

Estimating Returns to Education in Zambia

Chrispin Mphuka and Joseph Simumba

Zambia Institute For Policy Analysis and Research (ZIPAR)

December 4, 2012

Working Paper No. 11

Abstract

This paper estimates private returns to schooling among individuals aged 25 to 40 years old residing in Zambia in 2010. Using Instrumental Variable (IV) approach based on Two-Stage Least Squares (2SLS) and Generalised Method of Moments (GMM), we estimate average returns to schooling in the order of 15.1 percent. Proximity to primary and secondary school as well as per capita household education expenditure at district level in 1991 served as instruments. These instruments correspond to the period when the wage earners in the 2010 Living Conditions and Monitoring Survey (LCMS) were in school. These results provide important information on incentives regarding accumulation of human capital, efficiency of resource allocation in the education sector and the distributional consequences of differences in human capital.

1 Introduction

It is well known that education is an investment both to an individual and society. Like every investment, education accrues benefits and costs. Private costs to education, which are costs borne by an individual, include school fees, transport costs, school meals, uniform and book costs as well the income a person would have earned if they opted to work (opportunity cost). On the other hand, the private benefit to an individual's schooling mainly constitute the value of income they will earn when they enter into productive activities including the labour market. In view of the private benefits and costs, what amount of education will an individual choose?

Economic theory postulates that a rational individual will invest an additional amount into their education if the expected economic net benefit to the decision is positive. Therefore, private return to education serves as one of the most powerful determinants of the demand for schooling to an individual. However, it is worth to note that in the presence of liquidity constraints or negative cultural practices or limited availability of schools or poor quality of service provision, high returns to education may not necessarily translate into higher demand.

From a public sector point of view, knowledge of returns to education should help to rationalize public resources to areas where returns are low and leave private provision to high return areas where the public sector should merely retain a regulatory role. In this sense, returns to education provide a critical indicator that must guide the public to safeguard against under-provision or over-provision of education. From a private sector perspective, knowledge of return to investment forms a critical component in their investment decisions. To put the issues into perspective, Figure 1 below shows the evolution of education attainment between 1991 and 2010.

The figure shows the evolution in the highest grade attained by individuals who left

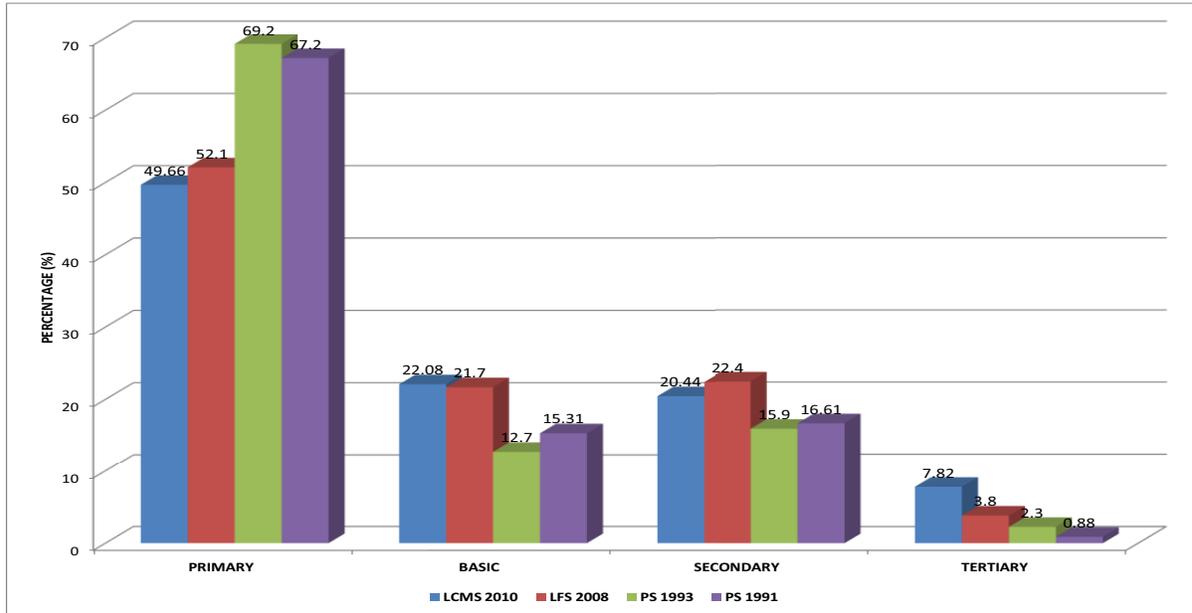


Figure 1: Education Attainment in Zambia: 1991-2010

school across four surveys, namely 1991 and 1993 Priority surveys, 2008 Labourforce survey and 2010 Living Conditions and Monitoring surveys. Generally, a majority of individuals continue to leave school at primary level though there has been a significant decline between 1991 and 2010. Similarly, the proportion of individuals leaving school with tertiary education increased over this period besides it still remaining the lowest. A key question concerning this trend is: can returns to education provide an explanation to this phenomenon? A first step to answer this question requires the measurement of the size of returns across the education levels. This is precisely the preoccupation of this paper.

Our motivation to undertake this study stems from the observation that returns to education in most less developed countries have been predominately high at low levels of education and reduce with a rise in education levels¹. Estimates of returns to education in Zambia, should help education stakeholders to evaluate whether focusing on basic education is still a worthwhile investment. A cursory look at the education budget over

¹A detailed review is provided in Psacharopoulos & Patrinos (2002) and Schultz (2003)

the last decade reveals that the major allocation was concentrated in basic education.

This has raised concerns that while enrollment and progression rates have improved in the basic education sector, the pupils usually fail to progress into secondary and tertiary levels. UNDP (2011) reports that net enrollment of children in primary education increased from 80 percent in 1990 to 102 percent in 2009 while primary completion rates increased from 64 percent in 1990 to 91.7 percent in 2009. This means that Zambia has already met the Millennium Development Goal of reaching 100 percent primary enrollment. But the fundamental problem of critical skill shortage, alluded to in Sixth National Development Plan, can not be met sufficiently if people leave schooling at the basic education level which is void of industrial skills.

