposition technique developed by Blinder (1973) and Oaxaca (1973) to the independent double hurdle model[2]. We then use the proposed Blinder-Oaxaca decomposition of the independent double hurdle model to conduct a decomposition of the gap in household expenditure on education between the two areas. The decomposition isolates how much of the differential in expenditure can be attributed to characteristics (*characteristic effect*) and how much is due to differences in returns to those characteristics (*coefficient effect*), which we interpret as the difference due to behavioural differences. The two effects give us an aggregated picture of the reasons for the expenditure gap, and to move on from this black box explanation of the expenditure gap, we further propose a disaggregated decomposition of the *characteristic effect* of the independent double hurdle model[3]. This detailed decomposition enables us to pinpoint the major factors behind the spending gap. For example, a detailed decomposition of the *characteristic effect* provides an understanding of the role of household income in the rural-urban spending differential. From a policy standpoint, while it is important to know whether these expenditure differences arise due to differences in characteristics of the households or whether they are attributable to behavioural differences, it is even more critical that we have knowledge of which individual characteristics are vital in driving the spending gap.

Our empirical results for the two areas of residence show that different factors influence household expenditure on primary education differently. The level of household income in rural and urban areas positively and significantly impacts both the participation and expenditure decisions. Computed elasticities indicate that spending on education by rural households is more sensitive to changes in income compared to urban households, suggesting that spending on education in rural areas is a luxury good. We find that a father's and mother's employment status has a bigger impact on spending (at both decision levels) in rural areas compared to urban areas. For both areas, a mother's employment and education has a larger impact on spending compared to a father's. Urban households compared to their rural counterparts are more sensitive to the quality of access of primary schools as measured by the distance to nearest primary school. We find no evidence of gender bias in school spending in urban areas, but rural households exhibit bias in favour of boys.

Results from the proposed aggregated decomposition indicate that at least 66% of the expenditure differential is as a result of differences in characteristics and about 34% arises from behavioural differences (*coefficient effect*) between rural and urban households. This suggests that an equalization of household characteristics (behavior) would lead to about

[2]Al-Samarrai and Reilly (2000) conduct a decomposition of school enrolment gaps between rural and urban areas in Tanzania.

[3]Owing to interpretational problems of the *coefficient effect*, we do not undertake a disaggregated decomposition of the same in this paper.

66% (34%) of the spending gap disappearing. Results from the disaggregated decomposition of the *characteristic effect* indicate that household income, parental education and employment, and quality of access of primary schools are the key factors driving the spending gap. The rural-urban difference in household income is found to be the largest contributor to the spending gap, followed by quality of access of primary schools. Besides, rural-urban differences in mothers education and employment are found to contribute more to the spending differential relative to the same for fathers.

The rest of the chapter proceeds as follows. Section 2.2 looks at the education sector in Malawi. Section 2.3 presents the theoretical underpinnings on which the study is based as well possible explanations regarding the gaps in spending between rural and urban households. In Section 2.4 we discuss the model specification, variables used, estimation issues, and data and descriptives. Econometric results are the focus of Section 2.5. The extension of the Blinder-Oaxaca decomposition technique and results are discussed in Section 2.6. We finally conclude in Section 2.7.

2.2 Education in Malawi

The formal education system in Malawi is composed of three levels namely; primary, secondary, and post secondary. Education at all three levels is not compulsory. The official entry age at the primary school level is about six years. Primary school is made of standards one to eight; which is divided into infant (standards 1-2), junior (standards 3-5), and senior (standards 6-8). Since 1994, the government introduced free primary education (FPE), which entailed that parents no longer had to pay fees for the primary education of children who attend government schools. Private primary schools however continue to charge fees. At the end of the eight years of primary education, pupils sit for the primary school leaving certificate examination (PSLCE). This is a national exam which determines eligibility of entry into secondary school. Secondary school education takes four years; the Junior Certificate level (Forms 1 and 2), and the Malawi School Certificate level (Forms 3 and 4). Parents pay for the secondary education of their children. So the primary-secondary education cycle takes twelve years[4]. The length of post secondary education depends on the type of education programme. University education takes about three years for a diploma, four to five years for a degree. In the recent past, Malawi has experienced a mushrooming of private providers of education at all three levels of education.

In 2005, four out of five pupils attending primary education were in government schools. The next highest providers of primary education were religious institutions. Almost seventeen percent of pupils attending primary school were in religious institutions (National

[4]This could however be longer with repetition.

Statistical Office 2005). Although government is the dominant provider of secondary education, the rate is slightly lower compared to that of primary education. In 2005, government was providing secondary education to 65 percent of all the pupils attending secondary education relative to 80 percent in primary education. The situation is different for private schools. More secondary school pupils attended private schools relative to those in primary education. Nearly one in every three pupils attending secondary education were at private institutions. In terms of area of residence in 2005, 81% of primary school pupils in urban and rural areas attended government schools. This suggests that the majority of primary school pupils in the two areas are in government schools. There is however a marked difference in attendance at secondary level, with 42% and 76% of pupils attending government secondary schools in urban and rural areas respectively. It has also been noted that the substitution by households for private providers is highest for those in the upper expenditure quantiles (National Statistical Office 2005). At the university level, government remains a major provider, until 1998 the University of Malawi was the only university.

2.3 Theoretical framework

2.3.1 Human capital theory

The theoretical framework on which this study is based is the human capital theory (Becker 1981; Becker and Tomes 1976). Under human capital theory, consideration is made of the fact that these investments are generally not made by the primary beneficiaries but by their care givers. Thus, there are issues not only of the efficiency of the investment, but also of the intrahousehold allocation of the expected benefits (Alderman and King 1998). Parents' decision to educate children is done both for its own sake as a consumption good, and as an investment good. The theory suggests that parents will invest time which is a direct input, money which is an indirect input, and other resources in their children's education because they get utility from doing that, and it is also an investment which will give them returns in future. Parents will invest in the education of their children up to a point where the marginal benefit and the marginal cost of investing are equal (Becker 1981; Becker and Tomes 1976). The theory also postulates that the human capital of a child also depends on the genetic endowments which are passed on to children from parents. Becker and Tomes (1986) argue that these endowments from parents to children regress to the mean. They thus argue; "children with well endowed parents tend also to have above average endowments but smaller relative to the mean than their parents', whereas children with poorly endowed parents tend also to have below average endowments but larger relative to the mean than their parents'"(Becker and Tomes 1986, p 5). Thus human capital theory suggests that investments in children's human capital are related to parental characteristics, characteristics of the children, and parental preferences (Becker and Tomes 1986; Hanushek 1992). Expenditures on children's education, skills, health, and abilities are an indirect input into their children's human capital (Becker and Tomes 1986). It is also worth noting that if schooling is a pure investment good i.e. without current consumption aspects, and there are no credit constraints, then income would not affect schooling decisions. However, in many developing countries credit constraints are prevalent (Behrman and Knowles 1999).

Within the human capital theory framework, others explain gender discrimination regarding parental investment in the education of their children (Behrman *et al.* 1986; Alderman and Gertler 1997; Alderman and King 1998; Pasqua 2005; Yeuh 2006). This part of human capital theory deals with why parents may invest more in the education of their boys than girls or vice versa. This strand of literature identifies four possible sources from which gender differences in education may originate. Firstly, a girl will receive less schooling if the cost (direct and indirect) of educating her is higher than that of a boy. This is possible when one considers that the opportunity cost of a girl going to school might be higher as she is more likely to help in caring for younger siblings or fetching firewood and water (Pasqua 2005; Gertler and Glewwe 1992). Secondly, there will be less schooling investment in a girl relative to a boy if the returns to education for a girl are lower. The returns to schooling for a girl can be lower as a result of gender bias in the labour market. Kingdon (1998) for example, finds significant gender differences in returns to education in India. Thirdly, there will be schooling bias against a girl if the expectation/belief of how much the boy child will transfer in old age is higher than that of the girl child. This is quite possible under a patrilineal system where a woman has to leave her family when she gets married and become a member of her husband's family. Finally, the girl child will have less schooling if parents have preference bias against the education of a girl in favour of a boy. That is, there will be gender schooling bias against girls, if parents get more utility from a boy's education even when the education level is the same as that for a boy. We utilize this theoretical framework while focussing on the rural-urban differences in household school investment on primary school children. In the next subsection, we present possible explanations for differentials in investment in education between rural and urban households.

2.3.2 Explaining differences in school investment between urban and rural households

Broadly, the reasons for why there may be differences in investment in schooling between rural and urban households can be put into two categories[5]. The first category relates to explanations which attribute the difference to differences in characteristics between rural and urban households. The second category comprises explanations which ascribe the difference to differences in returns to the characteristics. That is, the characteristics between rural and urban households may be the same, but the returns to (or effectiveness of) those characteristics may be different.

We start with the first broad category. Differences in characteristics of urban and rural households may explain the gaps in school investment between the two areas. There may be differences in characteristics with respect to school quality such as distance to schools, pupil teacher ratio where these statistics are generally bad for rural areas. In most developing countries there is an urban bias in terms of general infrastructure including school facilities. This is well expressed by Lipton (1977) when he observes;

> "*The most important class conflict in the poor countries of the world today is not between labor and capital. Nor is it between foreign and national interests. It is between rural classes and urban classes....Scarce investment, instead of going into water-pumps to grow rice is wasted on urban motorways. Scarce human skills administer, not clean village wells and agricultural extension services, but world boxing championships in showpiece stadia*"(Lipton 1977, p1)

There are several reasons for why rural areas may not be favoured in terms of facilities[6]. It could be due to the fact that the provision of urban public goods is cheaper (Arnott and Gersovitz 1986). It could also arise from the influence and lobbying power of the urban elite (Lipton 1977). The disparity could also be due to the fact that urban households have an information advantage. Majumdar *et al.*(2004) contends;

> "*Urban residents have an information advantage that may arise due to several factors: greater average wealth, higher education, better access to the media as well as a stronger urban focus in media coverage. Even if both rural and urban residents observe public good outcomes equally well, this information advantage implies that urban residents are better positioned to evaluate the role*

[5] See Al-Samarrai and Reilly (2000) for a similar categorization.

[6] Majumdar *et al.*(2004) document some evidence of urban bias in public goods provision in developing countries.

of the government's ability in achieving a given outcome" (Majumdar *et al.* 2004, p 139).

To the extent that access, availability, and quality of school facilities influence parental investment in education of children[7], this urban bias may explain the differences in schooling between the two areas. The urban bias in terms of access and availability of other facilities such clinics, water facilities may also explain the rural-urban differences in households' investment in schooling. For example, children are generally involved in fetching water, and if water facilities are very far (as is the case in rural areas) this may affect children's schooling as they dedicate more time to fetching water. Kabubo-Mariara and Mwabu (2007) find a negative relationship between time taken to fetch water and the likelihood of primary school enrolment in Kenya. In addition to the community/area level disparities in favour of urban areas, we can also have characteristic differences at the household level between rural and urban areas. Rural households tend to have larger families than urban households, and assuming a quantity-quality trade-off, this should entail lower schooling in rural areas. Parental education is different between the two areas, rural parents are generally less educated than their urban counterparts. This may have implications on schooling, for instance the cost of helping with homework may be less for more-schooled parents than for less-schooled parents (Behrman and Knowles 1999).We now turn to the differences in returns to characteristics story. Al-Samarrai and Reilly (2000) argue that the perceived expected rate of return to education may be different between rural and urban areas due to differences in return between the formal sector (mostly urban) and the agricultural sector (mostly rural). The implication of this is that a household's expected return to investing in education may be different between the two areas, and this would be reflected in differentials in school investment between urban and rural households. To the extent that there may be cultural differences between rural areas (which tend to be traditional) and urban areas (which tend to be modern) this would be reflected in parental preferences for education. The opportunity cost of schooling between the two areas may also be different, in rural areas children are more likely to work in the field or indeed be sent off to work as child labourers to supplement family income. Thus, in rural areas the opportunity cost of sending a child to school is higher relative to the urban areas. We later propose a decomposition technique which enables us to calculate which of these broad categories is the predominant explanation for the rural-urban school spending differential in Malawi. In addition, we develop a disaggregated decomposition technique which helps us to look at each individual characteristic's contribution to the rural-urban education spending disparity. Before talking about the decomposition, we first present in the next section, the econometric model on which the decomposition is

[7]Studies by Case and Deaton (1999), Lavy (1996), and Al-Samarrai and Reilly (2000), find school quality to be positively related with school enrolment.

based.

2.4 Methodology

2.4.1 Model specification

As discussed earlier, the study is based on direct expenditures that households make on the primary education of their children. One underlying feature of expenditure data is that it contains excess zeros, and the choice of a statistical technique used to deal with the zeros is important, as an inappropriate treatment of zeros can lead to biased and inconsistent estimates (Greene 1981). The Tobit model (Tobin 1958) has been widely used to model outcomes which have excess zeros. The Tobit model is derived from an individual optimization problem and views zeros as corner solution outcomes. The major drawback of the Tobit model is that it assumes that the same stochastic process determines both the extensive and intensive margins, that is the decision whether or not to spend (participation decision) and how much (expenditure decision), are treated as the same. This assumption is restrictive. A model which corrects this limitation of the Tobit model is the Double Hurdle model (DH hereafter)[8]. The DH model, originally formulated by Cragg (1971), assumes that households make two decisions with regard to spending, each of which is determined by a different underlying stochastic process (Blundell and Meghir 1987). Following Jones (1989), the DH is formally specified as follows;

The participation equation (the first hurdle) is given as;

$$\begin{aligned} D_i^* &= Z_i'\alpha + \varepsilon_i \\ D_i &= \begin{cases} 1 \; if \; D_i^* > 0 \\ 0 \; otherwise \end{cases} \end{aligned} \tag{2.1}$$

The expenditure equation (the second hurdle) is given as follows;

$$\begin{aligned} Y_i^* &= X_i'\beta + \nu_i \\ Y_i^{**} &= \max(0, Y_i^*) \end{aligned} \tag{2.2}$$

[8]The DH has been used in various areas in the literature and without purporting to be exhaustive it has been used to model; expenditure on alcohol (Yen and Jensen 1996), expenditure on cigarettes (Yen 2005; Jones 1989), time use (Daunfeldt and Hellström 2007), expenditure on food away from home (Jensen and Yen 1996; Newman *et al* 2003), expenditure on cheese (Yen and Jones 1997), and expenditure on education (Mauldin *et al.* 2001).

Observed expenditure(Y_i);

$$Y_i = D_i Y_i^{**} \tag{2.3}$$

where; D_i^* is a latent variable describing the household's decision to participate (spend or not) on children's education, Y_i^* is a latent variable describing household expenditure on children's education, Z_i' is a vector of variables explaining the participation decision, X_i' is a vector of variables explaining the expenditure decision. ε_i, ν_i are independent random errors with the following properties; $\varepsilon_i \sim N(0,1)$ and $\nu_i \sim N(0,\sigma^2)$, and i denotes household. The assumption of independence is quite common when using the DH (Mauldin *et al.* 2001; Jensen and Yen 1996; Su and Yen 1996). In words, it means that all the unmeasurable and unobserved factors influencing whether or not to spend are unrelated to the unmeasurable and unobserved factors affecting how much is spent. The alternative would be to assume that the errors are dependent. However, Smith (2003) shows that there is little statistical information to support the estimation of a DH with dependent errors even when dependence exists.The parameter vectors are α, β assumed to be linear.

For a positive level of expenditure on education to be observed, two hurdles (hence the name double hurdle) have to be overcome; firstly, the household must be a potential spender (i.e. $D_i = 1$) and secondly, it must actually spend on education (i.e. $Y_i^{**} = Y_i^*$). In the DH model, observed zeros in expenditure on education may arise either from participation or consumption decisions and potential spenders may have zero expenditure on education[9].

Using 0 to represent zero expenditure and + to denote positive expenditure, the sample likelihood equation for the independent double hurdle model can be written as follows;

$$L = \prod_0 \left[1 - \Phi\left(Z_i'\alpha\right) \left(1 - \Phi\left(\frac{X_i'\beta}{\sigma} \right) \right) \right] \prod_1 \Phi\left(Z_i'\alpha\right) \frac{1}{\sigma} \phi \left(\frac{Y_i - X_i'\beta}{\sigma} \right) \tag{2.4}$$

Where $\Phi(.)$and $\phi(.)$denote the standard normal cumulative density function (CDF), and the standard normal probability density function (PDF) respectively. The likelihood function above (equation 2.4), reduces to that of a Tobit when $\Phi\left(Z_i'\alpha\right) = 1$. A closer look at the likelihood function (equation 2.4) reveals that it is simply a product of the likelihood functions of a probit model and a truncated regression model where truncation is at zero. In other words, the log likelihood of the independent DH is the sum of log likelihood functions of a probit model and a truncated regression model where truncation is at zero. This is quite useful as it implies that the independent DH can be estimated by estimating the probit and truncated regressions separately [10]. Accordingly, a likelihood ratio test can

[9]This is unlike the Heckman model (Heckman 1979), where zeros in expenditure would arise only through participation.

[10]It worth pointing out that in the independent DH unlike the dependent DH exclusion restrictions are

be used to test the Tobit model versus the independent DH[11].

2.4.2 Variables used

As said earlier, the DH model is estimated separately by area of residence (rural and urban). The dependent variable is the share of total annual household expenditure on the education of primary school children in total annual consumption expenditure[12]. In order to account for price variability across areas and time, both expenditure items are deflated by using the Malawi National Statistical Office's spatial and temporal deflator with base national, and February/March 2004. The expenditure items include; fees (tuition and boarding), books and other materials, school uniform, contributions to school building and maintenance, parental association fees, and other school related expenses. In coming up with the factors which influence household investment in the education of children, we are guided by human capital theory as discussed in the theoretical framework as well as other empirical studies which have looked at parental investment in education.

We include the age of the youngest primary school going child in the household; this is motivated by the fact that as children get older education expenditures increase. Age of the child may also reflect the opportunity cost of home production which increases with age. We include the square of age of the youngest child to measure possible nonlinearities. Household permanent income as proxied by the log of total household per capita expenditure[13] has been found to affect spending on education (e.g. Song *et al.*2006; Yueh 2006; Kingdon 2005). The expectation as intimated in the theoretical literature is that if schooling is a pure investment good and capital markets are perfect then income should not influence spending on education, however income will influence spending on education if it is a consumption good and/or it is an investment good but there are credit constraints. We also include a variable which captures proportion of children who are day scholars in a

not needed to identify the parameters.

[11] The log likelihood ratio test statistic (LR) is computed as follows: $LR = -2\left[LL_T - (LL_P + LL_{TR})\right] \sim \chi^2_k$; where LL_T = log likelihood for the Tobit model; LL_P = log likelihood for the Probit model; LL_{TR}= log likelihood for the truncated regression model. LR follows a Chi-square distribution with degress of freedom k equal to the number of independent variables in the equations.

[12] One could alternatively use absolute expenditure on education as the dependent variable. We use the Engel curve approach in keeping with similar studies looking at household expenditure on education e.g. Kingdon (2005), Yeuh (2006), and Song *et al.*(2006).

[13] We use consumption expenditure other than income for two reasons. First, particularly in an agricultural economy such as Malawi, income is often very lumpy. Farming households receive a large amount of cash income in May and June after the harvest, and receive very little the rest of the year. In contrast, households are constantly expending their income and consuming. Consumption expenditure is a smoother measure of welfare through time than is income. In other words, consumption can be viewed as realized welfare, whereas income is more a measure of potential welfare (Murkhejee and Benson 2003). Second, in Malawi much of household income is derived from self-employed business or subsistence-oriented agricultural production. Assigning income values to the proceeds of these enterprises is often problematic (Hentschel and Lanjouw 1996).

household. This variable is defined as the number of day scholars divided by the number of children in the household.

The number of children in a household may also affect whether or not a household spends on their education, and if so how much. In the literature there are basically two opposite findings regarding the impact of number of children on investment in human capital. The first finding which confirms the quantity-quality trade off is that having more children negatively impacts on investment in human capital (Gertler and Glewwe 1990). The other finding is that having more children actually increases human capital formation as it ensures that each child requires less time for home production (Al-Samarrai and Reilly 2000). Additionally, we include the square of number children in the household to measure the possibility that expenditures diminish with more children.

Employment status of parents may be positively related to expenditures on a child's primary education as it may influence their perception of the relationship between human capital investments and returns on those investments. Studies by Haveman *et al.* (1991) and Ribar (1993) in the US, find a significant and positive relationship between mother's employment during a child's teenage years, and high school completion but find no significant effect on the same of the father's employment. In this study, we measure the employment status of both parents by whether they work for a wage or not. The educational level of parents is expected to have a positive effect on investment in education. The theoretical explanation of this expectation is that parents with higher levels of education are more likely to perceive greater future benefits or returns on investing in their children's education and, thus may be willing to sacrifice more for these future returns. More educated parents expect that their children will exhibit greater promise and thus will be more willing to invest in their child's education (Becker 1981; Becker and Tomes 1976). At the empirical level, several studies which look at the relationship between attainment and parental education support this human capital perspective (e.g. Ray 2000; Gertler and Glewwe 1990; Song *et al.* 2006; Kabubo-Mariara and Mwabu 2007).

Parental age may influence expenditures on children's primary education. Age reflects experience, and the expectation is that with age comes the ability to appreciate the benefits and returns on investments in education. As argued by Mauldin *et al.* (2001), if parents are older at the time their children are in primary and secondary schools, they will be more financially secure as well and be more willing to sacrifice a larger proportion of income for their children's education. We thus include the age of the mother and father as well as the square of ages for both parents to measure the possibility of nonlinearities. Studies by Case and Deaton (1999), Lavy (1996), and Al-Samarrai and Reilly (2000) have found significant negative effects of distance to the nearest primary school. Distance to the nearest primary school can be a measure of the quality of access of primary schools, it can also reflect the direct cost of primary education. Households are less likely to invest in the education of their children if for example schools are very far. In this study, distance to nearest primary school measured in kilometres is set equal to zero if there is a primary school in the community.

As has been discussed in the theoretical literature, there may be bias in spending against a particular sex. Besides, some empirical studies have found evidence of son preference in spending for example, Song *et al.*(2006) and Yueh (2006) for China and Kingdon (2005) for India. In order to capture the possibility of gender bias in spending, we construct a variable defined as; $\sum_{i=1}^{10} \frac{H_g}{H}$, where H_g is the number of household members in age-gender group g and H is the household size[14]. We distinguish ten age and gender categories; ages 0-6, 7-15, 16-19, 20-55, and over 55 for each gender. Since we are using aggregate household education expenditure data, this variable can give an indirect test of gender bias in spending. In particular, to check for evidence of differences in spending between primary school going boys and girls we are concerned with the coefficients of the age-gender variable for the ages 7-15 for both sexes. If the coefficients are significant and different that is evidence of preference for a particular sex in spending[15]. We control for regional fixed effects by including a three class regional dummy for the north, centre, and south.

2.4.3 Estimation issues

The log of per capita expenditure is potentially endogenous, and this may lead to biased and inconsistent results. One possible channel of endogeneity is that the log of per capita expenditure and spending on education can be jointly determined through labour supply decisions in the sense that a decision to send children to school may be jointly determined with a decision to send the children to work to supplement household income. Another route for endogeneity would be that parents with a good taste for the education of their children may work harder so they are able to pay for their schooling (Kingdon 2005).

We address this problem in both the participation and expenditure decision equations. In the participation equation we use the Rivers and Vuong (1989) procedure for discrete choice models, and in the expenditure equation we use the Smith and Blundell (1986) procedure for limited dependent variable models. The two procedures are analogous and they are done in two stages. In the first stage, a reduced form regression of an endogenous variable is regressed using ordinary least squares (OLS) on exogenous variables including instruments, and residuals are predicted. In the second stage, the predicted residuals are

[14] In the estimation the age-gender category over 55 for males is omitted to avoid multicollinearity since the categories sum up to one in each household.

[15] Testing for equality of coefficients in both participation and expenditure equations for all groups of household is done by using a Wald test. This approach to testing for gender bias was first proposed by Deaton (1989).

included in the participation equation (Rivers and Vuong procedure) and the expenditure equation (Smith and Blundell procedure) including the endogenous variable. A simple t-test of the coefficient on the residual tests the null hypothesis of exogeneity. We use household assets namely hectares of land, and its square as instrumental variables for log of per capita expenditure[16].Similar instruments are used by Glewwe and Jacoby (1994), Glewwe and Ilias (1996), and Kingdon (2005). An instrumental variable (IV) must be correlated with the endogenous variable (log of per capita expenditure in our case), but uncorrelated with the error term for the participation equation or the expenditure equation i.e. the IV must be redundant in the participation equation or the expenditure equation once log of per capita expenditure is included. Thus, the effect of the IV on school spending must work through log of per capita expenditure only. As is shown later, land and its square are correlated with log of per capita expenditure. Land is an illiquid asset, and therefore is unlikely to be sold in the short term to cover schooling expenses (Kingdon 2005). One can argue that land cannot work as IV for urban areas. It has to said that as of 1998, nearly eight out of 10 households in Malawi owned land where agricultural crops were grown, and 15.1% of urban households owned land for agricultural crops (NSO 1998)

2.4.4 Data and descriptives

The data used in the study come from the Second Malawi Integrated Household Survey (IHS2). This is a nationally representative sample survey designed to provide information on the various aspects of household welfare in Malawi. The survey was conducted by the National Statistical Office from March 2004 -April 2005. The survey collects information from a nationally representative sample of 11,280 households. In addition, the survey collects information from a nationally representative sample of 564 communities. During the survey a "community" was defined as the village or urban location surrounding the selected enumeration area, which most residents recognize as being their community. The households were sampled from these communities. This data contains detailed information on socioeconomic and demographic characteristics of the households. The survey also collects annualized household education information which includes household expenditure on primary, secondary, and tertiary education, for household members aged 5 and above. The expenditure items are; school fees (tuition and boarding), books and other materials, school uniform, contributions to school building and maintenance, parental association fees, and other school related expenses. In this study, we use husband-wife and single-parent families with at least one child in primary school. We do this for two reasons. Firstly, the survey does not record the parental characteristics of children who do not live with their parents, thus this restriction allows us to examine the impact of parental characteristics as discussed in section 2.4.2. Secondly, schooling decisions are cumulative

[16] Similar instruments are used by Glewwe and Jacoby (1994), and Glewwe and Ilias (1996).

in nature such that the circumstances in which a person was raised in as a child are more relevant than current ones (Glick and Sahn 2000). This restriction may potentially lead to a non random sample (i.e. a selected sample), which may bias our results. Specifically, if children are fostered out or older children leave the house to marry or work, this may lead to a selected sample of children who are different from those that have left. Since fostering increases with age and the likelihood of children leaving to marry or work also increases with age, by focussing on primary education, we somewhat mitigate the fears of selection bias.

Descriptive statistics of all the variables used in the analysis for families with nonzero expenditures and for the full sample by area of residence are presented in Tables 2.1 and 2.2. The full sample comprises households with primary school going children, with zero expenditures and nonzero expenditures on education. In Table 2.1, we report sample means of annual household expenditure on primary education (absolute expenditure) and the share of annual expenditure on primary education in total household consumption expenditure; our dependent variable. The table also presents results of tests of statistical significance of the differences in expenditure between rural and urban households. The results show that there are differences between rural and urban households. In terms of absolute expenditure, rural households spend less on average compared to urban households. The share of education spending out of total household consumption expenditure for rural households is lower than that of urban households. These differences hold for both the full and spending samples. Additionally, the differentials are statistically significant. Looking at the various components of expenditure on education, we notice that urban spending on all items is significantly higher than that of rural households. We also observe that for urban households tuition takes up a big part of spending, whereas for rural households most of the spending is done on uniforms.

Table 2.2, presents results of summary statistics of explanatory variables used in the study by area of residence for the full sample and the sample of households which actually spend on education. The table also reports whether the differences in the variables are statistically significant. With the rural-urban demarcation of the sample, we have 3739 rural households and 676 urban households with primary school going children. Of these full samples, 2782 rural households (74.4% of sample) and 548 urban households (81.1% of sample) have nonzero expenditures on primary school children. Thus suggesting that compared to rural areas, there are more households in urban areas with positive expenditures on education. In terms of the proportion of children going to day schools, the results show that rural households have a higher number (90%) compared to 87% for urban households. The difference is statistically significant. Urban households have generally significantly better parental characteristics. Specifically, in urban areas a significantly higher proportion of both mothers and fathers work for a wage, and have more years of schooling compared to their rural counterparts. The results show that the urban households have significantly nearer schools compared to rural ones. Looking at the age-gender demographics for the primary school going age (7-15), the results suggest that there are differences between the two areas with rural households having a significantly higher proportion of boys (16%) compared to 13% for urban households. In terms of the proportion of girls of the schooling going age, we find no significant difference between the two areas. Essentially, we observe that just like expenditure on education discussed earlier; there are differences in the characteristics across area of residence. We have also reported enrolment rates by age for rural and urban areas (see appendix Table A2.1), again here we get a picture which is consistent with that from education spending. We discuss the econometric results in the next section.

2.5 Econometric results

The descriptive statistics show that there are differences in expenditure on primary education as well as characteristics between rural and urban households. In the light of this, we formally test the hypothesis that households in rural and urban areas are not different with respect to their investment in children's education[17]. We essentially seek to investigate whether or not coefficients for the different variables are the same for rural and urban households. This is done by conducting a pooling test; a failure of pooling between the two groups would indicate that they are different. To conduct the pooling test, we use the likelihood ratio (LR) test. For comparison, the hypothesis is tested using both the DH and the Tobit models. The unrestricted regression is estimated with separate urban and rural households, and the restricted regression with the pooled sample using an area of residence dummy variable 'rural'. If we denote the log-likelihoods for the urban, rural and pooled samples respectively as $LL_{urban}, LL_{rural}, LL_{pooled}$ with corresponding number of parameters $k_{urban}, k_{rural}, k_{pooled}$, then the LR statistic which follows a Chi-square distribution with degrees of freedom $(k_{rural} + k_{urban}) - k_{pooled}$ is given by;

$$LR = -2\left[LL_{pooled} - (LL_{rural} + LL_{urban})\right] \sim \chi^2_{(k_{rural}+k_{urban})-k_{pooled}} \tag{2.5}$$

Results of the pooling tests are presented in Table 2.3. The results for both the DH and Tobit models show that rural and urban households are different, and thus pooling the rural and urban households is inappropriate. This means that the DH model or the Tobit

[17] Since we do not have information on whether the expenditures are on private or government primary schools, in our preliminary estimations we dropped tuition fees as we figured this may be a major factor between urban and rural areas, in the sense that there is a predominance of private schools which tend to be expensive in urban areas. However, our econometric results were by and large unaffected by this exclusion, so we retained tuition fees in all estimations in the study.

model should be estimated separately for the two areas. The next issue that we address is whether the DH or Tobit is the right model for our data. Basically, we seek to ascertain by using the LR test whether there is another censoring mechanism as represented by the participation equation. Results of the tests are reported in Table 2.4. The LR test results favour the use of the independent DH as opposed to the Tobit model. This implies that there are two decision processes underlying spending on education; households decide whether or not to spend, and if yes, how much. We therefore discuss results of the DH for the two groups of households.

As discussed earlier the log of per capita expenditure is potentially endogenous, we tested for this using the Rivers and Vuong procedure for the participation equation and the Smith and Blundell procedure for the expenditure equation as outlined earlier. We find that the log of per capita expenditure is endogenous in the expenditure equation only for rural households. To ensure comparability in terms of number of variables, we included residuals from the reduced form regression for urban households in the urban expenditure equation as well. The reduced form regressions of log of per capita expenditure for both areas reported in the appendix Table A2.1, show that the instrumental variables land and its square perform reasonably well as they are significantly correlated with the log of per capita expenditure.

The final maximum likelihood results of the DH are presented in Table 2.5. Since the Tobit model has been rejected in favour of the DH, our discussion of the results is based on the DH but we show results of the Tobit model (Table A2.3 in the appendix) for comparison. The results generally show that some variables are significant for one group but insignificant for another; an indication of the rural-urban differences alluded to earlier. The age of the youngest child is significant and negative only in the participation equation for rural households. This suggests that parents in rural areas are less likely to spend on the education of children as they get older. This perhaps reflects the opportunity cost of sending children to school, that is as they get older they can be a source of labor for agriculture, and other income generating activities to supplement parental income. This opportunity cost may not be as high in urban areas. The level of income as proxied by the log of per capita expenditure significantly increases the likelihood of spending on education and how much is spent for both rural and urban households. The results therefore suggest that income matters at both the extensive and intensive margins for the two groups of households. Mauldin *et al.* (2001) also find that income has positive and significant effect on household spending on education at both decision levels in the US. We cannot compare the magnitudes of these coefficients of income in the two areas, but later in the next section we compare the magnitudes of the coefficients by computing elasticities. Suffice to say that the positive and significant effect of income indicates that spending on education is considered a normal consumption good. It may also indicate the presence of credit constraints in both areas.

For rural households, having a higher proportion of children going to day schools significantly increases the probability of spending on them but lowers the share of education expenditure. For urban households having more day scholars lowers the chance of spending on primary education but it has no impact on the share of education expenditure in total expenditure. We find that the number of children influences positively and significantly the share of education expenditure for rural households, but does not significantly affect the likelihood of spending on education[18]. For urban households having more children increases the likelihood that a household will spend on their education but does not affect the share of expenditure. This positive effect conforms with the argument by Al-Samarrai and Reilly (2000) that the more children a household has, the less is the time needed for household production activities, and hence the higher will be the investment in their education. This however, contradicts an argument by Gertler and Glewwe (1990) that larger families may derive less utility from sending an additional child to school if some are already enrolled. This lower enrolment resulting from having many children could be reflected in lower spending. This also runs counter to the expectation that with more children there is more competition for resources.

In terms of parental employment, the results show that for rural and urban households a father's and a mother's employment significantly increases the share of expenditure on education as well as the chance that they will spend on children. This suggests that holding other things constant, employed parents will invest more on their children. With respect to education, we find that the education of both the mother and the father positively and significantly affects the decision whether or not to spend as well as how much to spend on the primary education of their children in both rural and urban areas. Thus, *ceteris paribus* the higher is the parental human capital, the higher will be the investment in schooling of children. These results are in line with findings by Song *et al.* (2006) for rural China where they found that the educational level of both parents positively impacts household spending on education. We cannot compare the magnitudes of the DH coefficients of the employment and education for parents in the two areas, however this issue is taken up later in the next section where we compute elasticities. These comparisons allow us to say something about the possible differences in the impact of the two variables between parents and between the two areas.

The quality of access of primary schools as proxied by distance to the nearest primary

[18] It is worth recognizing that the number of children is potentially endogenous, if there is a quantity-quality trade off where parents prefer fewer children with a good education. Besides, if there is son preference which affects expenditure on children's education, this may also affect family size. We control for the possibility of son preference as discussed earlier. Since we have no valid instruments; we addressed the simultaneity problem arising from the quantity-quality trade by re-estimating the DH models for all groups without number of children; our results largely remained unchanged thus giving us confidence that our results may not be biased due to simultaneity.

school has a negative impact on the participation and the expenditure decisions of both rural and urban households[19]. This suggests that households will be less likely to spend on primary education if the schools are far away and if they do actually decide to spend, the amount spent will be lower[20]. In terms of the age-gender demographics, the results suggest that having more primary school going boys (i.e.proportion of males aged 7-15) and girls (i.e.proportion of females aged 7-15) significantly and positively impacts on the participation and the expenditure decision levels of rural households. The same is true for urban households. We investigate further to check evidence of gender bias against girls by conducting Wald tests of the equality of the coefficients for proportion of males and females aged 7-15 in the two areas. Results of the tests are shown at the bottom of Table 2.5. The test results indicate that for rural households there is gender bias against girls at both the participation and expenditure decision levels. For urban households, the Wald test results indicate that there are no statistically significant gender differences at both the intensive and extensive margins. Thus, the Wald tests show evidence of gender bias in favour of boys in rural areas only. Interestingly, we observe that when the Tobit model is used (see Table A2.3 in the appendix), there is no evidence of gender bias in both areas. This is in conformity with a finding by Kingdon (2005) who shows that when a variant[21] of the DH model is used more evidence of gender bias in school spending is found in India as compared to using a single equation model. This underlines the importance of the participation decision when modelling a dependent variable with excess zeros. We complement the Wald tests results by comparing the magnitude of elasticities for the proportion of males and females aged 7-15 in the next section.

