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

Mussa, Richard

Economics Department, Chancellor College, University of Malawi

27 January 2009

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

Richard Mussa†

June 21, 2009

Abstract

The paper 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.

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

*I would like to thank Martin Wittenberg for his invaluable comments. Financial support from the African Economic Research Consortium (AERC) is appreciated. The standard disclaimer applies.

†Department of Economics, Chancellor College, University of Malawi, Box 280, Zomba, Malawi, rimussa@yahoo.co.uk

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 4 that there are differences in average expenditure between households

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

in rural and urban areas. 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). And 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 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.

[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.

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 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 paper proceeds as follows. Section 2 looks at the education sector in Malawi. Section 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 4 we discuss the model specification, variables used, estimation issues, and data and descriptives. Econometric results are the focus of Section 5. The extension of the Blinder-Oaxaca decomposition technique and results are discussed in Section 6. We finally conclude in Section 7.

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 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.

[4]This could however be longer with repetition.

3 Theoretical framework

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.

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,*

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

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 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. And 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).

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

[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.

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. And 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 based.

4 Methodology

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

[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

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;

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

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

$$\begin{aligned} Y_i^* &= X_i'\beta + \nu_i \\ Y_i^{**} &= \max(0, Y_i^*) \end{aligned} \tag{2}$$

Observed expenditure(Y_i);

$$Y_i = D_i Y_i^{**} \tag{3}$$

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[9] 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 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[10].

Using 0 to represent zero expenditure and + to denote positive expenditure, the sample

(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).

[9]The assumption of independence is quite common when using the DH (Mauldin *et al.* 2001; Jensen and Yen 1996; Su and Yen 1996). Further, Smith (2003) shows that there is little statistical information to support the estimation of a DH with dependent errors even when dependence exists.

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

likelihood equation for the independent double hurdle model can be written as follows;

$$L = \prod_{0} \left[1 - \Phi\left(Z_i'\alpha\right) \left(1 - \Phi\left(\frac{X_i'\beta}{\sigma} \right) \right) \right] \prod_{1} \Phi\left(Z_i'\alpha\right) \frac{1}{\sigma} \phi \left(\frac{Y_i - X_i'\beta}{\sigma} \right) \tag{4}$$

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 4), reduces to that of a Tobit when $\Phi\left(Z_i'\alpha\right) = 1$. A closer look at the likelihood function (equation 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 [11]. Accordingly, a likelihood ratio test can be used to test the Tobit model versus the independent DH[12].

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[13]. 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

[11] It worth pointing out that in the independent DH unlike the dependent DH exclusion restrictions are not needed to identify the parameters.

[12] The log likelihood ratio test statistic (LR) is computed as follows: $LR = -2\left[LL_T - \left(LL_P + LL_{TR}\right)\right] \sim \chi^2_k$; 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.

[13] 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).

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[14] 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 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

[14] 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).

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[15]. 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[16].

We control for regional fixed effects by including a three class regional dummy for the north, centre, and south. The variables are formally defined in the appendix Table A1.

[15]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.

[16]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).

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 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[17]. 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).

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. This data contains detailed information on socioeconomic and demographic characteristics of the households. The

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

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 4.2. Secondly, schooling decisions are cumulative 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 1 and 2. The full sample comprises households with primary school going children, with zero expenditures and nonzero expenditures on education. In Table 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, 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 discuss the econometric results in the next section.

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[18]. 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

[18] 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.

a Chi-square distribution with degrees of freedom $(k_{rural} + k_{urban}) - k_{pooled}$ is given by;

$$LR = -2\left[LL_{pooled} - (LL_{rural} + LL_{urban})\right] \sim \chi^2_{(k_{rural}+k_{urban})-k_{pooled}} \quad (5)$$

Results of the pooling tests are presented in Table 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 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 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, 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 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 A3 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[19]. 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*

[19] 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.

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 school has a negative impact on the participation and the expenditure decisions of both rural and urban households[20]. 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[21]. In terms of the age-gender demographics, the results suggest that having more primary school going boys (mal15h) and girls (fem15h) 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 mal15h and fem15h in the two areas. Results of the tests are shown at the bottom of Table 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 A3 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[22] 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 of mal15h and fem15h in the next section.

