Education and earnings in Malawi: panel data evidence

ANDERSON SAWIRA GONDWE

Department of Economics, Stellenbosch University, Private Bag X1,7602, Matieland, South Africa.
Email: asgondwe@gmail.com

Preliminary draft paper

Abstract

Using the Malawian Integrated Household Panel Survey data, the paper finds large and positive returns to education in Malawi suggesting that education is a good investment. Our sample is limited to economically active individuals aged between 15 and 64 years with positive earnings. The returns increase with the levels of education. Interestingly, females have higher returns to education than males with similar skills at all levels of education. The paper sheds light on the importance of distinguishing between formal and informal employment sectors when estimating rates of return on education in developing countries. Furthermore, studying Malawi's informal sector is important as it accounts for 89% of total employment. Based on the findings, the study recommends education policies that improve female education and higher education in general. Our results are robust to different model specifications and compare favourably with those observed in previous studies in Malawi and other Sub-Saharan African countries.

JEL Classification: J31

Key words: Education, Employment, Malawi

1. INTRODUCTION

1.1 Background

Malawi remains a poor country despite registering gains in poverty reduction over the past two decades. The incidence of poverty as measured by household percapita consumption has marginally fallen from 40% in 2010 to 39% in 2013. In 2004, the poverty rate was 52.4%. Rural areas, which make up about 85% of the population, have significantly higher poverty rates than rural areas although the gap between the two is closing over time. The Gini coefficient shows that inequality increased from 0.390 in 2005 to 0.452 in 2011 before falling to 0.390 in 2013.

The labour participation rate for Malawi, defined for the share of the population aged 15 and above working or seeking employment, stands at 88%. On the other hand, education levels are low. About 74% of the population aged 15 years and above do not have any qualification at all and 21% of have reported to have never attended education. Literacy remains a challenge in Malawi. As at 2011, the literacy rate (defined as the ability to read and write with understanding in any language) amongst people aged 15 years and above stood at 65% which is an insignificant improvement from 64% reported in 2005 (National Statistical Office, 2012).

The Malawi Growth and Development Strategy (MGDS) identifies education as one of the themes necessary for growth and socio-economic development. Malawi's formal education system consists of primary, secondary and tertiary or post-secondary education. The Country's education policy has been focussed towards increasing access to primary and secondary education. According to National Statistical Office (2005, 2012), as a response to these policy changes, the net primary enrolment rate has increased to 86% in 2011 from 80% in 2004 while the primary dropout rate has dropped from 5% to 1% over the same period.

The Malawian labour market can be categorised into the formal and informal sectors. Just like in many other developing countries, the Malawian formal sector only absorbs a small percentage of the labour force. Consequently, most people are involved in either self-employment activities or paid employment in the informal sector (Chirwa & Matita, 2009).

1.2 Study motivation

Education has been identified as a tool for poverty reduction in Malawi (National Statistical Office, 2012). Through education attainment, the poor are said to be empowered and equipped for better opportunities in national development. In recognition of this, primary school education was universally made free in 1994 for all government schools. In addition to this, tertiary education in universities and colleges is subsidized in order to make it more affordable.

It is against this background that this study seeks to look at the role of education in poverty reduction in Malawi. The link between poverty and education is identified through the labour market. A number of studies that link education to poverty reduction have been conducted. Da Maia (2012) looks at the link between education and poverty reduction in Mozambique. The paper estimates the probability of an individual getting employment in any of the given sectors conditioned on education and also models the relationship between education and earnings. Chirwa & Matita (2009), Chirwa & Zgovu (2002), Psacharopoulos (1994, 2002), Becker (1975) and Mincer (1974), among others, look at the role played by education in earnings.

Previous studies on Malawian labour markets such as Chirwa & Matita (2009), Matita & Chirwa (2009) and Chirwa & Zgovu (2002) have explored the link between education and earnings using cross-sectional data sets. This study expands on the available literature by taking advantage of the newly released panel data set. Use of panel data has many advantages. Firstly, we can control for unobservable individual heterogeneity. As shown in the literature, failure to control for individual specific effects leads to bias in results. Secondly, panel data contains rich information about cross-sectional variations and dynamic behaviour. Thirdly, with panel data, we are able to identify time effects which cannot be identified with cross-sectional data. Baltagi (2013) and Hsiao (1986) provide a detailed discussed on the advantages and limitations of panel data.

In addition to using panel data, this study distinguishes between the formal and informal sectors of the economy to see if the role played by education in earnings differs by sector. Malawi's informal sector is important as it accounts for 89% of total employment. Furthermore, the study analyses the returns to education for different age groups. This analysis allows one to observe generational and life-cycle differences. People in the same age-groups experience similar labour and macroeconomic conditions. Currently, there is ongoing debate on youth unemployment in Malawi. This study is, therefore, timely and meaningful basis for further research.