We have assessed the impact of different regressors on expenditure, and found some to be significant in the levels equation only while others are significant in the participation equation only or both the levels and participation equations. Further to that, some variables have been found to have opposite signs in the two decision levels. As noted by Yen (2005), when examining the impact of explanatory variables, the presence of parameter estimates with opposite signs in the participation and level equations complicate the interpretation of the estimated effects. Thus, the impact of explanatory variables can be better explored

[19]Distance to the nearest primary school can be endogenous, for example some communities may have a leadership which values education and is more vocal and progressive. This may affect both household schooling decisions as well as placement of schools. Another possible source of endogeniety is that parents with high aspirations for their children may "vote with their feet" by moving to areas where schools are nearer. And this unobserved high aspiration by parents may affect both distance to schooling and schooling decisions. We don't have valid instruments for distance to nearest primary school, so we re-estimated the models without distance to nearest primary school and our results were marginally different from those with distance to nearest primary school thus giving us some level of assurance about the reliability of our results.

[20]If the distance to the nearest primary school is thought of as a measure of the direct cost of primary education, then the result suggests that households will be less likely to spend on primary education if costs are high and if they do actually decide to spend, the amount spent will be lower.

[21]The model used by Kingdon (2005) assumes that once a household decides to spend there are no zero expenditures. Essentially, implying that the first hurdle dominates the second hurdle.

by computing elasticities. It is worth noting that the elasticities unlike the coefficients we have just discussed also allow us to talk about the economic significance of the variables used.

2.5.1 Elasticities in the independent DH

The interpretation of coefficients in limited dependent variable models is complicated, and to overcome this the effect of explanatory variables on the unconditional expectation of the dependent variable (Y_i) as measured by elasticities is decomposed into an effect on the probability of a positive expenditure and an effect on conditional expenditure (Yen 2005)[22].

The unconditional expectation of Y_i in the independent DH is given as;

$$E(Y_i) = \Pr(Y_i > 0)E(Y_i|Y_i > 0) \tag{2.6}$$

Where the probability of expenditure is given by;

$$\begin{aligned} \Pr(Y_i &> 0) = \Pr(Z_i'\alpha + \varepsilon_i > 0, X_i'\beta + \nu_i > 0) \\ &= \Pr\left(\varepsilon_i > -Z_i'\alpha, \nu_i > -X_i'\beta\right) \\ &= \Phi\left(Z_i'\alpha\right)\Phi\left(\frac{X_i'\beta}{\sigma}\right) \end{aligned} \tag{2.7}$$

The conditional expectation of Y_i is expressed as[23];

$$\begin{aligned} E(Y_i|Y_i &> 0) = X_i'\beta + E\left(\nu_i|\varepsilon_i > -Z_i'\alpha, \nu_i > -X_i'\beta\right) \\ &= X_i'\beta + \left[\Phi\left(Z_i'\alpha\right)\Phi\left(\frac{X_i'\beta}{\sigma}\right)\right]^{-1}\sigma \times \phi\left(\frac{X_i'\beta}{\sigma}\right)\Phi\left(Z_i'\alpha\right) \\ &= X_i'\beta + \frac{\sigma\phi\left(\frac{X_i'\beta}{\sigma}\right)}{\Phi\left(\frac{X_i'\beta}{\sigma}\right)} \end{aligned} \tag{2.8}$$

The elasticities of the unconditional expectation of Y_i with respect to the continuous regressors are computed by differentiating equations 2.7 and 2.8, and using the adding up property, equation 2.6. Formally, the elasticity of a continuous variable j which appears

[22] This follows a proposed decomposition by McDonald and Moffit (1980) for Tobit models on the effect of a regressor on the unconditional expectation.

[23] The probability of positive expenditure and conditional expectation of expenditure are based on the error term properties given earlier. See Yen (2005) for details of the same when errors are dependent.

in both the participation and the expenditure equations is written as follows:

$$\begin{aligned}
\eta_j^{UC} &= \frac{\partial E(Y_i)}{\partial X_{ij}} \frac{X_{ij}}{E(Y_i)} \\
&= \frac{\partial \Pr(Y_i > 0)}{\partial X_{ij}} \frac{X_{ij}}{\Pr(Y_i > 0)} + \frac{\partial E(Y_i | Y_i > 0)}{\partial X_{ij}} \frac{X_{ij}}{E(Y_i | Y_i > 0)} \\
&= \eta_j^P + \eta_j^C
\end{aligned} \tag{2.9}$$

Equation 2.9, shows that the elasticity of the unconditional expectation of Y_i with respect to a continuous variable j which appears in both the participation and the expenditure equations (η_j^{UC}), is simply a sum of the elasticity of the probability of observing a positive expenditure (η_j^P) and the elasticity of conditional expenditure (η_j^C).

These elasticities of the probability, conditional level and unconditional level for continuous variables are computed at the sample means of the regressors. Table 2.6 reports the elasticities for the probability, conditional and unconditional levels of some selected variables for the DH. For comparison, we present the elasticities for the probability, conditional and unconditional levels of some selected variables for the Tobit model in the appendix Table A2.4. The elasticity of probability for both rural and urban households with respect to the log of per capita expenditure which proxies permanent income is positive and significant implying that spending on education is considered a normal item. The same holds true for the elasticity of conditional and unconditional levels for the log of per capita expenditure. It is worth noting that rural households have greater than one elasticities of the probability, conditional level and unconditional level compared to urban households. We test whether the income elasticities are statistically greater than one for rural areas and less than one for urban areas[24]. For rural areas with t-statistics 4.25, 2.64, and 5.04 respectively for probability, conditional level and unconditional level, we reject the null that the elasticities are equal to one and conclude that they are greater than one. For urban areas with t-statistics, -10.2, -6.9, and -7 respectively for probability, conditional level and unconditional level, we reject the null that the elasticities are equal to one and conclude that they are less than one. This means that for rural households spending on the schooling of children is more sensitive to income compared to urban households, and thus schooling is a luxury good in rural areas[25].

The elasticities of probability, conditional level and unconditional level with respect to parental employment and education are positive and significant in both areas. However,

[24]The hypothesis is done using the following t-statistic; $t = \frac{\hat{\eta} - 1}{se(\hat{\eta})}$. Where; $\hat{\eta}$ is an estimated elasticity (probability, conditional level, and unconditional level), and $se(\hat{\eta})$ is the corresponding standard error.

[25]We do not address the possibility that the elasticity of expenditure on education with respect to income may vary non-monotonically i.e. the income elasticities peak in the middle-income categories (have a value of greater than one), and diminish for the lower and upper ends of the income distribution (For details on this possibility see Hashimoto and Heath 1995).

we note two things, firstly the elasticities for parental employment and education are higher for rural areas, and secondly, the elasticities for mothers employment and education are higher than those of fathers in both areas. These findings indicate that parental characteristics have a bigger impact on spending in rural areas, and that a mothers characteristics have a larger impact on spending compared to a father's. If one thinks of the employment status and education of the mother as a reflection of the bargaining power of the mother in the household, this would imply that children's education benefits from an improvement in the bargaining position of the mother. Besides, this result has intergenerational implications for human capital formation in that more female education entails more educated mothers, and hence more education for children.

The elasticities of probability, conditional level and unconditional level with respect to the distance to the nearest primary school are negative and statistically significant for both areas. We observe that the elasticities are larger for urban areas as compared to rural areas suggesting that urban households are more sensitive to the quality of access of primary schools. The elasticities of probability, conditional level and unconditional level with respect to the proportion of primary school going boys (proportion of males aged 7-15) and girls (proportion of females aged 7-15) are positive, statistically significant and economically substantial for rural and urban households. In addition, we also note that for rural households the elasticities of probability, conditional level and unconditional level for boys are larger than those for girls suggesting a bias against girls. The computed elasticities for urban households are not noticeably different. These elasticities therefore reinforce evidence shown earlier using Wald tests that boys are favored when it comes to whether or not to spend as well as how much to spend in rural households, but there is no evidence of school spending gender bias in urban households. Just like the raw coefficients discussed earlier for the Tobit model, we find that the elasticities (see appendix Table A2.4) are both statistically insignificant and economically not very different from each other. Thus, when a single equation model is used we find no evidence of gender bias in spending in both rural and urban households.

Both the descriptive and econometric results show that there are differences in household investment in the human capital of primary school children. Specifically, the results indicate rural and urban households are different both in terms of how much they spend and the effect of different characteristics on their spending behavior. We therefore know that there are these differences, but we don't know why there are these differences. Are these differences largely due to differences in characteristics or due to differences in behavior? The next section addresses this issue.

2.6 Extending the Oaxaca-Blinder decomposition to the independent DH model

The observed rural-urban differences in household investment in the education of primary school children call for an understanding of what explains these differences. This section therefore provides a comprehensive analysis of the rural-urban differential in household expenditure on education. To achieve this, we propose an extension of the decomposition technique proposed independently by Oaxaca (1973) and Blinder (1973) for linear models to the independent DH model, which is a nonlinear model. The technique has almost exclusively been used in the labour economics literature to study gender wage discrimination (e.g. Appleton *et al.* 1999; Sicillian and Grossberg 2001; Neuman and Oaxaca 2004), and to the best of our knowledge our study is the first to apply the technique to study household expenditure. The proposed decomposition isolates the expenditure gap into a *characteristic effect*, which is a part of the differential explained by differences in social-economic characteristics, and a *coefficient effect* which is the part of the gap which is due to differences in coefficients. In this study, we interpret the *coefficient effect* as part of gap which is due to household behavior[26].

As will be demonstrated later, the standard Blinder-Oaxaca decomposition method cannot be used to decompose the DH as it is strictly meant for linear models. For nonlinear models; Fairlie (1999, 2005) has proposed the Blinder-Oaxaca decomposition for logit and probit models, Bauer and Sinning (2005, 2008) have proposed an extension of the same for Tobit models. To derive the Blinder-Oaxaca decomposition for the independent DH; consider the DH as expressed in equation 2.3, which is estimated separately for two groups of households, $m = (U, R)$, where; U =urban and R =rural households. We want to decompose the gap in average expenditure share between urban and rural households, $\triangle^{DH}= E(Y_U) - E(Y_R)$, by using the following sample counterpart $\hat{\triangle}^{DH} = \bar{Y}_U - \bar{Y}_R$. The sample average expenditure share for group m is given as $\bar{Y}_m = \frac{\sum_{i=1}^{N_m} \hat{Y}_{im}}{N_m}$; where N_m is the sample size for group m. The "hat" denotes sample estimates. The Blinder-Oaxaca decomposition of the independent DH similar to that for the Tobit by Bauer and Sinning (2005, 2008) is expressed in terms of unconditional expectations of the dependent variable (Y_i). The unconditional expectation for the two groups estimated separately is expressed as follows[27];

$$E(Y_{im}) = \Pr(Y_{im} > 0)E(Y_{im}|Y_{im} > 0) \tag{2.10}$$

[26] The coefficient effect in the labor economics literature is interpreted as a measure of discrimination.
[27] For ease of exposition, we have reproduced equation 2.6.

Where the probability of expenditure is given by;

$$\Pr(Y_{im} > 0) = \Phi\left(Z'_{im}\alpha_m\right)\Phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right) \tag{2.11}$$

The conditional expectation of Y_i is expressed as;

$$E(Y_{im}|Y_{im} > 0) = X'_{im}\beta_m + \frac{\sigma_m\phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right)}{\Phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right)} \tag{2.12}$$

Three things need to be noted about equation 2.10. Firstly, the unconditional expectation $E(Y_{im})$ is not equal to $E(X_{im})'\beta_m$ as is the case in linear models on which the standard Blinder-Oaxaca decomposition is based[28,29]. As discussed earlier, imposing a linear model on a dependent variable with excess zeros leads to biased and inconsistent coefficients, and therefore using coefficients from the linear model would give a misleading decomposition as well. Secondly, the unconditional expectation is not equal to that of Tobit as it has another censoring mechanism, $\Phi\left(Z'_{im}\alpha_m\right)$ which represents participation; this means that we cannot use the Blinder-Oaxaca decomposition for Tobit models as developed by Bauer and Sinning (2005, 2008). Finally, equation 2.10 shows that the unconditional expectation has the standard error of the error term of the expenditure equation,σ_m. This may affect the magnitude of the decomposition and therefore has to be included in the decomposition. As a result, there are several possible decompositions of the mean difference $\triangle^{DH}$, depending on which σ_m is used in the counterfactual part of the decomposition.

We therefore derive two possible decompositions for the independent DH[30]:

$$\begin{aligned}\triangle^{DH}_{R_1} = & \left[E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iU}\right) - E_{\alpha_U,\beta_U,\sigma_R}\left(Y_{iR}\right)\right] \\ & + \left[E_{\alpha_U,\beta_U,\sigma_R}\left(Y_{iR}\right) - E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iR}\right)\right]\end{aligned} \tag{2.13}$$

[28] It is worth noting the difference in terminology used here; the conditional expectation in linear models is given by $E(Y|X)$ while the conditional expectation in limited dependent variable models (e.g. Tobit, Truncated, DH models) is expressed as $E(Y|Y > 0)$.

[29] Assuming a linear model $Y_{im} = X'_i\beta + \nu_i$ for illustration; the standard Blinder-Oaxaca decomposition is based on the property of linear models with an intercept that the mean of a dependent variable is equal to the mean of the regressors evaluated at their respective estimated coefficients i.e. $\bar{Y}_{im} = \bar{X}_{im}\hat{\beta}_m$. Hence, the standard Blinder-Oaxaca decomposition is given as; $\bar{Y}_U - \bar{Y}_R = \left(\bar{X}_U\hat{\beta}_U - \bar{X}_R\hat{\beta}_R\right) = \left(\bar{X}_U - \bar{X}_R\right)\hat{\beta}_U + \left(\hat{\beta}_U - \hat{\beta}_R\right)\bar{X}_R$.

Where the "overbars" denote sample means and the "hats" denote sample estimates.

[30] These two possibilities are similar to that of Bauer and Sinning (2005) for the Tobit.

and

$$\triangle_{U_1}^{DH} = \left[E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iU}\right) - E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iR}\right)\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iR}\right) - E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iR}\right)\right] \tag{2.14}$$

Where $E_{\alpha_m,\beta_m,\sigma_m}\left(Y_{im}\right)$ denotes the unconditional expectation of Y_{im} evaluated at the parameter vectors α_m, β_m and the standard error σ_m. The difference between the two decompositions is that equation 2.13 treats the standard error as part of the variables while equation 2.14 treats it as part of the coefficients.

The above decompositions use the urban coefficients in the counterfactual; this implies that if there was no gap in average expenditure share, the expenditure profile of the urban would prevail. We can alternatively use the rural coefficients; this implies that if there was no gap in average expenditure, the expenditure structure of the rural areas would exist. When the rural coefficients are used the two possibilities are written as[31] :

$$\triangle_{U_2}^{DH} = \left[E_{\alpha_R,\beta_R,\sigma_U}\left(Y_{iU}\right) - E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iR}\right)\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iR}\right) - E_{\alpha_R,\beta_R,\sigma_U}\left(Y_{iR}\right)\right] \tag{2.15}$$

and

$$\triangle_{R_2}^{DH} = \left[E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iU}\right) - E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iR}\right)\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}\left(Y_{iR}\right) - E_{\alpha_R,\beta_R,\sigma_R}\left(Y_{iR}\right)\right] \tag{2.16}$$

The first term in the decompositions (equations 2.13 -2.16) captures part of the average expenditure share gap between the urban and rural households attributable to differences in covariates. This is the *characteristic effect*. This basically is the part of the gap in average expenditure share between the two groups of households assuming that both types had the same coefficients (behavior) but different endowments. Thus, this is a part of the gap explained by differences in characteristics. The last term in equations 2.13 -2.16, measures the difference in average expenditure share between the two groups which is due to differences in coefficients. This is the *coefficient effect*. It is part of the gap which is unexplained by the differences in characteristics. Essentially, it is part of the gap assuming that urban and rural households had the same characteristics but different

[31] This provides a robustness check of our results to choice of reference group. When decompositions give different conclusions depending on the reference group used, an index number problem is said to obtain. Various attempts have been made in the literature to resolve the index number problem for linear models (e.g. Reimers 1983; Cotton 1988; Neumark 1988; Oaxaca and Ransom 1994).

coefficients (behavior). So for example, assuming that rural and urban households have the same income levels, this income may be a more important factor (implying a bigger coefficient) to rural households as compared to urban ones in their spending decisions.

In order to conduct the Blinder-Oaxaca decomposition as given in equations 2.13 to 2.16, the following sample equivalent of the unconditional expectation (equation 2.10) is employed;

$$T\left(\hat{\alpha}_m, \hat{\beta}_m, Z_{im}, X_{im}, \hat{\sigma}_m\right) = N_m^{-1} \sum_{i=1}^{N_m} \left\{ \begin{array}{c} \Phi\left(Z'_{im}\hat{\alpha}_m\right) \Phi\left(\frac{X'_{im}\hat{\beta}_m}{\hat{\sigma}_m}\right) \\ \times \left(X'_{im}\hat{\beta}_m + \frac{\sigma_m \phi\left(\frac{X'_{im}\hat{\beta}_m}{\sigma_m}\right)}{\Phi\left(\frac{X'_{im}\hat{\beta}_m}{\hat{\sigma}_m}\right)} \right) \end{array} \right\} \tag{2.17}$$

Where $\hat{\alpha}_m, \hat{\beta}_m$, and $\hat{\sigma}_m$ denote sample estimates. With this sample counterpart of the unconditional expectation, equation 2.13 is estimated by;

$$\begin{aligned} \hat{\Delta}_{R_1}^{DH} = \quad & \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{2.18}$$

Equation 2.14 is estimated by;

$$\begin{aligned} \hat{\Delta}_{U_1}^{DH} = \quad & \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{2.19}$$

Equation 2.15 is estimated by;

$$\begin{aligned} \hat{\Delta}_{U_2}^{DH} = \quad & \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right)\right] \end{aligned} \tag{2.20}$$

Finally, equation 2.16 is estimated by;

$$\begin{aligned} \hat{\Delta}_{R_2}^{DH} = \quad & \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{2.21}$$

If there is only one censoring mechanism, that is $\Phi\left(Z'_{im}\hat{\alpha}_m\right) = 1$, decompositions 2.13 to 2.16 reduce to that of a Tobit with censoring from below at zero, as proposed by Bauer and Sinning (2005, 2008) for Tobit models. If expenditure is uncensored at zero, decomposition 2.13 and 2.14 are equal and reduce to the standard Blinder-Oaxaca decomposition with urban coefficients used in the counterfactual. Similarly, decompositions 2.15 and 2.16 are equal and reduce to the standard Blinder-Oaxaca decomposition with rural coefficients used in the counterfactual.

2.6.1 Detailed decomposition of the independent DH

The decomposition we have just derived gives us the overall or aggregate *characteristic effect* and *coefficient effect* of the independent DH. This while important gives us only a black box explanation of the differences in education spending between rural and urban households. It does not for example address the issue of how much of the *characteristic effect* arises from differences in household income. Similarly, it does not show how much of the unexplained gap is due to differences in household income. So a detailed decomposition which further disaggregates the two effects is important in pinpointing the major factors driving the spending gap. Knowledge of the major drivers of the spending gap is important for policy interventions aimed at closing or reducing the gap.

Owing to the difficulty in interpreting the detailed decomposition of the *coefficient effect,* this study only dwells on the detailed decomposition of the *characteristic effect* (see Jones 1983 for more details on the interpretational problems)[32]. In deriving the detailed decomposition of the *characteristic effect* of the independent DH[33], we use the average predicted gaps given in equations 2.18 to 2.21.

A detailed decomposition of the *characteristic effect* denoted as CE for the j^{th} variable $(j = 1, ..K)$ corresponding to equation 2.18 is given as;

$$CE_{R_1}^{DH} = \sum_{j=1}^{K} W_1^j \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right) \right] \tag{2.22}$$

for equation 2.19 it is expressed as;

$$CE_{U_1}^{DH} = \sum_{j=1}^{K} W_2^j \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right) \right] \tag{2.23}$$

for equation 2.20 it is represented as;

$$CE_{U_2}^{DH} = \sum_{j=1}^{K} W_3^j \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right) \right] \tag{2.24}$$

[32] In addition to the interpretational problems, a detailed decomposition of the *coefficient effect* for dummy variables may suffer from an invariance problem in the sense that the detailed coefficients effect attributed to dummy variables is not invariant to the choice of the base category (Oaxaca 1999). Solving this problem involves the estimation of a normalized regression (see Suits 1984; Gardeazabal and Ugidos 2005; Yun 2005).

[33] It should be noted that the Blinder-Oaxaca decomposition for Tobit models proposed by Bauer and Sinning (2005, 2008) does not go as far as the detailed decomposition.

and finally, for equation 2.21 it is denoted as;

$$CE_{R_2}^{DH} = \sum_{j=1}^{K} W_4^j \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \tag{2.25}$$

where the weights (W^j) are given as ;

$$\begin{aligned}
W_1^j &= \frac{\left(\left(\overline{X}_U^j \frac{\hat{\beta}_U^j}{\hat{\sigma}_U}\right)\left(\bar{Z}_U^j \hat{\alpha}_U^j\right) - \left(\overline{X}_R^j \frac{\hat{\beta}_U^j}{\hat{\sigma}_R}\right)\left(\bar{Z}_R^j \hat{\alpha}_U^j\right)\right)}{\left(\left(\overline{X}_U \frac{\hat{\beta}_U}{\hat{\sigma}_U}\right)\left(\bar{Z}_U \hat{\alpha}_U\right) - \left(\overline{X}_U \frac{\hat{\beta}_U}{\hat{\sigma}_R}\right)\left(\bar{Z}_U \hat{\alpha}_U\right)\right)} \\
W_2^j &= \frac{\left(\left(\overline{X}_U^j \hat{\beta}_U^j\right)\left(\bar{Z}_U^j \hat{\alpha}_U^j\right) - \left(\overline{X}_R^j \hat{\beta}_U^j\right)\left(\bar{Z}_R^j \hat{\alpha}_U^j\right)\right)}{\left(\left(\overline{X}_U \hat{\beta}_U\right)\left(\bar{Z}_U \hat{\alpha}_U\right) - \left(\overline{X}_U \hat{\beta}_U\right)\left(\bar{Z}_U \hat{\alpha}_U\right)\right)} \\
W_3^j &= \frac{\left(\left(\overline{X}_U^j \frac{\hat{\beta}_R^j}{\hat{\sigma}_U}\right)\left(\bar{Z}_U^j \hat{\alpha}_R^j\right) - \left(\overline{X}_R^j \frac{\hat{\beta}_R^j}{\hat{\sigma}_R}\right)\left(\bar{Z}_R^j \hat{\alpha}_R^j\right)\right)}{\left(\left(\overline{X}_U \frac{\hat{\beta}_R}{\hat{\sigma}_U}\right)\left(\bar{Z}_U \hat{\alpha}_R\right) - \left(\overline{X}_U \frac{\hat{\beta}_R}{\hat{\sigma}_R}\right)\left(\bar{Z}_U \hat{\alpha}_R\right)\right)} \\
W_4^j &= \frac{\left(\left(\overline{X}_U^j \hat{\beta}_R^j\right)\left(\bar{Z}_U^j \hat{\alpha}_R^j\right) - \left(\overline{X}_R^j \hat{\beta}_R^j\right)\left(\bar{Z}_R^j \hat{\alpha}_R^j\right)\right)}{\left(\left(\overline{X}_U \hat{\beta}_R\right)\left(\bar{Z}_U \hat{\alpha}_R\right) - \left(\overline{X}_U \hat{\beta}_R\right)\left(\bar{Z}_U \hat{\alpha}_R\right)\right)}
\end{aligned} \tag{2.26}$$

and

$$\sum_{j=1}^{K} W_1^j = \sum_{j=1}^{K} W_2^j = \sum_{j=1}^{K} W_3^j = \sum_{j=1}^{K} W_4^j = 1 \tag{2.27}$$

The contribution of each variable to the *characteristic effect* is computed by replacing the value of one group of households (rural or urban) with that of the other group of households sequentially one by one[34]. Assuming that there is only one censoring mechanism, the detailed decompositions in equations 2.22 to 2.25 reduce to that of a Tobit model with censoring at zero. Further, if expenditure is uncensored, detailed decompositions 2.22 to 2.25 reduce to that of linear models. The corresponding weights for both the Tobit and linear models reduce to the single equation weights as proposed by Yun (2004).

2.6.2 Results of the decomposition

We present the results and discussion of the aggregated decomposition in subsection 2.6.2, this is followed up by results and discussion of the detailed decomposition in subsection 2.6.2.

[34] The sequential replacement of each variable does not lead to path dependency i.e. it is insensitive to order of switching (see for example Yun 2004).

Results of the aggregate decomposition

Results of the proposed aggregate decomposition are presented in Table 2.7. For comparison we also show in Table A2.5 in the appendix results of the decomposition for the Tobit model. In both tables, we have also presented the actual average expenditure share gap for the full sample from Table 2.1. The results indicate that the DH model compared to the Tobit model has a lower approximation error, implying that it predicts spending more accurately. The gap in the predicted average share of primary education expenditure between rural and urban households is largely due to differences in characteristics. For example, looking at the expenditure differential when urban coefficients are used in the counterfactual, and we also use the urban variance in the counterfactual, 66% of the gap is due to differences in characteristics of the households, and 34% of the gap is explained by differences in estimated coefficients, hence due to behavioural differences. The two aggregate effects are statistically significant at 1%. This result means that if rural and urban household characteristics were to be equalized, 66% of the spending gap would vanish. On the other hand, if the behavior of rural and urban households was equalized, 34% of the spending gap would disappear. Similarly, when the urban coefficients and the rural variance are used in the counterfactual, the results indicate that the *characteristic effect* is 67.6% and that 32.4% of the expenditure gap is attributable to differences in coefficients. Both effects are statistically significant. In this case 67.6% (32.4%) of the spending gap would vanish if household characteristics (behavior) were equalized.

The picture that is emerging from the DH decomposition results is that the gap in spending between rural and urban households largely arises from differences in their characteristics. The same conclusion is arrived at when we ignore the participation equation and use Tobit model (see appendix Table A2.5). It is however worth noting that decomposition results for the Tobit consistently give a higher (lower) measure of the *characteristic effect* (*coefficient effect*); which suggests that when we when do not account for the fact that spending is made in two stages, we overestimate (underestimate) the *characteristic effect* (*coefficient effect*). In a nutshell, the DH and Tobit results suggest that the rural-urban gap in expenditure is mainly due to differences in characteristics; and this finding is robust to choice of both variance and coefficients[35] used in the counterfactual as well as ignoring the participation equation as a censoring mechanism.

Results of the detailed decomposition

The aggregated decomposition results presented in the preceding show that the rural-urban spending gap is predominantly due to differences in characteristics, however this

[35]The robustness of the decomposition results to choice of counterfactual implies that we do not have an index number problem.

does not tell us which characteristics are key. In Tables 2.8 and 2.9, we present results of the disaggregated decomposition of the DH. For comparison, we also report results of the same for the Tobit model in Tables A2.6 and A2.7 in the appendix. We have reproduced the *characteristic effect* in the top panel of the tables for ease of exposition. The detailed decomposition results of the DH show that a big part of the *characteristic effect* is taken up by six variables namely; household income, fathers and mothers education, fathers and mothers employment status, and the distance to the nearest primary school. This conclusion is robust to choice of variance and coefficients used in the counterfactual. For example, when we use the urban variance and the urban coefficients (rural coefficients) in the counterfactual, we find that these six variables constitute 83.59% (90.45%) of the *characteristic effect*, and the remainder,16.41% (9.55%) is taken by the other variables. This implies that these six variables are the major factors behind the rural-urban spending difference, and that an equalization of these six variables jointly between rural and urban households would wipe out 83.59% (90.45%) of the *characteristic effect*.

In terms of the specifics, and when we use the urban variance and the urban coefficients (rural coefficients) in the counterfactual, the results show that differences in household income as proxied by the log of per capita annual consumption take up 34.38% (36.36%) of the *characteristic effect*, and that this effect of income is statistically significantly different from zero. Thus, if household income alone was to be the same between the two areas, this would take off 34.38% (36.36%) of the *characteristic effect*. When we change the variance and coefficients used in the counterfactual, we get a similar story. This suggests that differences in household income are the largest factor in driving the rural-urban spending differential. This result conforms to a finding by Al-Samarrai and Reilly (2000) in Tanzania where they found differences in income to be the largest and statistically significant driver of rural-urban enrolment differences. In terms of policy interventions, this result suggests that efforts aimed at reducing the rural-urban poverty gap would have a significant contractionary effect on the spending differential.

When we use the urban variance and the urban coefficients (rural coefficients), the results also show that differences in the quality of access of primary schools as proxied by the distance to the nearest primary school have the second largest impact of 17.19% (25.76%) on the spending gap. So 17.19% (25.76%) of the *characteristic effect* would be knocked off as a result of closing the quality of access gap between the two areas. We get a similar picture when the rural variance and urban or rural coefficients are used the counterfactual. Thus, reducing the differences in the quality of access of primary schools between the two areas would go a long way in reducing the spending gap[36]. Interestingly, the results which are robust to choice of variance and coefficients used in the counterfactual,

[36] If the distance to the nearest primary school is thought of as a measure of the direct cost of primary education, then the result means that reducing the differences in cost of primary education between the two areas would go a long way in reducing the spending gap.

show that differences in mothers characteristics in terms of education and employment contribute more to the *characteristic effect* compared to the same for fathers. Hence, targeting mothers education and employment would have a bigger impact as compared to the same for fathers in narrowing or closing the spending gap between the two areas. It is also noteworthy that mother's education has a larger contribution to the gap than mother's employment. Similar to the econometric results (subsection 2.5.1), this finding has intergenerational implications for reducing or closing the rural-urban gap in spending. Educating more girls entails more educated mothers in future, who would then have a larger effect on the rural-urban spending gap.

When we ignore the fact that the spending decisions are done in two stages and use the Tobit model (see Tables A2.6 and A2.7 in the appendix), we get conclusions similar to the DH, albeit with generally higher effects for the six variables, again implying that we overestimate the impact of the variables when the participation decision is not accounted for. Again, these conclusions are robust to choice of variance or coefficients used in the counterfactual. In summary, both results from the DH and the Tobit models show the six variables to be the major drivers of the spending gap. Thus, policy interventions to narrow or close the rural-urban household spending gap should focus on reducing the poverty gap, school quality gaps, men's and women's education and employment gaps, especially the women's education gap.

2.7 Conclusions

Using the Second Malawi Integrated Household Survey (IHS2) data, the chapter has looked at household expenditure on the education of own primary school children. We make a distinction between rural and urban households. With this distinction in mind, we have looked at two issues. Firstly, we have investigated the factors which influence a household's expenditure decision. Specifically, here we have looked at two interrelated questions; what factors influence a household's decision to spend or not (the participation decision), and then what factors influence how much is spent (the expenditure decision). We have found that there are differences in the impact of factors by area of residence. It has been established that the level of household income in rural and urban areas positively and significantly impacts both the participation and expenditure decisions. Computed elasticities have shown that spending on education by rural households is more sensitive to changes in income compared to urban households, suggesting that spending on education in rural areas is a luxury good. We have found that a father's and mother's employment has a bigger impact on spending (at both decision levels) in rural areas compared to urban areas. For both areas, a mother's employment and education has been found to exert a bigger influence on spending compared to a father's. We have shown that urban households compared to their rural counterparts are more sensitive to the quality of access of primary schools as measured by the distance to the nearest primary school. The study has found evidence of gender bias in school spending in rural areas only.

The second issue addressed in the study relates to why there are these differences between rural and urban households, and we have dealt with this issue by conducting a decomposition analysis. We have proposed an extension of the Blinder-Oaxaca decomposition technique to the independent DH. The extension has been done at two levels namely; the aggregated decomposition which shows just how much of the spending gap is due to differences in characteristics (*characteristic effect*) and how much is due to differences in the estimated coefficients (*coefficient effect*), and the disaggregated decomposition of the *characteristic effect* which shows the contribution of each variable to the *characteristic effect*. Results from the aggregated decomposition show that at least 66% of the expenditure differential arises from differences in characteristics and about 34% is due to behavioural differences (estimated coefficients) between rural and urban households. This conclusion is robust to choice of coefficients and variance used in the counterfactual. It is also robust to assuming that the zeros in expenditure are entirely a result of a corner solution. The results from the disaggregated decomposition show that household income, parental education and employment, and quality of access of primary schools are the major factors behind the spending gap. It has been shown that the difference in household income between the two areas is the largest contributing factor, followed by quality of access of primary schools. Further, it has been demonstrated that differences in mothers employment and education have a larger effect relative to the father's on the spending differential.

Our empirical analysis has a a number of caveats which are worth pointing out. Firstly, we have not taken into account the possibility that sending children to school and sending them to work (i.e. child labour) are joint decisions. Secondly, the study makes the implicit assumption that more spending on education entails more schooling either in levels or in quality. However, more spending on education can translate in more schooling or more quality of schooling rather imperfectly. Finally, there may be a possibility of spatial sorting which may hold if most hard working parents or those with the strongest taste for education are more likely to move to urban areas which may result in urban households spending more on education. We unable to address this selection on unobservables. Our conclusions should therefore be taken with these caveats in mind.

In this chapter, we have focussed on the parental role in the human capital formation of own primary school children in Malawi. In many African countries however, parents are not only responsible for the education of their own children, but they are also responsible for the investments in education of non-biological children who are put in their care. With the increasing number of HIV/Aids orphans in many African countries, the role that parents play in educating non-biological children in addition to their own takes on an added significance. This blending of biological and non-biological children in families may give rise to schooling discrimination, in the sense that non-biological children may receive less schooling relative to biological ones. The next chapter, examines this schooling bias against non-biological children that parents may exhibit.

Table 2.1: Annual primary education expenditure

	Rural		**Urban**			
	Full sample	**Spending sample**	**Full sample**	**Spending sample**	**Gap 1**	**Gap 2**
Expenditure	**(1)**	**(2)**	**(3)**	**(4)**	**(3)-(1)**	**(4)-(2)**
Absolute	379.97	510.68	4696.00	6863.38	6352.70***	4316.03***
Share	0.004	0.005	0.014	0.022	0.01***	0.017***
Disaggregated absolute **expenditure of full sample**						
Tuition	35.48		2945.85		2910.27***	
Books	74.81		250.63		175.82***	
Uniform	160.64		343.62		182.98***	
Boarding	13.05		124.06		111.01***	
Building	53.78		82.78		28.89*	
PTA	10.55		233.23		222.68***	
Other	31.46		715.83		684.37***	

Notes: The full sample is made up of all households with school going children, and the spending sample is made of households with nonzero expenditure on education. Absolute is the absolute expenditure while share is absolute expenditure divided by household annual consumption expenditure. We use two-tailed tests to test the significance of the differences (gaps) in expenditure between rural and urban. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01.Expenditure is measured in Malawi Kwacha (MK).

Table 2.2: Sample descriptives of explanatory variables

	Rural		Urban			
	full sample	spending sample	full sample	spending sample	gap 1	gap 2
Variable	(1)	(2)	(3)	(4)	(3)-(1)	(4)-(2)
Household characteristics						
last child's age	7.97	7.65	9.30	8.57	1.3***	0.92**
consumption expenditure	9.64	9.65	10.07	10.19	0.43***	0.53***
government scholars	0.79	0.80	0.69	0.70	-0.10***	-0.10**
children	3.52	3.68	3.59	3.71	0.04	0.07
Parental characteristics						
father works	0.71	0.72	0.82	0.81	0.10**	0.11***
mother works	0.23	0.24	0.30	0.30	0.05	0.07*
father's education	2.02	2.01	5.40	5.75	3.74***	3.38***
mother's education	0.79	0.76	2.90	3.31	2.55***	2.11***
father's age	47.77	47.78	47.88	47.87	0.09	0.11
mother's age	43.62	43.23	43.11	42.13	-1.1	-0.51
School characteristics						
distance primary	2.75	2.95	1.99	1.30	-1.65	-0.76
Age-gender composition of household						
males aged 0-6	0.11	0.11	0.09	0.09	-0.02*	-0.02
males aged 7-15	0.16	0.16	0.13	0.13	-0.03**	-0.03**
males aged 16-19	0.05	0.05	0.05	0.04	-0.003	0.0006
males aged 20-55	0.15	0.15	0.18	0.18	0.03***	0.02**
males above 55	0.03	0.03	0.04	0.04	0.007	0.01
females aged 0-6	0.11	0.11	0.09	0.10	-0.005	-0.03
females aged 7-15	0.15	0.16	0.16	0.15	-0.008	0.002
females aged 16-19	0.03	0.03	0.04	0.04	0.01	0.01
females aged 20-55	0.18	0.18	0.20	0.21	0.03***	0.02**
females above 55	0.03	0.02	0.02	0.01	-0.01	-0.01
Region						
north	0.18	0.16	0.29	0.19	0.02	0.12
centre	0.40	0.46	0.44	0.51	0.05	0.04
south	0.42	0.38	0.27	0.31	-0.07	-0.15***
Sample size	3739	2782	676	548		

Notes: The full sample is made up of all households with school going children, and the spending sample is made of households with nonzero expenditure on education. We use two-tailed tests to test the significance of the differences (gaps) in regressors between rural and urban. For continuous regressors we use mean differences, and for dummies we use proportional differences. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01.