[20] 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.

[21] 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.

[22] 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.

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 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.

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)[23].

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

$$E(Y_i) = \Pr(Y_i > 0)E(Y_i|Y_i > 0) \tag{6}$$

Where the probability of expenditure is given by;

$$\begin{aligned} \Pr(Y_i &> 0) = \Pr(Z_i'\alpha + \varepsilon_i > 0, X_i'\beta + \nu_i > 0) \\ &= \Pr\left(\varepsilon_i > -Z_i'\alpha, \nu_i > -X_i'\beta\right) \\ &= \Phi\left(Z_i'\alpha\right)\Phi\left(\frac{X_i'\beta}{\sigma}\right) \end{aligned} \tag{7}$$

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

And the conditional expectation of Y_i is expressed as[24];

$$\begin{aligned} E(Y_i|Y_i &> 0) = X_i'\beta + E\left(\nu_i|\varepsilon_i > -Z_i'\alpha, \nu_i > -X_i'\beta\right) \\ &= X_i'\beta + \left[\Phi\left(Z_i'\alpha\right)\Phi\left(\frac{X_i'\beta}{\sigma}\right)\right]^{-1}\sigma \times \phi\left(\frac{X_i'\beta}{\sigma}\right)\Phi\left(Z_i'\alpha\right) \\ &= X_i'\beta + \frac{\sigma\phi\left(\frac{X_i'\beta}{\sigma}\right)}{\Phi\left(\frac{X_i'\beta}{\sigma}\right)} \end{aligned} \tag{8}$$

The elasticities of the unconditional expectation of Y_i with respect to the continuous regressors are computed by differentiating equations 7 and 8, and using the adding up property, equation 6. Formally, the elasticity of a continuous variable j which appears in both the participation and the expenditure equations is written as follows:

$$\begin{aligned} \eta_j^{UC} &= \frac{\partial E(Y_i)}{\partial X_{ij}}\frac{X_{ij}}{E(Y_i)} \\ &= \frac{\partial \Pr(Y_i > 0)}{\partial X_{ij}}\frac{X_{ij}}{\Pr(Y_i > 0)} + \frac{\partial E(Y_i|Y_i > 0)}{\partial X_{ij}}\frac{X_{ij}}{E(Y_i|Y_i > 0)} \\ &= \eta_j^P + \eta_j^C \end{aligned} \tag{9}$$

Equation 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 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 A4. 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. This means that for rural households spending on the schooling of children is more sensitive to income compared to urban households, and thus schooling

[24]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.

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, 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 (mal15h) and girls (fem15h) 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 A4) 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

[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).

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.

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 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

[26]The coefficient effect in the labor economics literature is interpreted as a measure of discrimination.

as follows[27];

$$E(Y_{im}) = \Pr(Y_{im} > 0)E(Y_{im}|Y_{im} > 0) \tag{10}$$

Where the probability of expenditure is given by;

$$\Pr(Y_{im} > 0) = \Phi\left(Z'_{im}\alpha_m\right)\Phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right) \tag{11}$$

And the conditional expectation of Y_i is expressed as;

$$E(Y_{im}|Y_{im} > 0) = X'_{im}\beta_m + \frac{\sigma_m\phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right)}{\Phi\left(\frac{X'_{im}\beta_m}{\sigma_m}\right)} \tag{12}$$

Three things need to be noted about equation 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 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]:

[27] For ease of exposition, we have reproduced equation 6.

[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.

$$\triangle_{R_1}^{DH} = \left[E_{\alpha_U,\beta_U,\sigma_U}(Y_{iU}) - E_{\alpha_U,\beta_U,\sigma_R}(Y_{iR})\right] + \left[E_{\alpha_U,\beta_U,\sigma_R}(Y_{iR}) - E_{\alpha_R,\beta_R,\sigma_R}(Y_{iR})\right] \quad (13)$$

and

$$\triangle_{U_1}^{DH} = \left[E_{\alpha_U,\beta_U,\sigma_U}(Y_{iU}) - E_{\alpha_U,\beta_U,\sigma_U}(Y_{iR})\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}(Y_{iR}) - E_{\alpha_R,\beta_R,\sigma_R}(Y_{iR})\right] \quad (14)$$