The rest of the paper is organised as follows. Section 2 provides a discussion of the literature. The data and descriptive statistics are discussed in Section 3. Section 4 presents the econometric results. Section 5 gives conclusions and policy discussions.

2. LITERATURE REVIEW

2.1 Definitions and concepts

2.1.1 Formal sector

This primarily includes salaried employment in the private and government sectors as well as non-governmental organisations (NGOs). In this sector, the relationship between the employer and employees is governed by formal labour laws including employee benefits and income taxation.

2.1.2 Informal sector

We use the expanded statistical definition of informal employment which was endorsed by the International Labour Conference (2000) and the International Conference on Labour Statisticians (2003). This definition has two components, namely self-employment activities and informal wage employment.

a. **Self-employment:** This mainly includes employers in informal enterprises, own account workers in informal enterprises, contributing family workers (in informal and formal enterprises) and members of informal producers' cooperatives. Informal self-employment includes enterprises that are not registered under any national legislative authority and not engaged in agricultural activities.

b. **Informal wage employment:** employees of informal enterprises, casual or day labourers, temporary or part-time workers, paid domestic workers, contract workers, unregistered or undeclared workers and industrial outworkers (also called homeworkers). Of great importance to the Malawian economy is casual employment (locally known as ganyu[1]). It is largely seasonal and very important in both urban and rural areas.

2.2 Education and earnings

Mincerian earning functions are the standard approach for estimating returns to education in labour markets. The methodology started with the work of Mincer (1974) and has been widely used in the literature (Psacharopoulos, 1994). The methodology is based on the human capital theory which argues that investment in education improves workers' skills resulting in high productivity and, therefore, higher earnings (Mincer, 1974). The basic model is summarised below:-

$$\ln(Y_{it}) = \alpha_0 + \alpha_1 S_i + \beta_1 E_{it} + \beta_2 E_{it}^2 + \sum_{k=3}^{n} \beta_k Z_{it} + \varepsilon_{it} \quad (1)$$

Where, for any individual (i), at time (t), Y is the earnings of individual, S is the number of years of schooling, E is the experience, Z is a vector of control variables and ε is the error term. The coefficient (α) is interpreted as the private rate of return to education (RORE) and $(\alpha * 100)$ gives the percentage return to one additional year of schooling.

The classic model presented above has been improved in the literature in two main ways, namely to account for the fact that returns to education may be heterogeneous rather than homogeneous and correct for selection bias caused by using non-random data for analysis.

2.2.1 Heterogeneous returns to education:

the basic model disregards the differences in the level of educational achieved by looking at a single overall education level- years of schooling. This approach is called the one factor or homogeneous model since it assumes that there are no differential trends in the returns to education for different levels of education. There is little statistical evidence and causal empiricism for the

[1] Ganyu is the dominant form of employment in the informal sector.

homogenous model. The alternative approach, called the multiple factor approach or heterogeneous model, looks at the different levels of education as having separate effects on earnings. Using this model specification, we, therefore, replace S with an educational dummy variable to represent the different educational categories. Model (1) then becomes:-

$$\ln(y_{it}) = \alpha_0 + \sum_{i=1}^{n} \alpha_i D_{ji} + \beta_1 E_{it} + \beta_2 E_{it}^2 + \sum_{k=3}^{n} \beta_k Z_{it} + \varepsilon_{it} \quad (2)$$

Where: $(D)$is the education dummy (j) and all the other variables are as previously defined. Four educational qualifications are captured in the data, namely "None", "Primary Education", "Secondary education" and "Tertiary education".

Our data set does not have information on number of years of experience. Following the standard procedure used in the literature (e.g. Chirwa (2009), Kahyarara & Teal (2008) and Appleton, Bigsten & Manda (1999)), we estimate the number of potential years of experience as age less years of schooling less preschool age. This assumes that once people complete their education, they immediately enter the labour market. Those without education are assumed to enter the labour market at the lowest labour market entry age of 15 years.