On the contrary, if returns to schooling increase at more advanced levels of education, poorer families who are on average educating their children at primary school level will face low returns, while richer families who are on average educating their children to secondary or post-secondary school level face much higher return (Schultz 2003). Thus with such distribution of returns to education, there is an element that the poor may become poorer since poor parents may choose to take children up to primary school and end there or if they may rationalize to take to school only those children who performing better and hold a better promise of success.

The returns to education in Zambia have not be estimated for a very longtime. Using 1993 Priority survey data Nielson & Westergard-Nielson (2001) estimate the returns to primary schooling in the order of 9.7 percent and 12.2 percent for males and females respectively that are employed in rural areas. Strangely, they report nil returns for males in formal employment in urban areas regardless of their education level. A similar strange result was also reported among females that attained a higher than primary level and were employed in rural areas. For employed males with higher than primary school in rural areas, returns were estimated at 20.2 percent.

There are a number of varied reasons to doubt the validity of these results. Importantly, this study failed to remedy some of the measurement errors in education captured by the number of years spent in school. It also failed to control for factors that simultaneously affect individual incomes and years spent in school but cannot be measured such as individuals ability, tastes for schooling and incomes and social networks. Unlike their study, we use emerging methodologies to expunge the effect of all unobserved factors on schooling to obtain a better measure of the returns to education in Zambia.

The rest of the paper is organized as follow: Section II presents the descriptive summaries and distribution of education attainment and gross monthly earnings. Section III presents a review of the literature on returns to education. Section IV describes the methodology. Section V presents the findings and their discussion. Section VI concludes the paper.

2 Profile of education, employment and earnings

3 Literature on Returns to Education

3.1 Conceptual Foundations of Returns to Education Literature

The underlying interpretation of most recent studies of the return on education can be illustrated in the framework of the human capital model by Becker (1967). In such a model individuals face a market opportunity locus that gives the level of earnings associated with the alternative schooling choices, and reach an optimal schooling decision by balancing the benefits of higher schooling (which are earned over one's life time) against the costs (which are born early on when in school).

Usually, in this model, it is assumed that individuals seek to maximize the discounted present value of earnings, net of schooling costs. This is appropriate if individuals can

borrow or lend at a fixed interest rate, and if they are indifferent between attending school or working during their late teens and early twenties. More generally individuals may have different tastes and aptitudes for schooling relative to work, and this variation lead to variations in the optimal level of schooling across individuals(Card 2001).

Building on this conceptual framework, Card (2001) derives an optimal schooling choice path for individuals which underpins the econometric models that are used in much of the returns regression models. Assuming that individuals maximize a present discounted value of lifetime utility subject to an intertemporal budget constraint, Card (2001) concludes that individuals will invest in schooling until the marginal return is equal to the rate of interest at which they borrow or lend in order to go to school. From this baseline finding, he incorporates further assumptions that underlie the presence of heterogeneous returns to schooling that have been observed in most empirical studies. The section that follows reviews the rather expansive literature on the economic returns to schooling focusing on what is relevant to this current study.

3.2 Empirical Literature

Schultz (2003) examined wage differentials by education of men and women from African households to suggest private wage returns to schooling. He found that rates are similar between male and females. He also found that in six African countries, the private returns to education are highest at the secondary and post-secondary levels. Such a finding would suggest that large public subsidies for post secondary and secondary education are not necessary because there would be demand for such higher levels due to higher private returns.

Only one study has ever been done to estimate returns to schooling in Zambia.

3.2.1 Modelling Issues When Estimating Returns to Education

Empirical studies of the private returns to education do suggest, in line with theoretical literature, that education confers significant advantage to individuals (Blundell, Dearden, Meghir & Sianesi 1999). This entails that individuals with higher levels of education are more likely to earn higher wages. Studies estimating returns to education use the human capital earnings function of Mincer (1974). In this model, the log of individual earnings (y) in a given time period can be decomposed into an additive function of a linear education term and a quadratic experience term:

$$\log y = a + bS + cX + dX^2 + e \quad (1)$$

where S represents years of completed education, X represents the number of years an individual has worked since completing schooling, and e is a residual or error term. Education in this model is measured as a single variable and the coefficient of S is, thus, interpreted as each additional year of schooling has a proportional effect on the earnings, holding constant years in the labour market

Most of the early studies used ordinary least squares method to estimate returns to education. These studies ignored the problem of ability and measurement errors. OLS has been found to be biased due to omitted ability and measurement error. The omission of ability over-estimates the returns to education while the measurement error bias attenuates or reduces the estimate by not more than 10 percent (Card 2001). Due to these biases, OLS estimates of returns to education are normally regarded as mere correlations that do not suggest causation.

Psacharopoulos & Patrinos (2002) gives an expansive review of the early studies. Although the estimates of returns to education presented in this review are diverse across countries and time, the authors found the average rate of return to one year of schooling

to be 10 percent. They also found that the highest return are for low and middle income countries. Regionally, they found returns to schooling to be highest for Latin America, the Caribbean region and Sub-Saharan Africa. Based on estimates for a few countries in each region, the estimated returns to schooling in Sub-Saharan Africa at 7.3 percent compared to 8.4 in Asia, 9.0 in the developed countries. A number of African countries including Zambia have estimated returns in the review. Appendix ??, these estimates. As already mentioned, these results can not be relied upon when one wants to establish the causal link between education and returns because the majority of the studies reviewed ignored serious biases by using ordinary least squares methods.

Controlling for ability and measurement error bias, most recent studies use instrumental variable estimation methods or instrumental variable methods using control function or matching methods. Card (1999) reviews studies that use the instrumental variable method and use one of three types of instruments based on: 1) institutional features of the education system such as proximity to college; 2) family background; and 3) earnings of twins. Reviewed studies in this paper that used institutional features as instruments were limited to developed countries and showed that instrumental variable estimates were higher than ordinary least squares by as much as 30 percent or more. Griliches (1977) and Angrist & Krueger (1991) argue that the ability bias in OLS estimates reflect the downward bias attributable to measurement error. On the contrary, Card (1999) argues that this observation is because variables such as compulsory schooling or proximity of schools are most likely to affect schooling choices of individuals who would otherwise have relatively low schooling. If this individuals have higher than average marginal returns to schooling, then instrumental variables estimators based on compulsory schooling or school proximity might be expected to yield higher estimated returns than OLS estimates. Card (1999) found that studies (based on developed countries) that used proximity to school as instrument estimated returns to education that ranged from 5 to 15 percent. The same

author gives a selected review of studies that use other types of variables but these are not included here because they are not used in this paper.