Where $E_{\alpha_m,\beta_m,\sigma_m}(Y_{im})$ denotes the unconditional expectation of Y_{im} evaluated at the parameter vectors α_m, β_m and the error standard error σ_m. The difference between the two decompositions is that equation 13 treats the standard error as part of the variables while equation 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] :

$$\triangle_{U_2}^{DH} = \left[E_{\alpha_R,\beta_R,\sigma_U}(Y_{iU}) - E_{\alpha_R,\beta_R,\sigma_R}(Y_{iR})\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}(Y_{iR}) - E_{\alpha_R,\beta_R,\sigma_U}(Y_{iR})\right] \quad (15)$$

and

$$\triangle_{R_2}^{DH} = \left[E_{\alpha_R,\beta_R,\sigma_R}(Y_{iU}) - E_{\alpha_R,\beta_R,\sigma_R}(Y_{iR})\right] + \left[E_{\alpha_U,\beta_U,\sigma_U}(Y_{iR}) - E_{\alpha_R,\beta_R,\sigma_R}(Y_{iR})\right] \quad (16)$$

The first term in the decompositions (equations 13 -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

[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).

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 13 -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 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 13 to 16, the following sample equivalent of the unconditional expectation (equation 10) is employed;

$$T\left(\hat{\alpha}_m, \hat{\beta}_m, Z_{im}, X_{im}, \hat{\sigma}_m\right) = N_m^{-1} \sum_{i=1}^{N_m} \left\{ \begin{array}{c} \Phi\left(Z'_{im}\hat{\alpha}_m\right) \Phi\left(\frac{X'_{im}\hat{\beta}_m}{\hat{\sigma}_m}\right) \\ \times \left(X'_{im}\hat{\beta}_m + \frac{\sigma_m \phi\left(\frac{X'_{im}\hat{\beta}_m}{\sigma_m}\right)}{\Phi\left(\frac{X'_{im}\hat{\beta}_m}{\hat{\sigma}_m}\right)} \right) \end{array} \right\} \tag{17}$$

Where $\hat{\alpha}_m, \hat{\beta}_m$, and $\hat{\sigma}_m$ denote sample estimates. With this sample counterpart of the unconditional expectation, equation 13 is estimated by;

$$\begin{aligned} \hat{\Delta}_{R_1}^{DH} = & \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{18}$$

Equation 14 is estimated by;

$$\begin{aligned} \hat{\Delta}_{U_1}^{DH} = & \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{19}$$

Equation 15 is estimated by;

$$\begin{aligned} \hat{\Delta}_{U_2}^{DH} = & \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right)\right] \end{aligned} \tag{20}$$

Finally, equation 16 is estimated by;

$$\begin{aligned} \hat{\Delta}_{R_2}^{DH} = & \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \\ & + \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right)\right] \end{aligned} \tag{21}$$

If there is only one censoring mechanism, that is $\Phi\left(Z'_{im}\hat{\alpha}_m\right) = 1$, decompositions 13 to 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 13 and 14 are equal and reduce to the standard Blinder-Oaxaca decomposition with urban coefficients used in the counterfactual. Similarly, decompositions 15 and 16 are equal and reduce to the standard Blinder-Oaxaca decomposition with rural coefficients used in the counterfactual.

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 18 to 21.

A detailed decomposition of the *characteristic effect* denoted as CE for the j^{th} variable $(j = 1, ..K)$ corresponding to equation 18 is given as;

$$CE_{R_1}^{DH} = \sum_{j=1}^{K} W_1^j \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \quad (22)$$

for equation 19 it is expressed as;

[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.

$$CE_{U_1}^{DH} = \sum_{j=1}^{K} W_2^j \left[T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_U, \hat{\beta}_U, Z_{iR}, X_{iR}, \hat{\sigma}_U\right)\right] \tag{23}$$

for equation 20 it represented as;

$$CE_{U_2}^{DH} = \sum_{j=1}^{K} W_3^j \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_U\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \tag{24}$$

and finally, for equation 21 it is denoted as;

$$CE_{R_2}^{DH} = \sum_{j=1}^{K} W_4^j \left[T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iU}, X_{iU}, \hat{\sigma}_R\right) - T\left(\hat{\alpha}_R, \hat{\beta}_R, Z_{iR}, X_{iR}, \hat{\sigma}_R\right)\right] \tag{25}$$