Assuming that growth is linear, the rate of return to schooling becomes:-

$$RORE_j = \frac{\exp(\alpha_j) - 1}{y_j} \quad (3)$$

Where: $RORE_i$ is the return on education of education category (j) and (y) is the number of years of foregone earnings for that level. The alternative is to compute the rate of return on education for level (j) relative to the immediate lower education level (h) and this is given as:

$$RORE_{jh} = \frac{\exp(\alpha_j - \alpha_h) - 1}{y_j - y_h} \quad (4)$$

2.2.2 Selection bias: The bias arises from either self-selection into different employment categories or non-random attrition in a panel data. The Heckman (1979) two-step procedure has been used to correct for selection bias. In the first step, we estimate the probability of an individual selecting into an economic sector or attriting through a probit model given as:-.

$$c_{it} = \beta x_{it} + v_{it} \quad (5)$$

Where: i represents an individual, c denotes the occupational choice or attrition equal to 1 if an individual is in formal sector or exits the sample and 0 otherwise and x be set of regressors. The error term is given by v .

From equation (5), we obtain the inverse of the Mills ratio and use it as an explanatory variable in the estimation of the wage equation (2). If the coefficient of the inverse Mills ratio is statistically significant, then we are justified in correcting for selection bias. We then proceed to correct for and report heteroskedasticity consistent standard errors in the first stage, otherwise not.

2.3 Modelling unobserved heterogeneity

Let us consider the following model when time, $T = 1,2$

$$y_{it} = \beta x_{it} + u_{it} \tag{6}$$

Where y_{it} is the log of earnings for individual i at time t. Suppose that the error term is made up of two components as follows:-

$$u_{it} = n_i + v_{it} \tag{7}$$

Where n_i is time invariant and correlated with x_{it}, v_{it} is time varying and uncorrelated with x_{it}. If the exogeneity assumption is violated, i.e. when $E(x_{it} n_i) \neq 0$, the OLS estimator will be biased in cross-section. On the other hand, panel estimators can be used to control for unobserved individual time invariant heterogeneity and this allows us to obtain unbiased estimates of β. Even when unobserved correlated effect is not time invariant, using panel data techniques could reduce the magnitude of the bias.

The most commonly estimated models with panel data are the fixed effects and random effects models and several considerations will affect the choice between the two. We discuss these considerations in the following paragraphs.

a. **Nature of the variables omitted from the model:** If we think that there are no omitted variables or that the omitted variables are uncorrelated with the explanatory variables in the model, then a random effects model is the best. A random effects model under these assumptions will produce unbiased estimates of the coefficients, use all the data available, and yield the smallest standard errors. However, it is more likely that omitted variables will produce at least some bias in the estimates. If there are omitted variables and these variables are correlated with the variables in the model, then a fixed effects model provides a means for controlling for omitted variable bias. In a fixed-effects model, subjects serve as their own controls. For this to work, the omitted variables must have time-invariant values with time-invariant effects. For example, gender does not change overtime and its effect on the outcome in wave 1 is the same as the effect of gender in wave 2.

b. **Amount of variability within subjects:** If subjects do not change much, or not at all, across time, then fixed effects models may not work very well or even at all. There is need to have within-subject variability in the variables if we are to use subjects as their own controls. When there is little variability within subjects, the standard errors from a fixed effects model may be too large to tolerate. Conversely, random effects models will often have smaller standard errors.

c. **What effects are we interested in studying?** In fixed effects models, we are not interested in estimating the effects of variables that do not change or change very little over time. Rather, we control for them or "partial them out." On the other hand, with random effects models, we are able to estimate the effects of time-invariant variables such as gender although the method is no longer controlling for omitted variables.

Given the above considerations, we choose to use the random effects model for three main reasons. First, that education, whose effect we are interested in measuring, is generally a slow changing variable especially for a three year period over which we have data. Secondly, given that our panel is short (only two periods) there is not much within subject variables in most of our variables. Third, a random effects model allows us to estimate the effects of time invariant variables such as gender which are important aspects in Malawi. With fixed effects, this is not possible since the variable gets dropped off after demeaning.

Using a random effects model naturally comes at a cost and the trade-off is that their coefficients are more likely to be biased than the fixed effects estimates. Nevertheless, according to Wooldridge (2002), panel data techniques reduce the magnitude of bias compared to ordinary least squares (OLS).

3. DATA AND DESCRIPTIVE STATISTICS

3.1 Data

This section provides a brief description of the Malawian Integrated Household Panel Survey (IHPS) data, a two wave panel conducted in 2010 and 2013. The survey was implemented by the National Statistical Office (NSO) of Malawi. The 2010 wave was part of the third nationally representative Integrated Household Survey (IHS3) conducted between March 2010 and March 2011 during which 3,247 households were selected as a panel subcomponent to be resurveyed in 2013. The second wave, carried out between April and December 2013, saw the panel sample increase to 4,000 households because split-off members who formed new households were also brought into the sample. The IHPS data is nationally representative.