Apart from using instrumental variables methods to recover the causal relationship between education and earnings, some studies have used a controls function method or matching method. These methods use much relaxed assumptions than IV estimators and may in certain situations be preferable particularly when returns to education differ (i.e. heterogenous returns) from one individual to another. In a recent study, Blundell, Dearden & Sianesi (2005) estimate the returns to education from OLS, IV, control function and propensity score matching methods. Based on their results, they argue that sequential multiple-treatment model (propensity score matching) is suited for the education returns formulation since education qualification levels in formal schooling tend to be cumulative. They prefer this method in that it is able to answer specific policy questions. For example, they find that those completing higher education in their data were earning an average return of 27 percent more than those without. They also find that compared with stopping at 16 years of age without qualification, the average return to O-levels was 18 percent, to A-levels was 24 percent and to higher education was 48 percent. In essence this result contradicts the finding in Psacharopoulos & Patrinos (2002) that returns reduce with the increase in the level of education.

With the increase in the availability of survey data in Africa, more studies have been done to estimate the returns to education. Schultz (2003) presents a review of six African² countries whether returns to education have been estimated. One regularity this author observes is that returns to education in these African countries increase as with the increase in the levels of education. Although the coverage of countries with this result is limited, Schultz (2003) goes on to argue that with this pattern of returns to education there is no empirical justification for the large public subsidies for postsecondary educa-

²These include Kenya, Ghana, Ivory Coast, South Africa, Nigeria and Burkina Faso.

tion in Africa. Other studies on Africa include Soderbom, Wambugu & Kahyarara (2005) on Tanzania and Kenya and Okuwa (2004). Soderbom et al. (2005) estimate the average marginal return of education at all levels of education at 14 percent and 9 percent in Kenya and Tanzania respectively. They also find that returns to education are convex and this results are robust to endogeneity. On the other hand, Okuwa (2004) finds the results like in Schultz (2003) that returns are higher at higher levels of education attainment.

As already mentioned in section Nielson & Westergard-Nielson (2001) is the only study of returns to education in Zambia. Although this study goes to greater length to control for sample selection bias, it fails to deal with of ability bias and measurement bias. Nonetheless, the results of this study show that in 1993, returns to primary education were positive in rural areas but nil in urban areas. Returns to higher than primary education were found to positive and higher in urban than in rural areas.

3.3 Summary

In summary, this section has reviewed the conceptual framework and empirical estimations of the private return to schooling. What comes out is that there is an expansive literature globally that estimates returns to education. However, this literature is diverse both in methodology and results. Whereas earlier studies, ignored ability bias which usually leads to upward bias of OLS and measurement error that leads to attenuation bias most recent studies have adopted either instrumental variables method, or control function method or propensity score matching to recover the causal impact of education on earnings. The section also shows that there is no clear pattern on returns to schooling as one moves from lower to higher levels of education. Some studies find that, the returns to schooling increase with the level to schooling while others find the opposite. This then suggests that the pattern of returns to schooling remains an empirical one which is most likely context

or country specific.

4 Methodology

4.1 Data

The study used the 2005 and 2008 Labour Force Surveys, the 1991, and 1993 priority surveys and the LCMS survey of 2010. All these surveys were conducted by the Central statistical office, are all nationally representative and use the similar sampling design. In all these surveys, a two-stage clustered sampling strategy is used to choose the final sample of households. Whereas the 1991 and 1993 priority surveys are used to help establish proximity to a primary school and a secondary school, when those for whom returns to education are estimated from the 2010 LCMS or the 2008 LFS. The justification of using proximity to instrument for endogenous schooling is explained in the next sub-section.

4.2 Model

Recent studies³ on the returns to education have demonstrated clearly that there are serious biases in OLS estimates of returns to education due to omitted ability or measurement error since education is only measured through an imperfect measure, schooling or due to heterogeneity in the observed returns from one person to another(Blundell et al. 2005, Card 1999). Although there are a number of alternative methodologies⁴, in this study we use instrumental variables approach to control for the biases induced by omitted ability and measurement error of education. Following Card (1993), we instrument education using the geographic proximity of primary and secondary schools at the

³See for example ?

⁴See Blundell et al. (2005) for a recent detailed comparison between estimated returns from Instrumental variable method, control function method and matching method and Carneiro, Heckman & Vytlačil (2010) use local instrumental variable estimators to estimate the impact of marginal policy changes on the returns to education

time when those we are observing 2010 returns in the LCMS were in colleges model is as follows: We consider the following human capital model⁵

$$\log y_i = \alpha_i + \beta S_i + \varepsilon_i \quad , \quad i = 1, 2, 3, \dots, n \quad (2)$$

where y_i is the earnings of individual i , α_i is a person specific constant of integration, β is the homogenous return to schooling, S_i is the years of schooling by individual i , and ε_i is the error term which accounts for all unobserved factors such as ability. If schooling and omitted ability were not correlated implying that S_i is not correlated to ε_i and schooling was not measured with error, then an ordinary least squares estimator of β in equation 2 would be unbiased. In reality, this is not the case. Usually omitted ability results into an upward bias while measuring schooling with errors leads to a downward bias β leading to a net bias of about 10 percent or more(Card 2001). This implies that oLS estimates of β will be biased and this formulation it implies that suggested in the literature that are used to remedy this problems not indicate the causal effect of schooling on earnings.