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

$$\begin{aligned}
W_1^j &= \frac{\left(\left(\overline{X}_U^j \frac{\hat{\beta}_U^j}{\hat{\sigma}_U}\right)\left(\bar{Z}_U^j \hat{\alpha}_U^j\right) - \left(\overline{X}_R^j \frac{\hat{\beta}_U^j}{\hat{\sigma}_R}\right)\left(\bar{Z}_R^j \hat{\alpha}_U^j\right)\right)}{\left(\left(\overline{X}_U \frac{\hat{\beta}_U}{\hat{\sigma}_U}\right)\left(\bar{Z}_U \hat{\alpha}_U\right) - \left(\overline{X}_U \frac{\hat{\beta}_U}{\hat{\sigma}_R}\right)\left(\bar{Z}_U \hat{\alpha}_U\right)\right)} \\
W_2^j &= \frac{\left(\left(\overline{X}_U^j \hat{\beta}_U^j\right)\left(\bar{Z}_U^j \hat{\alpha}_U^j\right) - \left(\overline{X}_R^j \hat{\beta}_U^j\right)\left(\bar{Z}_R^j \hat{\alpha}_U^j\right)\right)}{\left(\left(\overline{X}_U \hat{\beta}_U\right)\left(\bar{Z}_U \hat{\alpha}_U\right) - \left(\overline{X}_U \hat{\beta}_U\right)\left(\bar{Z}_U \hat{\alpha}_U\right)\right)} \\
W_3^j &= \frac{\left(\left(\overline{X}_U^j \frac{\hat{\beta}_R^j}{\hat{\sigma}_U}\right)\left(\bar{Z}_U^j \hat{\alpha}_R^j\right) - \left(\overline{X}_R^j \frac{\hat{\beta}_R^j}{\hat{\sigma}_R}\right)\left(\bar{Z}_R^j \hat{\alpha}_R^j\right)\right)}{\left(\left(\overline{X}_U \frac{\hat{\beta}_R}{\hat{\sigma}_U}\right)\left(\bar{Z}_U \hat{\alpha}_R\right) - \left(\overline{X}_U \frac{\hat{\beta}_R}{\hat{\sigma}_R}\right)\left(\bar{Z}_U \hat{\alpha}_R\right)\right)} \\
W_4^j &= \frac{\left(\left(\overline{X}_U^j \hat{\beta}_R^j\right)\left(\bar{Z}_U^j \hat{\alpha}_R^j\right) - \left(\overline{X}_R^j \hat{\beta}_R^j\right)\left(\bar{Z}_R^j \hat{\alpha}_R^j\right)\right)}{\left(\left(\overline{X}_U \hat{\beta}_R\right)\left(\bar{Z}_U \hat{\alpha}_R\right) - \left(\overline{X}_U \hat{\beta}_R\right)\left(\bar{Z}_U \hat{\alpha}_R\right)\right)}
\end{aligned} \tag{26}$$

and

$$\sum_{j=1}^{K} W_1^j = \sum_{j=1}^{K} W_2^j = \sum_{j=1}^{K} W_3^j = \sum_{j=1}^{K} W_4^j = 1 \tag{27}$$

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 22 to 25 reduce to that of a Tobit model

[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).

with censoring at zero. Further, if expenditure is uncensored, detailed decompositions 22 to 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).

6.2 Results of the decomposition

We present the results and discussion of the aggregated decomposition in subsection 6.2.1, this is followed up by results and discussion of the detailed decomposition in subsection 6.2.2.

6.2.1 Results of the aggregate decomposition

Results of the proposed aggregate decomposition are presented in Table 7. For comparison we also show in Table A5 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 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 A5). 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.

6.2.2 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 does not tell us which characteristics are key. In Tables 8 and 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 A6 and A7 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.

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

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, 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 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 A6 and A7 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.

7 Conclusions

Using the Second Malawi Integrated Household Survey (IHS2) data, the paper 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

[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.

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 differenti

References

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.

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.

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. S. (1981), *A treatise on the family.* Cambridge, MA: Harvard University Press.