Table 2.3: LR test of differences in expenditure on education

	Log likelihood value (number of parameters)					
Model	Pooled	Rural	Urban	LR statistic	df	p-value
DH	-8306.19(66)	-6167.27(64)	-2075.03(64)	127.78	62	0.00
Tobit	-8398.78(33)	-6211.47(32)	-2107.53(32)	159.56	31	0.00

Table 2.4: LR test of Tobit against the independent DH

	Model				
Group	Independent DH	Tobit	LR statistic	df	p-value
Rural	-6167.27	-6211.47	88.4	32	0.00
Urban	-2075.03	-2107.53	65.0	32	0.00

Table 2.5: Results of the independent DH by area of residence

Variable	Rural participation	Rural level	Urban participation	Urban level
Household characteristics				
last child's age	-0.05488***	-0.01060	0.30450	0.00072
	(0.01614)	(0.00883)	(0.23719)	(0.00365)
last child's age^2	0.00104***	0.00022	-0.00767	-0.00009
	(0.00038)	(0.00021)	(0.00710)	(0.00009)
consumption expenditure	0.23207***	0.05227***	0.56821***	0.03576***
	(0.05461)	(0.00355)	(0.01981)	(0.01283)
government scholars	0.76890***	-0.11241**	-2.86960**	0.02045
	(0.10820)	(0.04691)	(1.34249)	(0.01268)
children	0.03128	0.04425***	1.63650**	-0.00016
	(0.04129)	(0.01411)	(0.64050)	(0.00804)
children^2	-0.00132	-0.00132**	-0.14069**	0.00046
	(0.00398)	(0.00056)	(0.07013)	(0.00084)
Parental characteristics				
father works	0.00650***	0.00756***	0.04989***	0.02324***
	(0.00138)	(0.00155)	(0.00224)	(0.00296)
mother works	0.20134***	0.02023***	0.64032***	0.02352***
	(0.05835)	(0.00219)	(0.07506)	(0.00132)
father's education	0.00677***	0.01142***	0.02940***	0.00121***
	(0.0017)	(0.00259)	(0.00354)	(0.00019)
mother's education	0.00683***	0.00865***	0.03234***	0.00231***
	(0.00101)	(0.00148)	(0.00609)	(0.00026)
father's age	0.03908	0.01789	0.90438**	-0.01001
	(0.03091)	(0.01648)	(0.37120)	(0.00628)
father's age^2	-0.00019	-0.00018	-0.00825**	0.00010
	(0.00027)	(0.00015)	(0.00341)	(0.00006)
mother's age	0.04274	0.05428***	-0.51033**	-0.01090
	(0.03055)	(0.01915)	(0.21431)	(0.01311)
mother's age^2	-0.00045	-0.00048***	0.00328*	0.00015
	(0.00028)	(0.00018)	(0.00181)	(0.00014)
School characteristics				
distance primary	-0.00699***	-0.00908***	-0.56579***	-0.02440***
	(0.00024)	(0.00084)	(0.00536)	(0.00158)
Age-gender composition of household				
males aged 0-6	1.11652**	-0.27137	-8.03960	0.20918*
	(0.55346)	(0.21291)	(5.46235)	(0.11936)
males aged 7-15	1.94601***	0.23238***	6.16139***	0.18465***
	(0.54091)	(0.0095)	(0.09781)	(0.00321)
males aged 16-19	1.05852*	0.30828	-11.41691*	0.26668*
	(0.57515)	(0.22514)	(6.57410)	(0.14466)
males aged 20-55	0.43034	0.14724	-12.28649**	0.17875*
	(0.50640)	(0.18605)	(5.65072)	(0.10662)
females aged 0-6	0.87586	-0.58748**	-7.37600	0.26562**
	(0.54953)	(0.25236)	(5.68817)	(0.12612)
females aged 7-15	1.82512***	0.25020***	7.70956***	0.29012**
	(0.2362)	(0.0093)	(0.40535)	(0.12162)
females aged 16-19	0.33254	0.36888*	-8.92036	0.31863***
	(0.59034)	(0.22344)	(5.63699)	(0.12218)
females aged 20-55	0.63089	0.30406	-3.77753	0.12265
	(0.59596)	(0.23380)	(4.96818)	(0.10029)
females above 55	1.47903**	0.51368*	-4.29883	0.41458**
	(0.71350)	(0.28441)	(6.11054)	(0.18104)
Region				
north	0.17206***	0.13929*	-1.56052*	0.04564
	(0.06414)	(0.07396)	(0.87341)	(0.03108)
centre	0.70344***	0.01791	-0.73286	-0.02001*
	(0.05767)	(0.02882)	(0.61972)	(0.01063)

Table 2.5: continued

Variable	Rural participation	Rural level	Urban participation	Urban level
Controls for endogeneity				
residualcons		-0.19670**		-0.02123
		(0.08155)		(0.01426)
constant	-5.71966***	-1.95478*	-9.54081	-0.12453
	(1.40696)	(1.16150)	(12.48037)	(0.33370)
sigma		0.01358***		0.01182***
		(0.00258)		(0.00160)
Log-likelihood	-6167.27		-2075.03	
P-values of equality of coefficients of males aged 7-15 and females aged 7-15:				
	0.007	0.002	0.52	0.36

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors. Residualcons is the residual from the reduced form of log per capita consumption expenditure.

Table 2.6: Elasticities with respect to selected regressors for the DH

	DH					
	Rural			Urban		
Variable	Prob	Cond	Uncond	Prob	Cond	Uncond
consumption expenditure	1.889***	1.145***	2.33***	0.154***	0.177***	0.331***
	(0.209)	(0.055)	(0.264)	(0.083)	(0.12)	(0.095)
father works	0.164***	0.143***	0.307***	0.124***	0.132***	0.256***
	(0.006)	(0.007)	(0.013)	(0.003)	(0.025)	(0.028)
mother works	0.272***	0.312***	0.584***	0.205***	0.206***	0.411***
	(0.003)	(0.004)	(0.007)	(0.064)	(0.114)	(0.178)
father’s education	0.174***	0.137***	0.311***	0.114***	0.0856***	0.1996***
	(0.007)	(0.032)	(0.039)	(0.014)	(0.0021)	(0.0161)
mother’s education	0.441	0.318***	0.759***	0.166***	0.224***	0.390***
	(0.030)	(0.073)	(0.03)	(0.017)	(0.023)	(0.04)
distance primary	-0.018***	-0.047***	-0.065***	-0.296**	-0.854***	-1.15***
	(0.002)	(0.005)	(0.007)	(0.115)	(0.268)	(0.383)
males aged 7-15	0.120***	0.66***	0.780***	0.314***	0.320***	0.634***
	(0.033)	(0.055)	(0.088)	(0.077)	(0.036)	(0.113)
females aged 7-15	0.014***	0.070***	0.084***	0.317***	0.321***	0.638***
	(0.003)	(0.004)	(0.007)	(0.024)	(0.013)	(0.037)

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors.

Table 2.7: Blinder-Oaxaca decomposition of the independent DH

	Using the urban variance	
Actual expenditure share gap	0.01	0.01
Predicted expenditure share gap	0.0097*** (0.0012)	0.0097*** (0.0012)
Characteristic effect	0.0064*** (0.0011)	0.0066*** (0.0002)
% of raw gap	66%	68.43%
Coefficient effect	0.0032*** (0.00041)	0.0031*** (0.00063)
% of raw gap	34%	31.57%
Counterfactual coefficients	urban	rural
Approximation error	0.0003	0.0003
	Using the rural variance	
Actual expenditure share gap	0.01	0.01
Predicted expenditure share gap	0.0097*** (0.0012)	0.0097*** (0.0012)
Characteristic effect	0.006*** (0.00057)	0.0069*** (0.0015)
% of raw gap	67.6%	71.13%
Coefficient effect	0.0031*** (0.0002)	0.0028*** (0.00082)
% of raw gap	32.4%	28.87%
Counterfactual coefficients	urban	rural
Approximation error	0.0003	0.0003

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. The actual expenditure share gap is for the full sample reproduced from Table 1. Approximation error is the difference between the actual expenditure share gap and the predicted expenditure share gap.

Table 2.8: Detailed decomposition of the characteristic effect of the DH using the urban variance

	Urban coefficients			**Rural Coefficients**		
CE	0.0064***			0.0066***		
Of which:						
consumption expenditure	0.0022***	(0.00072)	[34.38%]	0.0024***	(0.00012)	[36.36%]
father works	0.0003***	(0.000033)	[4.69%]	0.00011***	(0.00007)	[1.67%]
mother works	0.00071***	(0.00008)	[11.09%]	0.00067***	(0.000086)	[10.15%]
father’s education	0.0002***	(0.000065)	[3.13%]	0.00019***	(0.000052)	[2.88%]
mother’s education	0.00084***	(0.000029)	[13.13%]	0.0009***	(0.000061)	[13.64%]
distance primary	0.0011***	(0.000076)	[17.19%]	0.0017***	(0.000063)	[25.76%]
other variables	0.00105		[16.41%]	0.00063		[9.55%]

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. In square brackets are percentage contributions of each variable to the characteristic effect (CE). Other variables comprise the remaining variables. We have not computed the standard error for these remaining variables.

Table 2.9: Detailed decomposition of the characteristic effect of the DH using the rural variance

	Urban coefficients			**Rural Coefficients**		
CE	0.006***			0.0069***		
Of which:						
consumption expenditure	0.0019***	(0.00017)	[31.67%]	0.0023***	(0.0001)	[33.33%]
father works	0.00019***	(0.00003)	[3.17%]	0.00017***	(0.000031)	[2.46%]
mother works	0.00071***	(0.000064)	[11.83%]	0.00069***	(0.00002)	[10.00%]
father’s education	0.00015***	(0.000047)	[2.50%]	0.0001***	(0.00002)	[1.45%]
mother’s education	0.00093***	(0.000042)	[15.50%]	0.001***	(0.0001)	[14.49%]
distance primary	0.0012***	(0.0001)	[20.00%]	0.0016***	(0.00012)	[23.19%]
other variables	0.00092		[15.33%]	0.00104		[15.07%]

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. In square brackets are percentage contributions of each variable to the characteristic effect (CE). Other variables comprise the remaining variables. We have not computed the standard error for these remaining variables.

2.8 Appendix to Chapter 2

A2.1 Current enrolment rates by age and area of residence

Age	Rural	Urban
6-8	76	86
9-11	73	86
12-14	68	79
15-17	63	77
18-20	53	62

Table A2.2: Reduced form regressions of log per capita consumption

Variable	Rural	Urban
last child's age	0.090***	0.121***
	(0.005)	(0.045)
last child's age^2	-0.002***	-0.002
	(0.000)	(0.001)
father's age	-0.143***	-0.007
	(0.009)	(0.081)
father's age^2	0.001***	0.002
	(0.000)	(0.001)
mother's age	-0.084***	-0.071
	(0.010)	(0.047)
mother's age^2	0.01***	0.01
	(0.002)	(0.05)
north	-0.001	-0.120
	(0.022)	(0.158)
centre	0.214***	0.159
	(0.017)	(0.149)
Land	0.023***	0.036***
	(0.003)	(0.001)
Land^2	-0.12***	-0.24***
	(0.002)	(0.014)
constant	14.815***	10.741***
	(0.263)	(2.211)
F-test of joint significance of instruments:		
F-stat	111	9.64
Prob> F-stat	0.00	0.00
F-test of overall significance:		
F-stat	122	19.68
Prob> F-stat	0.00	0.00
R-squared	0.2988	0.4564

Notes: The instruments for per capita consumption expenditure are land, its square. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A2.3: Results of the Tobit by area of residence

Variable	Rural	Urban
Household characteristics		
last child's age	-0.00044***	0.00191
	(0.00016)	(0.00176)
last child's age^2	0.00001**	-0.00008
	(0.00000)	(0.00005)
consumption expenditure	0.00171***	0.01259**
	(0.00018)	(0.00611)
government scholars	0.00337***	0.01023
	(0.00072)	(0.00670)
children	-0.00009	0.00293
	(0.00025)	(0.00351)
children^2	0.00006**	0.00009
	(0.00002)	(0.00040)
Parental characteristics		
father works	0.00134***	0.0164***
	(0.00032)	(0.00368)
mother works	0.02277**	0.01213***
	(0.00034)	(0.00129)
father's education	0.00151***	0.01206***
	(0.00004)	(0.00037)
mother's education	0.00431***	0.01074***
	(0.00007)	(0.00049)
father's age	0.00059*	-0.00024
	(0.00030)	(0.00216)
father's age^2	-0.00000*	0.00000
	(0.00000)	(0.00002)
mother's age	0.00077***	-0.00259*
	(0.00023)	(0.00144)
mother's age^2	-0.00001***	0.00002*
	(0.00000)	(0.00001)
School characteristics		
distance primary	-0.00042*	-0.01039***
	(0.00022)	(0.00295)
Age-gender composition of household		
males aged 0-6	0.00150	-0.02059
	(0.00346)	(0.03275)
males aged 7-15	0.00871	0.01916
	(0.337)	(0.23)
males aged 16-19	0.00763**	-0.02771
	(0.00361)	(0.03614)
males aged 20-55	0.00291	-0.04197
	(0.00316)	(0.02973)
females aged 0-6	-0.00124	-0.00743
	(0.00345)	(0.03255)
females aged 7-15	0.00867	0.07097
	(0.334)	(0.977)
females aged 16-19	0.00640*	-0.00056
	(0.00374)	(0.02995)
females aged 20-55	0.00419	-0.00840
	(0.00378)	(0.03112)
females above 55	0.00874*	0.05879
	(0.00448)	(0.04469)

Table A2.3: continued

Variable	Rural	Urban
Region		
north	0.00097	0.01347
	(0.00125)	(0.01054)
centre	0.00266***	-0.00272
	(0.00050)	(0.00388)
Controls for endogeneity		
residualcons	-0.00246*	-0.00538
	(0.00132)	(0.00683)
constant	-0.04727**	-0.05142
	(0.02112)	(0.11718)
sigma	0.00813***	0.01340***
	(0.00011)	(0.00102)
Log-likelihood	-6211.47	-2107.53
P-values of equality of coefficients of males aged 7-15 and females aged 7-15:		
	0.2315	0.5768

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors. Residualcons is the residual from the reduced form of log per capita consumption expenditure.

Table A2.4: Elasticities with respect to selected regressors for the Tobit

	Tobit					
	Rural			Urban		
Variable	Prob	Cond	Uncond	Prob	Cond	Uncond
consumption expenditure	1.524***	1.399***	2.923***	.450***	0.461***	0.911***
	(0.043)	(0.071)	(0.114)	(0.098)	(0.020)	(0.120)
father works	0.23***	0.17***	0.40***	0.16***	0.12***	0.28***
	(0.007)	(0.005)	(0.001)	(0.015)	(0.001)	(0.016)
mother works	0.713***	0.64***	1.35***	0.449**	0.412**	0.861***
	(0.004)	(0.003)	(0.007)	(0.070)	(0.051)	(0.120)
father’s education	0.431***	0.532***	0.963***	0.252***	0.407***	0.659***
	(0.046)	(0.034)	(0.08)	(0.033)	(0.064)	(0.097)
mother’s education	0.869***	0.761***	1.630***	0.785***	0.658***	1.443
	(0.068)	(0.056)	(0.124)	(0.025)	(0.066)	(0.091)
distance primary	-0.089*	-0.068*	-0.157*	-0.988***	-0.728***	-1.716***
	(0.046)	(0.035)	(0.080)	(0.303)	(0.211)	(0.504)
males aged 7-15	0.104	0.079	0.182	-0.116	-0.086	-0.202
	(0.40)	(0.30)	(0.71)	(0.186)	(0.136)	(0.322)
females aged 7-15	0.102	0.078	0.181	-0.082	-0.061	-0.143
	(0.040)	(0.030)	(0.070)	(0.223)	(0.164)	(0.388)

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A2.5: Blinder-Oaxaca decomposition of the Tobit

	Using the urban variance	
Actual average expenditure share gap	0.01	0.01
Predicted average expenditure share gap	0.0059*** (0.001)	0.0059*** (0.001)
Characteristic effect	0.0044*** (0.0004)	0.0046*** (0.0001)
% of raw gap	74.6%	77.97%
Coefficient effect	0.0015*** (0.00021)	0.0013*** (0.00041)
% of raw gap	25.4%	22.03%
Counterfactual coefficients	urban	rural
Approximation error	0.0041	0.0041
	Using the rural variance	
Actual average expenditure share gap	0.01	0.01
Predicted average expenditure share gap	0.0059*** (0.001)	0.0059*** (0.001)
Characteristic effect	0.0048*** (0.00021)	0.0045*** (0.00037)
% of raw gap	81.56%	76.27%
Coefficient effect	0.0011*** (0.00026)	0.0014*** (0.00022)
% of raw gap	18.64%	23.73%
Counterfactual coefficients	urban	rural
Approximation error	0.0041	0.0041

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. The actual expenditure share gap is for the full sample reproduced from Table 1. Approximation error is the difference between the actual expenditure share gap and the predicted expenditure share gap.

Table A2.6: Detailed decomposition of the characteristic effect of the Tobit using the urban variance

	Urban coefficients			Rural Coefficients		
CE	0.0044***			0.0046***		
Of which:						
Consumption expenditure	0.0012***	(0.0001)	[27.27%]	0.0016***	(0.00036)	[34.78%]
father works	0.0002***	(0.000031)	[4.55%]	0.0004***	(0.00004)	[8.70%]
mother works	0.00032***	(0.00032)	[7.27%]	0.00048***	(0.000043)	[10.43%]
father's education	0.0005***	(0.000052)	[11.36%]	0.0002***	(0.00004)	[4.35%]
mother's education	0.00082***	(0.000047)	[18.64%]	0.0004***	(0.000013)	[8.70%]
distance primary	0.0009***	(0.000073)	[20.45%]	0.0008***	(0.00008)	[17.39%]
other variables	0.00046		[10.45%]	0.00072		[15.65%]

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. In square brackets are percentage contributions of each variable to the characteristic effect (CE). Other variables comprise the remaining variables. We have not computed the standard error for these remaining variables.

Table A2.7: Detailed decomposition of the characteristic effect of the Tobit using the rural variance

	Urban coefficients			Rural Coefficients		
CE	0.0048***			0.0045***		
Of which:						
Consumption expenditure	0.002***	(0.0005)	[41.67%]	0.002***	(0.0003)	[44.44%]
father works	0.00023***	(0.000061)	[4.79%]	0.00017***	(0.000041)	[3.78%]
mother works	0.00038***	(0.000032)	[7.92%]	0.00041***	(0.000037)	[9.11%]
father's education	0.00016***	(0.000037)	[3.33%]	0.00013***	(0.00001)	[2.89%]
mother's education	0.00026***	(0.000042)	[5.42%]	0.0002***	(0.000029)	[4.44%]
distance primary	0.0009***	(.000033)	[18.75%]	0.0005***	(0.00007)	[11.11%]
other variables	0.00087		[18.13%]	0.00109***		[24.22%]

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are bootstrapped (1000 replications) standard errors. In square brackets are percentage contributions of each variable to the characteristic effect (CE). Other variables comprise the remaining variables. We have not computed the standard error for these remaining variables.

Chapter 3

Household economic status, schooling costs, and schooling bias against non-biological children in Malawi

3.1 Introduction

Parents as primary care givers of children play an important role in the formation of human capital which is vital for the economic development of any country. They do this by investing in the health and the education of their children. The role of parents in providing for the education of both their own off spring as well as non-biological children in developing countries especially sub-Saharan African countries is more critical now given the impact of HIV/AIDS. An estimated 1.7 million people were infected with HIV in 2007, bringing to 22.5 million the total number of people living with the virus in sub-Saharan Africa (UNAIDS 2007). This entails more orphans in the future who have to be educated by extended family or non-relatives. A number of studies find evidence of schooling bias against non-biological children within households in sub-Saharan Africa, suggesting that parents discriminate against non-biological children who stay with them. Case *et al.* (2004), using 19 Demographic and Health Survey (DHS) datasets from 10 African countries, find that orphans are less likely to be enrolled than nonorphans with whom they live. Gundersen *et al.* (2004) and Kabubo-Mariara and Mwabu (2007) find a similar result for Zimbabwe and Kenya respectively. Shapiro and Tambashewe (2001) find that children living in households headed by someone other than their father or mother tend to have somewhat lower educational attainment in the Democratic Republic of Congo. They find this to be more evident for the ages 10–14[1].

Thus, while there is a plethora of economic studies which show evidence of discrimination against non-biological children, the literature is scanty on the possible sources of this discrimination. For example, Case *et al.* (2004) show that the probability of school enrollment is inversely proportional to the degree of relatedness of the child to the household head, regardless of whether the child is an orphan or not. There has been an almost exclusive focus on gender bias by studies which attempt to offer sources of intrahousehold bias in schooling (e.g. Behrman *et al.* 1986; Davies and Zhang 1995; Alderman and Gertler 1997; Alderman and King 1998; Echevarria and Merlo 1999; Rose 2000; Pasqua 2005). In addition to the paucity of economic studies on sources of schooling bias against non-biological children, to the best of our knowledge there is no study which addresses the issue of what happens to schooling bias following household income and cost changes. The contributions of the study are threefold. Firstly, the study proposes a theoretical model which offers possible sources of schooling bias against non-biological children. The second contribution of the study is that it theoretically demostrates in the presence of discrimination, how households respond to changes in household income and school costs, and how the household's response to cost changes varies with household income. The final contribution of the study is that it empirically investigates the theoretical predictions. Specifically, the empirical analysis seeks to examine using Malawian data, how households respond to changes in household income and school costs, and how a household's response to cost changes varies with household income. The theoretical and empirical analyses conducted by the study are significant. They not only contribute specifically to the understanding of schooling bias against non-biological children, but they also add to the available literature on intrahousehold schooling bias in general. Knowledge of the factors which cause intrahousehold schooling bias as well as how schooling changes as household income and school cost change, and how a household's response to cost changes varies with household income would no doubt go a long way in the formulation of strategies to fight schooling bias at the household level. More crucially, this knowledge takes on an added significance in the light of the increasing number of orphans due to HIV/AIDS in sub Saharan Africa who mostly end up living with extended family and other non-relatives[2].

To understand why there may be schooling bias against a non-biological child in a household, we construct a two period model of the family in which parents work in the first period and retire in the second period. In the first period, they allocate their income between consumption and investment in the schooling of a biological child and non-biological

[1] While in a number of countries there is evidence of schooling bias, in others households exhibit no discrimination. For example, Zimmerman (2003) finds that South African households treat foster children as they do their own children in terms of human capital investment.

[2] In 2005, the estimated adult (age 15-49) HIV prevalence rate for Malawi was 14.1%. With this prevalence rate, Malawi was ranked number eight in the world (UNAIDS 2006).

child. In the second period, parents consume from the income transfers that the two children make when they are adults. The income transfered depends on the schooling invested in the two children in the first period. Thus, there are both investment and consumption motives to educating the children. The two children are assumed to be of the same sex. The model predicts two broad sources of schooling bias against a non-biological child, one attributable to non-preference based conditions, and the other due to a pure preference bias by parents. In terms of discrimination coming from the non-preference based conditions, the model shows that there will be schooling bias against a non-bioligical child if the cost (direct and opportunity costs) of educating the non-biological child are higher; if the returns to education of the non-biological child are lower; and if the subjective belief about how much will be transferred by the own child when the parents retire is higher compared to the non-biological child's. Further to that, the model predicts that the schooling gap between the biological child and the non-biological child gets wider as the relationship distance between the non-biological child and the parents gets wider. The model also predicts schooling bias arising from preference bias, where parents get more utility from the income of their own child.

The model shows that the impact of a change in costs and income on the amount of schooling investment is bigger for the non-biological child. That is, an increase in the cost (income) leads to a larger reduction (increase) in schooling of the non-biological child relative to the biological child. This suggests that households respond asymmetrically to changes in costs and household income. The model further predicts that the gap between the two children following these changes gets wider the more distantly related the non-biological child is. The model also shows that the change in schooling due to a change in costs falls with income, and falls faster for the non-biological child who is distantly related to the parents. This suggests that an increase in cost of schooling leads to a bigger reduction in schooling for poor households, and that the difference in the impact of cost changes between the biological and the non-biological child declines as household economic status improves i.e. there is convergence.

Using blended households in Malawi, that is households with both biological and non-biological children of school going age, and measuring schooling either as current enrolment or as grade attainment, the empirical results confirm the theoretical predictions. Specifically, we find that when both measures of schooling are used the price and income elasticities of schooling are larger for non-biological children. Further, we find that non-biological children who are non-relatives have higher price and income elasticities. The empirical analysis also indicates that households in the lowest income quintile (the poorest) have the largest price elasticities, and households in the highest income quintile (the wealthiest) have the smallest price elasticities. The study also finds that the price elasticities for biological and non-biological children converge as we move from the lowest income quintile to the highest income quintile, and that the convergence is faster for non-biological children who are non-relatives.

The rest of the chapter is structured as follows. Section 3.2 sets up a model of intrahousehold schooling discrimination and discusses its implications as well as conducts comparative static exercises. In Section 3.3 we discuss the theoretical predictions to be checked, the econometric model used, variables used, estimation issues, and data and descriptives. Econometric results are presented in Section 3.4. We conclude in Section 3.5.

3.2 A model of intrahousehold schooling bias

We adapt a model structure used by Alderman and Gertler (1997), and Alderman and King (1998), to study gender schooling bias in households. Consider a society in which parents live in two periods, indexed by $t = 1, 2$ respectively. They work in the first period, and retire in the second period. In the first period, they give birth to a child (b), and in the same period they have a non-biological (nb) child moving into the family[3]. Throughout, we use subscripts $j = b, nb$ to refer to the two children. The children (biological and non-biological child) are of the same sex[4], and approximately of the same age[5]. The parents' consumption in the first period is their income less the investment on schooling of the two children[6]. In the first period parents can also use income from child labour, that is instead of sending children to school they may decide to send them to work. For simplicity we don't model this joint decision between sending children to school (future benefit) and sending them to work (current benefit). In the second period (during retirement), their consumption depends on the income transfered by the two children, which in turn depends on the schooling investment that the parents made in the first period. Thus, parents' decision to educate children is done both for its own sake as a consumption good, and as an investment good. This entails that there are both investment and consumption reasons for investing in the education of the two children. There is a trade off between current consumption and second period consumption, in that less consumption in period 1, means more schooling for the children, and hence more consumption during retirement. We assume that there are no savings, and no old age pension. We also assume that only parents are responsible for the schooling of the children in the first period, that is there are no private or public scholarships. We assume that there is complete and perfect information meaning that there is no uncertainty[7].

[3] For simplicity, we assume that the movement of a non-bilogical child into the family is exogenous.

[4] This allows us to focus on schooling differences in the household which are not due to gender bias.

[5] This allows us control for the children's future level of earnings.

[6] For simplicity, we don't allow for overlapping generations in which the parents' also transfer part of their income to their parents.

[7] One could also allow for parental uncertainty in the transfers that the children would make when they are adults.

The life time utility function of parents is given as follows;

$$U = G(C_1) + \delta V(W_{2b}, W_{2nb}, C_2) \tag{3.1}$$

Where; C_1 is their first period consumption, C_2 is their second period consumption, W_{2b} is the income of their child in period 2 (retirement period), W_{2nb} the income of the non-biological child in period 2 (retirement period), and δ is the discount rate or subjective rate of time preference. This utility function says that parents get utility from consumption in the two periods, and also they get utility from the income of the two children in the second period. We assume that the utility function is twice continuously differentiable and has the following conventional properties;

$$\begin{aligned} G' &> 0, V' > 0 \\ G'' &< 0, V'' < 0 \end{aligned} \tag{3.2}$$

Thus, the utility functions are concave meaning that utility is increasing but diminishing.

Since we have assumed that there are no pensions and savings, parents' second period consumption is given as;

$$C_2 = \theta_k W_{2b} + \theta_k W_{2nb} \tag{3.3}$$

That is, the resources available for consumption by parents in retirement come from the transfers that the two children make when they are adults. The parents have a subjective belief θ_k, about how much of each child's income will be transferred to them when they retire. Where $\theta_k \in [0, 1]$, and k is a measure of the degree of relatedness of the child. For a biological child $k = 1$, and for a non-biological child $k > 1$. A higher k denotes a more distantly related non-biological child. For ease of exposition, we assume that k takes positive integer values. There are a number of reasons why transfers may be made by children when they are adults to their retired parents. As argued by Cigno (1993), there may exist in a society a social norm according to which adults give a fraction of their income to their old parents. The workings of the social norm are aptly explained by López-Calva and Miyamoto (2004, p 491) when they say;

> *"The adult has to decide on whether to transfer money to her retired parents or not. An informal intergenerational contract exists, which can only be enforced through "social punishment." The social perception of the adult's decision shall determine the optimal reaction of her own child and thus whether she herself is going to get a transfer when retired."*

Even in the absence of the said social norm, the non-availability of pension schemes in poor countries entail that old people rely on the resources received from their adult children (Pasqua 2005)[8].

The two children's income when they are adults is;

$$W_{2b} = \pi_b S_{1b} + I_b \tag{3.4}$$

and

$$W_{2nb} = \pi_{nb} S_{1nb} + I_{nb} \tag{3.5}$$

where $\pi_b S_{1b}$ ($\pi_{nb} S_{1nb}$) is the biological (non-biological) child's labour income, and I_b (I_{nb}) is the own (non-biological) child's non-labour income. The labour income for the biological child is a linear function of the level of schooling (S_{1b}) invested by the parents in period 1. Similarly, the labour income for the non-biological child is a linear function of the level of schooling (S_{1nb}) invested by the parents in period 1. π_b and π_{nb} are the returns to education for the biological and non-biological child respectively. Though the innate ability of each child might also affect his/her earnings, for simplicity it not included[9].

In period 1, the parents face the following budget constraint;

$$C_1 + P_{1b} S_{1b} + P_{1nb} S_{1nb} = Y \tag{3.6}$$

where P_{1b} and P_{1nb} are indirect and direct costs of schooling of the biological and non-biological child respectively, and Y is parental labour and non-labour income. We normalize the price of consumption to one. The budget constraint, equation 3.6 says that in period 1, parents allocate their income on current consumption and the schooling investment of their own child and the non-biological child.

3.2.1 Equilibrium

Parents choose the level of schooling of the two children, S_{1b} and S_{1nb} to maximize their life time utility as given by equation 3.1, subject to transfers that the two children will make in retirement represented by equation 3.3, and subject to the budget constraint 3.6. Substituting equations 3.3 to 3.6 into equation 3.1, the utility maximization problem of the parents is formally expressed as;

[8] We assume here that the subjective belief is exogenous.

[9] Apart from investing in schooling, the parents may also indirectly affect the earnings potential of the two children through other child specific inputs such health and nutrition. We do not model this indirect channel.

$$\begin{aligned}\underset{S_{1b},S_{1nb}}{Max}\,U &= G(Y - P_{1b}S_{1b} - P_{1nb}S_{1nb}) \\ &+\delta V(\pi_b S_{1b} + I_b, \pi_{nb}S_{1nb} + I_{nb}, \theta_k\left(\pi_b S_{1b} + I_b\right) + \theta_k\left(\pi_{nb}S_{1nb} + I_{nb}\right))\end{aligned} \tag{3.7}$$

The first order conditions are (compare with Alderman and Gertler 1997, and Alderman and King 1998);

$$\frac{\partial G}{\partial C_1}P_{1b} = \frac{\partial V}{\partial W_{2b}}\pi^b + \frac{\partial V}{\partial C_2}\pi_b\theta_k \tag{3.8}$$

and

$$\frac{\partial G}{\partial C_1}P_{1nb} = \frac{\partial V}{\partial W_{2nb}}\pi_{nb} + \frac{\partial V}{\partial C_2}\pi_{nb}\theta_k \tag{3.9}$$

We assume that $\delta = 1$ for simplicity. The two first order conditions suggest that parents will invest in the education of each child until the marginal cost of sacrificing consumption in period 1 (left hand side) is equal to the marginal benefit in period 2. The marginal benefit in period 2 is equal to the utility the parents derive from a marginal increase in each child's human capital plus the marginal utility of second-period consumption multiplied by the subjective belief about future transfers per unit of school investment.

3.2.2 Implications

We now turn to the implications of the model for intrahousehold schooling bias against a non-biological child. The model predicts two broad sources of intrahousehold schooling bias against a non-biological child. Bias could be due to non-preference based factors favouring the biological child, and secondly, it could arise from preference bias against the non-biological child. These two predictions are formally expressed in the next two propositions.

Proposition 1 *If parents exhibit no preference bias against the non-biological child, the non-biological child will receive less schooling i.e.* $S_{1b} > S_{1nb}$ *when at least one of the following holds:*

i) *Direct and indirect costs of educating the non-biological child are higher than those of the biological child i.e.* $P_{1b} < P_{1nb}$.

ii) *Returns to education of the non-biological child are lower than those of the biological child i.e.* $\pi_b > \pi_{nb}$.

iii) *The subjective belief about how much the non-biological child will transfer in old age is lower than that of the biological child i.e.* $\theta_1 > \theta_k$ for $k > 1$. *Further to this, the schooling bias worsens, if the subjective belief decreases as the relationship with the non-biological child becomes distant i.e.* $\theta_1 > \theta_2 > \theta_3, \ldots$

Proof. Assuming that the marginal benefit (right hand side of 3.8 and 3.9) is the same, $S_{1b} > S_{1nb}$ holds only if $P_{1b} < P_{1nb}$. No preference bias means $\frac{\partial V}{\partial W_{2b}} = \frac{\partial V}{\partial W_{2nb}}$, and $\frac{\partial^2 V}{\partial W_{2b}^2} = \frac{\partial^2 V}{\partial W_{2nb}^2}$, when $S_{1b} = S_{1nb}$. Now assuming that the marginal cost (left hand side of 3.8 and 3.9) is equal, means that $\frac{\partial V}{\partial W_{2b}}\pi_b + \frac{\partial V}{\partial C_2}\pi_b\theta_k = \frac{\partial V}{\partial W_{2b}}\pi_{nb} + \frac{\partial V}{\partial C_2}\pi_{nb}\theta_k$. With $\pi_b > \pi_{nb}$, this equality holds only if $S_{1b} > S_{1nb}$, due to the concavity of schooling in the parents utility function. Similarly, with $\theta_1 > \theta_k$ for $k > 1$, and concavity of schooling, this equality prevails only if $S_{1b} > S_{1nb}$. Thus, $S_{1b} > S_{1nb}$ (bias) occurs if i) $P_{1b} < P_{1nb}$ or ii) $\pi_b > \pi_{nb}$ or iii) $\theta_1 > \theta_k$. Further to this, when we have $\theta_1 > \theta_2 > \theta_3, \ldots$ the schooling gap must be widening. ■

The cost especially the indirect opportunity cost of educating a non-biological child may be higher than that for the biological child owing to the possibility that a non-biological child is more likely to be sent out to work (child labour) to supplement household income. The forgone income from the non-biological child makes it more costly to send him/her to school. The returns to schooling for the non-biological child may be lower in Africa where most of the non-biological children are orphans due to parental death caused by HIV/ AIDS. For example, Case *et al.* (2004, p 484) argue that;

> *"orphans may also be more likely than nonorphans to have HIV/AIDS because of maternal-child transmission, which could depress schooling. In addition, the returns to schooling could be reduced by the experiences surrounding the death of a parent, including time lost from school during the parent's illness and death and emotional scarring that may compromise the child's ability to learn."*

The stigma that often follows those children whose parents died of HIV/AIDS may leave emotional scars which could quite possibly affect their future returns to education. As Gachuhi (1999) contends, the psychological effects of having to cope with pervasive illness and death, and the debilitating impact of the stigmatization associated with HIV/AIDS can be a detriment to learning. Since parents cannot observe *a priori* a child's future transfers to them in old age, they form a subjective belief that due to the biological ties, their own offspring will transfer a higher fraction of his/her income, than the non-biological child with whom they have weaker ties, and the ties with the non-biological child get weaker and weaker the more distant is the relationship. The idea that biological relatedness matters is aptly expressed by evolutionary biologist Hamilton (1964a,b) in what is called Hamilton' rule which is expressed as follows;

> "*The social behavior of a species evolves in such a way that in each distinct behavior evoking situation the individual will seem to value his neighbors' fitness against his own according to the coefficients of relationship appropriate to that situation.*"(1964b, p 19.)