To remedy this problem, we use the instrumental variable approach. We choose a variable called an instrument for schooling. This variable must be such that it is correlated to schooling , S_i ,but uncorrelated to earnings, y_i . We assume that more than one instruments are available, then these two condition would imply that we estimate an auxiliary equation (supposing two instruments for illustrative purposes), we then have:

$$\widehat{S}_i = \delta_0 + \delta_1 z_1 + \delta_2 z_2 \quad (3)$$

where \widehat{S}_i is the predicted schooling for individual i , z_1 and z_2 are the two instrumental variables. For these two to be considered as instruments, they then it must be they are uncorrelated to schooling implying that $\delta_1 \neq 0$ and $\delta_2 = 0$. This is tested through a

⁵Although not clearly indicated here we do control for sector and district of employment.

standard t-test or F-test or any other equivalent test. The second condition that the two instruments must not be correlated to earnings is much more hard to test but we adopt certain tests as explained in the results section to indeed establish that our instruments are valid. The question is what can really be used as instruments? Valid instruments are indeed to find. Recent studies have focused more on institutional sources of variation in schooling attributable to features such as the minimum school leaving age, tuition costs for higher education, or the geographic proximity of schools (Card 1999). In line with this class of studies, we use proximity to a primary school and proximity to a secondary school as measured by distance to these facilities in 1991 when majority of the current income earners were in school. We get this these proximity measures by estimating the average, median and 75th percentile of the distribution of distance at the district level⁶. We then attach these to the current income earners in the 2010 LCMS. , this study follows the approach of instrumental variables. as demonstrated in the previous section. Using the predicted schooling in equation 3, we estimate the equation:

$$\log y_i = \alpha_i + \beta \widehat{S}_i + v_i \quad , \quad i = 1, 2, 3, \dots, n \quad (4)$$

where equation 4 is normally referred to as the second stage in the 2-stage least squares method. Given that z_1 and z_2 are indeed valid instruments, the estimate of β in equation 4 called the instrumental variable estimator. Based on the validity of the instruments, the IV estimator recovers the true causal effect of schooling on education. It is clear now that given the biasedness of the OLS estimator, the IV estimator is a preferred estimator.

⁶Ofcourse, the strong assumption we make here is that people remain in the same districts in 2010 when they are observed income earners as they were in 1991. This is indeed not true for a substantial number of people although migration can not easily be established from the current series of independent cross sections of surveys in Zambia. To mitigate this problem, we assume that most of those who have moved have done so largely from similar districts in terms of distance to schools to where they are observed in 2010. In the absence of longitudinal data this assumption can not be validated yet we do take it that this is true so long as our test for instruments is found to hold in the final result.

This is why this study focuses on using the IV estimator while also estimating OLS as a benchmark case.

However, a related estimator of β in equation 4 is what is referred to as the Instrumental Variable Generalized Method of Moments (IV-GMM). This estimator uses all instruments in their raw form without estimating the auxiliary regression of equation 3. In addition, the IV-GMM uses an appropriate weighting matrix to ensure this improves the efficiency of the estimator in terms of having smaller variances. Standard econometric text books will generally recommend that when there is an endogenous variable and there is heteroskedasticity of an arbitrary nature, we use the robust standard errors in the IV estimator but since IV-GMM optimizes a different problem and is more efficient, in the face of presence of arbitrary heteroskedasticity and endogeneity, we must prefer the IV-GMM because it is more efficient. Another advantage of the IV-GMM estimator is that, it is able to control for intra-cluster correlation. This makes the IV-GMM estimator preferable.

However, it is important to mention that in the literature it has been found that the returns to schooling are not homogenous⁷. They differ from one individual to another. Then model 2 becomes altered to

$$\log y_i = \alpha_i + \beta_i S_i + \varepsilon_i, \quad i = 1, 2, 3, \dots, n \quad (5)$$

where β_i shows that the returns to schooling vary from one individual to another. In this case, IV estimation remains valid only under stronger assumptions that may not be easy to hold. The alternative to this becomes two other methods: the control function method and the matching method⁸. This study is still in the process of using these estimators for purposes of resolving this problem of heterogenous returns to schooling.

⁷For arguments on this and more citations see ?

⁸Blundell et al. (2005) compares between IV, control function and the matching estimators.

4.3 Characterization of Returns to Education and the Education Variable

4.3.1 Returns to Education

As already mentioned in earlier in section 1, we focus on estimating the private returns to education. We do this by looking at the monthly earning of the individual. So we do focus on those persons in active employment and not those that are in self employment per se. In fact due to the uncertain nature of income under self employment, we find that the majority of the individuals under this category have a zero income. This is most likely because the income of such persons may be irregular or it is hard for these individuals to characterize what they earn as income.

It is important to reiterate that ,though, the focus of this study is private returns to education, there are two equally important measures of returns to education. These include social returns and labour productivity returns. The social returns include any spill-overs that is private returns plus the net of government transfers and taxes. Labour productivity returns refer to the gross increase in labour productivity (Blundell et al. 2005).

4.3.2 The Measure of Education

As is the tradition, we measure education as a single factor represented by the number of years of schooling. Although this is a very restrictive measure of education because it assumes that the proportional earnings to a change in a year of schooling remains the same regardless of the level of qualification or education reached(Blundell et al. 2005). In this regard a better measure of schooling is to define education as a multi-category variable for instance based on the qualification attained. By so doing, one is able to distinguish the earnings to education by different levels of schooling. For example, one is able to

say what the returns to secondary schooling as opposed to tertiary schooling are. , once the schooling variable has more than one schooling category, it makes it hard to use IV methods to instrument for all those endogenous variables. We thus adopt the approach of estimating returns to schooling in each category- primary, secondary and tertiary by creating cells and limiting the estimating of equation 4 within that cell or people with that level of education.

5 Results on Returns to Education in Zambia

5.1 Introduction

This section presents the estimated returns to education in Zambia. We estimate the returns based on the IV estimation methods either using the two-stage least squares method to largely control for endogeneity. of education and the generalized method of moments which controls both for endogeneity. of education and also the non-constant variance in the earnings of individuals. As mentioned in the previous section, the instrumental variable we use is the distribution of proximity to primary schools and proximity to secondary schools as instruments. We implement various tests of these instruments to establish their validity and that the cure is actually no worse than the disease. This is normally the case when instruments are weak one OLS would perform even better than IV.

5.2 Returns to Education in Zambia: Estimates at National Level

Table 1 presents the national estimates of returns to investment in education in Zambia. Column one of the table shows the OLS results. The OLS results show that the returns to education are 17.8 percent. This result is statistically significant at 1 percent level of

significance. This would be interpreted as a one additional year of schooling would lead to a 17.8 percent increase in the average earnings. Due to the reasons of the biases of OLS elaborated in section 4, we only regard this as an indicator case. Column 2 of the same table shows the two-stage least squares (IV) estimator of returns to education while column 3 shows the Generalized Method of Moments (GMM) estimator. In both cases we use the proximity variable distance to primary school and secondary school within a district as instruments. We see that IV and GMM estimator of returns to schooling are very close and both lower than the OLS estimator indeed suggesting the OLS estimator might have an upward bias. According the IV estimator, the returns to education in Zambia is 14.7 percent. This result is statistically significant at 5 percent level. On the other hand, the GMM estimator suggests that the returns to education in Zambia stand at 15.1 percent. This suggests that a one additional year of schooling in Zambia might lead to a 14.7 percent increase in monthly earnings.