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

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

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.

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.

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

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.

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

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.

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

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.

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

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.

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.

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.

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.

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.

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.

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.

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

Malawi National Statistics Office (NSO) (2005), "Integrated Household Survey 2004-2005 report" [online] available at http://www.nso.malawi.net/data_on_online/economics/ihs/IHS2/IHS2_report.pdf.

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.

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.

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.

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.

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

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.

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.

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

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

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.

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, 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.

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

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.

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

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.

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

Table 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: 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						
agelast	7.97	7.65	9.30	8.57	1.3***	0.92**
lnrexpa	9.64	9.65	10.07	10.19	0.43***	0.53***
daysratio	0.90	0.93	0.87	0.88	-0.03	-0.044*
children	3.52	3.68	3.59	3.71	0.04	0.07
Parental characteristics						
fathwage	0.71	0.72	0.82	0.81	0.10**	0.11***
mothwage	0.23	0.24	0.30	0.30	0.05**	0.07**
edufath	2.02	2.01	5.40	5.75	3.74***	3.38***
edumoth	0.79	0.76	2.90	3.31	2.55***	2.11***
agefath	47.77	47.78	47.88	47.87	0.09	0.11
agemoth	43.62	43.23	43.11	42.13	-1.1	-0.51
School characteristics						
distprimary	2.75	2.95	1.99	1.30	-1.65***	-0.76***
Age-gender composition of household						
mal6h	0.11	0.11	0.09	0.09	-0.02*	-0.02
mal15h	0.16	0.16	0.13	0.13	-0.03**	-0.03**
mal19h	0.05	0.05	0.05	0.04	-0.003	0.0006
mal55h	0.15	0.15	0.18	0.18	0.03***	0.02**
malover55h	0.03	0.03	0.04	0.04	0.007	0.01
fem6h	0.11	0.11	0.09	0.10	-0.005	-0.03
fem15h	0.15	0.16	0.16	0.15	-0.008	0.002
fem19h	0.03	0.03	0.04	0.04	0.01	0.01
fem55h	0.18	0.18	0.20	0.21	0.03***	0.02**
femover55h	0.03	0.02	0.02	0.01	-0.01	-0.01

Table 2: continued

	Rural		Urban			
	Full sample	Spending sample	Full sample	Spending sample	Gap 1	Gap 2
Variable	(1)	(2)	(3)	(4)	(3)-(1)	(4)-(2)
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 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 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 5: Results of the independent DH by area of residence

	Rural		Urban	
Variable	participation	level	participation	level
Household characteristics				
agelast	−0.05488***	−0.01060	0.30450	0.00072
	(0.01614)	(0.00883)	(0.23719)	(0.00365)
agelast2	0.00104***	0.00022	−0.00767	−0.00009
	(0.00038)	(0.00021)	(0.00710)	(0.00009)
lnrexpa	0.23207***	0.05227***	0.56821***	0.03576***
	(0.05461)	(0.00355)	(0.01981)	(0.01283)
daysratio	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)
children2	−0.00132	−0.00132**	−0.14069**	0.00046
	(0.00398)	(0.00056)	(0.07013)	(0.00084)
Parental characteristics				
fathwage	0.00650***	0.00756***	0.04989***	0.02324***
	(0.00138)	(0.00155)	(0.00224)	(0.00296)
mothwage	0.20134***	0.02023***	0.64032***	0.02352***
	(0.05835)	(0.00219)	(0.07506)	(0.00132)
edufath	0.00677***	0.01142***	0.02940***	0.00121***
	(0.0017)	(0.00259)	(0.00354)	(0.00019)
edumoth	0.00683***	0.00865***	0.03234***	0.00231***
	(0.00101)	(0.00148)	(0.00609)	(0.00026)
agefath	0.03908	0.01789	0.90438**	−0.01001
	(0.03091)	(0.01648)	(0.37120)	(0.00628)
agefath2	−0.00019	−0.00018	−0.00825**	0.00010
	(0.00027)	(0.00015)	(0.00341)	(0.00006)