The rule suggests that the degree of altruism is an increasing function of biological relatedness, that is a child would care more about his parents than a distant relative[10]. At the empirical level, Case *et al.* (2004) show that the probability of school enrollment is inversely proportional to the degree of relatedness of the child to the household head, regardless of whether the child is an orphan or not[11]. Thus, intrahousehold discrimination against a non-biological child can arise due to the aforementioned non-preference based conditions. Schooling bias in households can also arise if parents have preference bias against the non-biological child, this is formally expressed in the following proposition.

Proposition 2 *If parents exhibit preference bias against the non-biological child, the non-biological child will receive less schooling i.e.* $S_{1b} > S_{1nb}$.

Proof. Preference bias means $\frac{\partial V}{\partial W_{2b}} > \frac{\partial V}{\partial W_{2nb}}$, and $\frac{\partial^2 V}{\partial W_{2b}^2} > \frac{\partial^2 V}{\partial W_{2nb}^2}$, when $S_{1b} = S_{1nb}$. Now assuming that the marginal cost (left hand side of 3.8 and 3.9) is equal, and $\pi_b = \pi_{nb}$ means that $\frac{\partial V}{\partial W_{2b}}\pi_b + \frac{\partial V}{\partial C_2}\pi_b\theta_k = \frac{\partial V}{\partial W_{2nb}}\pi_b + \frac{\partial V}{\partial C_2}\pi_b\theta_k$. Therefore, with preference bias this equality holds only if $S_{1b} > S_{1nb}$, since utility is concave in school investment. ■

Thus if parents get more satisfaction from the income of their own child relative to the biological child, the non-biological child receives less schooling. In summary, the model predicts four possible sources of intrahousehold schooling bias against a non-biological child. Firstly, a non-biological child will receive less schooling if the cost (direct and indirect) of educating him/her is higher than that of the own child. Secondly, there will be less schooling investment in a non-biological child relative to a biological one if the returns to education for a non-biological child are lower. Thirdly, there will be schooling bias against a non-biological child if the belief of how much the own child will transfer in old age is higher than that of the non-biological child. Further, schooling bias against a non-biological child is worse, the more distantly related he/she is to the parents. Finally, the non-biological child will have less schooling if parents exhibit preference bias against him/her, in the sense that they get more utility from the income of the own child if the income is the same.

[10]It should be pointed that the standard Hamilton coefficient of relatedness works the other way i.e. low values of the coefficient imply low values of biological relatedness. Our reformulation where higher k denotes a more distantly related non-biological child does not affect the spirit of the Hamilton's rule.

[11]Biological relatedness transcends schooling, for example, Bishai *et al.* (2003), find that reduced biological relatedness is associated with reduced child survival in Uganda. Case *et al.* (2000), find that households in which a child is raised by an adoptive, step or foster mother, less is spent on food in the US and South Africa.

3.2.3 Comparative statics

As discussed earlier, the study seeks to investigate differences in the human capital formation (schooling) of the biological child and non-biological child following changes in the cost of schooling as well as changes in household income. In addition, the study examines the differences in the relationship between household income and the change in school investment due to cost changes for the two children. We answer these questions by conducting comparative static exercises. The results of the comparative exercise are summarized in the next two propositions.

Proposition 3 *Assuming that $P_{1b} = P_{1nb} = P_1$, and the bias sources discussed earlier prevail, the following holds;*

i) *an increase in costs leads to a bigger reduction in schooling for the non-biological child relative to the biological one.*

ii) *an increase in household income leads to a bigger increase in schooling for the non-biological child relative to the biological one.*

iii) *and the gap between the two children following these changes gets wider the more distantly related the non-biological child is.*

Proof. Using the implicit function theorem to differentiate the first order conditions as given in equations 3.8 and 3.9, with respect to cost P_1

we get $\frac{\partial S_{1b}}{\partial P_1} = \frac{\frac{\partial G}{\partial C_1}}{\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_b\theta_k)^2} < 0,$

and $\frac{\partial S_{1nb}}{\partial P_1} = \frac{\frac{\partial G}{\partial C_1}}{\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_{nb}\theta_k)^2} < 0$. Therefore, with preference bias or $\pi_b > \pi_{nb}$ or $\theta_1 > \theta_k$ for $k > 1$, the following is true $|\frac{\partial S_{1b}}{\partial P}| < |\frac{\partial S_{1nb}}{\partial P}|$. Similarly, differentiating 3.8 and 3.9, with respect to household income Y, we get $\frac{\partial S_{1b}}{\partial Y} = \frac{\frac{\partial^2 G}{\partial C_1^2}P_1^2}{\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_b\theta_k)^2} > 0$

and $\frac{\partial S_{1nb}}{\partial Y} = \frac{\frac{\partial^2 G}{\partial C_1^2}P_1^2}{\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_{nb}\theta_k)^2} > 0$.Therefore, with preference bias or $\pi_b > \pi_{nb}$ or $\theta_1 > \theta_k$ for $k > 1$, the following holds $\frac{\partial S_{1b}}{\partial Y} < \frac{\partial S_{1nb}}{\partial Y}$. It must be the case that the gap following these changes widens when we have $\theta_1 > \theta_2 > \theta_3, ...$ ■

Thus, when there is intrahousehold schooling bias against a non-biological child originating from the sources discussed earlier, the schooling of a non-biological child is more sensitive to changes in costs and income. This suggests an asymmetry in the way a family would respond to cost and income shocks to the household. That is, if a family experiences a shock to their income, the schooling of the non-biological child in the house will suffer more relative to the biological children. The same implication holds with respect to cost shocks. Further, this non-neutrality in household response to income and cost changes for the two children gets more asymmetric the more distantly related the non-biological child is. This suggests that a non-biological child who is not related to the parents would have his/her schooling suffer more following cost increases and a decline in a household's economic status. In terms of policy interventions, the theoretical predictions imply that efforts aimed at fighting poverty would go a long way in improving the schooling of non-biological children.

Proposition 4 *Assuming that* $P_{1b} = P_{1nb} = P_1, \frac{\partial^3 G}{\partial C_1^3} > 0$*,*

$$\frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2 > \frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}\left(\pi_b\theta_k\right)^2\right)\text{,and}$$

$$\frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2 > \frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_{nb}^2 + \frac{\partial^2 V}{\partial C_2^2}\left(\pi_{nb}\theta_k\right)^2\right),$$

and the bias sources discussed earlier prevail, the change in schooling due to a change in costs P_1;

i) *falls with household income.*

ii) *and the fall is faster for the non-biological child relative to the biological child i.e.there is convergence in the sense that the difference in the impact of cost changes between the biological and the non-biological child declines as household income increases.*

iii) *the more distantly related the non-biological child is, the faster the convergence.*

Proof. Using the result from the preceding proof, it can be shown that

$$\frac{\partial\left(\frac{\partial S_{1b}}{\partial P_1}\right)}{\partial Y} = \frac{\frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_b\theta_k)^2\right) - \frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2}{\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_b\theta_k)^2\right)^2} < 0$$

if $\frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2 > \frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2b}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}\left(\pi_b\theta_k\right)^2\right)$, and

$$\frac{\partial\left(\frac{\partial S_{1nb}}{\partial P_1}\right)}{\partial Y} = \frac{\frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_b^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_{nb}\theta_k)^2\right) - \frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2}{\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_{nb}^2 + \frac{\partial^2 V}{\partial C_2^2}(\pi_{nb}\theta_k)^2\right)^2} < 0 \text{ if}$$

$$\frac{\partial G}{\partial C_1}\frac{\partial^3 G}{\partial C_1^3}P_1^2 > \frac{\partial^2 G}{\partial C_1^2}\left(\frac{\partial^2 G}{\partial C_1^2}P_1^2 + \frac{\partial^2 V}{\partial W_{2nb}^2}\pi_{nb}^2 + \frac{\partial^2 V}{\partial C_2^2}\left(\pi_{nb}\theta_k\right)^2\right).$$

It therefore must be the case that $\frac{\partial\left(\frac{\partial S_{1b}}{\partial P_1}\right)}{\partial Y} < \frac{\partial\left(\frac{\partial S_{1nb}}{\partial P_1}\right)}{\partial Y}$, if there is preference bias or $\pi_b > \pi_{nb}$ or $\theta_1 > \theta_k$ for $k > 1$. The fall gets bigger with decreasing biological relatedness i.e. when we have $\theta_1 > \theta_2 > \theta_3, ...$ ■

This result implies that when there is intrahousehold schooling bias against a non-biological child emanating from the sources discussed earlier, the impact of changes in costs on schooling is bigger for low income households compared with high income households. There is therefore an asymmetry between poor households and non poor households regarding how they respond to cost changes. That is, an increase in cost of schooling leads to a bigger reduction in schooling for poor households. The model further suggests that the difference in the impact of cost changes between the biological and the non-biological child declines as household income increases. That is, as households become richer, the impact of cost changes on the schooling of the two children converges. Besides, the model suggests that this convergence as household income increases is faster the more distantly related the non-biological child is. This has policy significance, in that improvements in the economic status of households would lead to a reduction in bias against non-biological children. The assumption that $\frac{\partial^3 G}{\partial C_1^3} > 0$ deserves some comment. A positive third derivative of consumption implies that $\frac{\partial G}{\partial C_1}$ is a convex function of consumption. The value $\frac{\frac{\partial^3 G}{\partial C_1^3}}{\frac{\partial^2 G}{\partial C_1^2}}$ is similar to the coefficient of relative prudence by Kimball (1990). In Kimball's theory of precautionary saving under uncertainty, a higher coefficient of relative prudence implies that economic agents become more prudent following an increase in their income by reducing consumption, and hence engange in precautionary saving. In our model, we argue by analogy that the assumption of a positive third derivative of current consumption entails that parents act "prudently" by reducing current consumption and investing more in the education of children following an increase in income. It has to be pointed out though that the analogy is hardly perfect.

3.3 Empirical analysis

3.3.1 Predictions checked

The purpose of the empirical analysis is to check the predictions of the preceding theoretical model using empirical data. We specifically check the following predictions:

1. *An improvement in household economic status has a larger increase in the schooling of non-biological children relative to own children, and the increase is larger the more distantly related the non-biological children are.*

2. *An increase in the cost of schooling leads to a bigger decrease in the schooling of non-biological children relative to own children, and the increase is larger the more distantly related the non-biological children are.*

3. *The decrease in schooling due to cost increases is negatively related to household economic status.*

4. *As household economic status improves, the fall in schooling as a result of cost increases for non-biological children converges to that of biological children, and converges faster the more distantly related the non-biological children are.*

3.3.2 Model specification

In order to check these predictions, we use two measures of schooling outcome namely; the highest grade attained and current school enrolment. These two measures represent different conceptualizations of schooling. They capture different dimensions of schooling, and therefore can be viewed as complementary. Each measure of schooling outcome is modeled using a different econometric model.

Grade attainment

The advantage of grade attainement over current school enrolment is that it represents the cumulative investment in a child's education, that is current school enrolment ignores the fact that current schooling depends on previous levels of schooling. We use the censored ordered probit to model grade attainment. The censored ordered probit model was originally developed by King and Lillard (1987) to study grade attainment. It has subsequently been used to study grade attainment by among others; Glewwe and Jacoby (1994), Alderman *et al.* (1996), Behrman *et al.* (1997), Holmes (1999), and Maitra (2003). The censored ordered probit addresses three problems which are inherent in grade attainment. Firstly, the model allows for the fact that grade attainment represents ordered discrete choices i.e. whether to move to the next grade or withdraw. Secondly, it accommodates the possibility that grade attainment often exhibits a large mass point at zero years of schooling and similar probability spikes at primary and secondary completion levels where graduating to the next grade is impeded by fees or entrance examinations (Holmes 1999). Finally, it addresses the problem that grade attainment is right censored. Right censoring occurs because for those children who are still in school, their final grade attained is unknown and to treat their grade as being equal to those who have stopped at that grade would lead to biased estimates of the effects of regressors on true grade attainment (Glick and Sahn 2000).

Following Holmes (1999) and Maitra (2003), the censored ordered probit is formally expressed as follows;

$$S_i^* = X_i'\beta + \varepsilon_i \tag{3.10}$$

where S_i^* is a continuous, and unobserved latent variable representing desired level of schooling for child i, X_i' is a vector of variables which explain schooling, β is a vector of parameters including a constant, and ε_i is an error term. The observed level of completed schooling outcomes S_i, has J discrete possible outcomes, $S_i = 0, 1, 2..J$ which follow a natural ordering i.e. grade 2 is higher than grade 1, etc. For those children who have completed schooling (uncensored observations), the observed level of completed schooling S_i is given as[12];

$$\begin{aligned} S &= 0 \text{ if } S^* \leq 0 \\ S &= 1 \text{ if } 0 < S^* \leq \mu_1 \\ S &= 2 \text{ if } \mu_1 < S^* \leq \mu_2 \\ &\vdots \\ &= J \text{ if } \mu_{J-1} \leq S^* \end{aligned} \tag{3.11}$$

μs are threshold parameters (cut off points)[13] which denote a transition from one grade to the next where the next grade is higher than the previous, and J denotes the highest attainable schooling grade. For those with no schooling, we know only that the latent variable falls below the lowest threshold, i.e. $S^* < 0$, and for those with the maximum level of schooling, we know that $\mu_{J-1} \leq S^*$. Under the assumption that the error term ε follows a standard normal distribution[14], the conditional probability of observing each schooling outcome is;

$$\begin{aligned} \Pr(S=0) &= \Phi(-X'\beta) \\ \Pr(S=1) &= \Phi(\mu_1 - X'\beta) - \Phi(-X'\beta) \\ \Pr(S=2) &= \Phi(\mu_2 - X'\beta) - \Phi(\mu_1 - X'\beta) \\ &\vdots \\ \Pr(S=J) &= 1 - \Phi\left(\mu_{J-1} - X'\beta\right) \end{aligned} \tag{3.12}$$

For all probabilities to be positive, the following condition is imposed;

$$0 < \mu_1 < \mu_2 < \cdots \mu_{J-1} \tag{3.13}$$

The likelihood function for the uncensored observations (L_U) is expressed as;

[12] We have suppressed subcript i to avoid notational clutter.

[13] We can alternatively assume that the parameter vector β does not include a constant, and then include it as the first cut off point i.e. the zero is replaced with a constant.

[14] Assuming that the error term follows a logistic distribution, would give us a censored ordered logit.

$$L_U = \begin{cases} \Phi\left(-X'\beta\right) & \text{for } S = 0 \\ \Phi\left(\mu_S - X'\beta\right) - \Phi\left(\mu_{S-1} - X'\beta\right) & \text{for } S = 1, ...J-1 \\ 1 - \Phi\left(\mu_{S-1} - X'\beta\right) & \text{for } S = J \end{cases} \tag{3.14}$$

If there is no right censoring, the likelihood function L_U, is equivalent to that of the standard ordered probit model. For children who are still enrolled in school (censored observations), the highest grade attained is unknown. However, we know that a currently enrolled student will ultimately attain at least his or her current grade. Thus, the current grade level represents a lower bound, which means that the desired level of schooling S^* is bounded from below i.e. $S^* \geq \mu_{S-1}$. The probability of achieving at least the current grade is therefore;

$$1 - \Phi\left(\mu_{S-1} - X'\beta\right) \quad \text{for } S = 0, 1, 2..J \tag{3.15}$$

Thus the likelihood function for censored observations (L_C) is;

$$L_C = 1 - \Phi\left(\mu_{S-1} - X'\beta\right) \tag{3.16}$$

The likelihood function for the sample (L) is therefore given as[15];

$$L = \prod L_U \prod L_C \tag{3.17}$$

If there is no right censoring and $J = 1$, the likelihood function L reduces to that of a probit model. So the probit model is a special case.

Since the theoretical predictions we are checking rely on the magnitude of coefficients as well as their direction, we use elasticities. The elasticity of probability for each grade in the censored ordered probit is expressed as;

[15] The estimation of the censored ordered probit was done by using Stata code written by Haaga O (2003), and available online at
http://www.stata.com/statalist/archive/2003-08/msg00426.html

$$
\begin{aligned}
\frac{\partial E(S|X)}{\partial X}\frac{X}{E(S|X)} &= \left(\frac{\partial \Pr(S=0)}{\partial X}\frac{X}{\Pr(S=0)}\right)*0 \\
&+\left(\frac{\partial \Pr(S=1)}{\partial X}\frac{X}{\Pr(S=1)}\right)*1 \\
&+\left(\frac{\partial \Pr(S=2)}{\partial X}\frac{X}{\Pr(S=2)}\right)*2 \\
&\vdots \\
&+\left(\frac{\partial \Pr(S=J)}{\partial X}\frac{X}{\Pr(S=J)}\right)*J
\end{aligned}
\tag{3.18}
$$

In this study, the elasticities of probability are computed at the sample means of the regressors.

Current school enrolment

The advantage of current enrolment status over grade attainment as a measure of schooling is that it allows us to accommodate time varying effects. Using current school enrolment enables us to better capture the contemporaneous effect of household structure and income which change overtime (Glick and Sahn 2000). In terms of household structure, the arrival of new children either through new births or fostering, may alter the allocation of time to schooling and household work, and this may affect schooling outcomes of the children. Besides, since we are looking at the schooling of non-biological children (in relation to biological children), their grade attainment may not reflect the schooling investment of their current care givers. In this study, we model the enrolment decision using a probit model.

3.3.3 Variables used

For highest grade attained, in this study we have four discrete and ordered categories defined for each child as;

$$
S = \begin{cases} 0 \text{ if no education attained} \\ 1 \text{ if highest education attained is junior primary} \\ 2 \text{ if highest education attained is senior primary} \\ 3 \text{ if highest education attained is secondary} \end{cases} \tag{3.19}
$$

Junior primary corresponds to standards 1 to 5, and senior primary corresponds to standards 6 to 8. For secondary, we have merged junior secondary (forms 1 and 2) with senior secondary (forms 3 and 4), because in our dataset there are few children we have gone as far secondary school[16]. Current school enrolment is a dummy variable defined for each child as;

$$Enrol = \begin{cases} 1 \text{ if currently enrolled} \\ 0 \text{ otherwise} \end{cases} \tag{3.20}$$

What this effectively means compared to grade attainment, is that every child who is in school is given a one, and a zero is given to those who should be in school but are not. Both grade attainment and enrolment status are defined by age, we discuss the details later.

The key explanatory variables for this study are annual household income and school cost. We use the log of per capita annual consumption expenditure as our measure of household economic status other than actual household income[17]. To measure the cost of schooling, we use local child wages prevailing in the area, that is we use the average community level child wage (measured in Malawian Kwacha)[18]. Local child wages represent the opportunity cost (indirect cost) of sending children to school if the alternative is to work in the farm, family business or other market work (Tansel 1997)[19]. Other explanatory variables included in the two models are; age and sex of the child, mother's and father's employment status, mother's and father's education, mother's and father's age, household size. We include a rural dummy to control for possible rural-urban differences. We also control for regional fixed effects by including regional dummies.

3.3.4 Estimation issues

The log of per capita expenditure is potentially endogenous, and this may lead to biased and inconsistent results. One possible channel of endogeneity is that the log of per capita expenditure and spending on education can be jointly determined through labour supply decisions in the sense that a decision to send children to school may be jointly determined with a decision to send the children to work to supplement household income. Another route for endogeneity would be that parents with a good taste for the education of their children may work harder so they are able to pay for their schooling (Kingdon 2005).

We address this problem in both the probit and censored ordered probit by using the

[16]The structure of the Malawian education system is discussed in the previous chapter, in section 2.2.

[17]For a justification of why consumption expenditure and not income is more appropriate, see footnote 14 in Chapter 2.

[18]This is measured as follows; for those children who work, the survey data has information on the wages that they get, we use these wages (after annualising them to ensure comparability with the other variables) to then compute a community level average wage.

[19]The two types of children face the same local child wages. Recall, that our comparative static exercise was based on the assumption that the biological and non-biological child face the same school costs i.e $P_{1b} = P_{1nb} = P_1$.

Rivers and Vuong (1989) procedure[20]. We use household assets namely hectares of land, and its square as instrumental variables for log of per capita expenditure[21].

3.3.5 Data and descriptives

The data used in the study is obtained from the Second Malawi Integrated Household Survey (IHS2)[22]. The survey collects information on the education status of all children such as the highest grade attained and current enrollment status. It records the relationship of each child to the household head. This allows us to distinguish in each household, biological children from non-biological ones. It further enables us to separate the non-biological children into whether they are related or unrelated to their care givers. Since we are focussing on intrahousehold schooling bias against non-biological children, we restrict our sample to blended households, that is households which have both biological and non-biological children. For grade attainment, we restrict the children's ages to between 10 and 19. This restriction is necessitated by the fact that this enables us to some extent to separate out non-enrolment from late enrolment which is common in Malawi. Non-enrolment could either be due to late entry into school or parents not deciding to send a child to school at all. The lower age limit 10 therefore ensures that a child who has not enrolled in school by the age of 10 will never do so. The upper age limit of 19 is driven by the fact that this helps us to some extent to mitigate the problem of sample selection where older children are absent from home[23]. For current school enrolment, we restrict children's ages to between 6 and 19. We have a total of 10241 children of whom 8347 (representing 82%) are biological, and 1894 (representing 18%) are non-biological. Of the non-biological children, 1534 (representing 81%) are relatives, and the remainder 360 (representing 19%) are non-relatives.

We now look at the descriptive statistics. In Table 3.1, we report grade attainment rates of biological and non-biological children, and the results show that biological children have consistently higher attainments at all schooling levels. The results also indicate that non-biological children that are not related to the people who keep them fare badly in terms of attainment as compared to those who are related. We also notice that attainment declines with age, and the decline is more pronounced for non-biological children who are not relatives. For instance, the results indicate that 7.8% of biological children attained secondary education compared to 2.17% of non-biological children who are not relatives. The widening gap in attainment between biological and non-biological children as they get

[20] Details of the procedure are discussed in Chapter 2, section 2.4.3.

[21] Since the rationale for using the instruments is the same as that given in the previous chapter, in section 2.4.3, we do not repeat it here to conserve space.

[22] More details about IHS2 data are discussed in the previous chapter, in section 2.4.4.

[23] Similar sample restrictions are imposed in other studies e.g. Glick and Sahn (2000), Maitra (2003), Kabubo-Mariara and Mwabu (2007).

older may be a reflection of early withdrawals from school by the non-biological children or grade repetition. The withdrawals may increase with age due to the fact that as the children get older they can be a source of labor for agriculture and other income generating activities to supplement household income. This need for child labour is stronger for non-biological children especially those who are not relatives. There may be a direct cost dimension to this as well, in the sense that at lower ages (coinciding with primary school) education is free in Malawi, and at higher ages (coinciding with secondary school) parents have to pay fees among other things which might discourage attainment of secondary education. Either way, this may suggest bias against non-biological children, and that this bias gets worse when a child is not a relative. When we use current enrolment to measure schooling (see Table 3.2), a similar picture emerges. For example, for the age range 15-19, the results show that 78.% of biological children are still in school compared to just 63.8% and 51.2% of non-biological children who are relatives and non-biological children who are not relatives respectively.

Table 3.3, reports enrolment rates by income quintile. We observe that children in the wealthiest households have higher enrolment rates regardless of whether they are biological or not. For all quintiles, the results consistently show that biological children have higher enrolment rates. For instance, for the lowest quintile and comparing biological children with non-biological children, the results indicate that biological children have a higher enrolment rate of 94% compared to 76.9% for non-biological children. Looking at the highest quintile and comparing biological children with non-biological children, the results show that biological children have a higher enrolment rate of 96.2% compared to 93.1% for non-biological children. The results also show that non-biological children who are not relatives have consistently lower enrolment rates across all quintiles. The relationship between enrolment rates and the opportunity cost of schooling as measured by local child wages is presented in Table 3.4. The results show that the enrolment rate for all children declines as the opportunity cost of schooling increases. Comparing the lowest cost bracket (0-100) with the highest cost bracket (801+), we find that the enrolment rate of non-biological children drops more sharply compared to that of biological children[24]. Descriptive statistics of the explanatory variables used in the econometric analysis are presented in the appendix Table A3.1.

3.4 Econometric results

The descriptive results show that the schooling of non-biological children however measured is worse than that of biological children. We pursue this matter further by checking

[24] Results for descriptive analysis of grade attainment versus income or opportunity cost of schooling are similar, we have therefore not reported them to conserve space.

the predictions highlighted earlier. As earlier discussed, the log of per capita expenditure is potentially endogenous, we therefore conducted exogeneity tests in the probit as well as the censored ordered probit models using the Rivers and Vuong procedure outlined before. We reject the null hypothesis of exogeneity of the log of per capita expenditure in all probit models estimated. For all censored probit models estimated, we find that the log of per capita expenditure is exogenous. The reduced form regressions of log of per capita expenditure reported in the appendix Table A3.2, show that the instrumental variables land and its square perform reasonably well as they are significantly correlated with the log of per capita expenditure.

We now empirically look at the four predictions outlined earlier. We check these predictions by computing elasticities of probability for the censored ordered probit and the probit models estimates using children who stay in blended households. The elasticities of probability are computed at the sample means of the regressors. While controlling for the parental and household characteristics of the children's care givers, we estimate separate regressions for all biological children, and all non-biological children, who are further demarcated into non-biological children who are relatives and non-biological children who are non-relatives. The relationship between raw coefficients ($\hat{\beta}s$) and elasticities of probability in an ordered probit model deserves some mention. The elasticity of probability of the first outcome (no education) with respect to any regressor has the opposite sign to that of the regressor's coefficient. The elasticity of probability of the highest outcome (secondary education) with respect to any regressor has the same sign as the regressor's coefficient. For the intermediate outcomes there is no simple relationship between the elasticities of probability with respect to any regressor and corresponding regressor coefficients. Thus, for the lowest and highest education outcomes, the relationship between the elasticity of probability with respect to any regressor's coefficient is unambiguous[25].

Table 3.5, presents computed elasticities of probability with respect to income as proxied by the log of per capita expenditure for the censored ordered probit and probit models, to examine whether the schooling of non-biological children increases more relative to that of biological children following an increase in household income (first prediction). For the non-biological children, we further separate them into whether they are related to the care givers they stay with or not. This separation allows us to further investigate differences in schooling responsiveness to income changes between children who are relatives and those who are not. The raw coefficients for the censored ordered probit and probit models which are used to compute these elasticities are presented in Tables A3.3 and A3.4 respectively in the appendix. Using grade attainment as our measure of schooling, the results show that non-biological children have higher income elasticities (in absolute value terms) compared to non-biological children. We observe that the grade attainment of non-biological children

[25] These relationships can be seen by partially differentiating equation 3.18 with respect to any variable.

who are not relatives is more income elastic compared to non-biological children who are relatives. For example, at the secondary school level, non-biological children who are relatives have an income elasticity of 1.77 compared to 1.82 for non-biological children who are not relatives. A closer look at the magnitude of the elasticities indicates that non-biological children have greater than one elasticities and biological children have less than one elasticities. This implies that the education of non-biological children is considered a luxury good. Education becomes more luxurious if the non-biological children are non-relatives. The computed elasticities for all children also show that the income elasticity increases as we move up the education hierarchy i.e. from no education to secondary education. This means that higher education levels are considered a luxury. Further to that, the income elasticity as one moves up the education system are consistently largest for the non-biological children who are non-relatives, thus suggesting that for children who are non-relatives their further education is considered more of a luxury good relative to non-biological children who are relatives and own children. We get a similar picture when we use current enrolment to measure schooling. These results therefore confirm the prediction that an improvement in household economic status has a larger increase in the schooling of non-biological children relative to own children, and the increase is larger the more distantly related the non-biological children are. This finding is invariant to choice of schooling measure. This finding suggests that policies aimed at fighting poverty have the potential to improve the schooling outcomes of non-biological children.

In Table 3.6, we report censored ordered probit and probit elasticities to examine whether the schooling of non-biological children falls more relative to that of biological children following an increase in costs (second prediction). Again for the non-biological children, we further separate them into whether they are related to the care givers they stay with or not. The raw coefficients for censored ordered probit and probit models which are used to compute these elasticities are presented in Tables A3.3 and A3.4 respectively in the appendix. When we use grade attainment as a measure of schooling the computed price elasticities (in absolute value terms) indicate consistently for all grade levels that the schooling of non-biological children is more price elastic compared to that of biological children. Further to that, the results show that unrelated biological children have higher price elasticities relative to non-biological children who are relatives. For instance, when we look at the secondary level, non-biological children who are relatives have a price elasticity of -0.35 compared to -0.38 for non-biological children who are not relatives. We also note that for all children the price elasticities increase as we move from the lowest educational level (no education) to the highest (secondary education). This suggests that households become more responsive to school cost changes as a child goes up the education ladder. This responsiveness is more pronounced for non-biological children who are not relatives. When we use current enrolment as our measure of schooling we get similar conclusions. So our results confirm the prediction that an increase in the cost of schooling leads to a bigger decrease in the schooling of non-biological children relative to own children, and the increase is larger the more distantly related the non-biological children are. This conclusion is robust to how schooling is measured. This finding has policy relevance in the sense that interventions to end child labour would benefit non-biological children a lot.

Results of predictions 3 and 4, are reported in Tables 3.7 and 3.8 for the censored ordered probit and probit models respectively. Here we estimate elasticities of probability with respect to price for the two models at different consumption expenditure quintiles to ascertain the relationship between price elasticities as we move up the income ladder i.e. moving from the poorest households to the wealthiest households. Like before, for the non-biological children we further separate them into whether they are related to the care givers they stay with or not. The corresponding raw coefficients for censored ordered probit and probit models which are used to compute these elasticities are presented in Tables A3.5-A3.8 and A3.9-A3.12 respectively in the appendix. Using grade attainment as our measure of schooling, the results for all children show that the computed price elasticities fall as we move from the poorest households (1^{st} quintile) to the wealthiest households (5^{th} quintile). For example, looking at senior primary education for biological children, the results show that the 1^{st} quintile has a price elasticity of -0.29 compared with -0.15 for the 5^{th} quintile. This indicates an asymmetry in schooling responsiveness following changes in costs between poor households and rich households in the sense that poor households are more price elastic compared to rich households. The results also show that the price elasticities of non-biological children move towards those of biological children as we move up from the poorest households (1^{st} quintile) to the wealthiest households (5^{th} quintile), thus suggesting a convergence of price elasticities as household economic status improves. We further note that this convergence is faster for the non-biological children who are non-relatives. When school enrolment is used, we arrive at similar conclusions. Essentially, these findings confirm predictions 3 and 4. That is, the decrease in schooling due to cost increases is negatively related to household economic status, and as household economic status improves the fall in schooling as a result of cost increases converges, and converges faster the more distantly related the non-biological children are. This conclusion is independent of how the schooling of children is measured. This conclusion has policy implications in that efforts to improve the economic status of households would lead to an improvement in schooling of non-biological children.

3.5 Conclusions

The chapter has looked at the intrahousehold schooling bias against non-biological children in a family at both the theoretical and empirical levels. At the theoretical level, we have looked at the possible sources of schooling bias against non-biological children and how schooling responds to changes in household economic status and schooling costs. This has been done by constructing a two period model in which parents work and invest in a biological and non-biological child in the first period, and retire in the second. The parents survive on remittances from both children in old age. The model predicts discrimination against a non-biological child can stem from either non-preference based conditions which favour the biological child, and/or can originate from a pure preference bias against a non-biological child. Specifically, the model shows that parents will invest more in the education of their own child if costs, especially opportunity cost of schooling are higher for the non-biological child, or if returns to education of the own child are higher than those for the non-biological child, or if the subjective belief that parents have about future transfers during retirement is lower for the non-biological child relative to the biological child. Further to that, the model predicts that the schooling gap between the biological child and the non-biological child gets wider as the relationship distance between the non-biological child and the parents gets wider. We have also shown that schooling against non-biological children in a household can be a result of pure preference bias by parents against them, in the sense that they get more satisfaction from the income of their own child relative to the non-biological child.

The model also shows that there is an asymmetry in the impact of changes in costs and income on schooling in the sense that the impact is larger for the non-biological child. We have also shown that an increase in cost of schooling leads to a bigger reduction in schooling for poor households, and that the difference in the impact of cost changes between the biological and the non-biological child declines as income increases. An empirical investigation of these predictions using the Second Malawi Integrated Household Survey (IHS2) data has shown that when current enrolment and grade attainment are used to measure schooling, the price (measured as the opportunity cost of schooling) and income elasticities of schooling are larger for non-biological children. It has been found that non-biological children who are unrelated to their care givers have higher price and income elasticities. The empirical analysis has also indicated that households in the lowest income quintile (the poorest) have the largest price elasticities, and households in the highest income quintile (the wealthiest) have the smallest price elasticities. It has been demostrated that the price elasticities for biological and non-biological children converge as we move from the lowest income quintile to the highest income quintile, and that the convergence is faster for non-biological children who are non-relatives.

Our empirical analysis has not taken into account the possibility that sending children to school and sending them to work (i.e. child labour) are joint decisions, our results should therefore be interpreted with this caution in mind. Further to that the integrated household survey data presents some limitations to the study which are worth mentioning.

The survey did not collect information on quality of children such as IQ scores, which might affect grade attainment and school enrolment of biological and non-biological children. Secondly, the survey data does not have information on the past schooling performance of the children. Past schooling performance of the non-biological children before moving to their current care givers may also affect their current schooling in the sense that current care givers may have little incentive to send a non-biological child to school if s/he was struggling academically before coming to them. Since we do not control for these factors in our empirical analysis, our conclusions should be taken with due cognizance of these limitations.

In the previous two chapters, we have shown how household economic status as measured by the log of per capita consumption affects parental spending on primary school education of own children, and how it affects schooling bias against non-biological children in Malawi. In Chapter 2, we have found that improvements in household economic status positively affect spending on education in both rural and urban areas. The gap in poverty (household economic status) between rural and urban areas is the largest driver of the differential in spending on education between rural and urban areas. In Chapter 3, we have demonstrated both theoretically and empirically that improvements in household economic status have a bigger impact on the schooling of non-biological children. As households become less poor, the difference in the impact of cost changes between biological and non-biological children declines. These findings beg the question, how does the number of children impact on household economic status? In the next chapter, we answer this question.