However, the IV and GMM results in this table are only valid if and only if the instrument used, proximity to primary and secondary schooling, is valid. The diagnostic tests in the table help us establish that fact. For the IV-GMM, the null for the Hansen J-test is that there is over identification in the model. This implies that the instruments are not correlated to the error terms under the null hypothesis. In table 1, the Hansen J-Statistic is 16.50 with a p-value of 0.1236 implying that we fail to reject the null. The other equivalent test is that of under identification called the Kleibergen-Paap rk LM test. Under the null-hypothesis we postulate that the model is under identified In the table the statistic which measures this is the idstat which is 31.31 with a p-value of 0.0018. So we reject the null of under, identifying meaning that the model could either be just identified or over identified. These two tests confirm the validity of our instruments, primary school and secondary school proximity.

To test whether the education variable is indeed endogenous, we use the endogeneity.

test. Under the null-hypothesis, this test says that education is indeed endogenous. The test statistic is estat which is 3.75 in the table with a P-value of 0.501 signifying that we fail to reject the null and conclude that education is indeed an endogenous variable.

This cache of tests does indeed suggest that the education variable is endogenous and so OLS results are likely to have a bias Hence the need to use either IV or IV-GMM estimates. It also suggests that our instrumental variables are indeed valid.

5.3 Estimated Returns to education by Gender

It is important that to estimate the returns to education by gender. Once these are know education may indeed be used to close any wage gaps that may be existent between the genders. Table 2 presents the returns to education according to whether one is male or female. Columns one and two of the table shows OLS results and IV-GMM results for males respectively. Columns three and four present the OLS and IV-GMM results for females.

In all these cases the coefficient of schooling (educyrs) is statistically significant at 5 percent level. For both males and females, the OLS results are higher than the IV-GMM results suggesting an upward bias in the OLS results. Moreover, the IV-GMM results have lower standard errors than their respective OLS coefficients. This is on account that IV-GMM is more efficient than OLS.

We therefore focus interpretation on the IV-GMM coefficients only. The IV-GMM results suggest that the returns to education for men (i.e., row 1 and column 2) are 13.2 percent. This implies that one additional year of education for a man translates would increase his earnings by 13.3 percent while for a female (row 1 and column 4) one additional year spent in school increases her earnings by 16.7 percent holding all other factors constant.

The coefficients of controls which include age and age squared (age_sq) and rural-

urban dummy are not significant suggesting that experience or location may not be a key determinant of some one's returns to education. In this model, we include age in quadratic form to proxy for experience assuming that log earnings are a concave function of years of schooling.

5.4 Education Returns to Schooling in Rural Areas Compared to Urban Areas

Table 3 shows the results of returns to education to those leaving in rural areas and those residing in the urban areas. Columns one and two show the results for OLS and IV-GMM respectively in rural areas and columns three and four show the estimated returns to education using OLS and IV-GMM respectively for urban areas. Just like in the previous cases the coefficients for age and age squared and those for gender are all insignificant. In this particular case the OLS estimates are larger than the GMM estimates. The reason for this could be because of the small samples that we end up with. For it is clear that the IV-GMM perform less efficiently in small samples.

The OLS estimates suggest that the returns to schooling are the same in the rural areas and in urban areas. The IV-GMM results show a small difference. The IV-GMM estimate shows that the returns to education in urban areas are 34.5 percent while those in rural areas are 21.4 percent⁹

5.5 Returns to Different Levels of Education

Table 4 shows the results for returns to education by the levels of education. In this case education is considered as a multifactor variable. We are still trying to estimate IV-GMM results so only OLS results are reported here. Column one presents the regression results

⁹I am still verifying this result because of the small samples. Will update by Monday

for primary education, column two for basic or grades nine and eight, column three for senior secondary and column four college and higher. The coefficient of education is not statistically significant at basic level. The results at all other levels are significant but with varying levels of significance. Contrary to the generally held notion that in developing countries, returns to schooling are higher at primary level and lower at higher levels of education, we find that returns are lowest at primary level and keep on increasing at higher levels. This is in line with what Schultz (2003) in six other African countries.

The table shows that the returns to education are estimated at 5.6 percent, 15.5 percent and 37.7 percent at tertiary level. Clearly, the returns are highest at tertiary level implying there should be sufficient demand at that level so that the private sector is the one to play a critical role of service provision with government focusing much of its attention to lower levels of education. Schultz (2003)

6 Conclusion and Policy Recommendations

This paper has estimated the returns to education in Zambia using the instrumental variable generalized method of moments approach. For a long time, Zambia has had no studies estimating returns to education despite this being a critical factor in knowing what is driving the demand to education, particularly different levels of education.

Based on the instrumental variable generalized method of moments approach, the study finds that the returns to education in Zambia are 15.1 percent. This means that an additional year spent in school would lead to an increase in earnings of 15.1 percent. It is also found that females have a higher return to education than males. It is also found that returns to education are higher in urban areas than in rural areas. This clearly entails that urban areas have sufficient incentive to drive the demand for education while rural areas lag behind. It may imply that government focuses more in rural areas to invest

in more facilities and other necessary supply side aspects.