Table 5: continued

Variable	Rural participation	Rural level	Urban participation	Urban level
agemoth	0.04274	0.05428***	-0.51033**	-0.01090
	(0.03055)	(0.01915)	(0.21431)	(0.01311)
agemoth2	-0.00045	-0.00048***	0.00328*	0.00015
	(0.00028)	(0.00018)	(0.00181)	(0.00014)
School characteristics				
distprimary	-0.00699***	-0.00908***	-0.56579***	-0.02440***
	(0.00024)	(0.00084)	(0.00536)	(0.00158)
Age-gender composition of household				
mal6h	1.11652**	-0.27137	-8.03960	0.20918*
	(0.55346)	(0.21291)	(5.46235)	(0.11936)
mal15h	1.94601***	0.23238***	6.16139***	0.18465***
	(0.54091)	(0.0095)	(0.09781)	(0.00321)
mal19h	1.05852*	0.30828	-11.41691*	0.26668*
	(0.57515)	(0.22514)	(6.57410)	(0.14466)
mal55h	0.43034	0.14724	-12.28649**	0.17875*
	(0.50640)	(0.18605)	(5.65072)	(0.10662)
fem6h	0.87586	-0.58748**	-7.37600	0.26562**
	(0.54953)	(0.25236)	(5.68817)	(0.12612)
fem15h	1.82512***	0.25020***	7.70956***	0.29012**
	(0.2362)	(0.0093)	(0.40535)	(0.12162)
fem19h	0.33254	0.36888*	-8.92036	0.31863***
	(0.59034)	(0.22344)	(5.63699)	(0.12218)
fem55h	0.63089	0.30406	-3.77753	0.12265
	(0.59596)	(0.23380)	(4.96818)	(0.10029)
femover55h	1.47903**	0.51368*	-4.29883	0.41458**
	(0.71350)	(0.28441)	(6.11054)	(0.18104)

Table 5: continued

	Rural		Urban	
Variable	participation	level	participation	level
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)
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 mal15h and fem15h:				
	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 6: Elasticities with respect to selected regressors for the DH

	DH					
	Rural			Urban		
Variable	Prob	Cond	Uncond	Prob	Cond	Uncond
lnrexpa	1.889***	1.145***	2.33***	0.154***	0.177***	0.331***
	(0.209)	(0.055)	(0.264)	(0.0083)	(0.0012)	(0.0095)
fathwage	0.164***	0.143***	0.307***	0.124***	0.132***	0.256***
	(0.006)	(0.007)	(0.013)	(0.003)	(0.025)	(0.028)
mothwage	0.272***	0.312***	0.584***	0.205***	0.206***	0.411***
	(0.003)	(0.004)	(0.007)	(0.064)	(0.114)	(0.178)
edufath	0.174***	0.137***	0.311***	0.114***	0.0856***	0.1996***
	(0.007)	(0.032)	(0.039)	(0.014)	(0.0021)	(0.0161)
edumoth	0.441	0.318***	0.759***	0.166***	0.224***	0.390***
	(0.030)	(0.073)	(0.03)	(0.017)	(0.023)	(0.04)
distprimary	−0.018***	−0.047***	−0.065***	−0.296**	−0.854***	−1.15***
	(0.002)	(0.005)	(0.007)	(0.115)	(0.268)	(0.383)
mal15h	0.120***	0.66***	0.780***	0.314***	0.320***	0.634***
	(0.033)	(0.055)	(0.088)	(0.077)	(0.036)	(0.113)
fem15h	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 standard errors.

Table 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 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:						
lnrexpa	0.0022***	(0.00072)	[34.38%]	0.0024***	(0.00012)	[36.36%]
fathwage	0.0003***	(0.000033)	[4.69%]	0.00011***	(0.00007)	[1.67%]
mothwage	0.00071***	(0.00008)	[11.09%]	0.00067 ***	(0.000086)	[10.15%]
edufath	0.0002***	(0.000065)	[3.13%]	0.00019***	(0.000052)	[2.88%]
edumoth	0.00084***	(0.000029)	[13.13%]	0.0009***	(0.000061)	[13.64%]
distprimary	0.0011***	(0.000076)	[17.19%]	0.0017***	(0.000063)	[25.76%]
other vars	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 vars comprises the remaining variables. We have not computed the standard error for these remaining variables.