Table 3.1: Grade attainment of children (age 10-19)

Grade	Biological	Non-biological		
		All	Related	Unrelated
No education	11.21	21.25	18.2	24.3
Junior primary	48.3	45.63	46.54	44.72
Senior primary	32.7	29.34	29.87	28.81
Secondary	7.8	3.78	5.38	2.17
Total	100	100	100	100

Table 3.2: Current school enrolment rates of children (age 6-19)

Age	Biological	Non-biological		
		All	Related	Unrelated
6-8	0.947	0.88	0.913	0.847
9-11	0.948	0.863	0.894	0.832
12-14	0.857	0.738	0.762	0.713
15-19	0.783	0.575	0.638	0.512

Table 3.3: Current school enrolment rates of children (age 6-19) by expenditure quintile

Quintile	Biological	Non-biological		
		All	Related	Unrelated
1st	0.940	0.769	0.783	0.754
2nd	0.918	0.812	0.831	0.793
3rd	0.937	0.886	0.916	0.856
4th	0.944	0.922	0.921	0.923
5th	0.962	0.931	0.934	0.927

Table 3.4: Current school enrolment rates of children (age 6-19) and opportunity cost of schooling

Cost	Biological	Non-biological		
		All	Related	Unrelated
0-100	0.882	0.84	0.864	0.816
101-400	0.891	0.802	0.837	0.767
401-800	0.887	0.687	0.712	0.662
801+	0.844	0.563	0.631	0.494

Table 3.5: Income elasticities from probit and censored ordered probit models

	Biological	Non-biological		
		All	Related	Unrelated
Current enrolment				
	0.89	1.26	1.25	1.29
Grade attainment				
No education	-0.75	-1.34	-1.33	-1.37
Junior primary	0.79	1.42	1.41	1.43
Senior primary	0.86	1.64	1.62	1.67
Secondary	0.97	1.78	1.77	1.82

Table 3.6: Price elasticities from probit and censored ordered probit models

	Biological	Non-biological		
		All	Related	Unrelated
Current enrolment				
	-0.24	-0.32	-0.31	-0.36
Grade attainment				
No education	0.12	0.24	0.23	0.27
Junior primary	-0.18	-0.26	-0.25	-0.30
Senior primary	-0.22	-0.34	-0.32	-0.36
Secondary	-0.28	-0.36	-0.35	-0.38

Table 3.7: Price elasticities from censored ordered probit for different income quintiles

Quintile	No education	Junior primary	Senior primary	Secondary
	Biological			
1st	0.24	-0.27	-0.29	-0.32
2nd	0.20	-0.25	-0.26	-0.29
3rd	0.17	-0.21	-0.23	-0.25
4th	0.13	-0.18	-0.20	-0.21
5th	0.09	-0.12	-0.15	-0.17
	Non-biological: All			
1st	0.29	-0.31	-0.34	-0.38
2nd	0.24	-0.29	-0.30	-0.35
3rd	0.20	-0.22	-0.27	-0.30
4th	0.15	-0.17	-0.19	-0.24
5th	0.10	-0.13	-0.17	-0.19
	Non-biological: Related			
1st	0.37	-0.38	-0.45	-0.48
2nd	0.27	-0.33	-0.41	-0.43
3rd	0.22	-0.29	-0.34	-0.36
4th	0.13	-0.20	-0.22	-0.23
5th	0.11	-0.13	-0.16	-0.18
	Non-biological: Unrelated			
1st	0.39	-0.42	-0.46	-0.48
2nd	0.30	-0.35	-0.38	-0.39
3rd	0.19	-0.21	-0.24	-0.26
4th	0.11	-0.18	-0.20	-0.21
5th	0.09	-0.12	-0.15	-0.17

Table 3.8: Price elasticities from probit for different income quintiles

Quintile	Biological	Non-biological All	Non-biological Related	Non-biological Unrelated
1st	-0.26	-0.33	-0.34	-0.35
2nd	-0.19	-0.27	-0.28	-0.27
3rd	-0.17	-0.21	-0.20	-0.17
4th	-0.13	-0.15	-0.14	-0.12
5th	-0.11	-0.12	-0.13	-0.11

3.6 Appendix to Chapter 3

Table A3.1: Sample means (standard errors) of explanatory variables

Variable	(1)		(2)	
child's age	10.828	(4.058)	12.166	(4.409)
child's sex	0.521	(0.500)	0.473	(0.499)
consumption expenditure	9.568	(0.551)	9.664	(0.597)
household size	6.794	(2.405)	6.250	(3.013)
father works	0.233	(0.423)	0.192	(0.394)
mother works	0.050	(0.218)	0.043	(0.202)
father's education	2.015	(3.947)	1.843	(3.951)
mother's education	0.715	(2.439)	0.888	(2.780)
father's age	45.007	(10.069)	47.306	(16.588)
mother's age	38.392	(9.271)	48.027	(19.247)
community wage	97.394	(136.178)	99.297	(154.182)
rural	0.908	(0.289)	0.901	(0.299)
north	0.166	(0.372)	0.195	(0.396)
centre	0.427	(0.495)	0.386	(0.487)
south	0.408	(0.491)	0.419	(0.494)

Notes: Column 1 corresponds to the current enrolment sample, and Column 2 corresponds to the grade attainment sample.

Table A3.2: Reduced form regressions of log per capita consumption

Variable	(1)	(2)
child's age	0.007***	0.003*
	(0.001)	(0.001)
child's sex	-0.004	0.005
	(0.010)	(0.011)
father works	0.060***	0.063***
	(0.013)	(0.013)
mother works	0.027	-0.009
	(0.026)	(0.029)
father's education	0.026***	0.026***
	(0.001)	(0.002)
mother's education	0.042***	0.043***
	(0.002)	(0.003)
father's age	-0.014***	-0.016***
	(0.004)	(0.005)
father's age^2	0.012***	0.000***
	(0.001)	(0.000)
mother's age	-0.015***	-0.015***
	(0.004)	(0.006)
mother's age^2	0.024**	0.031**
	(0.003)	(0.004)
land	0.015***	0.011***
	(0.002)	(0.004)
land2	-0.022***	-0.017***
	(0.001)	(0.007)
north	0.019	0.016
	(0.014)	(0.016)
centre	0.304***	0.301***
	(0.011)	(0.012)
constant	9.955***	10.012***
	(0.053)	(0.085)
F-test of joint significance of instruments:		
F-stat	32.71	27.43
Prob> F-stat	0.00	0.00
F-test of overall significance:		
F-stat	22.54	37.39
Prob> F-stat	0.00	0.00
R-squared	0.4532	0.4117

Notes: Column 1 corresponds to the current enrolment sample, and Column 2 corresponds to the grade attainment sample. The instruments for per capita consumption expenditure are land, its square. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.3: Censored ordered probit results of grade attainment by type of child

	Biological	Non-biological		
Grade		All	Related	Unrelated
child’s age	-0.246***	-0.146***	-0.246***	-0.147***
	(0.005)	(0.007)	(0.005)	(0.007)
child’s sex	0.080***	0.077	0.081***	0.078
	(0.028)	(0.052)	(0.028)	(0.052)
father works	0.070**	0.023	0.067*	0.019
	(0.035)	(0.072)	(0.035)	(0.072)
Mother works	0.172***	0.066	0.015	0.055
	(0.003)	(0.137)	(0.073)	(0.137)
father’s education	0.050***	0.040***	0.050***	0.040***
	(0.004)	(0.008)	(0.004)	(0.008)
mother’s education	0.026***	0.037***	0.027***	0.037***
	(0.007)	(0.011)	(0.007)	(0.011)
father’s age	-0.003	-0.014	-0.002	-0.014
	(0.012)	(0.015)	(0.012)	(0.015)
father’s age^2	-0.001	0.003	-0.001	0.003
	(0.003)	(0.006)	(0.004)	(0.002)
mother’s age	0.019	0.021	0.018	0.021
	(0.015)	(0.016)	(0.015)	(0.016)
mother’s age^2	-0.002	-0.002	-0.003	-0.001
	(0.004)	(0.006)	(0.003)	(0.006)
consumption expenditure	0.370***	0.142***	0.371***	0.142***
	(0.029)	(0.052)	(0.029)	(0.052)
community wage	-0.015***	-0.023***	-0.000***	-0.000
	(0.001)	(0.004)	(0.000)	(0.000)
household size	0.033***	0.015	0.034***	0.015
	(0.006)	(0.010)	(0.006)	(0.010)
rural	-0.181***	-0.253***	-0.175***	-0.264***
	(0.050)	(0.088)	(0.050)	(0.088)
north	0.359***	0.546***	0.351***	0.538***
	(0.041)	(0.072)	(0.041)	(0.072)
centre	-0.162***	0.103*	-0.164***	0.099
	(0.032)	(0.061)	(0.033)	(0.062)
Log likelihood	-4321	-5353	-3722	-6167

Notes: Threshold parameters not reported. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.4: Probit results of current enrolment by type of child

Grade	Biological	Non-biological All	Non-biological Related	Non-biological Unrelated
child's age	-0.214***	-0.267***	-0.218***	-0.220***
	(0.023)	(0.036)	(0.039)	(0.017)
child's sex	0.129**	0.430***	0.718***	0.920***
	(0.051)	(0.092)	(0.102)	(0.101)
father works	0.444**	0.289	0.102	0.068
	(0.203)	(0.318)	(0.348)	(0.151)
mother works	0.652***	0.108	0.091	0.070
	(0.141)	(0.259)	(0.300)	(0.272)
father's education	0.138***	0.118	0.072	0.058***
	(0.006)	(0.130)	(0.143)	(0.017)
mother's education	0.246*	0.106	0.012	0.007
	(0.136)	(0.204)	(0.223)	(0.024)
father's age	0.087*	0.047	0.028	0.021
	(0.050)	(0.072)	(0.083)	(0.036)
father's age^2	-0.011**	0.001	0.026	0.013
	(0.001)	(0.001)	(0.001)	(0.001)
mother's age	0.105*	0.031	0.008	0.001
	(0.056)	(0.078)	(0.087)	(0.036)
mother's age^2	-0.01*	-0.001	-0.003	-0.021
	(0.001)	(0.001)	(0.001)	(0.000)
Consumption expenditure	0.353***	0.192***	0.664***	0.961***
	(0.053)	(0.007)	(0.008)	(0.008)
Community wage	-0.217***	-0.321***	0.334***	0.413***
	(0.004)	(0.001)	(0.001)	(0.001)
Household size	0.019*	0.032*	-0.003	-0.002
	(0.011)	(0.017)	(0.018)	(0.018)
rural	-0.055***	-0.039***	-0.056***	-0.052***
	(0.009)	(0.001)	(0.002)	(0.004)
north	0.185*	0.341**	0.311*	0.303**
	(0.096)	(0.149)	(0.168)	(0.147)
centre	-2.236**	0.980	0.271	0.121
	(0.988)	(1.488)	(1.624)	(0.119)
residcons	6.663**	8.562***	9.515***	7.149***
	(3.252)	(1.911)	(1.365)	(1.06)
constant	-66.159**	28.339	7.804	2.636**
	(32.356)	(48.877)	(53.408)	(1.042)
Log likelihood	-6339	-2167	-4486	-5413

Notes: residcons is the residual from the reduced form of per capita consumption expenditure. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.5: Censored ordered probit results of grade attainment of biological children by expenditure quintile

Quintile	1	2	3	4	5
child's age	0.227***	0.234***	0.226***	0.280***	0.314***
	(0.009)	(0.010)	(0.010)	(0.013)	(0.014)
child's sex	-0.016	-0.089	-0.010	-0.101	-0.231***
	(0.057)	(0.059)	(0.060)	(0.068)	(0.075)
father works	-0.029	-0.026	0.135*	0.223***	0.061
	(0.077)	(0.072)	(0.075)	(0.086)	(0.095)
mother works	0.087	0.355**	-0.360**	-0.322*	-0.007
	(0.174)	(0.146)	(0.176)	(0.174)	(0.182)
father's education	0.050***	0.041***	0.036***	0.052***	0.068***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
mother's education	0.032	0.014	0.019	0.028*	0.034***
	(0.023)	(0.020)	(0.016)	(0.015)	(0.013)
father's age	-0.047*	0.022	0.009	-0.021	-0.010
	(0.024)	(0.026)	(0.029)	(0.030)	(0.035)
father's age^2	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mother's age	0.081***	-0.017	-0.015	-0.001	0.020
	(0.028)	(0.031)	(0.034)	(0.036)	(0.042)
mother's age^2	-0.001***	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
consumption expenditure	0.450***	0.651**	0.135	1.124***	0.286**
	(0.142)	(0.328)	(0.369)	(0.355)	(0.116)
community wage	0.127***	-0.023***	-0.367***	-0.345***	0.413***
	(0.001)	(0.002)	(0.001)	(0.004)	(0.003)
household size	0.023**	0.016	0.043***	0.073***	0.040**
	(0.011)	(0.016)	(0.016)	(0.017)	(0.016)
rural	-0.046	0.055	-0.194	-0.192*	-0.230**
	(0.124)	(0.127)	(0.120)	(0.111)	(0.103)
north	0.465***	0.538***	0.320***	0.214*	0.251**
	(0.078)	(0.086)	(0.096)	(0.113)	(0.119)
centre	-0.155**	-0.229***	-0.223***	-0.153*	-0.003
	(0.072)	(0.069)	(0.068)	(0.079)	(0.092)
Log likelihood	-6567	-5448	-4321	-3389	-5117

Notes: Threshold parameters not reported. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.6: Censored ordered probit results of grade attainment of all non-biological children by expenditure quintile

Quintile	1	2	3	4	5
child's age	0.158***	0.170***	0.126***	0.158***	0.160***
	(0.017)	(0.019)	(0.015)	(0.016)	(0.016)
child'sex	0.240*	0.072	0.066	0.165	-0.081
	(0.123)	(0.134)	(0.119)	(0.116)	(0.116)
father works	-0.426*	0.050	0.047	0.088	0.019
	(0.225)	(0.185)	(0.156)	(0.155)	(0.161)
mother works	0.154	0.406	0.766	-0.643**	0.103
	(0.408)	(0.371)	(0.808)	(0.308)	(0.219)
father's education	0.034	0.016	0.027	0.054***	0.043***
	(0.024)	(0.026)	(0.018)	(0.016)	(0.015)
mother's education	-0.186***	0.105**	0.025	0.074***	0.032*
	(0.071)	(0.041)	(0.029)	(0.025)	(0.018)
father's age	-0.055	-0.046	-0.031	-0.004	0.047
	(0.039)	(0.045)	(0.029)	(0.034)	(0.039)
father's age^2	0.001	0.000	0.000	0.000	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mother's age	0.033	0.042	-0.043	0.040	-0.003
	(0.038)	(0.048)	(0.031)	(0.037)	(0.038)
mother's age^2	-0.000	-0.000	0.001*	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
consumption expenditure	0.010	0.437	0.302	-0.363	-0.047
	(0.284)	(0.687)	(0.752)	(0.584)	(0.161)
community wage	-0.321***	-0.237***	-0.364**	-0.433**	-0.311***
	(0.001)	(0.004)	(0.001)	(0.001)	(0.001)
household size	0.012	-0.023	0.061***	-0.006	0.038
	(0.019)	(0.024)	(0.023)	(0.025)	(0.026)
rural	-0.616**	0.172	-0.269	-0.269	-0.208
	(0.242)	(0.252)	(0.225)	(0.206)	(0.164)
north	0.671***	0.592***	0.576***	0.398**	0.634***
	(0.161)	(0.193)	(0.165)	(0.161)	(0.177)
centre	0.063	-0.163	-0.072	0.184	0.249*
	(0.157)	(0.161)	(0.142)	(0.126)	(0.131)
Log likelihood	-2321	-5478	-6311	-4299	-2139

Notes: Threshold parameters not reported. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.7: Censored ordered probit results of grade attainment of non-biological (but related) children by expenditure quintile

Quintile	1	2	3	4	5
child's age	0.227***	0.234***	0.226***	0.281***	0.315***
	(0.009)	(0.010)	(0.010)	(0.013)	(0.014)
child's sex	-0.017	-0.088	-0.010	-0.104	-0.238***
	(0.057)	(0.059)	(0.060)	(0.068)	(0.075)
father works	-0.040	-0.026	0.135*	0.211**	0.054
	(0.077)	(0.072)	(0.075)	(0.086)	(0.095)
mother works	0.074	0.360**	-0.364**	-0.306*	-0.012
	(0.174)	(0.146)	(0.176)	(0.174)	(0.182)
father's education	0.049***	0.040***	0.036***	0.055***	0.069***
	(0.010)	(0.009)	(0.009)	(0.009)	(0.009)
mother's education	0.032	0.015	0.020	0.024	0.034***
	(0.023)	(0.020)	(0.016)	(0.015)	(0.013)
father's age	-0.045*	0.022	0.009	-0.019	-0.008
	(0.024)	(0.026)	(0.029)	(0.030)	(0.035)
father's age^2	0.000	-0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mother's age	0.079***	-0.019	-0.014	-0.003	0.021
	(0.028)	(0.031)	(0.034)	(0.036)	(0.042)
mother's age^2	-0.001***	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
consumption expenditure	0.451***	0.638*	0.137	1.121***	0.291**
	(0.142)	(0.328)	(0.369)	(0.356)	(0.116)
community wage	-0.223***	-0.314***	-0.421***	-0.023	-0.043
	(0.003)	(0.002)	(0.001)	(0.043)	(0.067)
household size	0.023**	0.016	0.043***	0.074***	0.038**
	(0.011)	(0.016)	(0.016)	(0.017)	(0.016)
rural	-0.081	0.069	-0.190	-0.166	-0.236**
	(0.126)	(0.129)	(0.120)	(0.112)	(0.104)
north	0.450***	0.549***	0.317***	0.175	0.252**
	(0.079)	(0.087)	(0.097)	(0.115)	(0.121)
centre	-0.164**	-0.219***	-0.224***	-0.161**	-0.001
	(0.072)	(0.069)	(0.068)	(0.080)	(0.092)
Log likelihood	-3329	-4412	-8342	-7613	-3979

Notes: Threshold parameters not reported. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.8: Censored ordered probit results of grade attainment of non-biological (unrelated) children by expenditure quintile

Quintile	1	2	3	4	5
child's age	0.160***	0.170***	0.125***	0.162***	0.160***
	(0.017)	(0.019)	(0.015)	(0.016)	(0.016)
child's sex	0.238*	0.055	0.084	0.181	-0.082
	(0.123)	(0.135)	(0.120)	(0.116)	(0.116)
father works	-0.452**	0.051	0.085	0.125	0.016
	(0.228)	(0.186)	(0.159)	(0.158)	(0.162)
mother works	0.150	0.384	0.796	-0.678**	0.099
	(0.409)	(0.374)	(0.810)	(0.310)	(0.220)
father's education	0.033	0.015	0.028	0.051***	0.042***
	(0.024)	(0.026)	(0.018)	(0.016)	(0.015)
mother's education	-0.188***	0.109***	0.022	0.078***	0.033*
	(0.072)	(0.041)	(0.029)	(0.025)	(0.018)
father's age	-0.051	-0.042	-0.034	-0.011	0.046
	(0.039)	(0.045)	(0.029)	(0.035)	(0.040)
father's age^2	0.000	0.000	0.000	0.000	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
mother's age	0.031	0.035	-0.043	0.051	-0.003
	(0.038)	(0.048)	(0.031)	(0.038)	(0.038)
mother's age^2	-0.000	-0.000	0.001*	-0.000	0.001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
consumption expenditure	-0.011	0.486	0.227	-0.313	-0.046
	(0.288)	(0.693)	(0.759)	(0.587)	(0.162)
community wage	-0.313***	-0.411***	-0.261***	-0.428***	-0.415***
	(0.001)	(0.003)	(0.001)	(0.001)	(0.001)
household size	0.013	-0.021	0.060**	-0.003	0.038
	(0.020)	(0.024)	(0.024)	(0.025)	(0.026)
rural	-0.615**	0.227	-0.285	-0.330	-0.210
	(0.250)	(0.262)	(0.225)	(0.213)	(0.165)
north	0.673***	0.604***	0.571***	0.406**	0.611***
	(0.161)	(0.194)	(0.166)	(0.164)	(0.180)
centre	0.075	-0.159	-0.083	0.207	0.232*
	(0.158)	(0.162)	(0.143)	(0.131)	(0.133)
Log likelihood	-5388	-3419	-4358	-6931	-4266

Notes: Threshold parameters not reported. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.9: Probit results of current enrolment of biological children by expenditure quintile

Quintile	1	2	3	4	5
child's age	-0.269***	-0.267***	-0.213***	-0.146***	-0.336**
	(0.067)	(0.047)	(0.044)	(0.041)	(0.134)
child's sex	0.106	0.167	0.192*	0.132	0.266
	(0.103)	(0.103)	(0.115)	(0.131)	(0.180)
father works	0.962*	0.878**	-0.520	0.547	-1.326
	(0.568)	(0.404)	(0.373)	(0.381)	(1.174)
mother works	0.842**	0.485*	-1.389***	-0.155	-0.844
	(0.328)	(0.281)	(0.287)	(0.399)	(0.635)
father's education	0.308	0.366**	-0.100	0.141	-0.439
	(0.247)	(0.171)	(0.156)	(0.149)	(0.498)
mother's education	-0.581	-0.537**	-0.127	0.216	-0.829
	(0.398)	(0.274)	(0.246)	(0.235)	(0.796)
father's age	0.153	0.135	0.155	-0.052	0.266
	(0.137)	(0.103)	(0.097)	(0.092)	(0.275)
father's age^2	-0.002	-0.002	-0.002*	0.001	-0.003
	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)
mother's age	0.229	0.265**	0.060	-0.167	0.384
	(0.157)	(0.113)	(0.103)	(0.115)	(0.309)
mother's age^2	-0.002	-0.002**	-0.000	0.002	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
consumption expenditure	0.385*	0.109***	0.218	-0.089	0.989***
	(0.223)	(0.005)	(0.679)	(0.678)	(0.352)
community wage	-0.321***	-0.219***	-0.365***	0.413***	-0.451***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
household size	0.070***	-0.011	-0.022	-0.017	0.008
	(0.022)	(0.026)	(0.028)	(0.027)	(0.030)
rural	0.160	0.398**	-0.018	-0.131	-0.037
	(0.186)	(0.188)	(0.227)	(0.245)	(0.228)
north	0.162	0.029	-0.146	0.502*	0.281
	(0.211)	(0.192)	(0.213)	(0.282)	(0.450)
centre	-4.397	-4.867**	-1.713	1.059	-6.088
	(2.894)	(1.983)	(1.788)	(1.684)	(5.807)
residcons	14.041***	14.820**	8.687***	13.671***	19.261***
	(2.503)	(6.521)	(1.866)	(2.596)	(2.108)
constant	-139.659	-144.169**	-46.604	43.518	-199.378
	(94.693)	(65.244)	(58.351)	(56.283)	(190.196)
Log likelihood	-3221	-4771	-7834	-6389	-7227

Notes: residcons is the residual from the reduced form of per capita consumption expenditure. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.10: Probit results of current enrolment of all non-biological children by expenditure quintile

Quintile	1	2	3	4	5
child's age	-0.296*	-0.235***	-0.256***	-0.304***	-0.308*
	(0.177)	(0.081)	(0.066)	(0.093)	(0.173)
child's sex	0.303	0.187	0.455**	0.253	1.015***
	(0.251)	(0.227)	(0.222)	(0.229)	(0.242)
father works	-0.701	0.006	1.248*	0.252	1.632
	(1.563)	(0.695)	(0.644)	(0.797)	(1.531)
mother works	-0.609	0.985	0.375	0.117	0.390
	(0.878)	(0.776)	(0.243)	(0.322)	(0.636)
father's education	-0.197	0.204	0.463	0.155	0.583
	(0.662)	(0.289)	(0.368)	(0.512)	(1.017)
mother's education	0.427	0.403	-0.182	-0.025	-0.179
	(1.090)	(0.462)	(0.132)	(0.184)	(0.341)
father's age	0.143	-0.232	0.002	0.000	0.002
	(0.355)	(0.166)	(0.002)	(0.002)	(0.004)
father's age^2	-0.002	0.002	-0.126	0.039	-0.167
	(0.004)	(0.002)	(0.144)	(0.193)	(0.381)
mother's age	0.230	0.030	0.001	-0.000	0.001
	(0.387)	(0.186)	(0.001)	(0.001)	(0.003)
mother's age^2	-0.001	-0.000	-0.777	1.100	-0.690**
	(0.003)	(0.001)	(1.406)	(1.178)	(0.291)
consumption expenditure	0.212	-0.222	0.002	-0.002	-0.002
	(0.475)	(1.203)	(0.001)	(0.003)	(0.002)
community wage	-0.124***	-0.412***	0.335***	-0.179***	-0.472***
	(0.001)	(0.002)	(0.003)	(0.060)	(0.049)
household size	0.025	0.002	-0.222	-0.536	0.094
	(0.033)	(0.041)	(0.425)	(0.421)	(0.320)
rural	0.049	0.194	0.728**	0.072	0.679
	(0.378)	(0.450)	(0.335)	(0.406)	(0.516)
north	0.051	0.521	4.389	0.321	4.410
	(0.544)	(0.384)	(2.699)	(3.736)	(7.390)
centre	-2.780	2.213	-12.581	-2.212	-13.612
	(7.725)	(3.318)	(8.913)	(12.283)	(24.411)
residcons	10.380***	-17.633***	24.127***	-0.220***	0.354***
	(1.333)	(2.983)	(2.212	(0.008)	(0.002)
constant	-103.381	82.100	136.409	13.959	147.080
	(252.640)	(111.196)	(92.730)	(122.303)	(242.940)
Log likelihood	-9332	-8178	-1977	-2256	-4038

Notes: residcons is the residual from the reduced form of per capita consumption expenditure. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.11: Probit results of current enrolment of non-biological (but related) children by expenditure quintile

Quintile	1	2	3	4	5
child's age	-0.174***	-0.167***	-0.186***	-0.171***	-0.205***
	(0.014)	(0.015)	(0.018)	(0.020)	(0.027)
child's sex	0.045	0.105	0.179	0.147	0.180
	(0.093)	(0.097)	(0.114)	(0.129)	(0.156)
father works	-0.118	0.016	-0.251*	0.325*	-0.173
	(0.124)	(0.122)	(0.139)	(0.191)	(0.207)
mother works	-0.465**	-0.097	-1.215***	-0.253	-0.337
	(0.230)	(0.219)	(0.253)	(0.366)	(0.364)
father works	0.059***	0.025	0.030*	0.045**	0.065***
	(0.020)	(0.017)	(0.018)	(0.019)	(0.022)
mother works	0.002	0.075	0.069*	0.064	-0.030
	(0.045)	(0.046)	(0.041)	(0.044)	(0.030)
father's age	-0.041	-0.070	0.091*	-0.001	-0.001
	(0.044)	(0.051)	(0.051)	(0.057)	(0.073)
father's age^2	0.000	0.001	-0.001**	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)
mother's age	0.015	0.043	-0.016	-0.111	0.093
	(0.051)	(0.057)	(0.059)	(0.075)	(0.085)
mother's age^2	-0.000	-0.001	0.000	0.001	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
consumption expenditure	0.382*	0.080	0.286	-0.089	0.995***
	(0.223)	(0.547)	(0.688)	(0.678)	(0.354)
community wage	-0.345***	0.421***	-0.231	-0.453***	-0.413***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)
household size	0.071***	-0.012	-0.025	-0.017	0.007
	(0.022)	(0.026)	(0.028)	(0.027)	(0.030)
rural	0.166	0.345*	-0.048	-0.131	-0.033
	(0.187)	(0.192)	(0.228)	(0.246)	(0.228)
north	0.426***	0.276*	-0.090	0.434	0.648*
	(0.144)	(0.161)	(0.195)	(0.273)	(0.342)
centre	-0.118	-0.379***	-0.327**	-0.058	-0.219
	(0.111)	(0.111)	(0.128)	(0.145)	(0.189)
residcons	-0.322**	-0.021***	-0.042***	0.324***	-0.237***
	(0.001)	(0.001)	(0.004)	(0.012)	(0.004)
constant	0.142	3.673	-0.289	6.975	-7.707*
	(2.205)	(5.229)	(6.783)	(6.922)	(3.986)
Log likelihood	-8789	-5339	-2778	-4456	-6729

Notes: residcons is the residual from the reduced form of per capita consumption expenditure. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Table A3.12: Probit results of current enrolment of non-biological (unrelated) children by expenditure quintile

Quintile	1	2	3	4	5
child's age	-0.227***	-0.292***	-0.351***	-0.320***	-0.400***
	(0.031)	(0.039)	(0.045)	(0.041)	(0.050)
child's sex	0.263	0.159	0.533**	0.286	1.071***
	(0.210)	(0.228)	(0.225)	(0.227)	(0.225)
father works	-0.084	-0.458	0.579*	0.134	0.804**
	(0.365)	(0.297)	(0.315)	(0.324)	(0.316)
mother works	-0.327	0.917	0.045	0.054*	0.035
	(0.575)	(0.759)	(0.039)	(0.030)	(0.026)
father's education	0.074*	-0.001	-0.068	0.066	0.018
	(0.042)	(0.044)	(0.052)	(0.053)	(0.032)
mother's education	0.004	0.097	-0.013	0.006	0.008
	(0.352)	(0.070)	(0.049)	(0.062)	(0.072)
father's age	0.002	-0.104	0.000	0.000	-0.000
	(0.062)	(0.080)	(0.000)	(0.001)	(0.001)
father's age^2	-0.000	0.001	0.068	0.071	0.040
	(0.001)	(0.001)	(0.055)	(0.071)	(0.073)
mother's age	0.071	0.115	-0.001	-0.000	-0.000
	(0.060)	(0.086)	(0.001)	(0.001)	(0.001)
mother's age^2	-0.000	-0.001	-0.676	1.073	-0.691**
	(0.001)	(0.001)	(1.422)	(1.176)	(0.291)
consumption expenditure	0.214	0.065	0.002	-0.002	-0.002
	(0.475)	(1.226)	(0.002)	(0.003)	(0.002)
community wage	-0.320***	-0.453***	-0.338***	-0.182***	-0.475***
	(0.001)	(0.003)	(0.043)	(0.060)	(0.049)
household size	0.025	0.006	-0.246	-0.511	0.090
	(0.033)	(0.042)	(0.418)	(0.425)	(0.320)
rural	0.044	0.281	0.526*	0.066	0.422
	(0.378)	(0.457)	(0.316)	(0.356)	(0.341)
north	0.249	0.468	0.562**	-0.305	0.260
	(0.269)	(0.345)	(0.270)	(0.276)	(0.262)
centre	0.386	-0.023	-0.000	0.000	0.000
	(0.283)	(0.264)	(0.000)	(0.000)	(0.000)
residcons	-0.237***	-0.172***	0.248***	-0.324***	-0.218***
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)
constant	-0.093	3.411	10.420	-7.847	11.618***
	(4.299)	(11.466)	(14.001)	(11.673)	(3.337)
Log likelihood	-8189	-9747	-2015	-2321	-4775

Notes: residcons is the residual from the reduced form of per capita consumption expenditure. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors.

Chapter 4

Impact of fertility on objective and subjective poverty in Malawi

4.1 Introduction

Research looking at the relationship between poverty and fertility at the micro level on the African continent remains scarce. Ironically, Africa has the highest rates of poverty and fertility. A lack of data has often been given as the reason for the paradoxical lack of studies on the continent. Empirically, there has not been any consensus as to the nature of the relationship between fertility and poverty. The mixed empirical results include; no relationship between fertility and poverty in Botswana (Chernichovsky 1984), a positive relationship in Sierra Leone and Ethiopia (Ketkar 1979), a negative relationship in Burkina Faso (Langani 1997) and in Southern Sudan (Cohen and House 1994). Further to that, Noumbissi and Sanderson (1998) find that in Cameroon where fertility rates are very high, the relationship takes the inverse "J shape", implying that both low and high-income households have lower rates of fertility, whereas medium level income households have higher fertility. The J shape is explained by the fact very low income households tend to be landless farmers; as a result they don't depend on children as cheap labor, whereas those with the highest income have lower fertility due to higher investment in child quality. The middle income families are landholding farms which depend on cheap labor, and therefore have a higher demand for child quantity.

The common thing about all the cited studies is that they treated fertility as an exogenous variable. By doing that, these studies ignored the fact that fertility can influence poverty, and at same time be affected by it. That is, causality can run in both directions. Technically, they did not take into account the simultaneity that exists between the two variables. Further to that, they also ignored the fact that there are unobserved factors which influence both variables; that is unobserved heterogeneity[1] was not accounted for. Another shortfall of these studies is that they only focussed on poverty defined in the objective monetary sense which is a narrower definition of household welfare. Subjective measures of welfare better capture the multidimensional nature of poverty. They are likely to include a household's feelings of relative deprivation, exclusion from services and institutions, as well as feelings of marginalization related to household or individual status (such as ethnicity, or marital status) (Devereux *et al.* 2006). It is therefore also interesting to see how fertility impacts on poverty when poverty is conceived multidimensionally.

Disregarding simultaneity and unobserved heterogeneity leads to biased and inconsistent estimates. It is therefore important for the reliability of results of any econometric analysis that they be accounted for. It is also worth noting that despite the poverty-fertility relationship being a demographic issue as well as an economic one; most of the studies on the continent have been done by demographers. These studies have mostly been descriptive in nature. As discussed before the results have been divergent with some studies finding no relationship, while other studies find a negative or positive relationship. The only study we are aware of which accounts for the two effects was done in India by Gupta and Dubey (2006). With respect to Malawi, there have been a few studies which have looked at factors which influence objective poverty (Mukherjee and Benson 2003; Bokosi 2007) but none of these has looked at the impact of fertility on objective poverty let alone subjective poverty. The questions that this study therefore seeks to answer are twofold. Firstly, taking into account the simultaneity and unobserved effects, how does fertility impact on objective poverty in Malawi? Secondly, taking into account the simultaneity and unobserved effects, how does fertility impact on subjective poverty in Malawi? Answering these questions is significant in the sense that it will go a long way in contributing to the literature on poverty and fertility in Malawi as well as the African continent at large. Additionally, by using a methodology that captures the problems that the previous studies have ignored, we will be making a contribution with respect to how the two variables should be conceptualized and modeled. Further to that, by using subjective poverty, the study will shed some light on the impact of fertility on a broader definition of household welfare.

After accounting for endogeneity of fertility by using a natural experiment, two girls first as our instrumental variable, the study finds a positive relationship between fertility and objective poverty. That is, having a large family increases the likelihood of being objectively poor. This effect is robust for all poverty lines used. It is also robust to accounting for economies of scale and household composition as well as assuming that poverty is continuous. We also find that when fertility is treated as an exogenous variable its impact is underestimated. When poverty is defined more broadly by using self rated

[1] Manski (1993) calls this the correlated effect.

assessments of welfare, the results are opposite to those of objective poverty. We find that fertility lowers the likelihood of feeling poor, that is having more than two children (a large family) lowers the probability of feeling poor. The study also finds that fertility is exogenous with respect to subjective poverty.

The rest of the chapter is organized as follows. In Section 4.2 we present arguments for why causality between fertility and poverty is bidirectional. Section 4.3 focuses on the measurement of objective and subjective poverty, and fertility. Section 4.4 dwells on the specification of the empirical model, data, and descriptives. Econometric results are the focus of section 4.5. Our conclusions are presented in Section 4.6.

4.2 The poverty-fertility nexus

In this section, we give explanations which have been given in the literature which show that causality between fertility and poverty is bidirectional. The link between poverty and fertility may run from fertility to poverty. Poor households with big families have large dependency ratios, as a result investments in the human capital of children, which improve the future prospects of the children may be sacrificed to more immediate household needs such as food. This conflict is especially likely when the opportunity cost of certain investments in children (such as education) is high because of the associated loss of child labor in agriculture or home work (Birdsall and Griffin 1988).

Family size may have a negative impact on child development and human capital formation, and hence their future economic status. For instance, studies in both developed and developing countries find that children in big families tend to be shorter, less intelligent, and are even less likely to survive (e.g. Birdsall 1977; Bielicki 1986; Casterline *et al.*1987). Birdsall (1980) found that though total household spending on education tends to rise with family size, expenditures per child on education tend to be lower in large families for all income and education levels of the parents. Rosenzweig and Wolpin 1980, in a study of families with twins in India found that the additional unexpected child represented by twins reduced enrollment levels of all children in the household. Using Malaysian data, Rosenzweig and Schultz 1987 show that couples with a higher biological propensity to have births, also have lower schooling attainment for their children. A child's ability to learn is influenced by the amount and quality of attention received from parents and other adults in the first few years, and that is generally less in large families. Hence, children from large families are more likely not to be very educated and this makes them to be more prone to poverty (Birdsall and Griffin 1988). U.S. studies show that women with large families put in no more time on child care; educated women succeed in spending more time with each child principally by having fewer children (see Birdsall 1977).

On the other hand, the link between poverty and fertility may run from poverty to fertility.

Parents whose children die may try to replace them, and since high mortality is generally high in poor households, parents may try to insure themselves against possible child loss by having more children than they would otherwise want. Olsen (1987 cited in Birdsall and Griffin 1988) found that parents in Colombia directly replaced at least 0.2 of children that had died, but further compensated by having on average about 0.14 extra children. Similar results were found for Malaysia (Olsen 1983 cited in Birdsall and Griffin 1988). Thus, hoarding by having extra children can be interpreted as an insurance strategy by parents in the presence of high infant mortality. As the risk of infant mortality diminishes, hoarding becomes unnecessary (Birdsall and Griffin 1988). Related to this, is that poor households may decide to have more children as a source of support in old age given the absence of life insurance markets and social security in many developing countries.

According to the quantity-quality theory of Becker and Lewis (1973), increases in income lead to an increase in demand for quality of children and a fall in the quantity of children. Thus, as households become wealthier, they will tend to have fewer children. Additionally, Willis (1973) argues that increases in women's wages (and therefore income) leads to fewer children, as this increases the opportunity cost of having more children.