References

- Angrist, J. D. & Krueger, B. A. (1991). Does compulsory school attendance affect schooling and earnings, *Quarterly Journal of Economics* **106**: pp 979–1014.
- Becker, G. (1967). *Human Capital and the Personal Distribution of Income*, University of Michigan Press.
- Blundell, R., Dearden, L., Meghir, C. & Sianesi, B. (1999). Human capital investments: The returns from education and training to the individual, the firm and the economy, *Fiscal Studies* **20**(1): pp 1–23.
- Blundell, R., Dearden, L. & Sianesi, B. (2005). Evaluating the effects of education on earnings: Models, methods and results from the national child development survey, *Journal of the Royal Statistical Society* **168**(3): 473–512.
- Card, D. (1993). Using geographic variation in college proximity to estimate the return to schooling, NBER Working Paper No. 4483.
- Card, D. (1999). The causal effects of education on earnings, *in* O. Ashenfelter & D. Card (eds), *Handbook of Labour Economics*, Vol. 3 of *Handbook of Labour Economics*, Elsevier Science, chapter 30, pp. 1802–1859.
- Card, D. (2001). Estimating the return to schooling: Progress on some persistent econometric problems, *Econometrica* **69**(5): 1127–1160.
- Carneiro, P., Heckman, J. J. & Vytlacil, E. (2010). Estimating marginal returns to education, NBER Working Paper Number 16474.
- Griliches, Z. (1977). Estimating the returns to schooling: Some econometric problems, *Econometrica* **45**: pp. 1–22.

- Mincer, J. (1974). *Schooling, Experience and Earnings*, Columbia University Press, New York.
- Nielson, H. S. & Westergard-Nielson, N. (2001). Returns to schooling in less developed countries, *Economic Development and Cultural Change* **49**(2): 365–394.
- Okuwa, B. O. (2004). Private returns to higher education in nigeria, AREC Research Working Paper No. 139. investigates returns to education using OLS.
- Psacharopoulos, G. & Patrinos, A. H. (2002). Returns to investment in education: A further update, World Bank Policy Research Working Paper WPS2881.
- Schultz, P. T. (2003). Evidence of returns to schooling in africa from household surveys: Monitoring and restructuring market for education, Economic Growth Centre Discussion Paper No. 875.
- Soderbom, M., Wambugu, A. & Kahyarara, G. (2005). The dynamics of the returns to education in kenyan and tanzanian manufacturing, Global Poverty Research Group Working Paper GPRG-WPS-017. Presents returns for Kenyan and Tanzanian manufacturing sector.
- UNDP (2011). Zambia millennium development goals progress report 2011, *Technical report*, UNDP.

Table 1: National Estimates of Returns to Education in Zambia

| | OLS | 2SLS | IV-GMM |
|----------------|-------------------------|-------------------------|-------------------------|
| ed yrs | 0.178*** (0.00704) | 0.147** (0.0471) | 0.151*** (0.0412) |
| age | 0.0376 (0.0507) | 0.0445 (0.0571) | 0.0428 (0.0512) |
| age_sq | -0.000378 (0.000772) | -0.000492 (0.000862) | -0.000493 (0.000781) |
| sex==2 | -0.121*** (0.0349) | -0.107** (0.0406) | -0.110** (0.0372) |
| rural_urban==2 | 0.0830 (0.0488) | 0.0995 (0.0740) | 0.105 (0.0638) |
| Constant | 10.53*** (0.841) | 10.54*** (1.018) | 10.54*** (0.890) |
| Adjusted R^2 | 0.531 | 0.496 | 0.496 |
| chi2 | | 1879.3 | |
| idstat | | | 31.31 |
| j | | | 16.50 |
| cdf | | | 10.64 |
| estat | | | 0.375 |
| Observations | 5109 | 5109 | 5109 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Author's Compilation

All the models controlled for districts and employment sector

A

Table 2: Estimates of Returns to Education By Gender

| | OLS_M | 2SLS_M | OLS_F | 2SLS_F |
|----------------|------------------------|------------------------|------------------------|------------------------|
| edysrs | 0.179*** (0.00836) | 0.132** (0.0402) | 0.179*** (0.00836) | 0.167** (0.0641) |
| age | 0.0736 (0.0667) | 0.0794 (0.0592) | 0.0736 (0.0667) | 0.0169 (0.0796) |
| age_sq | -0.000880 (0.00100) | -0.00103 (0.000899) | -0.000880 (0.00100) | -0.000116 (0.00123) |
| rural_urban==2 | 0.0766 (0.0531) | 0.164* (0.0679) | 0.0766 (0.0531) | 0.0572 (0.0949) |
| Constant | 9.706*** (1.126) | 10.08*** (1.035) | 9.706*** (1.126) | 10.80*** (1.321) |
| Adjusted R^2 | 0.479 | 0.462 | 0.479 | 0.562 |
| idstat | | 31.13 | | 17.87 |
| j | | 18.59 | | 16.31 |
| cdf | | 9.329 | | 2.679 |
| estat | | 1.846 | | 0.0355 |
| Observations | 3510 | 3510 | 3510 | 1599 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Author's Compilation

All the models controlled for districts and employment sector

Table 3: IV_GMM Estimates of Returns to Education By Location

| | OLS_R | 2SLS_R | OLS_U | 2SLS_U |
|----------------|----------------------|----------------------|----------------------|-------------------------|
| edyrs | 0.144*** (0.0140) | 0.214*** (0.0441) | 0.144*** (0.0140) | 0.345*** (0.0526) |
| age | -0.100 (0.142) | -0.0780 (0.110) | -0.100 (0.142) | 0.0729 (0.0564) |
| age_sq | 0.00166 (0.00212) | 0.00135 (0.00168) | 0.00166 (0.00212) | -0.000819 (0.000863) |
| _Isex_2 | -0.0949 (0.0815) | | -0.0949 (0.0815) | |
| Constant | 12.70*** (2.376) | 11.81*** (1.851) | 12.70*** (2.376) | 8.177*** (0.973) |
| Adjusted R^2 | 0.572 | 0.543 | 0.572 | 0.282 |
| idstat | | 25.27 | | 19.06 |
| j | | 8.408 | | 22.94 |
| cdf | | 7.371 | | 3.375 |
| estat | | 2.253 | | 5.702 |
| Observations | 890 | 890 | 890 | 4219 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Author's Compilation

All the models controlled for districts and employment sector

Table 4: OLS Estimates of Returns to Education By Level

| | Primary | Basic. | Sec. | Tertiary |
|----------------|-----------------------|-----------------------|-----------------------|----------------------|
| edys | 0.0562* (0.0230) | 0.0460 (0.0837) | 0.155** (0.0487) | 0.377*** (0.0513) |
| age | 0.118 (0.112) | 0.115 (0.0915) | -0.0231 (0.0752) | -0.0808 (0.124) |
| age_sq | -0.00164 (0.00172) | -0.00156 (0.00142) | 0.000700 (0.00117) | 0.00131 (0.00186) |
| Constant | 10.12*** (1.828) | 10.02*** (1.665) | 11.39*** (1.387) | 9.649*** (2.347) |
| Adjusted R^2 | 0.336 | 0.204 | 0.216 | 0.164 |
| idstat | | | | |
| j | | | | |
| cdf | | | | |
| estat | | | | |
| Observations | 592 | 798 | 1582 | 2137 |