Table 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:						
lnrexpa	0.0019***	(0.00017)	[31.67%]	0.0023***	(0.0001)	[33.33%]
fathwage	0.00019***	(0.00003)	[3.17%]	0.00017***	(0.000031)	[2.46%]
mothwage	0.00071***	(0.000064)	[11.83%]	0.00069***	(0.00002)	[10.00%]
edufath	0.00015***	(0.000047)	[2.50%]	0.0001***	(0.00002)	[1.45%]
edumoth	0.00093***	(0.000042)	[15.50%]	0.001***	(0.0001)	[14.49%]
distprimary	0.0012***	(0.0001)	[20.00%]	0.0016***	(0.00012)	[23.19%]
other vars	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 vars comprises the remaining variables. We have not computed the standard error for these remaining variables.

Appendix

Table A1: *Definition of variables*

Variable	Definition
Household characteristics	
agelast	age of the youngest child
agelast2	Square of age of the youngest child
lnrexpa	log of per capita household consumption expenditure
daysratio	ratio of day school going to all school going
children	number of children in the household
children2	square of number of children in the household
Parental characteristics	
fathwage	=1 if father works for a wage,0 otherwise
mothwage	=1 if father works for a wage,0 otherwise
edufath	Years of education of the father
edumoth	Years of education of the mother
agefath	Age of the father in years
agefath2	Square of the age of the father
agemoth	Age of the mother in years
agemoth2	Square of the age of the mother
School characteristics	
distprimary	Distance to nearest primary school in kilometres
Age-gender composition of household	
mal6h	proportion of males aged 0-6 in household
mal15h	proportion of males aged 7-15 in household
mal19h	proportion of males aged 16-19 in household
mal55h	proportion of males aged 20-55 in household
maleover55h[a]	proportion of males aged over 55 in household
fem6h	proportion of females aged 0-6 in household
fem15h	proportion of females aged 7-15 in household
fem19h	proportion of females aged 16-19 in household
fem55h	proportion of females aged 20-55 in household
femover55h	proportion of females aged over 55 in household
Region	
north	=1 if household is in the north,0 otherwise.
Centre	=1 if household is in the centre, 0 otherwise.
South[a]	=1 if household is in the south, 0 otherwise.

Notes: [a] denotes reference category.

Table A2: Reduced form regressions of log per capita consumption

Variable	Rural	Urban
agelast	0.090***	0.121***
	(0.005)	(0.045)
agelast2	-0.002***	-0.002
	(0.000)	(0.001)
agefath	-0.143***	-0.007
	(0.009)	(0.081)
agefath2	0.001***	0.002
	(0.000)	(0.001)
agemoth	-0.084***	-0.071
	(0.010)	(0.047)
agemoth2	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 A3: Results of the Tobit by area of residence

Variable	Rural	Urban
Household characteristics		
agelast	−0.00044***	0.00191
	(0.00016)	(0.00176)
agelast2	0.00001**	−0.00008
	(0.00000)	(0.00005)
lnrexpa	0.00171***	0.01259**
	(0.00018)	(0.00611)
daysratio	0.00337***	0.01023
	(0.00072)	(0.00670)
children	−0.00009	0.00293
	(0.00025)	(0.00351)
children2	0.00006**	0.00009
	(0.00002)	(0.00040)
Parental characteristics		
fathwage	0.00134***	0.0164***
	(0.00032)	(0.00368)
mothwage	0.02277**	0.01213***
	(0.00034)	(0.00129)
edufath	0.00151***	0.01206***
	(0.00004)	(0.00037)
edumoth	0.00431***	0.01074***
	(0.00007)	(0.00049)
agefath	0.00059*	−0.00024
	(0.00030)	(0.00216)
agefath2	−0.00000*	0.00000
	(0.00000)	(0.00002)
agemoth	0.00077***	−0.00259*
	(0.00023)	(0.00144)