4.3 Measurement of objective and subjective poverty and fertility

In this section, we discuss the measurement of both objective monetary poverty and subjective poverty as well as fertility. Objective poverty can be measured either by using household income or household consumption expenditure. Following Mukherjee and Benson (2003), we use a consumption expenditure based measure of poverty rather than income[2]. In the objective and monetary poverty analysis income or consumption is considered to be a measure of welfare. This approach reflects how most empirical work on poverty in Africa has been done. A household's subjective assessment of its well being is however much broader. Subjective well being (SWB) better captures the multidimensional nature of poverty. Subjective measures are likely to extend well beyond the narrow income or consumption needs, as they will include a household's feelings of relative deprivation, exclusion from services and institutions, as well as feelings of marginalization related to household or individual status (such as ethnicity, or marital status) (Devereux *et al.* 2006). Due to its broader scope, it is possible that some factors might affect the two poverty definitions differently. It is therefore imperative that we investigate how fertility impacts subjective poverty to complete the picture. There are three alternative "subjective" questions which are used to measure subjective wellbeing. Firstly, there is

[2]For reasons why a consumption based measure is better than an income one, see footnote 14 in Chapter 2.

what is called the Income Evaluation Question (IEQ) introduced by Van Praag (1971). The IEQ asks what level of income is regarded as 'very bad', 'bad','not good', 'not bad', 'good', 'very good'. The IEQ for example goes like "Taking into account my (our) present living circumstances, I would regard a net weekly/monthly/yearly (encircle period) family income as: excellent, good,. . . ,.bad, very bad." Secondly, there is the Minimum Income Question (MIQ). Here people are asked what they consider as a minimum level of income to make ends meet. The MIQ is for example phrased like "We would like you to tell us the absolute minimum income of money for a household such as yours – in other words, a sum below which you couldn't make ends meet." The MIQ has been criticized for its focus on income, in that the concept of income may be poorly defined for respondents particularly but not only in developing countries (Ravallion and Lokshin 2002). Both the IEQ and MIQ are based on income as a measure of welfare, and therefore they are not broad. A measure of subjective poverty which is much broader and open-minded is the Economic Ladder Question (ELQ). Here the respondents are asked a question framed as follows: "Please imagine a 6-step ladder where on the bottom, the first step, stand the poorest people, and on the highest step, the sixth, stand the rich. On which step are you today?" (Kalugina and Najman 2002). In the survey data we are using, this question was answered by the household head. Owing to its broadness relative to the IEQ and the MIQ, we employ the ELQ method to measuring subjective poverty[3]. In this study, we our measure of fertility is based on the number of children in a household[4].

4.4 Methodology

4.4.1 Motivation of the methodology

In order to take into account the fact that fertility and poverty (objective and subjective) are potentially endogenous, we use instrumental variable (IV) estimation. We use a natural experiment as our instrument[5]. For the IV to be valid it must be correlated with fertility but should be uncorrelated with poverty. Our use of a natural experiment is inspired by a number of studies in the labour supply literature. Bronars and Grogger (1994) use the

[3] The use of subjective wellbeing is advantageous in the sense that well being is self rated. These measures are however not completely perfect. For example, an individual's answers could be influenced by different factors, for instance, attitudes and anticipations. Individuals may estimate themselves by the means of comparison with socially accepted norms and rules, their group of reference etc (Kalugina and Najman 2002; Kingdon and Knight 2003).

[4] A Fertility measure normally used by demographers is the Total Fertility Rate (TFR) which is defined as the average number of children that would be born alive to a woman at the end of her reproductive period given the current specific fertility rate.

[5] A natural experiment is a naturally occurring random event or situation, which can be exploited as an instrumental variable. For a discussion on natural experiments in economics see Rosenzweig and Wolpin (2000).

incidence of twins in the 1970 and 1980 U.S. Census Public Use Microdata Samples to estimate the effect of an unplanned second child on labor force participation for unwed mothers. Angrist and Evans (1998) study the labor supply of married women with at least two children, using both twins and the gender mix of the first two children as instruments. Since parents tend to prefer having a mix of genders among their children, gender mix operates as an instrument because couples with two children of the same sex are more likely to have a third child than couples with one boy and one girl. It has to be said that using twins reduces usable data dramatically, and using sibling sex mix as an instrument applies in the US and probably other western countries. In most developing countries, parents tend to prefer sons to daughters. Gupta and Dubey (2006) in a study of the impact of fertility on poverty in India (which is probably the first to use IV estimation) use having two girls first as an IV on households with more than two children. *A priori* parents are more likely to have another child if the first two are girls.

Just like Gupta and Dubey (2006) we use two girls first as our IV. Sons are preferred in the developing world for a number of reasons. First, in many societies, old-age support is exclusively the task of male offspring by way of social practice and tradition. Even though female offspring may be just as able to offer support, there may be a stigma associated with receiving such support from daughters. Second, in societies where female employment is not in demand or undervalued, males may be potentially more productive future 'assets' (Gupta and Dubey 2006). Finally, sons may be preferred to daughters for the continuation of the family name.

4.4.2 Model specification

Following the motivation given in the preceding section, we specify a recursive bivariate probit model which nets out simultaneity and unobserved heterogeneity effects and therefore enables us to isolate the causal effect of fertility on poverty (objective and subjective). We have two dependent variables; poverty status and fertility which are binary. Our unit of analysis is a household.

Objective poverty

As discussed earlier, in this study we measure objective poverty using consumption, and a household is defined as poor if its total real annualized per capita consumption expenditure (Y_i) falls below the poverty line. Letting Y^{PL} be a poverty line, then household i is poor ($T_i = 1$) if $Y_i \leq Y^{PL}$ and non poor ($T_i = 0$) if $Y_i > Y^{PL}$. Parameter estimates of a probit change with the poverty line. This means that the effects of different variables on poverty are strictly speaking specific to that poverty line. To find out whether or not the effect of fertility on poverty is robust to choice of poverty line we use three poverty lines, and look for the presence of sign reversals in the impact of fertility on the alternative poverty lines. When there are no sign reversals i.e. monotonicity holds, then the results can be considered first order dominant, implying that the direction of the impact of a fertility variable on the probability of being poor remains the same regardless of poverty line selected (Ravallion 1996).

We use three poverty lines; two as defined by the National Statistical Office of Malawi (NSO), and the third as defined by the World Bank. The two NSO poverty lines are; one for those considered ultra poor which is 10029 Malawi Kwacha per year, and another for the poor which is 16165 Malawi Kwacha per year. The World Bank poverty line is the US \$1 per day (equivalent to an annualized figure of 11051 Malawi Kwacha after adjusting for purchasing power parity). The NSO poverty lines are based on the cost of basic needs approach. They are adjusted for interspatial and intertemporal price differences. It is worth pointing out at this stage, that there is an unsettled debate in the poverty literature regarding whether poverty should be modeled as a continuous variable by using a levels regression or as a dummy by using probit or logit models. The first advantage of the levels regression is that it uses all the information on the distribution of consumption expenditure, whilst the binary model loses important information by collapsing consumption expenditure into two values. Secondly, the binary variable is derived from an observed continuous variable, and this runs counter to the fundamental assumption on which the probit or logit is based. Specifically, the binary indictor models assume that there is an unobserved latent response variable which generates an observed binary variable (Ravallion 1996). However, the levels regression has a major shortcoming in that it imposes constant parameters over the entire distribution and thus assumes that the impact of various factors on welfare is constant across the expenditure distribution. That is, it assumes that there is no difference between the rich and the poor in terms of their characteristics. In reality, the poor face different constraints such access to credit and services. As Grootaert (1997) argues, the poor's ability to cope with these constraints can be envisaged as a latent variable which is a function of household characteristics which generates binary welfare outcomes. In this study, we use both approaches to check the robustness of our results to the poverty definition.

Fertility (Z_i) is defined as equal to one if a household has more than two children and zero if it has two[6]. Our study is essentially about large families versus small families. It should also be pointed out that two girls first would be more evident in the birth of the third child and not the second child since most families prefer having at least two children (Gupta and Dubey 2006). This implies that the two girls first IV only works in the transition from the second to the third child. We later (see subsection 4.5.8) change this definition of

[6] One can also quite plausibly assume that there is a latent variable which depends on personal and household characteristics which generates binary fertility outcomes i.e large family versus small family.

fertility, as a way of checking the sensitivity of our results to the definition of family size. The poverty and fertility equations are jointly estimated in a recursive bivariate probit which is formally specified below.

Consider the following levels regression;

$$Y_i = \beta' X_i + \delta Z_i + \varepsilon_i \tag{4.1}$$

then poverty status is defined as;

$$T_i = \begin{cases} 1 & if\ Y_i \leq Y^{PL} \text{ (poor)} \\ 0 & if\ Y_i > Y^{PL} \text{ (non poor)} \end{cases} \tag{4.2}$$

Consider the following levels regression for number of children (C_i);

$$C_i = \theta' X_i + \lambda M_i + \eta_i \tag{4.3}$$

then fertility is defined as;

$$Z_i = \begin{cases} 1 & \text{if } C_i > 2 \\ 0 & \text{if } C_i = 2 \end{cases} \tag{4.4}$$

The recursive bivariate probit is therefore defined as;

$$\Pr(T_i = 1, Z_i = 1 | X_i, Z_i, M_i) = \Phi_{i2}\left(Y^{PL} - (\beta' X_i + \delta Z_i), (\theta' X_i + \lambda M_i) - 2, \rho\right) \tag{4.5}$$

Where, $\Phi_{i2}(\cdots)$ is the bivariate normal cumulative density function, X_i is a vector of exogenous variables which influence both fertility and poverty, M_i is a zero-one dummy IV defined as equal to one if the first two children are girls and zero otherwise. β and θ are vectors of parameters to be estimated, and δ and λ are scalar parameters of the fertility dummy and the IV respectively. ε_i and η_i error terms with the following properties;

$$\rho = Cov\left(\varepsilon_i \eta_i\right) \tag{4.6}$$

$$E(\varepsilon_i | X_i, Z_i, M_i) = E(\eta_i | X_i, M_i) = 0 \tag{4.7}$$

$$Var(\varepsilon_i | X_i, Z_i, M_i) = Var(\eta_i | X_i, M_i) = 1 \tag{4.8}$$

The parameters β, θ, δ, λ, ρ are estimated by maximum likelihood (see Maddala 1983; Greene 2003; and Monfardini and Radice 2008 for more details).

The log likelihood to be maximized is[7];

$$
\begin{aligned}
L(\beta,\theta,\delta,\lambda,\rho) &= \sum \left[d_{11}\ln P_i^{11} + d_{10}\ln P_i^{10} + d_{01}\ln P_i^{01} + d_{00}\ln P_i^{00}\right] \qquad (4.9)\\
\text{where} &: \\
d_{11} &= T_iZ_i, d_{10} = T_i(1-Z_i), d_{01} = Z_i(1-T_i), d_{00} = (1-Z_i)(1-T_i)\\
P_i^{11} &= \Pr(T_i=1, Z_i=1|X_i,Z_i,M_i) = \Phi_{i2}(\beta' X_i + \delta, \theta' X_i + \lambda, \rho)\\
P_i^{10} &= \Pr(T_i=1, Z_i=0|X_i,Z_i,M_i) = \Phi_{i2}(-\beta' X_i - \delta, \theta' X_i + \lambda, -\rho)\\
P_i^{01} &= \Pr(T_i=0, Z_i=1|X_i,Z_i,M_i) = \Phi_{i2}(\beta' X_i + \delta, -\theta' X_i - \lambda, -\rho)\\
P_i^{00} &= \Pr(T_i=0, Z_i=0|X_i,Z_i,M_i) = \Phi_{i2}(-\beta' X_i - \delta, -\theta' X_i - \lambda, \rho)
\end{aligned}
$$

Testing the null that $\rho = 0$ using a Wald test amounts to testing for the exogeneity of fertility. The specified recursive bivariate probit corrects for simultaneity (through the IV) and at the same time controls for unobserved heterogeneity (by allowing correlation between the errors which capture unobserved factors among other things). Our two equation system is identified by way of exclusion restriction i.e. the poverty equation does not have M_i the IV as a regressor[8].

The coefficients in any limited dependent variable model can be misleading. Since the model is a probability model, the absolute level of a coefficient can convey a wrong picture of the impact of a regressor on the dependent variable. To overcome this problem, we compute marginal effects on the conditional mean function given by;

$$
\begin{aligned}
E\left[T_i|X_i,Z_i,M_i\right] &= \Pr\left[Z_i=1\right]E\left[T_i|Z_i=1,X_i,Z_i,M_i\right]\\
&\quad + \Pr\left[Z_i=0\right]E\left[T_i|Z_i=0,X_i,Z_i,M_i\right] \qquad (4.10)\\
&= \Pr(T_i=1,Z_i=1|X_i,Z_i,M_i) + \Pr(T_i=1,Z_i=0|X_i,Z_i,M_i)\\
&= \Phi_{i2}(\beta' X_i + \delta, \theta' X_i + \lambda, \rho) + \Phi_{i2}(-\beta' X_i - \delta, \theta' X_i + \lambda, -\rho)\\
&= P_i^{11} + P_i^{10}
\end{aligned}
$$

[7]For ease of exposition and in keeping with Maddala (1983), Greene (2003), and Monfardini and Radice (2008), we express the log likelihood assuming that the poverty and fertility thresholds are at zero. This simplification does not affect our analysis.

[8]It should however be pointed out that theoretically it is possible to achieve identification by functional form only i.e. without exclusion restrictions. This type of identification depends entirely on the bivariate normality of the errors. The exclusion restrictions help in making results robust to distributional misspecification (Monfardini and Radice 2008). Further, in our case the instrument allows us to check the robustness of our probit results to assuming that poverty is continuous.

The marginal effects are just the derivatives of this conditional mean function[9]. For variables which appear in both the fertility and poverty equations, the total marginal effect of these variables is decomposed into the direct effect (derivative of the second part of equation 4.10) and the indirect effect (derivative of the first part of equation 4.10). This indirect effect works through fertility. For example, education may affect poverty directly, but may also affect poverty indirectly through its impact on fertility. For binary explanatory variables, we do not take derivatives of equation 4.10 rather the marginal effect is just the difference in the conditional mean function with the dummy set equal to one and zero (Greene 2003). The marginal effect of fertility on poverty is calculated as follows;

$$\Pr(T_i = 1, |Z_i = 1|X_i, M_i) - \Pr(T_i = 1, |Z_i = 0|X_i, M_i) \tag{4.11}$$

The marginal effects in the fertility equation are just the derivatives of the marginal distribution quite like in a univariate probit.

Subjective poverty

As said earlier, this study uses the ELQ method to measure subjective poverty. Using this method, one can model subjective poverty using an ordered probit model (see for example Ravallion and Lokshin 2002), where the rungs of the ladder represent ordered outcomes. Following Devereux *et al.* (2006) and Kalugina and Najman (2002), we define a subjective poverty dummy as follows; households are subjectively poor if they fall on the bottom two rungs of the ladder and non-poor if they fall on rungs 3 to 6. So the impact of fertility on subjective poverty is modelled using the recursive bivariate probit presented in the preceding section for objective poverty.

In addition to the variables already discussed, for both objective and subjective poverty we include variables to capture household demographics, education, employment, agriculture, religion, and community level characteristics. We also control for regional effects by including regional dummies.

4.4.3 Data and descriptives

The data for this analysis come from the Second Malawi Integrated Household Survey (IHS2)[10]. The survey collects demographic information which *inter alia* includes; age, sex, together with the relationship of each household member to the household head. This information allows us to identify children and their birth orders, which we then

[9] If $\rho = 0$ then the two parts of equation 4.10 reduce to a product of marginals (Greene 2003).

[10] More details about IHS2 data are discussed in Chapter 2, in section 2.4.4.

use to generate the two girls first IV. The survey also collects information on subjective assessment of well-being. Out of a total of 11280 households, we focus on 9827 rural households (87%) of the total, as it is the rural areas where two girls first may be more evident. Because the survey does not track children across households; we impose the following restriction on the rural sample. The sample is limited to mothers aged 20-40, whose oldest child was less than 17 years of age at the time of the survey. Since we are focusing on households with at least two children, we would not expect many women younger than age 20 to have two children. Besides, it is to be expected that a child over age 17 has moved to a different household[11]. We therefore have about 3400 rural households constituting the restricted sample.

Table 4.1, presents objective poverty rates for the three poverty lines[12]. The results indicate that for all rural households, 56.4% are poor with a corresponding restricted sample head count rate of 52.5%. Additionally, 24.4% and 21.8% of all rural and restricted rural households respectively are ultra poor. This means that about one in five of the rural population (restricted and unrestricted) live in dire poverty such that they cannot even afford to meet the minimum standard daily recommended food requirement. In terms of the World Bank poverty line of US$1 a day, 30.6% and 27.7% of all Malawians residing in rural areas and those in the restricted rural sample respectively live on less than a dollar a day. Using the three poverty lines, we also find that the poverty gap measures are similar for the two samples. For instance, using the poor poverty line we find that all rural households (restricted rural households) have a poverty gap of 19.3% (17.7%) suggesting that they on average subsist on 19.3% (17.7%) less than K16165. What is therefore emerging from the results is that even though the poverty rates for all rural households are consistently higher than those for the restricted sample, the difference is not very big. This would imply that the restricted sample that we are using for this study is a reasonable representation of all rural households. In Table 4.2, we present results of the relationship between poverty headcounts and fertility measured as number of children. We find that for all poverty lines the poverty headcount rate is increasing with the number of children. For example, using the poor poverty line we find that for the unrestricted (restricted) 47.3% (46.2%) of households with less than three children are poor; this is in contrast to a headcount rate of 71.6% (74.5%) for those households with more than six children. This suggests that poverty and fertility might be positively related. This

[11] Similar restrictions are used by Angrist and Evans (1998), and Gupta and Dubey (2006). We later relax these age restrictions in subsection 4.5.7, to see if our results are not affected by the possiblity of sample selection.

[12] The poverty indices are based on the Foster, Greer and Thorbecke (FGT) measure given by $P_\alpha(c_i, z) = \frac{1}{n}\sum_{i=1}^{q}\left(\frac{g_i}{z}\right)^\alpha$. Where c_i is consumption of household i , z is the poverty line, and $g_i = z - c_i$ is a consumption shortfall. q is the number of poor households, α is a measure of poverty aversion. For $\alpha = 0$ we have the headcount index, for $\alpha = 1$ we have the poverty gap index, for $\alpha = 2$ we have the poverty severity index.

pattern holds for both samples; we should also note that the head counts are not very different for the two samples implying that our restricted sample represents quite well the rural population.

We now turn to the descriptive analysis of subjective poverty. We find that 84.8% of all rural households consider themselves to be subjectively poor with a corresponding figure of 83.5% for the restricted sample. These rates are very high as compared to the objective rates given in Table 4.1. In Table 4.3, we check the relationship between subjective poverty headcounts and the number of children. The results show an opposite relationship to that found under the objective poverty analysis (Table 4.2). Where as before we found that the more the number of children the higher the poverty rate; the results here show that the more the number of children a household has the lower the subjective poverty. This suggests that there might be a negative relationship between subjective poverty and fertility.

Table 4.4 summarizes results of the relationship between objective poverty and subjective poverty[13]. The results suggest that the objectively and subjectively poor are not the same people. This is evidenced by the fact that the off-diagonals (unshaded cells) are nonzero. This indicates that the matching of households between the two definitions of poverty is weak. For example, using the poor poverty line, we find that of 1359 households who are subjectively non-poor only 987 households are non-poor in the objective sense[14]. The Cramer's V statistics test the null hypothesis of no association between the two measures. A Cramer's V statistic of close to 1 (0) indicates strong (weak) association. The values are between 0 and 1, implying that there is a relationship albeit weak between the two and this is confirmed by the likelihood test (probabilities of the chi-square are zero.).

In Table 4.5, we report results of the descriptive analysis of the explanatory variables used in the study. The average number of children is 2.9. About three quarters of households have more than two children. Households which have two girls first make about 19% of our sample. This would suggest weak evidence of non random sex targeting since you expect the proportion of households with two girls first to be 25%. Education levels are low as is indicated by very low averages of numbers of people both male and female with some education be it primary or secondary. The averages are less than one suggesting very low numbers of people with education. In terms of education of parents, we note that fathers have more education than mothers as we move up the education ladder. For instance,13.6% of fathers have secondary as their highest education level as compared to just about 5% for mothers. The labour force participation for fathers is higher than that of mothers with 23% of fathers working for a wage compared to 4% for mothers.

[13] Since the results for the restricted and whole samples are similar, we only report results for all rural households

[14] A similar mismatch is found by Ravallion and Lokshin (2002) for Russia.

The average for number of enterprises is very low indicating that very few households engage in non-agricultural income generating activities. About a quarter of our sampled households grow tobacco which is a cash crop. The results indicate that close to two thirds of households have no clinics in their communities; in addition 2% of the households live in trading centres suggesting that most households are not close to markets. In Table 4.5, we also show descriptive statistics for all rural households[15]. This is done in order to check the representativeness of the variables used in our regression analyses. The results indicate that the restricted sample is generally not very different from the sample of all rural households; suggesting that it is a realistic representation of rural households. For example, the average number of children for the two or more children sample is slightly higher than that of all households, 2.9 against 2.4 for all households. In terms of employment, we also note that the sample of two or more children households has somewhat higher labor force participation rates for both fathers and mothers. Looking at all rural households their educational measures are to some extent lower. The same pattern emerges for religion and community characteristics.

4.5 Econometric results

In this section, we present econometric results of the impact of fertility on objective and subjective poverty. We start with the presentation and discussion of results for objective poverty, and this is followed later by results for subjective poverty.

4.5.1 Impact of exogenous fertility on objective poverty

In this section results (Table 4.6) of naïve univariate probit regressions which assume that fertility is exogenous for the three poverty lines are presented and discussed. These results serve as our base for comparison with the scenario where we assume that fertility is endogenous. For all the three poverty regressions, the chi-square statistics show all variables included in the models are jointly significant.

Similar to the findings of Gupta and Dubey (2006), the univariate probit results suggest a positive and statistically significant effect of fertility on poverty. This implies that exogenous fertility increases the likelihood of being poor. The size of the effect ranges from 11% to 23%, and these values are economically substantial. This relationship is monotonic, as it holds for all the three poverty lines, suggesting that our results are robust to choice of a poverty line and that the first order dominance assumption is not violated. We also observe that the impact of exogenous fertility on poverty increases as the poverty line increases i.e. moving from ultra poor to poor. This might indicate

[15] These are households with at least one child.

that children become more expensive as your income increases as households opt for good quality children. As expected, the dependency ratio is significantly and positively related to poverty across the three poverty lines. An increase in the dependency ratio *ceteris paribus* increases the probability of falling into poverty by 4%, 5.7% and 11% for ultra poor, World Bank, and poor poverty lines respectively. These effects are quite large economically. Having secondary education for mothers and fathers significantly affects poverty. The relationship as expected is negative. Holding other things constant, when a father (mother) has secondary education it lowers the likelihood of being poor by 5% (7%), 8.8% (6.6%), and 18.9% (18.4%) for ultra poor, World Bank, and poor poverty lines respectively. All the other education variables have the expected negative sign but their impact is statistically insignificant. It should however be pointed out that the magnitudes of the education variables are economically quite significant.

With respect to employment, we find that for fathers being employed for a wage lowers the likelihood of being poor with economically significant effects ranging from 4% to 6%. However, for mothers being employed for a wage is not statistically significant, probably reflecting the very low labour force participation rates for mothers (see Table 4.5). In terms of the magnitude of the effect, we note that they are quite large with values quite similar to those for fathers. The higher the number of non agricultural enterprises a household has, the lower the chance of being poor. For instance, using the poor poverty line we note that *ceteris paribus* having more enterprises increases the probability of being non poor by about 10%. The effect is increasing on successive poverty lines suggesting that the effect is more pronounced as the level of consumption increases. Accessing loans has the expected negative and significant effect on poverty.

Growing tobacco which is a cash crop has the expected negative relationship with poverty. The impact is both statistically and economically significant, with the magnitude ranging from 1% to 9.6%. We notice however that for the lowest poverty line (ultra poor), growing tobacco is not statistically significant suggesting that tobacco growing has no statistically significant effect on poverty at the lower end of the income distribution even though the effect seems to be economically large (about 1%). The statistical insignificance perhaps reflects the fact that due to its high cost nature very few ultra poor households can grow tobacco. Unsurprisingly, land which is a productive resource, statistically significantly increases the chance of being non poor. The magnitude of the effects suggests that it is economically significant. Although the importance of livestock as a means of livelihood is falling in Malawi, the results suggest that holding other things constant, owning livestock increases the probability of being non poor by 3%, 4% and 6% for the ultra poor, World Bank, and poor poverty lines respectively. Having a clinic in a community lowers the probability of being poor by 3%, 5%, and 7% in the ultra poor, World Bank, and poor poverty lines respectively. These effects are substantial from an economic as well as a statistical viewpoint. As might be expected, the presence of a clinic would imply easily accessible medical attention which would among other things improve the productivity of people in the area.

Our discussion above has been based on the assumption that fertility is exogenous, but as discussed before fertility might be endogenous. In the next section, we address this issue of endogeneity of fertility.

4.5.2 Controlling for endogenous fertility

As discussed before, to account for endogeneity we need an instrumental variable. In our case we are using two girls first as our IV. Before we go ahead to use the IV we first check two things. Firstly, we test using a hazard model whether indeed two girls first exists in rural Malawi[16]. Secondly, we then check the validity of two girls first as an IV. We address each one of these issues in the next subsections.

Having two girls first and fertility in rural Malawi

Since the focus of this study is not on measuring two girls first, we will not be too detailed about the methodology (for details see Haughton and Haughton 1998). In order to test for evidence of two girls first, we need to first define what we mean by two girls first. There are basically two concepts of two girls first. The first is called lexicographic preferences; also referred to as the threshold, fixed minima, or target view of two girls first, this approach assumes that the i^{th} household desires S_i sons, regardless of the number of daughters which it will need to have to achieve this goal. In practice, the target is likely to vary over households, and it may vary within a household over time, either way it is an unobservable quantity. The second concept of two girls first is what is called sequential preference. This obtains when for any given number of sons and daughters, parents prefer an additional son to an additional daughter.

To measure lexicographic preferences you need families which have stopped child-bearing i.e. complete families (Haughton and Haughton 1998). To measure sequential preference you can use incomplete families. Since in the IHS2 data there is no distinction between complete and incomplete families, we use the concept of sequential preference to measure two girls first. Sequential preference can be measured by using a hazard model. The hazard model estimates the risk (hazard) of having another child at any point in time. For an accelerated failure time (AFT) model, if the hazard is higher for families with a son (or sons), the implication is that two girls first is present. The dependent variable is the length of the interval (in months) between one birth and the next, a by-product of

[16] This can be viewed as complementary to the evidence of son preference in education spending in rural Malawi found in Chapter 2.

recording the birth dates of the children born in the household. Specifically, we focus on two intervals namely the transition from the second to the third child, and the transition from the third to the fourth child. In addition to the covariates included in the other regressions, we use the number of existing boys. in a household

If two girls first is present, we would expect the coefficient for number of existing boys to be positive, implying that the higher the number of boys, the longer the duration between births. It is supposed that households that do not have as many sons as they wish, will hurry to have another before it is too late. Underlying this idea is the notion that households may have sequential two girls first. Results in Table 4.7 are based on the accelerated failure time Weibull hazard model. For the two transitions, the coefficient for number of boys is positive and significant suggesting the presence of two girls first in rural Malawi[17]. Among other variables, we controlled for the employment of the father, and for the interval 2 to 3 we find that fathers' employment increases the duration of the birth interval though this effect is insignificant on the next birth interval. Having found that two girls first is present in rural Malawi the next thing to be done is to check if it is a valid instrument.

Two girls first as an instrumental variable

For a variable to be a good IV, it firstly must be uncorrelated with the error term, which in our case means that it must not be correlated with poverty. Secondly, it must be correlated with the endogenous variable. The consistency of our results may be affected by the possibility that the IV may be correlated with the error term, that is it may be endogenous. There are two possible scenarios in which this can happen[18]. Firstly, there is a possibility of using ultrasound services to know in advance the sex of a child which the rich can access, which can then be used to do prenatal sex screening. This would make our IV correlated with economic status (poverty). It however has to be said that while this is possible in rich countries where medical services are very advanced, this cannot be the case in rural Malawi where medical facilities are quite basic. Besides, abortion including sex selective abortion is illegal in Malawi[19]. The second issue which can lead to endogeneity is what Rosenzweig and Wolpin (2000) call the hand-me-down effect. They argue that the cost of children depends on sex composition and show that there is strong

[17] To complement these results, we tested (using a t-test) whether or not there is a difference in the average number of children between households with two girls first and those with two boys first. We find that households with two girls first have a significantly higher number of children with a mean difference (standard error) of 0.3036 (0.054).

[18] There is a possibility that poor households may prefer sons to work in the fields, this would also mean that the instrument would be correlated with poverty. We control for this by including a variable which captures whether or not children work outside the home.

[19] There is a possibility of using more traditional sex-targeting mechanisms such as female infanticide and extreme neglect of female children leading to their eventual death (Sen 1984). However, there is no reported evidence of this in Malawi.

evidence for a hand-me-down effect. This is an economies of scale effect where if you have children of the same sex you spend less because there are some things like clothing which can be used by the child coming after. Now if households with children of the same sex spend significantly less money than do households with children of different sexes, this difference in consumption may affect the poverty situation of the household. In this case therefore the IV is endogenous. In Table 4.8, we report results of two sample t-tests of mean differences to check for evidence of the hand-me-down effect. If the hand-me-down effect is present, we would expect there to be a statistically significant difference in expenditure on clothing and education by sex of the child. That is, if the hand-me-down effect is present, the expenditure on the two items should be significantly lower for the case where two girls or two boys are first than the case where there is a mix of a boy and a girl. However, we do not find a statistically significant difference in expenditure on the two items between households with two girls first or two boys first and those with a girl and a boy. The implication of this finding is that two girls first is not endogenous through the hand-me-down effect. We then need to check the second condition that two girls first and fertility are correlated.

We check for the relationship between fertility and our IV by estimating a reduced form univariate probit model of fertility. The results are presented in Table 4.9. Column 1 leaves out religion, column 2 leaves out region but includes religion, and column 3 has all covariates. Most of the variables have the expected signs. A father's education is a strong predictor of fertility though the education of mother does not have a significant effect on fertility. This probably reflects the fact in rural households a father has a final say on everything including for example contraceptives. The more educated a father is, the more likely is the family going to adopt family planning. We find that if children work at home it leads to more fertility. Of particular interest is the relationship between the IV and fertility. Having two girls first significantly increases the probability of having more than two children. The relationship holds for all the three specifications presented in Table 4.9. This suggests that fertility and the IV are correlated[20]. It is worth noting that whether or not we control for religion and region, the effect of having two girls first on the probability of having more than two children is not affected by religion or regional effects.

So far we have found that two girls first exists in rural Malawi, and that two girls first is a good IV in the sense that it is uncorrelated with poverty and it is correlated with fertility. We now test whether fertility is endogenous. We present two complementary tests of the endogeneity of fertility. Since the confidence intervals do not contain a zero, the cross equation error correlation (rho)[21] results in Table 4.10 suggest that the null hypothesis of exogenous fertility is rejected at the 5% significance level for all poverty lines. This

[20]This can in a sense be viewed as direct evidence of son preference.

[21]A low error correlation coefficient (rho) suggests that there no correlation between the error terms of the poverty and fertility equations i.e. poverty and fertility are exogenous.

conclusion is further confirmed by the Wald test results presented in Table 4.11. As said before failing to account for endogeneity of fertility would lead to biased and inconsistent results. We therefore present results of a recursive bivariate probit which jointly estimates fertility and objective poverty.

4.5.3 Impact of endogenous fertility on objective poverty

In Tables 4.12-4.14, we report the marginal effects of the recursive bivariate probit of the impact of fertility on poverty for the three poverty thresholds. For all the three poverty lines, the chi-square statistics suggest that the variables are jointly significant. The maximum likelihood results indicate that fertility and poverty are positively related. The effect is statistically significant. This implies that fertility increases the probability of being poor. The impacts are economically significant with values ranging from 0.139 to 0.304. This relationship is monotonic, as it holds for the three poverty thresholds, suggesting that just like in the base scenario where we assumed fertility to be exogenous, our results are robust to choice of a poverty line. This means that the first order dominance assumption is not violated. Just like the naïve results of exogenous fertility, the impact of endogenous fertility across the poverty lines increases as the poverty line increases i.e. moving from ultra poor to poor. We note however that the total effect of fertility on poverty is larger than the one we got when we assumed that fertility is exogenous. This implies that assuming that fertility is exogenous underestimates its impact on poverty. For all poverty lines, the underestimation is about 1.3 times. It should be pointed out that the statistically significant effect of endogenous fertility is not in conformity with a finding by Gupta and Dubey (2006) for India. They find that the impact of fertility on poverty is statistically insignificant after controlling for endogeneity. The finding that treating fertility as an endogenous variable increases its effect on poverty merits further speculative comment. If endogeneity was arising from reverse causation, then the IV would purge the reverse causality problem making the relationship less positive i.e. the effect of endogenous fertility should be smaller than that of exogenous fertility. This means that our result does not conform to this expectation. However, if the endogeneity problem arises from omitted variables which influence both fertility and poverty, then there might be downward omitted variable bias, and thus by purging out the endogeneity problem we get rid of this bias.

As expected, the dependency ratio is significantly and positively related to poverty across the three poverty lines. The total effect is almost equal to that from the base regressions. The recursive bivariate probit results show that having secondary education for mothers and fathers are statistically significant predictors of poverty. The relationship as expected is negative. In terms of the size of the impact, we note that in the case of the ultra poor model, for a father (mother) having secondary education *ceteris paribus* lowers the probability of being poor by 5.7% (6.6%), with the effect of a mother's education being slightly higher. As was the case with the univariate probit, most of the education variables have the expected negative sign but they are not statistically significant though they appear economically large. With respect to employment, we find that for a father all things being equal, being employed for a wage lowers the likelihood of being poor. The effect is almost the same as that for the base regressions for all poverty thresholds. However, as before wage employment for mothers has no statistically significant effect. The effect seems to be economically significant though, with the magnitudes of the effects similar to those for fathers. Similar to the results from the simple probits, we find that for the lowest poverty line (ultra poor), growing tobacco has no statistically significant effect though with a marginal effect of about 1% it would suggest that the effect is economically sizeable. Land and livestock increase the probability of being non poor. Having a clinic in a community lowers the probability of being poor by 7% in the poor poverty line model, and this effect is substantial from an economic as well as a statistical perspective.

Two things are coming out of our comparison between the base scenario regressions which assumed that fertility is exogenous and the recursive bivariate probit results. Firstly, all the variables which were significant in the simple probit regressions are also significant after accounting for endogeneity. Secondly, the total effects for the joint estimation of fertility and poverty are generally larger than those for the univariate probit regressions.

So far our analysis has been based on real per capita annualized consumption expenditure. This analysis does not take into account household composition and economies of scales. In the next section, we investigate whether or not the impact of fertility on objective poverty that we have found is robust to accounting for household composition and economies of scale.

4.5.4 Household composition and economies of scale

The use of per capita consumption expenditure is common in poverty studies; however this procedure has two problems. First, different individuals have different needs. For example, a young child typically requires less food than an adult. Second, there are economies of scale in consumption for such items as housing, kitchen utensils, and utilities such as electricity. It costs less to house two people than to house two individuals separately. Larger households can do bulk buying which can attract discounts. Some studies have shown that the impact of household size on poverty disappears once these two problems are addressed (e.g. Lanjouw and Ravallion 1995; White and Masset 2003). The solution to these problems is to use adult equivalent scales[22]. An adult equivalent scale measures the

[22] The implication of using per capita consumption for poverty analysis is that households with children are judged poorer on a per capita basis than they would be if their welfare level was measured on an adult equivalent basis. Besides, using the per capita measures overestimates the impact of number of children

number of adult males (typically) to which that household is deemed to be equivalent[23]. In this study, we use the arbitrary method to measure equivalence scales; in the literature there are different methods for measuring equivalence scale, none of them commands universal assent (see Deaton and Zaidi 2002). The number of adult equivalents (AE) is defined as follows:

$$AE = (A + \pi K)^{\kappa} \tag{4.12}$$

where A is the number of adults in the household, and K is the of number of children, parameter π is the cost of a child relative to that of an adult, and lies between 0 and 1. The parameter κ which also lies between 0 and 1, controls the extent of economies of scale; since the elasticity of adult equivalents with respect to "effective" size, $A + \pi K$ is κ, $(1 - \kappa)$ is a measure of economies of scale. When both π and κ are unity (the most extreme case with no discount for children or for size) the number of adult equivalents is simply household size, and deflation by household size is equivalent to deflating to a per capita basis. If κ is zero, then economies of scale are so extreme that welfare is the same for different households with the same total consumption expenditure regardless of household size. The choice of the values of the parameters π and κ is arbitrary, we use the following values $\pi = 0.65$ and $\kappa = 0.9$. For the cost of children parameter, our value is based on the one used for Zambia by the World Bank (2005). Being neighbors, we would expect the Zambian figures to be similar to Malawi's. With respect to the economies of scale parameter, our choice is motivated by the fact that in most developing countries food is major component of consumption. Food is largely a private good and therefore there are no economies of scale with food. This implies that a high value of κ should be used. The annualized real consumption expenditure for each household is divided by the adult equivalent (AE) to have consumption per adult equivalent. With this adjustment a household is considered poor if its annualized real consumption per adult equivalent is below the three poverty lines discussed before.