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Source: Author's Compilation

All the models controlled for districts and employment sector

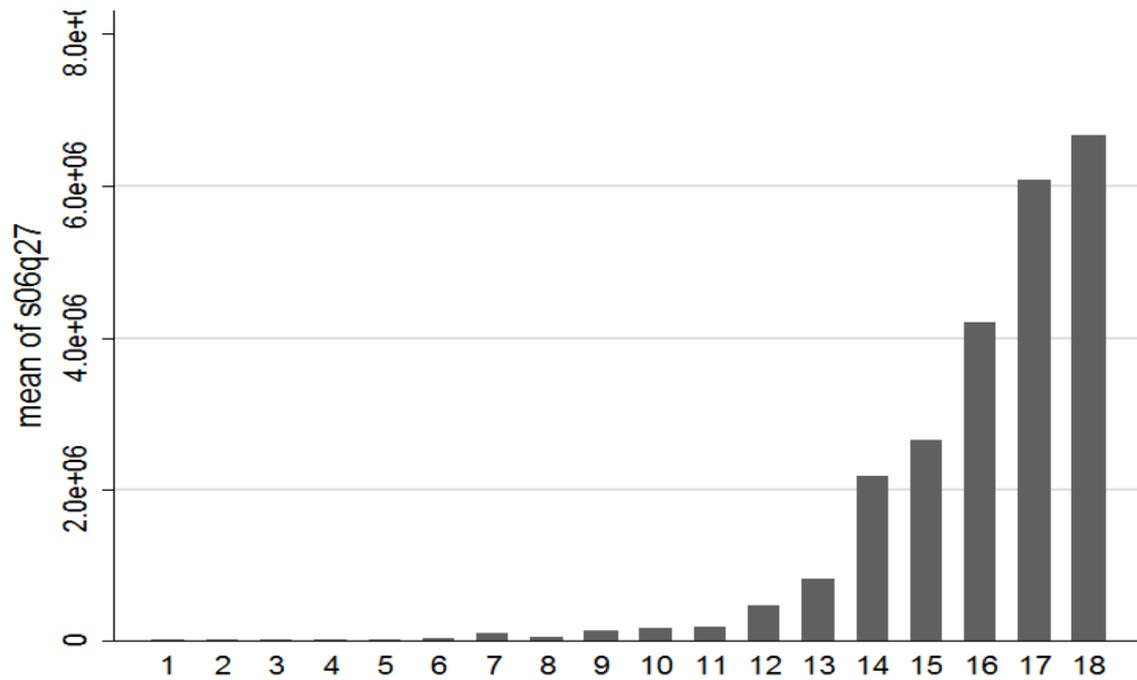
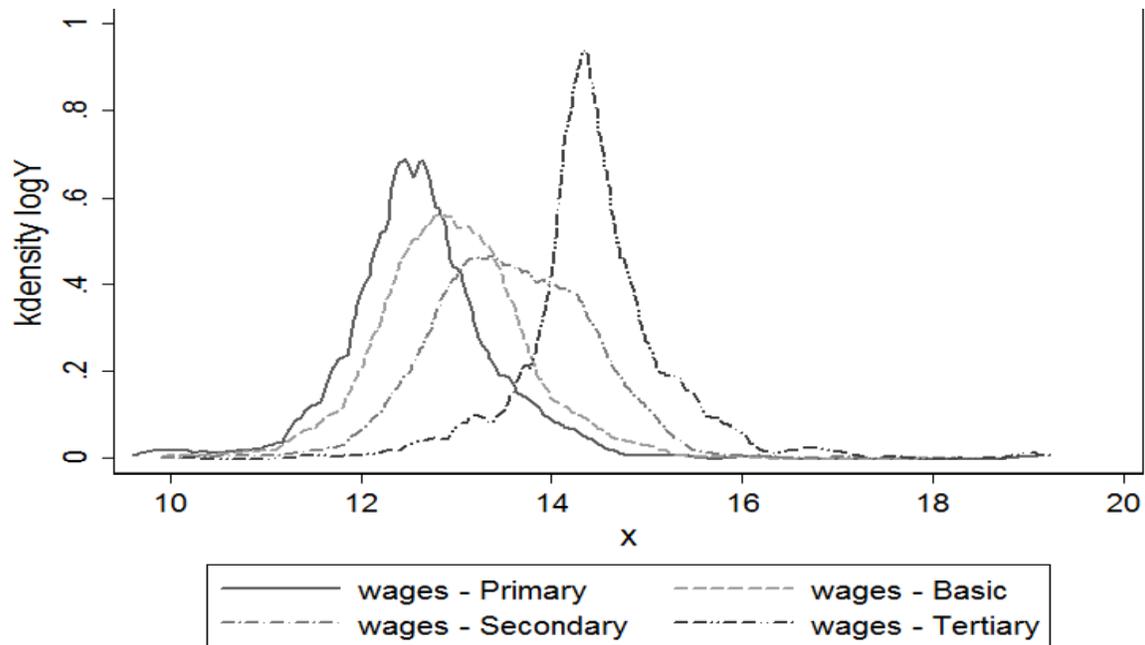