Table A3: continued

Variable	Rural	Urban
agemoth2	-0.00001***	0.00002*
	(0.00000)	(0.00001)
School characteristics		
distprimary	-0.00042*	-0.01039***
	(0.00022)	(0.00295)
Age-gender composition of household		
mal6h	0.00150	-0.02059
	(0.00346)	(0.03275)
mal15h	0.00871	0.01916
	(0.337)	(0.23)
mal19h	0.00763**	-0.02771
	(0.00361)	(0.03614)
mal55h	0.00291	-0.04197
	(0.00316)	(0.02973)
fem6h	-0.00124	-0.00743
	(0.00345)	(0.03255)
fem15h	0.00867	0.07097
	(0.334)	(0.977)
fem19h	0.00640*	-0.00056
	(0.00374)	(0.02995)
fem55h	0.00419	-0.00840
	(0.00378)	(0.03112)
femover55h	0.00874*	0.05879
	(0.00448)	(0.04469)
Region		
north	0.00097	0.01347
	(0.00125)	(0.01054)
centre	0.00266***	-0.00272
	(0.00050)	(0.00388)

Table A3: continued

Variable	Rural	Urban
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 mal15h and fem15h:		
	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 A4: Elasticities with respect to selected regressors for the Tobit

	Tobit					
	Rural			Urban		
Variable	Prob	Cond	Uncond	Prob	Cond	Uncond
lnrexpa	1.524***	1.399***	2.923***	.450***	0.461***	0.911***
	(0.043)	(0.071)	(0.114)	(0.098)	(0.020)	(0.120)
fathwage	0.23***	0.17***	0.40***	0.16***	0.12***	0.28***
	(0.007)	(0.005)	(0.001)	(0.015)	(0.001)	(0.016)
mothwage	0.713***	0.64***	1.35***	0.449**	0.412**	0.861***
	(0.004)	(0.003)	(0.007)	(0.070)	(0.051)	(0.120)
edufath	0.431***	0.532***	0.963***	0.252***	0.407***	0.659***
	(0.046)	(0.034)	(0.08)	(0.033)	(0.064)	(0.097)
edumoth	0.869***	0.761***	1.630***	0.785***	0.658***	1.443
	(0.068)	(0.056)	(0.124)	(0.025)	(0.066)	0.091)
distprimary	-0.089*	-0.068*	-0.157*	-0.988***	-0.728***	-1.716***
	(0.046)	(0.035)	(0.080)	(0.303)	(0.211)	(0.504)
mal15h	0.104	0.079	0.182	-0.116	-0.086	-0.202
	(0.40)	(0.30)	(0.71)	(0.186)	(0.136)	(0.322)
fem15h	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 A5: 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 A6: Detailed decomposition of the characteristic effect of the Tobit using the urban variance

	Urban coefficients			Rural Coefficients		
CE	0.0044***			0.0046***		
Of which:						
lnrexpa	0.0012***	(0.0001)	[27.27%]	0.0016***	(0.00036)	[34.78%]
fathwage	0.0002***	(0.000031)	[4.55%]	0.0004***	(0.00004)	[8.70%]
mothwage	0.00032 ***	(0.00032)	[7.27%]	0.00048 ***	(0.000043)	[10.43%]
edufath	0.0005***	(0.000052)	[11.36%]	0.0002***	(0.00004)	[4.35%]
edumoth	0.00082***	(0.000047)	[18.64%]	0.0004***	(0.000013)	[8.70%]
distprimary	0.0009***	(0.000073)	[20.45%]	0.0008***	(0.00008)	[17.39%]
other vars	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 vars comprises the remaining variables. We have not computed the standard error for these remaining variables.

Table A7: Detailed decomposition of the characteristic effect of the Tobit using the rural variance

	Urban coefficients			Rural Coefficients		
CE	0.0048***			0.0045***		
Of which:						
lnrexpa	0.002***	(0.0005)	[41.67%]	0.002***	(0.0003)	[44.44%]
fathwage	0.00023***	(0.000061)	[4.79%]	0.00017***	(0.000041)	[3.78%]
mothwage	0.00038***	(0.000032)	[7.92%]	0.00041 ***	(0.000037)	[9.11%]
edufath	0.00016***	(0.000037)	[3.33%]	0.00013***	(0.00001)	[2.89%]
edumoth	0.00026 ***	(0.000042)	[5.42%]	0.0002***	(0.000029)	[4.44%]
distprimary	0.0009***	(.000033)	[18.75%]	0.0005***	(0.00007)	[11.11%]
other vars	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 vars comprises the remaining variables. We have not computed the standard error for these remaining variables.