The previous regressions were re-estimated in order to check the robustness of our findings to accounting for household composition and economies of scale. We present results (Table 4.15) for both univariate probit which assumes exogenous fertility as well as the recursive bivariate probit. We have replicated the previous per capita results for comparison. For the univariate probit regression, the results indicate that the variables are jointly significant. The simple probit results indicate that when we account for economies of scale and household composition, fertility significantly increases the likelihood of being poor across the poverty lines. This effect is monotonic as before implying our results

on poverty.

[23] In keeping with other studies (e.g Lanjouw and Ravallion 1995; White and Masset 2003), in this study we do not make a gender distinction.

are robust to choice of poverty line, and that the first order dominance assumption is not violated. As was the case with the per capita poverty regressions, we find that for the adjusted regressions the impact of fertility across the poverty lines increases as the poverty line increases i.e. moving from ultra poor to poor. However, as expected adult equivalent scale adjustment reduces the impact of fertility on poverty. The reductions are economically substantial. For the ultra poor poverty line, the reduction is about 83.4%, for the World Bank US$1 line the reduction is 77.6%, and finally for the poor poverty line the reduction is 37.8%. This implies that the higher the consumption the lower the reduction in the impact of fertility after adult equivalent adjustments.

For the bivariate probit regressions a similar pattern emerges[24]. After accounting for the endogeneity of fertility as well as economies of scale and household composition, fertility significantly increases the likelihood of being poor across the poverty lines. Besides, the impact is not as economically significant as that for per capita models, as it ranges from 0.012 to 0.177, compared against a range of 0.139 to 0.304 for the unadjusted models. Our results are robust to choice of poverty line and the impact of fertility across the poverty lines increases as the poverty line increases i.e. moving from ultra poor to poor. Compared with the results from the simple probit models which account for economies of scale and household composition, we note that the impact of fertility on poverty is underestimated in the simple probit models. However, compared with the per capita bivariate probit results, the results show that the impact is reduced. For the ultra poor model the reduction is 91.4%, for the World Bank US$1 line the reduction is 77.8%, and finally for the poor poverty line the reduction is 35%. These are economically significant reductions. We also notice that these reductions are not very different from those found for univariate probit models. These findings suggest that it is quite possible that the impact of fertility would be economically insignificant with some values of the equivalent scale parameters[25].

The conclusion from these results is that accounting for economies of scale and household composition reduces the impact of fertility on poverty, and that these reductions are economically large. However, the impact of fertility is still statistically significant regardless of whether or not fertility is exogenous or endogenous. This far we have looked at the impact of fertility on poverty with poverty defined as a dummy. In the next section, we investigate the robustness of our results to treating objective poverty as a continuous variable.

[24] For the bivariate probit regression with adult equivalent adjustment we also find that fertility is endogenous, with the following Wald statistics (p values) for the three poverty lines; Ultrapoor 8.1933 (0.0042), World Bank 4.2567 (0.0391), Poor 4.392 (0.0361).

[25] Another way of accounting for economies of scale is to directly include household size and the square of household size in the poverty equation, however this approach ignores the interactions that may exist between economies of scale and other variables included in the model.

4.5.5 Impact of fertility on continuous objective poverty

With poverty treated as continuous, we estimated a Two Stage least Squares (2SLS) regression of the impact of fertility on poverty measured as the log of real annualized per capita and adult equivalent adjusted consumption expenditures[26]. For the 2SLS regression all right hand variables for the two regressions remain the same as in the bivariate probit[27]. In this framework, we also estimated a naïve regression which assumes that fertility is exogenous. This is done by using Ordinary least Squares (OLS). All right hand variables are the same as those for the univariate probit regressions. The results are presented in Table 4.16. The results indicate that for the OLS regression, fertility is negatively related to both per capita and adult equivalent adjusted consumption. This implies that having more than two children lowers consumption and hence increases poverty. This is similar to the finding earlier where poverty is defined as a dummy. The impact of fertility on poverty is lower when we account for economies of scale and household composition. Again this is similar to our earlier findings. The regression based Hausman test for endogeneity (see Woodridge 2002 for details) shows that fertility is endogenous. This implies that our OLS results may be biased and inconsistent.

The 2SLS results, which account for this endogeneity show that as is the case with the OLS results, fertility is significantly negatively related to both per capita and adult equivalent adjusted consumption. However, the impact of fertility on consumption is higher when we account for endogeneity of fertility. For example, the OLS results of the per capita regression underestimate the effect of fertility by about 2.3 times. The finding that accounting for endogeneity raises the impact of fertility on poverty is similar to the one before where poverty is defined as a dummy. Additionally, the impact of endogenous fertility is reduced when we account for economies of scale and household composition. The reduction after accounting for endogeneity of fertility (2SLS) is about 20%. We also note that this reduction though economically large is smaller than the reductions found for poverty defined as a dummy.

To conclude, these results suggest that our earlier findings are robust to a different conceptualization of objective poverty. Specifically, with objective poverty defined as a continuous variable; fertility increases the likelihood of being poor, that this effect is underestimated when the joint determination of the two is not accounted for, and that accounting for household composition and economies of scale diminishes the effect. Our analysis so far has looked at poverty in the objective and monetary sense which is a narrower definition of poverty. In the next section, we present econometric results of the impact of fertility

[26]The F-statistic of the first stage regression on the excluded instrument (girlIV) is 247.18 with a p-value of 0.000. The partial R^2 of the excluded instrument is 0.0376. Together these statistics suggests our IV is not weak.

[27]The coefficient for the instrument is .172 with a p-value of 0.000 suggesting that if the first two children are girls significantly increases the number of children.

on subjective poverty which is based on self reported well being.

4.5.6 Impact of fertility on subjective poverty

As before, we test for evidence of endogeneity between fertility and subjective poverty[28]. We find no evidence that fertility is endogenous with respect to subjective poverty. This is in stark contrast to the objective poverty analysis where we find that fertility is endogenous. We therefore present results in Table 4.17 of a univariate probit regression since fertility is exogenous.

Fertility is found to significantly lower the likelihood that a household will be subjectively poor. We find that fertility lowers the probability of being poor by about 3%. This result is however different from the objective poverty analysis where we find consistently that fertility increases the probability of being poor. This perhaps reflects rural Malawi's social cultural context where those households with more children are treated with respect and those with few or without children are looked at with some contempt. That is, having more children elevates your status in society and these intangible benefits feed into peoples' sense of wellbeing. Another possible explanation is that there is discounting taking place in the sense that households with more children expect to have a higher future discounted income and therefore feel less poor[29]. This finding can also be explained in terms of utility in the sense that couples get additional utility when they have additional children. The results also suggest that the higher the dependency ratio, the lower the subjective poverty, again we found an opposite effect for monetary poverty. In terms of the magnitudes, we note that the probability of feeling poor is lowered by 2.5%. The fact that having more dependents makes households feel less poor can also be explained by the cultural context that the more people depend on you the higher will be your social status. This intangible benefit is reflected in lower subjective poverty.

Interestingly, for all the other variables the results are similar in terms of the signs and statistical significance to those for monetary poverty. For example; education of the father, number of enterprises, loans, growing tobacco, land, and ownership of livestock lower the likelihood of being subjectively poor. We included a dummy variable marital status to capture some of the characteristics of the household head. We have three classes; monogamous (mono), polygamous (poly) and the base category is those who are not married i.e. widowed, divorced, or separated[30]. The results indicate that being married lowers the probability of feeling poor. We further note that the decrease in the likelihood of

[28] We estimated a recursive bivariate probit of subjective poverty and fertility, the Chi-square value (P-value) of the Wald test for exogeneity of fertility is 0.17 (0.6766).

[29] We would like to thank Erik Thorbecke for pointing out this possible explanation.

[30] In our sample 64% are monogamous, 9.6% are polygamous, and 26.4% are either widowed or divorced or separated.

feeling poor is higher for polygamous households than it is for monogamous households. Specifically, relative to being widowed, divorced, or separated, being polygamous lowers the probability of being subjectively poor by 7.6% as against 5.9% for monogamous households. The same cultural explanation can be given here where being married raises your status, and having more wives further increases the respect that people may give you. The level of per capita annualized real consumption is also included to capture household income status. We find that household economic status as measured by per capita consumption expenditure lowers the probability of feeling poor by about 10%.

4.5.7 Sample selection bias

As discussed earlier, our results are based on a restricted sample of women aged between 20 and 40, and the oldest child is under 17. This is motivated by the fact we need to have households which still have at least the first two children at home. However, this restriction may lead to a selected sample i.e. a non-random sample. Sample selection may bias our results. Sample selection may arise from; a) the possibility that older children may still be at home, b) the fact that women in developing countries tend to marry at a very young age, and c) the possibility that some women may start bearing children much later in life. To check the extent to which the restriction affects our results, we re-estimated the previous regressions with the mother's age relaxed to between 17 and 50, and the oldest child to under 20. With this relaxation, the sample size increases to 4572 rural households.

The results are similar to those found before, thus giving us confidence that our conclusions are invariant to the age restrictions. For example (compare with Table 4.15), when objective poverty is defined as a dummy and we control for household composition and economies of scale, the marginal effects (standard errors) of fertility for the recursive bivariate probit are 0.014 (0.003), 0.035 (0.006) and 0.148 (0.021) for ultra poor, World Bank U$S1 and poor poverty lines respectively. The same picture emerges when objective poverty is defined as a continuous variable (compare with Table 4.16), with 2SLS coefficients (standard errors) of fertility being -0.426(0.126) and -0.325 (0.116) for per capita and adult equivalent adjusted consumption respectively[31]. Similarly, the relaxation does not change our conclusions regarding the impact of fertility on subjective poverty. The marginal effect and standard error of fertility on subjective poverty (compare with Table 4.17) are -0.022 and 0.002 respectively.

[31] The results are also similar for exogenous fertility on objective poverty.

4.5.8 Two child versus one child families

In keeping with the literature (e.g. Angrist and Evans 1998; Gupta and Dubey 2006), our analysis has been based on the restriction of our sample to families with at least two children. We have argued that two girls first which we use as our IV, would be more evident in the birth of the third child and not the second child, since most families prefer having at least two children. With this restriction, the corresponding two girls first IV (two girls first) only works in causing exogenous variation in the transition from the second to the third child. However, with this restriction, the higher-order birth IV may be correlated with poverty. It may potentially be correlated with poverty in the sense that poorly nourished women may have difficulty conceiving three times. This possible correlation may make the IV invalid, and thus making our results inconsistent. To check if our results are affected by this potential problem, we re-estimated the previous regressions using a sample of families with at least one child. For these new regressions, fertility (Z_i) is re-defined as equal to one if a household has more than one child and zero if it has one[32]. The corresponding IV (M_i), is re-defined as a dummy equal to one if the first child is a girl, and zero if the first child is a boy[33]. With this relaxation, the sample size increases to 6595 rural households.

The pattern of results is generally similar to those found before albeit with higher marginal effects[34]. For example (compare with Table 4.15), when objective poverty is defined as a dummy and we control for household composition and economies of scale, the marginal effects (standard errors) of fertility for the recursive bivariate probit are 0.032 (0.002), 0.044 (0.003) and 0.203 (0.001) for ultra poor, World Bank U$S1 and poor poverty lines respectively. A similar conclusion is arrived at when objective poverty is defined as a continuous variable (compare with Table 4.16), with 2SLS coefficients (standard errors) of fertility being -0.612(0.014) and -0.511 (0.033) for per capita and adult equivalent adjusted consumption respectively. Interestingly, with the relaxation we find that the marginal effect of fertility on subjective poverty is lower. The marginal effect and standard error of fertility on subjective poverty (compare with Table 4.17) are -0.014 and 0.004 respectively. This suggests that while having a child makes a household feel less poor, higher-order births have an even bigger effect on subjective poverty. Since the general findings are not different from those found using a sample of families with at least two children, this reassures us that our conclusions are robust to the definition of family size.

[32] See equation 4.4, for comparison.

[33] See equation 4.3, for comparison.

[34] The mother's age was restricted to between 15 and 40, and the oldest child to under 17. It should be pointed out that alternative restrictions did not affect our conclusions.

4.6 Conclusions

In the chapter, we sought to find the impact of fertility on poverty while recognizing the fact that the two variables are jointly determined. The study uses data from the Second Malawi Integrated Household Survey (IHS2). By using a natural experiment, two girls first as our instrumental variable, we are able to use exogenous variation in number of children to uncover the causal effect of fertility on poverty of rural households in Malawi. First, we have looked at poverty defined in the monetary sense. A menu of three poverty lines has been used to check the sensitivity of our results to the choice of a poverty line. Results from the naïve probit models show that fertility increases the likelihood of being poor. Since fertility is found to be endogenous, we estimated a recursive bivariate probit where two girls first is used as an IV. For the bivariate probit models, it has been found that fertility increases the likelihood of being poor as well. However, this effect is larger for endogenous fertility, implying that when fertility is treated as exogenous its effect on poverty is underestimated. For both the base scenario of exogenous fertility and that of endogenous fertility, its impact has been found to be robust to choice of poverty line. The positive impact of fertility on objective poverty has also been shown to hold when household composition and economies are accounted for, though the effect tends to be reduced. It has also been demostrated that when objective poverty is conceptualized as a continuous variable this does not change our finding that fertility increases poverty and that its effect is higher when fertility is endogenous.

Second, we have looked at poverty defined more broadly by using self rated assessments of welfare. It has been shown that subjective poverty and objective poverty are related albeit weakly. Interestingly, fertility has been found to be exogenous with respect to subjective poverty, probably suggesting that the endogeneity of fertility is a monetary phenomenon. In terms of its impact on subjective poverty, it has been found to have the opposite effect to that found under objective poverty. That is having more than two children lowers the probability of feeling poor, probably reflecting the fact that having more children elevates your status in society and these intangible benefits feed into peoples' sense of wellbeing. This contradiction in the impact of fertility on the narrower objective poverty and the broader subjective poverty might be a possible explanation for why families in rural Malawi have many children (in spite of this making them poor in the objective monetary sense) as it makes them feel less poor.

Though the study is able to estimate a causal relationship between fertility and poverty, it is worth pointing out that the study is static in nature and therefore cannot capture dynamic aspects of the relationship between poverty and fertility. It should also be said that modernity can invalidate our instrumental variable in the sense that richer households may be less biased against female children for reasons related to modernity and modernity (which is unobserved) can influence both poverty as well as their gender preferences. Since we can not control for modernity, our conclusions from the chapter should be taken with that caveat in mind.

Table 4.1: Poverty lines and associated poverty rates

Poverty line name	Poverty index	Poverty line per year	Poverty measure	
			Restricted	All
Ultra poor	Headcount		21.8%	24.4%
	Poverty gap	MK10029	5.1%	5.8%
	Poverty severity		1.7%	2.0%
Poor	Headcount		52.5%	56.4%
	Poverty gap	MK16165	17.7%	19.3%
	Poverty severity		7.8%	8.7%
World Bank (US$1)	Headcount		27.7%	30.6%
	Poverty gap	MK11051	6.9%	7.9%
	Poverty severity		2.5%	2.9%

Notes: The ultra poor poverty line is a food poverty line. The poverty lines are expressed in Malawi Kwacha (MK).

Table 4.2: Poverty headcount and fertility

Poverty line name	Number of Children	Poverty Headcount	
		Restricted	All
Ultrapoor	Less than three	15.9%	17.5%
	Between three and six	33.1%	33.9%
	Greater than six	45.3%	42.8%
Poor	Less than three	46.2%	47.3%
	Between three and six	69.1%	69.8%
	Greater than six	74.5%	71.6%
World Bank (US$1)	Less than three	21%	22.8%
	Between three and six	40.8%	41.2%
	Greater than six	52.3%	48.8%

Notes: The ultra poor poverty line is a food poverty line.

Table 4.3: Subjective poverty and number of children

	Subjectively poor headcount	
Number of children	**Restricted**	**All**
Less than three	83.5%	85.38%
Between three and six	84.6%	85.12%
Greater than six	76%	76.51%

Table 4.4: Objective poverty and subjective poverty

Absolute poverty line		Subjective poverty: **Non poor**	**Poor**	**Total**
Ultrapoor	Non-Poor	1,261	6,759	8,020
	Poor	98	1,709	1,807
	Total	1,359	8,468	9,827
	Cramér's V = 0.1156 Chisquare= 157.9 Prob> Chisquare = 0.000			
Poor	Non-Poor	987	4,195	5,182
	Poor	372	4,273	4,645
	Total	1,359	8,468	9,827
	Cramér's V = 0.1596 Chisquare= 260 Prob> Chisquare = 0.000			
World Bank	Non-poor	1,224	6,283	7,507
	Poor	135	2,185	2,320
	Total	1,359	8,468	9,827
	Cramér's V = 0.1290 Chisquare= 191.5 Prob> Chisquare = 0.000			

Table 4.5: Sample statistics

	Restricted		All households	
Variable	**Mean**	**Standard Deviation**	**Mean**	**Standard Deviation**
Demographics				
number of children	2.910627	.0365515	2.423051	.0286636
two children or more	.7608118	.0082506	.6108187	.0066297
mother's age	29.10582	.1096724	37.50181	.240289
dependency ratio	1.361335	.0169433	1.187387	.0141034
two girls first	.188402	.0071974	.1373361	.00457
age at first birth	19.40541	3.838322	19.85218	3.949439
Education				
females with primary	.0983405	.0067846	.1014171	.0059358
males with primary	.1592433	.0108257	.1680409	.0084645
females with JCE	.0558154	.0046985	.0536465	.0037974
males with JCE	.1136867	.0065776	.1101352	.0050674
males with MSCE	.0438612	.0042456	.0372241	.0031862
father no education	.751661	.0023005	.807705	.0034122
mother no education	.891231	.0017032	.927293	.0062371
father primary	.1127936	.0063001	.0954657	.0041973
mother primary	.0589645	.0041884	.0425612	.0027699
father secondary	.1355454	.0073334	.096829	.0050333
mother secondary	.0498045	.0044107	.0300455	.0025036
Employment				
father works	.2328373	.010074	.1859403	.0076554
mother works	.0414332	.0042292	.0309512	.003083
children work at home	.504807	.010293	.5184155	.0073577
children work outside home	.0446999	.0041152	.0746205	.0039465
number of enterprises	.4597906	.016602	.4164337	.0131016
Agriculture				
loan	1050.168	98.66299	1635.488	582.397
grows tobacco	.2657778	.0119243	.2382278	.0097572
land	0.58458	2.652096	0.64926	1.424733
livestock	6.184791	.02887696	6.380213	.0250001
Religion				
muslim	.125529	.0077292	.1220689	.0066006
catholic	.257815	.0109174	.2680185	.0093857
protestant	.6340125	.0120363	.6403704	.0100509
Community				
has clinic	.2998732	.0222924	.2924817	.021051
lives in trading centre	.0224902	.0064503	.024335	.0067625
Region				
north	.1271692	.0045091	.1283621	.0034756
centre	.4070733	.0083056	.4092272	.0059133
Sample size	3402		6595	

Notes: Restricted rural households are those which; have a mother aged between 20 and 40, the oldest child under 17, and have at least two biological children. All households are rural households with at least one child.

Table 4.6: Marginal effects of the impact of exogenous fertility on poverty

Variable	Ultrapoor	Worldbank	Poor
Demographics			
two children or more	0.109***	0.134***	0.225***
	(0.013)	(0.016)	(0.027)
mother's age	0.004***	0.004**	0.006***
	(0.001)	(0.001)	(0.002)
dependency ratio	0.040***	0.057***	0.109***
	(0.007)	(0.009)	(0.016)
age at first birth	-0.010***	-0.010***	-0.017***
	(0.002)	(0.002)	(0.003)
Education			
females with primary	-0.028	-0.042	-0.123***
	(0.026)	(0.033)	(0.046)
males with primary	-0.013	-0.013	-0.013
	(0.019)	(0.023)	(0.038)
females with JCE	-0.017	-0.023	0.043
	(0.034)	(0.039)	(0.060)
males with JCE	0.041	0.033	0.055
	(0.026)	(0.033)	(0.051)
males with MSCE	-0.047	-0.049	-0.117
	(0.045)	(0.051)	(0.075)
father primary	-0.006	-0.008	-0.056
	(0.025)	(0.031)	(0.047)
mother primary	-0.010	0.017	0.084
	(0.035)	(0.048)	(0.065)
father secondary	-0.053**	-0.088***	-0.189***
	(0.024)	(0.028)	(0.053)
mother secondary	-0.070***	-0.066*	-0.184***
	(0.025)	(0.038)	(0.062)
Employment			
father works	-0.038***	-0.038**	-0.060**
	(0.013)	(0.016)	(0.024)
mother works	-0.034	-0.039	-0.012
	(0.025)	(0.031)	(0.047)
children work at home	0.008	0.026	0.020
	(0.014)	(0.017)	(0.025)
children work outside home	0.020	0.046	0.071
	(0.028)	(0.036)	(0.051)
number of enterprises	-0.020*	-0.033**	-0.095***
	(0.011)	(0.013)	(0.019)
Agriculture			
loan	-0.031**	-0.041*	-0.049**
	(0.002)	(0.021)	(0.024)
grows tobacco	-0.012	-0.036**	-0.096***
	(0.014)	(0.016)	(0.023)
land	-0.021***	-0.028***	-0.034**
	(0.004)	(0.003)	(0.017)
livestock	-0.032***	-0.040***	-0.062***
	(0.004)	(0.005)	(0.007)
Religion			
muslim	0.031	0.023	0.035
	(0.026)	(0.030)	(0.042)
catholic	0.029	0.032	0.018
	(0.019)	(0.024)	(0.033)
protestant	0.036**	0.021	0.001
	(0.017)	(0.022)	(0.032)

Table 4.6: continued

Variable	Ultrapoor	Worldbank	Poor
Community			
has clinic	-0.033***	-0.048***	-0.071***
	(0.012)	(0.014)	(0.021)
lives in trading centre	-0.028	-0.042	-0.093
	(0.032)	(0.038)	(0.062)
Region			
north	0.016	0.010	0.045
	(0.017)	(0.020)	(0.031)
centre	-0.105***	-0.144***	-0.190***
	(0.012)	(0.015)	(0.022)
Loglikelihood	-1274.66	-1470.5	-1904.2
Chisquare	511.47	624.9	892.5
Prob > Chisquare	0.00	0.00	0.00
Sample size	3402	3402	3402
McFadden R2	0.167	0.175	0.19

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The dependent variable is a poverty indicator based on annualized per capita real consumption expenditure. The coefficients are marginal effects evaluated as partial changes at the mean value of the continuous covariates. For dummy covariates, the partial changes are measured as a discrete change in the poverty indicator as the dummy covariate changes from 0 to 1. Numbers in parentheses are standard errors.

Table 4.7: Accelerated failure time Weibull model

Variable	Mean	Hazard ratio	Transition 2 to 3	Hazard ratio	Transition 3 to 4
number of boys	1.94	1.12	0.114*** (0.018)	1.11	0.113*** (0.023)
father works	.2328	2.71	0.100** (0.046)	0.92	-0.073 (0.062)
children work at home	.5048	0.90	-0.103** (0.048)	0.74	-0.299*** (0.081)
children work outside home	.0446	0.77	-0.259*** (0.074)	1.16	0.156* (0.081)
F-statistic			4.79		5.50
Prob > F-statistic			0.00		0.00
Sample size			2720		1651

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The dependent variables are birth intervals moving from 2 to 3 children, and moving from 3 to 4 children. In addition to the new variable number of boys, the models also include all the other covariates included in the previous models. The hazard ratio is an exponentiated coefficient. Numbers in parentheses are standard errors.

Table 4.8: Son preference and the hand-me-down effect

Variable	Sample mean (μ)	Mean difference ($\mu - \mu_{mix}$)
Two girls first		
Education	1458.875 (264.87)	322.929 (197.309) [0.1017]
Clothing	4431.883 (7022.024)	324.835 (199.979) [0.1043]
Two boys first		
Education	1199.512 (70.881)	63.566 (100.73) [0.5280]
Clothing	4173.29 (91.23)	66.249 (118.75) [0.5769]

Notes: Mean differences are defined as the sample means of rural households which have two girls first (two boys first) (μ) minus the sample of rural households which have a mix in the first two children i.e. boy and a girl (μ_{mix}). The means are weighted. Numbers in parentheses are standard errors. Numbers in square brackets are p-values. The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01.

Table 4.9: Marginal effects of reduced form univariate probit regressions of fertility

Variable	(1)	(2)	(3)
Demographics			
mother's age	0.001*	0.001*	0.001*
	(0.001)	(0.001)	(0.001)
dependency ratio	0.148***	0.149***	0.148***
	(0.023)	(0.023)	(0.023)
two girls first	0.102***	0.103***	0.102***
	(0.020)	(0.020)	(0.020)
age at first birth	-0.019***	-0.019***	-0.019***
	(0.004)	(0.004)	(0.004)
Education			
females with primary	-0.003	-0.004	-0.003
	(0.019)	(0.019)	(0.018)
males with primary	0.027	0.025	0.025
	(0.016)	(0.017)	(0.016)
females with JCE	0.020	0.023	0.019
	(0.025)	(0.024)	(0.024)
males with JCE	0.074***	0.074***	0.073***
	(0.020)	(0.020)	(0.020)
males with MSCE	0.084***	0.085***	0.083***
	(0.028)	(0.028)	(0.028)
father primary	-0.029	-0.029	-0.024
	(0.027)	(0.027)	(0.026)
mother primary	-0.005	-0.007	-0.007
	(0.025)	(0.025)	(0.026)
father secondary	-0.152***	-0.151***	-0.146***
	(0.052)	(0.052)	(0.051)
mother secondary	-0.044	-0.052	-0.046
	(0.043)	(0.044)	(0.043)
Employment			
father works	0.007	0.006	0.007
	(0.008)	(0.008)	(0.008)
mother works	0.017	0.018	0.017
	(0.011)	(0.011)	(0.011)
children work at home	0.103***	0.103***	0.104***
	(0.026)	(0.025)	(0.026)
children work outside home	0.026	0.028*	0.027*
	(0.016)	(0.016)	(0.016)
number of enterprises	-0.001	-0.001	-0.001
	(0.006)	(0.006)	(0.006)
Agriculture			
loan	0.003	0.005	0.002
	(0.006)	(0.004)	(0.002)
grows tobacco	0.011	0.016**	0.010
	(0.007)	(0.008)	(0.007)
land	-0.002	-0.001	-0.003
	(0.003)	(0.002)	(0.003)
livestock	-0.005*	-0.004	-0.005*
	(0.003)	(0.003)	(0.003)

Table 4.9: continued

Variable	(1)	(2)	(3)
Religion			
muslim		0.011	0.016
		(0.011)	(0.011)
catholic		0.001	-0.001
		(0.010)	(0.010)
protestant		0.016	0.015
		(0.011)	(0.011)
Community			
has clinic	0.007	0.008	0.007
	(0.007)	(0.007)	(0.007)
lives in trading centre	-0.001	-0.001	-0.001
	(0.022)	(0.022)	(0.021)
Region			
north	-0.004		-0.002
	(0.012)		(0.012)
centre	0.020**		0.022**
	(0.008)		(0.009)
Loglikelihood	-867.28	-870.62	-864.89
Chisquare	2006.61	2001.93	2013.39
Prob > Chisquare	0.00	0.00	0.00
Sample size	3402	3402	3402
McFadden R2	0.536	0.535	0.538

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The dependent variable is a dummy for more than two children. The coefficients are marginal effects evaluated as partial changes at the mean value of the continuous covariates. For dummy covariates, the partial changes are measured as a discrete change in the poverty indicator as the dummy covariate changes from 0 to 1. Numbers in parentheses are standard errors.

Table 4.10: Cross equation error correlation

Name of poverty line	Rho	95% Confidence Interval	
Ultra poor	-.2431234	-.3882092	-.0862774
Poor	-.1865539	-.3436128	-.0193431
World Bank (US$1)	-.177207	-.326869	-.0188744

Table 4.11: Wald test for exogeneity of fertility

Name of poverty line	Chi square	Prob. > Chi square
Ultra poor	9.05406	0.0026
Poor	4.76884	0.0290
World Bank (US$1)	4.79995	0.0285

Table 4.12: Impact of endogenous fertility on poverty (ultra poor)

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Demographics				
two children or more	0.139***		0.139***	
	(0.014)		(0.014)	
two girls first		0.011**	0.011**	0.105***
		(0.005)	(0.005)	(0.020)
mother's age	0.003***	0.001	0.004***	0.001
	(0.001)	(0.001)	(0.001)	(0.001)
dependency ratio	0.057***	-0.028***	0.043***	0.151***
	(0.008)	(0.006)	(0.007)	(0.023)
age at first birth	-0.011***	0.003***	-0.009***	-0.020***
	0.002)	(0.001)	(0.002)	(0.005)
Education				
females with primary	-0.027	-0.002	-0.028	-0.002
	(0.023)	(0.004)	(0.024)	(0.017)
males with primary	-0.010	-0.007*	-0.013	0.028*
	(0.015)	(0.004)	(0.016)	(0.016)
females with JCE	-0.014	-0.006	-0.017	0.020
	(0.031)	(0.006)	(0.033)	(0.024)
males with JCE	0.047**	-0.012**	0.042*	0.074***
	(0.024)	(0.005)	(0.025)	(0.021)
males with MSCE	-0.032	-0.022***	-0.043	0.085***
father primary -0.009	0.005	-0.006	-0.030	
	(0.020)	(0.006)	(0.022)	(0.028)
mother primary	-0.007	0.001	-0.007	-0.007
	(0.029)	(0.006)	(0.031)	(0.025)
father secondary	-0.063***	0.014	-0.057**	-0.146***
	(0.020)	(0.009)	(0.023)	(0.054)
mother secondary	-0.065***	-0.001	-0.066***	-0.048
	(0.022)	(0.006)	(0.024)	(0.043)
Employment				
father works	-0.035***	-0.005**	-0.037***	0.007
	(0.012)	(0.002)	(0.013)	(0.008)
mother works	-0.031	-0.006**	-0.034	0.019*
	(0.022)	(0.003)	(0.023)	(0.011)
children work at home	0.018	-0.021***	0.007	0.108***
	(0.014)	(0.006)	(0.014)	(0.025)
children work outside home	0.027	-0.006	0.024	0.030*
	(0.026)	(0.004)	(0.027)	(0.016)
number of enterprises	-0.019*	-0.002	-0.020*	-0.001
	(0.011)	(0.001)	(0.011)	(0.006)
Agriculture				
loan	-0.021***	-0.032***	-0.053***	0.001
	(0.001)	(0.002)	(0.002)	(0.001)
grows tobacco	-0.010	-0.003	-0.012	0.010
	(0.014)	(0.002)	(0.015)	(0.007)
land	-0.011***	-0.003	-0.012**	-0.001
	(0.001)	(0.003)	(0.002)	(0.001)
livestock	-0.030***	-0.002**	-0.031***	-0.005
	(0.004)	(0.001)	(0.005)	(0.003)

Table 4.12: Continued

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Religion				
muslim	0.031	-0.002	0.031	0.017
	(0.028)	(0.003)	(0.029)	(0.012)
catholic	0.025	0.002	0.027	0.000
	(0.019)	(0.003)	(0.020)	(0.011)
protestant	0.035**	-0.000	0.035**	0.016
	(0.015)	(0.003)	(0.016)	(0.012)
Community				
has clinic	-0.029**	-0.004*	-0.031**	0.006
	(0.014)	(0.002)	(0.014)	(0.007)
lives in trading centre	-0.027	-0.002	-0.029	-0.000
	(0.032)	(0.004)	(0.033)	(0.019)
Region				
north	0.012	0.001	0.013	-0.001
	(0.022)	(0.004)	(0.023)	(0.013)
centre	-0.093***	-0.013***	-0.100***	0.023**
	(0.015)	(0.004)	(0.015)	(0.009)
Chisquare	725.23			
Prob > Chisquare	0.00			
Sample size	3402			

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The coefficients are marginal effects. For variables which appear in both the poverty and fertility equations, the marginal effects are decomposed into two effects. The direct effect produced by its presence in the poverty equation, and an indirect effect which works through the fertility equation. The sum of the two makes the total effect. The total effect may not exactly equal the sum of the two effects due to rounding. Numbers in parentheses are standard errors. The poverty equation is based on annualized per capita real consumption expenditure.

Table 4.13: Impact of endogenous fertility on poverty (World Bank)

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Demographics				
two children or more	0.167***		0.167***	
	(0.020)		(0.020)	
two girls first		0.010*	0.010*	0.106***
		(0.005)	(0.005)	(0.020)
mother's age	0.003**	-0.002	0.003**	0.001
	(0.001)	(0.002)	(0.001)	(0.001)
dependency ratio	0.080***	-0.033***	0.059***	0.150***
	(0.010)	(0.007)	(0.010)	(0.023)
age at first birth	-0.013***	0.004***	-0.010***	-0.020***
	(0.002)	(0.001)	(0.002)	(0.005)
Education				
females with primary	-0.040	-0.002	-0.042	-0.003
	(0.034)	(0.005)	(0.035)	(0.017)
males with primary	-0.008	-0.008*	-0.013	0.027
	(0.017)	(0.004)	(0.018)	(0.016)
females with JCE	-0.018	-0.007	-0.023	0.020
	(0.036)	(0.007)	(0.038)	(.024)
males with JCE	0.043	-0.016***	0.033	0.075***
	(0.031)	(0.006)	(0.032)	(0.021)
males with MSCE	-0.030	-0.024***	-0.046	0.086***
	(0.049)	(0.009)	(0.051)	(0.029)
father primary	-0.012	0.006	-0.008	-0.028
	(0.025)	(0.007)	(0.027)	(0.028)
mother primary	0.017	0.003	0.019	-0.007
	(0.046)	(0.007)	(0.048)	(0.024)
father secondary	-0.099***	0.015	-0.091***	-0.147***
	(0.024)	(0.010)	(0.028)	(0.054)
mother secondary	-0.064*	0.003	-0.062*	-0.046
	(0.035)	(0.009)	(0.038)	(0.043)
Employment				
father works	-0.035**	-0.004*	-0.038**	0.007
	(0.016)	(0.002)	(0.016)	(0.008)
mother works	-0.036	-0.006**	-0.040	0.018
	(0.027)	(0.003)	(0.028)	(0.011)
children work at home	0.039**	-0.024***	0.024	0.108***
	(0.017)	(0.006)	(0.017)	(0.025)
children work outside home	0.053	-0.006	0.050	0.029*
	(0.034)	(0.005)	(0.035)	(0.016)
number of enterprises	-0.031**	-0.002	-0.033**	-0.001
	(0.013)	(0.002)	(0.013)	(0.006)
Agriculture				
loan	-0.026***	-0.022**	-0.046**	0.002
	(0.002)	(0.005)	(0.001)	(0.002)
grows tobacco	-0.032*	-0.005**	-0.035*	0.010
	(0.017)	(0.002)	(0.018)	(0.007)
land	-0.002*	-0.002	-0.004	-0.002
livestock	-0.038***	-0.001*	-0.039***	-0.005*
	(0.005)	(0.001)	(0.006)	(0.003)

Table 4.13: Continued

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Religion				
muslim	0.025	-0.003	0.023	0.017
	(0.033)	(0.004)	(0.034)	(0.012)
catholic	0.029	0.002	0.030	-0.000
	(0.023)	(0.003)	(0.024)	(0.011)
protestant	0.022	-0.002	0.021	0.015
	(0.021)	(0.003)	(0.022)	(0.012)
Community				
has clinic	-0.043**	-0.005*	-0.046**	0.007
	(0.018)	(0.002)	(0.019)	(0.007)
lives in trading centre	-0.040	-0.003	-0.042	-0.001
	(0.037)	(0.005)	(0.038)	(0.020
Region				
north	0.007	0.001	0.008	-0.001
	(0.027)	(0.004)	(0.029)	(0.013)
centre	-0.130***	-0.015***	-0.141***	0.024**
	(0.019)	(0.005)	(0.019)	(0.009)
Chisquare	828.37			
Prob > Chisquare	0.00			
Sample size	3402			

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The coefficients are marginal effects. For variables which appear in both the poverty and fertility equations, the marginal effects are decomposed into two effects. The direct effect produced by its presence in the poverty equation, and an indirect effect which works through the fertility equation. The sum of the two makes the total effect. The total effect may not exactly equal the sum of the two effects due to rounding. Numbers in parentheses are standard errors. The poverty equation is based on annualized per capita real consumption expenditure.