Figure 2: The Distribution of Real Consumption Per Adult Equivalent



Source: Authors Computations based on LCMS 2010

Figure 3: Density Distribution of Monthly Earnings by Education Level

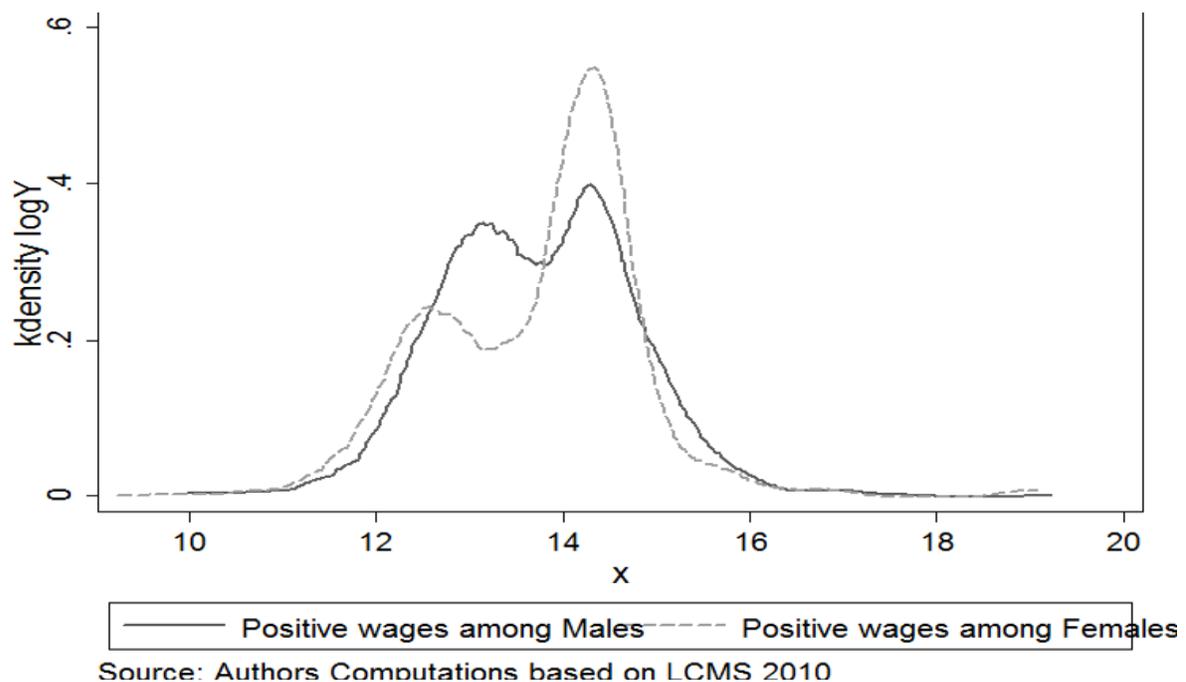


Figure 4: The Distribution of Real Consumption Per Adult Equivalent

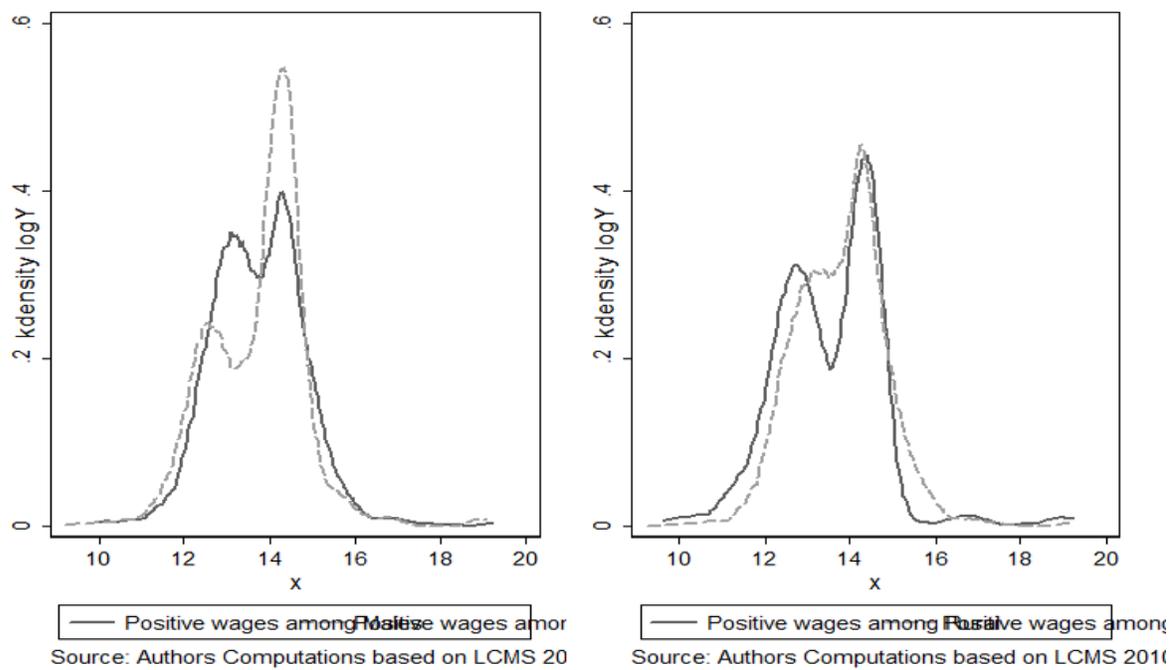


Figure 5: The Distribution of Real Consumption Per Adult Equivalent

B Returns to Education in Selected African Countries

| Country | Year | Private Returns to Schooling | | |
|-------------|------|------------------------------|-----------|--------|
| | | Primary | Secondary | Higher |
| Botswana | 1983 | 99.0 | 76.0 | 38.0 |
| Ethiopia | 1996 | 24.7 | 24.2 | 26.6 |
| Ghana | 1983 | 24.5 | 17.0 | 37.0 |
| Ivory Coast | 1985 | 25.7 | 30.7 | 25.1 |
| Kenya | 1980 | | 16.0 | |
| Lesotho | 1980 | 15.5 | 26.7 | 36.5 |
| Liberia | 1983 | 99.0 | 30.5 | 17.0 |
| Malawi | 1982 | 15.7 | 16.8 | 46.6 |
| Nigeria | 1966 | 30.0 | 14.0 | 34.0 |
| Senegal | 1985 | 33.7 | 21.3 | |
| Somalia | 1983 | 59.9 | 13.0 | 33.2 |
| Sudan | 1974 | | 13.0 | 15.0 |
| Tanzania | 1991 | 7.9 | 8.8 | |
| Gambia | 1997 | 37.1 | 12.7 | |
| Tunisia | 1980 | | 13.0 | 27.0 |
| Zambia | 1983 | | | 19.2 |
| Zimbabwe | 1987 | 16.6 | 48.5 | 5.1 |

Source: Psacharopoulos & Patrinos (2002) ,Table A1, pp 18-19