Table 4.14: Impact of endogenous fertility on poverty (poor)

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Demographics				
two children or more	0.304***		0.304***	
	(0.040)		(0.040)	
two girls first		0.015**	0.015**	0.106***
		(0.008)	(0.008)	(0.020)
mother's age	0.006***	-0.003	0.006***	0.001
	(0.002)	(0.002)	(0.002)	(0.001)
dependency ratio	0.169***	-0.081***	0.109***	0.148***
	(0.020)	(0.016)	(0.017)	(0.023)
age at first birth	-0.024***	0.010***	-0.016***	-0.019***
	(0.004)	(0.003)	(0.003)	(0.005)
Education				
females with primary	-0.119***	-0.005	-0.124***	-0.002
	(0.042)	(0.010)	(0.044)	(0.017)
males with primary	-0.001	-0.017	-0.014	0.028
	(0.038)	(0.011)	(0.041)	(0.017)
females with JCE	0.048	-0.011	0.041	0.022
	(0.055)	(0.014)	(0.059)	(0.024)
males with JCE	0.084*	-0.041***	0.054	0.075***
	(0.047)	(0.014)	(0.050)	(0.021)
males with MSCE	-0.074	-0.055***	-0.118	0.086***
	(0.070)	(0.019)	(0.074)	(0.029)
father primary	-0.064	0.013	-0.055	-0.029
	(0.043)	(0.016)	(0.046)	(0.028)
mother primary	0.079	0.009	0.088	-0.008
	(0.061)	(0.016)	(0.064)	(0.025)
father secondary	-0.228***	0.049*	-0.195***	-0.148***
	(0.043)	(0.025)	(0.050)	(0.054)
mother secondary	-0.185***	0.009	-0.181***	-0.047
	(0.057)	(0.018)	(0.063)	(0.042)
Employment				
father works	-0.055**	-0.007	-0.061**	0.007
	(0.024)	(0.004)	(0.025)	(0.007)
mother works	-0.008	-0.010	-0.016	0.016
	(0.046)	(0.007)	(0.047)	(0.011)
children work at home	0.062**	-0.060***	0.016	0.107***
	(0.028)	(0.015)	(0.027)	(0.025)
children work outside home	0.088*	-0.015	0.077	0.029*
	(0.049)	(0.011)	(0.050)	(0.016)
number of enterprises	-0.091***	-0.004	-0.095***	-0.001
	(0.020)	(0.004)	(0.020)	(0.006)
Agriculture				
loan	-0.028***	-0.001	-0.029***	0.003
	(0.002)	(0.004)	(0.005)	(0.007)
grows tobacco	-0.087***	-0.010**	-0.096***	0.011
	(0.028)	(0.005)	(0.029)	(0.007)
land	-0.002*	0.002	-0.004	-0.001
	(0.001)	(0.002)	(0.004)	(0.002)
livestock	-0.060***	-0.004	-0.061***	-0.005*
	(0.007)	(0.002)	(0.007)	(0.003)

Table 4.14: Continued

	Poverty Equation			Fertility Equation
Variable	**Direct effect**	**Indirect effect**	**Total effect**	**Total effect**
Religion				
muslim	0.039	-0.008	0.033	0.016
	(0.044)	(0.007)	(0.044)	(0.012)
catholic	0.015	0.001	0.016	-0.001
	(0.033)	(0.006)	(0.035)	(0.011)
protestant	0.006	-0.008	-0.001	0.014
	(0.031)	(0.007)	(0.033)	(0.012)
Community				
has clinic	-0.064**	-0.007	-0.070**	0.007
	(0.027)	(0.004)	(0.028)	(0.007)
lives in trading centre	-0.089	-0.005	-0.094	0.001
	(0.085)	(0.010)	(0.088)	(0.019)
Region				
north	0.040	0.002	0.042	-0.000
	(0.045)	(0.008)	(0.046)	(0.013)
centre	-0.171***	-0.022***	-0.190***	0.024**
	(0.029)	(0.007)	(0.029)	(0.009)
Chisquare	1076.35			
Prob > Chisquare	0.00			
Sample size	3402			

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The coefficients are marginal effects. For variables which appear in both the poverty and fertility equations, the marginal effects are decomposed into two effects. The direct effect produced by its presence in the poverty equation, and an indirect effect which works through the fertility equation. The sum of the two makes the total effect. The total effect may not exactly equal the sum of the two effects due to rounding. Numbers in parentheses are standard errors. The poverty equation is based on annualized per capita real consumption expenditure.

Table 4.15: Accounting for household composition and economies of scale (marginal effects)

	Ultrapoor		Worldbank		Poor	
Variable	**per capita**	**AES**	**per capita**	**AES**	**per capita**	**AES**
UNIVARIATE PROBIT						
two children or more	0.109***	0.018***	0.134***	0.030***	0.225***	0.147***
	(0.013)	(0.004)	(0.016)	(0.005)	(0.027)	(0.015)
All covariates	Yes	Yes	Yes	Yes	Yes	Yes
Chisquare	511.47	208.69	624.9	263.63	892.5	542.73
Prob > Chisquare	0.00	0.00	0.00	0.00	0.00	0.00
Sample size	3402	3402	3402	3402	3402	0.00
McFadden R2	0.167	0.18	0.175	0.176	0.19	0.153
BIVARIATE PROBIT						
two children or more	0.139***	0.026***	0.167***	0.037***	0.304***	0.177***
	(0.014)	(0.002)	(0.020)	(0.006)	(0.040)	(0.020)
Instrument	Yes	Yes	Yes	Yes	Yes	Yes
All covariates	Yes	Yes	Yes	Yes	Yes	Yes
Chisquare	725.23	4792.6	828.37	665.49	1076.35	779.68
Prob > Chisquare	0.00	0.00	0.00	0.00	0.00	0.00
Sample size	3402	3402	3402	3402	3402	3402

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors. The per capita poverty equations are based on annualized per capita real consumption expenditure, and AES poverty equations are based on annualized real consumption expenditure per adult equivalent scale and economies of scale. For the poverty equations in the bivariate probit we report the total marginal effects only. For brevity total marginal effects of the fertility equation for the bivariate probit are not reported. The per capita results are replicated from earlier regressions for comparison.

Table 4.16: OLS and 2SLS results of continuous fertility and poverty

	OLS		2SLS	
Variable	**per capita**	**AES**	**per capita**	**AES**
two children or more	-0.298***	-0.172***	-0.568***	-0.456***
	(0.026)	(0.026)	(0.124)	(0.122)
Instrument	-	-	Yes	Yes
All covariates	Yes	Yes	Yes	Yes
Hausman test	-	-	0.281**	0.295**
			(0.125)	(0.123)
Mean of dep variable	9.627	9.99	9.627	9.99
	(.578)	(.547)	(.578)	(.547)
F-stat	59.75	43.45	53.86	40.18
Prob> F-stat	0.00	0.00	0.00	0.00
Sample size	3402	3402	3402	3402

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. Numbers in parentheses are standard errors. The dependent variables for the per capita models are log of the annualized per capita real consumption expenditure. The dependent variables for the AES regressions are the log of the annualized real consumption expenditure per adult equivalent and economies of scale. The Hausman test is a regression based test of endogeneity of fertility.

Table 4.17: Impact of fertility on subjective poverty

Variable	
Demographics	
two children or more	-0.025**
	(0.012)
dependency ratio	-0.016***
	(0.005)
monogamous	-0.059***
	(0.011)
polygamous	-0.076***
	(0.022)
Education	
females with primary	-0.046***
	(0.014)
males with primary	-0.008
	(0.013)
females with JCE	-0.055***
	(0.019)
males with JCE	-0.012
	(0.014)
males with MSCE	-0.034
	(0.022)
father primary	-0.041*
	(0.022)
mother primary	0.038**
	(0.018)
father secondary	-0.134***
	(0.029)
mother secondary	-0.012
	(0.029)
Employment	
father works	-0.010
	(0.012)
mother works	-0.040
	(0.027)
children work at home	-0.048***
	(0.011)
children work outside home	0.042***
	(0.015)
number of enterprises	-0.008
	(0.005)
Agriculture	
loan	-0.002*
	(0.001)
grows tobacco	-0.040***
	(0.011)
land	-0.032***
	(0.003)
livestock	-0.024***
	(0.003)
Consumption	
Consumption expenditure	-0.099***
	(0.009)

Table 4.17: Continued

Variable	
Religion	
muslim	-0.087***
	(0.023)
catholic	-0.003
	(0.012)
protestant	0.001
	(0.012)
Community	
has clinic	-0.011
	(0.010)
lives in trading centre	0.045**
	(0.018)
Region	
north	-0.101***
	(0.017)
centre	-0.039***
	(0.011)
Chisquare	857.003
Prob > Chisquare	0.000
Sample size	3402
McFadden R^2	0.159

Notes: The significance asterisks are defined as follows: * p<0.10, ** p<0.05, *** p<0.01. The dependent variable is a subjective poverty indicator based on the Economic Ladder Question (ELQ. The coefficients are marginal effects evaluated as partial changes at the mean value of the continuous covariates. For dummy covariates, the partial changes are measured as a discrete change in the poverty indicator as the dummy covariate changes from 0 to 1. Numbers in parentheses are standard errors.

Chapter 5

Conclusion

Recognizing the important role that education plays in economic growth and the fight against poverty, the thesis has looked at three interrelated themes concerning household schooling investments in children and the impact of fertility on poverty in Malawi. We have used data from the Second Malawi Integrated Household Survey (IHS2) to conduct the empirical analyses. We have looked at factors which affect spending on primary education in both rural and urban Malawi as well as what explains the gap in spending in education between the two areas of residence. Mindful of the crucial role that parents play in the human capital formation of non-biological children in Africa, we have investigated the issue of schooling bias that parents may exhibit against non-biological children. We have also recognized that parents do not just make decisions on education investments of own and non-biological children, they also make choices on the size of the family to have. The thesis has made contributions on these rather intertwined issues.

Looking back, some of the take home messages from the thesis include the finding that parental characteristics have a bigger impact on spending in rural areas, and that a mother's characteristics have a larger impact on spending compared to a father's. We have argued that this finding has two important implications. First, if one thinks of the employment status and education of the mother as a reflection of the bargaining power of the mother in the household, this would imply that children's education benefits from an improvement in the bargaining position of the mother. Secondly, this result has intergenerational implications for human capital formation in that more female education entails more educated mothers, and hence more education for children. Besides, better education and employment for mothers has also been shown to be one of the significant factors in redressing the rural-urban spending gap. Improvements in household economic status have been shown to play an important role of increasing spending on education and reducing the rural-urban spending gap. We have also shown that as the economic status of a household gets better, intrahousehold schooling bias against non-biological children declines. So, improvements in the economic condition of a household not only have a positive effect on the education investments that parents make on their own children, but also lead to a reduction in schooling bias against non-biological children. There is however a catch in efforts to improve the welfare of a household and its human capital formation, in the sense that the number of children that a household has negatively impacts on a household's economic condition, but improvements in the economic status of a household lead to improvements in human capital accumulation. Thus, human capital formation at the household level would benefit from small family sizes in so far as small family sizes are good for a household's economic condition.

Critically, the finding that households feel less poor if they have more children suggests an apparent conflict in efforts to increase human capital formation, which as it has been argued is necessary for economic growth to take place. It further throws some light on why many households in Malawi have many children in spite of the negative economic consequences that this may have on their objective poverty as well as the education investments on children.

References

Alderman, H., Behrman, J., Khan, S. and Sabot, R. (1996), "Decomposing the regional gap in cognitive skills in rural Pakistan," *Journal of Asian Economics*, Vol. 7, pp. 49-76.

Alderman H, and Gertler P. (1997), "Family resources and gender differences in human capital investments: the demand for children's medical care in Pakistan," In: Haddad L, Hoddinott J, Alderman H (eds). *Intrahousehold resource allocation in developing countries: models, methods, and policy.* Baltimore: Johns Hopkins University Press for the International Food Policy Research Institute.

Alderman H and King E.M (1998), "Gender differences in parental investment in education," *Structural Change and Economic Dynamics,* Vol. 9, pp. 453-468.

Al-Samarrai, S., Reilly, B., (2000), "Urban and rural differences in primary school attendance: An empirical study for Tanzania," *Journal of African Economies*, Vol. 9, pp. 430-474.

Angrist, J. D. and Evans, W.N., (1998), "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review,* Vol. 88, pp. 450-477.

Appleton, S., Hoddinott, J., and Mackinnon, J. (1996), "Education and health in Sub-Saharan Africa," *Journal of International Development*, Vol. 8, pp. 307-339.

Appleton S ,Hoddinott J and Krishnan P (1999), "The Gender Wage Gap in Three African Countries," *Economic Development and Cultural Change*, Vol. 47, pp. 289-312.

Arnott, R.J. and Gersovitz, M.(1986), "Social welfare underpinnings of urban bias and unemployment," *Economics Journal*, Vol. 96, pp. 413-424.

Bauer T K, and Sinning M (2005), "Blinder-Oaxaca Decomposition for Tobit Models," RWI Discussion Papers, No 32 pp. 1-10.

Bauer T K,and Sinning M.(2008), "An extension of the Blinder–Oaxaca decomposition to nonlinear models," *AstA Advances in Statistical Analysis*, Vol. 92, pp. 197-451.

Becker, G, (1974), "A theory of social interactions," *Journal of Political Economy*, Vol. 82, pp. 1063-1093.

Becker, G. and Lewis H.G., (1973), "On the Interaction between the Quantity and Quality of Children," *Journal of Political Economy*, Vol. 81, pp. 279-288.

Becker, G. S., and Tomes, N. (1976), "Child endowments and the quantity and quality of children," *Journal of Political Economy*, Vol. 84, pp. 143-162.

Becker, G, (1981),*Treatise on the Family* Cambridge, MA: Harvard University Press.

Becker, G. S., and Tomes, N. (1986), " Human capital and the rise and fall of families," *Journal of Labor Economics*, Vol. 4, pp. 1-39.

Becker, G. (1996), *Accounting for Tastes.* Cambridge and London: Harvard University Press.

Behrman, J. R., Pollak, R. A., Taubman, P. (1986), "Do parents favor boys?" *International Economic Review*, Vol. 27, pp. 33-54.

Behrman, J. R., Khan, S., Ross D. and Sabot, R. (1997),"School quality and cognitive achievement production: a case study for rural Pakistan," *Economics of Education Review*, Vol. 16, pp. 127-142.

Behrman J.R. and J.C. Knowles (1999), "Household Income and Child Schooling in Vietnam," *World Bank Economic Review*, Vol. 13, pp. 211-256.

Beneito P, Ferri J, Molto M.L, and Uriel E (2001), "Determinants of the demand for education in Spain," *Applied Economics*, Vol. 33, pp. 1541-1551.

Bielicki, T. (1986), "Physical Growth as a Measure of Economic Well-being of Populations: The Twentieth Century," In *Human Growth: A Comprehensive Treatise*, 2nd ed. (F. Falkner and J. Tanner, Ms.). New York: Plenum, Vol. 3.

Birdsall, N. (1977), "Analytical Approaches to the Relationship of Population Growth and Development," *Population and Development Review,* Vol. 3, pp. 63-102.

Birdsall, N., (1980), "Population and Poverty in the Developing World," World Bank Staff Working Paper, 404, Washington D.C.: The World Bank.

Birdsall, N. and C. Griffin, (1988), "Fertility and poverty in developing countries, " *Journal of Policy Modeling,* Vol. 10, pp. 29-55.

Bishai, D., Suliman E.D. , H. Brahmbhatt F, Wabwire-Mangen G, Kigozi, N. Sewankambo, D. Serwaddaa, and Gray R (2003), "Does biological relatedness affect child survival?" Demographic Research, Vol 8(9), [on-line] available at www.demographic-research.org/volumes/vol8/9/8-9.pdf.

Blinder, A. S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, Vol. 8, pp. 436-455.

Blundell, R. W. and Meghir, C. (1987), "Bivariate alternatives to the univariate Tobit model," *Journal of Econometrics*, Vol. 34, pp. 179-200.

Bokosi F.K (2006), "Household Poverty Dynamics in Malawi: A bivariate probit analysis," *Journal of applied sciences*, Vol. 7, pp. 258-262.

Bronars, S.G., and Grogger, J., (1994), "The Economic Consequences of Unwed Motherhood: Using Twins as a Natural Experiment," *American Economic Review,* Vol. 84, pp. 1141-1156.

Case A, Deaton (1999), "A School Inputs and Educational Outcomes in South Africa," *The Quarterly Journal of Economics*, Vol. 114, pp. 1047-1084.

Case, A ., Lin I-F., and McLanahan S. (2000), "How hungry is the selfish gene?" *Economic Journal,* Vol. 110, pp. 781-804.

Case, A., Paxson, C., and Ableidinger, J. (2004), "Orphans in Africa: Parental death, poverty and school enrollment," *Demography,* Vol 41, pp. 483-508.

Casterline, J., Cooksey, E., and Ismail, A. (1987), "Household Income and Survival in Egypt," Population, Health, and Nutritional Technical Note 87- 12, The World Bank, Washington,D.C.

Chernichovsky, D., 1984, "Socioeconomic correlates of fertility behavior in rural Botswana," *Genus*, Vol. 40 , pp. 129-146.

Chun H and Oh J. (2002), "An instrumental variable estimate of the effect of fertility on the labour force participation of married women," *Applied Economics Letters,* Vol. 9, pp. 631- 634.

Cigno, A. (1993), "Intergenerational Transfers without Altruism, Family, Market and State," *European Journal of Political Economy*, Vol. 9, pp. 505-518.

Cohen, B. and W. House, (1994), "Demographic behavior and poverty: micro-level evidence from southern Sudan," *World Development*, Vol. 22, pp. 1031-1044.

Cotton, J, (1988), "On the decomposition of wage differentials," *Review of Economics and Statistics*, Vol. 70, pp. 236-243.

Cragg, J. G. (1971), "Some statistical models for limited dependent variables with applications to the demand for durable goods," *Econometrica*, Vol. 39, pp. 829-44.

Daunfeldt S, Hellström J (2007), "Intra-household Allocation of Time to Household Production Activities:Evidence from Swedish Household Data," *Labour,* Vol. 21, pp. 189-207.

Davies J.B , and Zhang J. (1995), "Gender Bias, Investments in Children, and Bequests," *International Economic Review*, Vol. 36, pp. 795-818.

Deaton A and Zaidi S (2002), "Guidelines for Constructing Consumption Aggregates For Welfare Analysis," Living Standards Measurement Study Working Paper No 135 Washington, D.C: The World Bank.

Deaton, A (1989), "Looking for Boy-Girl Discrimination in Household Expenditure Data," *World Bank Economic Review*, Vol. 3, pp. 1-15.

Devereux S, Baulch B, Phiri A, Sabates-Wheeler R (2006), *Vulnerability to Chronic Poverty and Malnutrition in Malawi: A report for DFID Malawi*, [online] available at www.aec.msu.edu/fs2/mgt/caadp/malawi_pva_draft_052606_final_draft.pdf

Doss C.R (1996), "Testing among Models of Intrahousehold Resource Allocation," *World Development*, Vol. 24, pp. 1597-1609.

Echevarria, C. and Merlo, A (1999), "Gender Differences in Education in Dynamic Household Bargaining Model," *International Economic Review*, Vol 40, pp. 265-286.

Fairlie, R. W. (1999), "The Absence of the African-American Owned Business:An Analysis of the Dynamics of Self-Employment," *Journal of Labor Economics*, Vol. 17, pp. 80-108.

Fairlie, R. W. (2005), "An extension of the Blinder-Oaxaca decomposition technique to logit and probit models," *Journal of Economic and Social Measurement*, Vol. 30, pp. 305-316.

Foster, J., Greer, J. and Thorbecke, E., (1984), "A Class of Decomposable Poverty measures," *Econometrica*, Vol. 52, pp. 761-766.

Gachuhi, D. (1999),"The impact of HIV/AIDS on education systems in the eastern and southern Africa regions and the response of education systems to HIV/AIDS: Life skills programmes," UNICEF.

Gertler P. and P. Glewwe (1990), "The Willingness to Pay for Education in Developing Countries," *Journal of Public Economics*, Vol. 42, pp. 251-275.

Gertler, P. J. and Glewwe P (1992) "The Willingness to Pay for Education for Daughters in Contrast to Sons: Evidence from Rural Peru," *The World Bank Economic Review,* Vol 6, pp. 171-188.

Glewwe P. and H. Jacoby. (1994), "Student Achievement and Schooling in low income countries:Evidence from Ghana," *The Journal of Human Resources*, Vol. 29, pp. 843-864.

Glewwe P and Ilias N (1996), "The determinants of school attainment in sub-Saharan Africa: A case study of Ghana," *Journal of International Development*, Vol. 8, pp. 395-413.

Glick, P. and Sahn, D. (2000), "Schooling of girls and boys in a West African country: the effects of parental education,income, and household structure," *Economics of Education Review*, Vol. 19, pp. 63-87.

Godlonton S and Keswell M (2005), "The impact of health on poverty: evidence from the south African: integrated family survey," *South African Journal of Economics,* Vol 73, pp. 133-148.

Green W.H (1981), " On the asymptotic bias of the ordinary least squares estimator of the Tobit model," *Econometrica,* Vol. 49, pp. 505-513.

Greene, W.H. (2003), *Econometric Analysis*, Upper Saddle River, NJ: Prentice Hall.

Grootaert, C. (1997), "The Determinants of Poverty in Cote d'lvoire in the 1980s," *Journal of African Economies*, Vol. 6, pp. 169-196.

Gundersen, C, Kelly,T, and Jemison ,K. (2004), "The Demand for Education for orphans in Zimbabwe," African Development and Poverty Linkage: The macro-Micro linkage, Forum paper.

Gupta, S., Davoodi, H., and Alonso-Terme, R. (1998), "Does corruption affect income inequality and poverty?" International Monetary Fund Working Paper WP/98/76, Washington DC.

Gupta N .D and Dubey A (2006), "Fertility and the Household's Economic Status: A Natural Experiment using Indian Micro Data," *Journal of Development Studies*, Vol. 42, pp. 110-138.

Haddad, L and Hoddinott, J. (1994) "Women's Income and Boy-Girl Anthropometric Status in Cote d'Ivoire," *World Development,* Vol 22, pp. 543-553.

Hamilton, W. D. (1964a), "The Genetical Evolution of Social Biology," *Journal of Theoretical Biology*, Vol 7, pp. 1-16.

Hamilton, W. D. (1964b), "The Genetical Evolution of Social Behavior," *Journal of Theoretical Biology*, Vol 7, pp. 17-52.

Hanushek, E. A. (1992), "The trade-off between child quantity and quality," *Journal of Political Economy,* Vol. 100, pp. 84-117.

Hashimoto K. and Heath J.A (1995), " Income Elasticities of Educational Expenditure by Income Class: The Case of Japanese Households," *Economics of Education Review*, Vol. 14, pp. 63-71.

Haughton J. and Haughton D. (1998), "Are simple tests of son preference useful? An evaluation using data from Vietnam," *Journal of population economics,* Vol. 11, pp. 495-516.

Haveman, R., Wolfe, B., and Spaulding, J. (1991), "Childhood events and circumstances influencing high school completion," *Demography,* Vol. 28, pp. 133-157.

Haveman, R., and Wolfe, B. (1995), "The determinants of children's attainment: A review of methods and findings," *Journal of Economic Literature*, Vol. 33, pp. 1829-1878.

Heckman, J. J. (1979), "Sample selection bias as a specification error," *Econometrica,* Vol. 47, pp. 153-61.

Hentschel, J., and Lanjouw, P., (1996), "Constructing an indicator of consumption for the analysis of poverty:Principles and illustrations with reference to Ecuador," Living standards measurement study working paper no. 124. The World Bank, Washington, DC.

Holmes, J. (1999),"School completion in Pakistan: analysis of censoring and selection bias," Working Paper # 794, Economic Growth Center, Yale University.

Jensen, H. and Yen, S. (1996), "Food Expenditures Away From Home by Type of Meal," *Canadian Journal of Agricultural Economics*, Vol. 44, pp. 67-80.

Jones F.L.(1983), "On Decomposing the Wage Gap: A Critical Comment on Blinder's Method," *Journal of Human Resources,* Vol. 18, pp. 126-130.

Jones, A. M. (1989), "A double-hurdle model of cigarette consumption," *Journal of Applied Econometrics,* Vol. 4, pp. 23-39.

Jones, A. M. and S. T. Yen (2000), "A Box-Cox Double-Hurdle model,"*The Manchester School,* Vol. 68, pp. 145-258.

Kabubo-Mariara and Mwabu, J (2007), "Determinants of school enrolment and education attainment:Empirical evidence from Kenya," *South African Journal of Economics,* Vol. 75, pp. 572-593.

Kalugina E. Najman B. (2002), "Labour and Poverty in Russia Self-Rated Perceptions and Monetary Evaluations," [online] available on http:// ecsoc-man.edu.ru/images/pubs/2004/02/10/0000148505/07_01_03.pdf

Ketkar, S. L., (1979),"Determinants of Fertility in a Developing Society: The Case of Sierra Leone," *Population Studies,* Vol. 33, pp. 479-488.

Kim, S., and Kim, Y. (2000),"Growth gains from trade and education ," *Journal of International Economics*, Vol. 50, pp. 519-545.

Kimball M.S (1990), "Precautionary saving in the small and in the large," *Econometrica,* Vol. 58 pp. 53-73.

King, E. M. and Lillard, L. (1987), "Education policy and schooling attainment in Malaysia and the Philippines," *Economics of Education Review,* Vol. 6, pp. 167-181.

Kingdon, G. G (1998), "Does the Labour Market Explain Lower Female Schooling in India?" *Journal of Development Studies,* Vol. 35, pp. 39-65.

Kingdon G.G and Knight J. (2003), "Well-being poverty versus income poverty and capabilities poverty ?" CSAE working paper series no. 16.

Kingdon G G (2005), "Where Has All the Bias Gone? Detecting Gender Bias in the Intrahousehold Allocation of Educational Expenditure," *Economic Development and Cultural Change,* Vol. 53, pp. 409-451.

Langani, Y.,(1997),"Groupes sociaux et fécondité au Burkina Faso. Une étude à partir de l'EDS de 1993," Mémoire de DEA. Institut de démographie. Louvain-la-Neuve, Université catholique de Louvain.

Lanjouw, P. F. and Ravallion, M. (1995), "Poverty and household size," *Economic Journal,* Vol. 105, pp. 1415-1434.

Lavy V (1996), "School supply constraints and children's educational outcomes in rural Ghana," *Journal of Development Economics,* Vol. 51, pp. 291-314.

Lazear, E. P., and Michael, R. T(1988), *Allocation of income within the household.* Chicago,IL:The University of Chicago Press.

Leibowitz, A. (1974), "Education and home production," *American Economic Review,* Vol. 64, pp. 243-250.

Leibowitz, A. (1977), "Parental inputs and children's achievement," *Journal of Human Resources,* Vol. 12, pp. 242-251.

Lipton, M., (1977), *Why Poor People Stay Poor: Urban Bias in World Development.* Harvard Univ. Press, Cambridge,MA.

Lokshin M., Umapathi N., Paternostro S. (2004),"Robustness of Subjective Welfare Analysis in a Poor Developing Country," World Bank Policy Research Working Paper No.3191.

López-Calva L.F and Miyamoto K. (2004), "Filial Obligations and Child Labor," *Review of Development Economics*, Vol. 8, pp. 489-504.

Maddala, G. (1983), *Limited Dependent and Qualitative Variables in Econometrics*, New York: Cambridge University Press.

Maitra, P (2003), "Schooling and Educational Attainment: Evidence from Bangladesh," *Education Economics*, Vol. 11, pp. 129-153.

Majumdar,S, Manib A, and Mukand S.W (2004), "Politics, information and the urban bias," *Journal of Development Economics*, Vol. 75, pp. 137-165.

Manski, C. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies*, Vol. 60, pp. 531-542.

Mauldin,T, Mimura, Y, and Lino, M (2001), "Parental Expenditures on Children's Education," *Journal of Family and Economic Issues,* Vol. 22, pp. 221-241.

McDonald, J. and Moffitt, R. (1980), "The Uses of Tobit Analysis," *The Review of Economics and Statistics,* Vol. 62, pp. 318-312.

Monfardini C and Radice R (2008), "Testing exogeneity in the bivariate probit model: a Monte Carlo study, " *Oxford Bulletin of Economics and Statistics,* Vol. 70, pp. 271-282.

Montgomery, M. Koume A, and Oliver R. (1995), "The trade-off between number of children and child schooling: Evidence from Cote d'Ivoire and Ghana" LSMS Working paper 112. Washington D.C: World Bank.

Mukherjee S and Benson T. (2003), "The Determinants of Poverty in Malawi," *World Development,* Vol. 31, pp. 339-358.

National Statistical Office (NSO)(1998),*Malawi poverty profile report*, Zomba, Malawi: National statistical office.

National Statistical Office (NSO)(1998),*Malawi population and housing census: Report of final census results*, Zomba, Malawi: National statistical office, December

National Statistical Office (NSO) (2005), *Integrated household survey* [online] available at http://www.nso.malawi.net/data_on_line/economics/ihs/IHS2/overviewIHS2.html

Neuman S and Oaxaca R (2004), "Wage decompositions with selectivity-corrected wage equations: A methodological note," *Journal of Economic Inequality*, Vol. 2, pp. 3-10.

Neumark, D. (1988), "Employers' Discriminatory Behavior and the Estimation of Wage Discrimination," *Journal of Human Resources,* Vol. 23, pp. 279-295.

Newman, C., M. Henchion, and A. Matthews (2003), "A double-hurdle model of Irish household expenditure on prepared meals," *Applied Economics,* Vol. 35, pp. 1053-1061.

Noumbissi, A. and J.-P. Sanderson, (1998), "Pauvreté et comportements démographiques au Cameroun," in F. Gendreau (ed.), Crises, pauvreté et changements démographiques-dans les pays du Sud. Paris: Estem, pp. 149-163.

Oaxaca, R. L (1999), "Identification in Detailed Wage Decompositions," *The Review of Economics and Statistics,* Vol. 81, pp. 154-157.

Oaxaca, R. L. (1973), "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review,* Vol. 14, pp. 693-709.

Oaxaca, R. L., and M. Ransom (1994), "On discrimination and the decomposition of wage differentials," *Journal of Econometrics,* Vol. 61, pp. 5-21.

Olsen, R.J. (1983), "Mortality Rates, Mortality Events, and Number of Births," *American Economic Review,* Vol. 73, pp. 29-32.

Olsen, R.J. (1987), "Cross-Sectional Methods for Estimating the Replacement of Infant Deaths," *Research in Population Economics* (T.P. Schultz, Ed.). Greenwich, Connecticut: JAI, Vol. 6.

Pasqua, S. (2005), "Gender Bias in Parental Investments in Children's Education: A Theoretical Analysis," *Review of Economics of the Household*, Vol. 3, pp. 291-314.

Pradhan M. and Ravallion M. (2000),"Measuring poverty using qualitative perceptions of welfare," *Review of Economics and Statistics*, Vol. 82, pp. 462-472.

Ranis, G., Stewart, F., and Ramirez, A. (2000), "Economic growth and human development," *World Development*, Vol. 28, pp. 197-219.

Ravallion, M. (1996), "Issues in Measuring and Modeling Poverty," *Economic Journal,* Vol. 106, pp. 1328-1343.

Ravallion M., and Lokshin M.(2002), "Self-Rated Economic Welfare in Russia," *European Economic Review*, Vol. 46, pp.1453-1473.

Ray R., (2000), "Child Labor, Child Schooling, and their Interaction with Adult Labor: Empirical Evidence for Peru and Pakistan," *World Bank Economic Review*, Vol. 14, pp. 347-367.

Reimers, C, (1983), "Labor market discrimination against hispanic and black men," *Review of Economics and Statistics*, Vol. 65, pp. 570-579.

Ribar, D. C. (1993), "A multinomial logit analysis of teenage fertility and high school completions," *Economics of Education Review*, Vol. 12, pp. 153-164.

Rivers, D., and Vuong, Q. H. (1988), "Limited information estimators and exogeneity tests for simultaneous probit models," *Journal of Econometrics,* Vol. 39, pp. 347-366.

Rose, E (2000), "Gender Bias, Credit Constraints and Time Allocation in Rural India," *The Economic Journal,* Vol. 110, pp. 738-758.

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

Rosenzweig, M.R., and Schultz, T.P. (1987), "Fertility and Investments in Human Capital:Estimates of the Consequences of lmperfect Fertility Control in Malaysia," Yale University Economic Growth Center Discussion Paper No. 534, New Haven,Connecticut.

Rosenzweig M. and Wolpin K. (2000), "Natural "Natural Experiments," *Journal of Economic Literature,* Vol. 38, pp 827-874.

Sen, A. (1984), *Resources, Values and Development,* Cambridge, M.A.: Harvard University Press.

Shapiro, D., Tambashe, O. (2001), "Gender, poverty, family structure and investments in children's education in Kinshasa, Congo," *Economics of Education Review*, Vol 20, pp. 359-373.

Sicilian P and Grossberg A.J (2001), "Investment in human capital and gender wage differences: Evidence from the NLSY," *Applied Economics*, Vol. 33, pp. 463-471.

Smith, M. D. (2003), "On dependency in Double-Hurdle models," *Statistical Papers,* Vol. 44, pp. 581-595.

Smith, R., and Blundell, W. R. (1986), "An exogeneity test for a simultaneous equation tobit model with an application to labor supply," *Econometrica,* Vol. 54, pp. 679-686.

Song L, Appleton S , and Knight J (2006), "Why Do Girls in Rural China Have Lower School Enrollment?" *World Development,* Vol. 34, pp. 1639-1653.

Su, S. and Yen S.T. (1996), "Microeconometric Models of Infrequently Purchased Goods: An Application to Household Pork Consumption," *Empirical Economics,* Vol. 21, pp. 513-533.

Suits, D. B. (1984), "Dummy Variables: Mechanics v. Interpretation," *The Review of Economics and Statistics,* Vol. 66, pp. 177-180.

Tansel, A. (1997), "School attainment, parental education and gender in Cote d'Ivoire and Ghana," *Economic Development and Cultural Change,* Vol. 45, pp. 825-856.

Tobin, J. (1958), "Estimation of relationships for limited dependent variables," *Econometrica,* Vol. 26, pp. 24-36.

UNAIDS (2006), *Report on the global AIDS epidemic,* Geneva

UNAIDS (2007), *AIDS epidemic update* [online] available at http://data.unaids.org/pub/EPISlides/2007/2007_epiupdate_en.pdf.

UNDP (2007), *Human development report* [online] http://hdr.undp.org/hdr2006/countries/country_-fact_sheets/cty_fs_MWI.html

Van Praag, B.M.S., (1971), "The welfare function of income in Belgium: An empirical investigation," *European Economic Review,* Vol. 2, pp. 337-369.

Van der Gaag, J. (1982), "On measuring the cost of children," *Children and Youth Services Review,* Vol. 4, pp. 77-109.

Vere J. (2006) "Dragon Children: Identifying the Causal Effect of the First Child on Female Labor Supply with the Chinese Lunar Calendar, " [online] available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=704241.

White, H. and Masset, E. (2003), "Constructing the poverty profile: An illustration of the importance of allowing for household size and composition in the case of Vietnam," *Development and Cultural Change*, Vol. 34, pp. 105-126.

Willis, R.J., (1973), "A New Approach to the Economic Theory of Fertility Behavior," *Journal of Political Economy*, Vol. 81, pp. S14-S64.

Wooldridge, Jeffrey M. (2002), *Econometric Analysis of Cross Section and Panel Data*, Cambridge, Mass: MIT Press.

World Bank (2005), "Zambia: Poverty and Vulnerability Assessment," Report No 32573-ZM [online] available at http://www.sarpn.org.za/documents/d0001457/index.php

Yen, S. T. and H. H. Jensen (1996), "Determinants of household expenditures on alcohol," *The Journal of Consumer Affairs*, Vol. 30 , pp. 48-67.

Yen, S. T. and Jones, A. M. (1997), "Household consumption of cheese: An inverse hyperbolic sine double-hurdle model with dependent errors," *American Journal of Agricultural Economics,* Vol. 79, pp. 246-51.

Yen, S T. (2005), "Zero observations and gender differences in cigarette consumption," *Applied Economics,* Vol. 37, pp. 1839-1849.

Yeuh, L (2006), "Parental investment in children's human capital in urban China," *Applied Economics,* Vol. 38, pp. 2089-2111.

Yun M, (2004), "Decomposing differences in the first moment," *Economics Letters,* Vol. 82, pp. 275-280.

Yun M, (2005), "A Simple Solution to the Identification Problem in Detailed Wage Decompositions," *Economic Inquiry*, Vol. 43, pp. 766-772.

Zimmerman F.J (2003), "Cinderella Goes to School: The Effects of Child Fostering on School Enrollment in South Africa," *The Journal of Human Resources,* Vol. 38, pp. 557-590.