The household formed the primary unit of analysis in the IHPS surveys. An attempt was made to track all baseline households as well as members that moved away from the baseline dwellings between 2010 and 2013. Servants and guests at the time of the IHS3 were excluded and only individuals who were projected to be at least 12 years of age and known to be residents in mainland Malawi[2] were tracked. Of the 3,247 households initially chosen in 2010, 20 could not be located while others split into new households. The rate of attrition at the household level was only 3.78 percent.

The 2010 baseline had 15,597 individuals, of which 14,232 are available in both waves, representing an overall attrition rate of 7.42 percent at the individual level. For purposes of this study, we are only interested in the household members available in both waves, i.e. the 14,232 individuals. Given these low rates of attrition, which also seem random, we pursue this issue no further because we believe the representativeness of the sample has not been affected[3].

3.1.1 Questionnaire on earnings

There are two sources of earnings as captured in the household questionnaire, namely self-employment activities and wage employment. Earnings from self-employment activities are provided as profit from enterprises over a period of 30 days while earnings from wage employment are given with an indication of the period over which earnings

[2] Excluding Likoma district which is an Island on Lake Malawi.
[3] In the literature, the most common ways of addressing attrition are Inverse Probability Weighting (IPW) and Heckman selection correction (see Wooldridge, 2002).

are earned, i.e. day, week, two weeks or month. Ganyu wages are given as daily earnings with an indication of the number of days worked in a week. All earnings are converted into real monthly earnings. The same questionnaire is used in both waves.

3.1.2 Dealing with outliers

Observations that are substantially different from the rest can make a difference to the regression results obtained. It is, therefore, important to not only investigate these unusual observations but also find ways of dealing with them. Wittenberg (2014), Burger and Yu (2007) provide a good discussion on dealing with outliers. Considering sample size issues, we used the second approach with results from the other three approaches presented as robustness checks (see Section 4.5).

a. **Remove millionaires:** The first approach is to take out millionaires. In our data set, there were 14 millionaires with average monthly earnings of K1,971,347.2 compared to K21,036.53 for the rest of the 6,676 individuals. However, the choice of millionaires is arbitrary and has a potential to remove genuine earners especially when considering that 85% of these millionaires have university education.

b. **Extreme regression residuals:** The second approach is to remove outliers by identifying observations with extreme regression residuals. In linear regression, an outlier is an observation with large residual. This is achieved by estimating a simple Mincerian type wage regression of the log of real monthly earnings on education, age, age squared, gender and occupation. After running this regression, studentised or standardised residuals were created. In this approach, studentised residuals with absolute values greater than five are flagged as extreme and corresponding observations dropped. Using this approach, only two observations were flagged out as extreme. We, therefore, reduced the cut-off to 4 resulting in 17 individuals being flagged out as having earnings that were too high or low for their characteristics.

c. **Robust regression:** When data is contaminated with outliers, using studentised residuals has been found to be insufficient in identifying the 'bad' observations. Robust regression is an alternative to least squares regression when this is the case, i.e. when data is contaminated with outliers or influential observations (Wittenberg (2014), Verardi & Croux (2009)). This approach is routinely handled in Stata and observations are given weights depending on whether they are outliers or not. Outliers are assigned zero weights and consequently identified as not belonging in the regression. In total, robust regression identified 26 observations as being extreme and this included all the outliers also identified through the studentised residuals.

d. **Remove observations in the 100th percentile:** We generated a new variable containing percentiles of real monthly earnings. This was used to identify and then drop 70 observations in the 100th percentile with average real monthly earnings of K811,561.35 compared to K220,676.72 in the 99th percentile. However this results in a loss of 70 observations which is deemed excessive. One can equally consider dropping observations in the bottom percentile but we are more concerned with outliers in the higher percentiles. The median of earnings is low while the mean is high suggesting the data distribution is skewed by the presence of large outliers.

3.1.3 Missing, negative and zero earnings

Missing earnings, for example because individuals refused to answer or the respondent did not know, are not imputed. Negative and zero earnings were dropped since their natural logs are undefined. After cleaning out, we remain with 6,678 individuals with positive real earnings. As is the practice in the literature, we only look at economically active individuals aged between 15 and 64 years.

3.2 Descriptive statistics

3.2.1 Earnings and changes in employment status between waves

Table 1 shows that overall, average real monthly earnings have increased by 45% from MWK 13,702 in 2010 to MWK19,826 in 2013. In the ensuing paragraphs, we attempt to breakdown and explain the sources of the increase. The first step is to examine how the employment status of individuals has changed between waves. Conditional on missing earnings, we came up with four employment statuses, unemployed in both waves (not shown in table), employed in either 2010 or 2013 only, and employed in both periods.

a. **Employed in either wave:** From the table, we can see that part of the increase in earnings is explained by the new entrants into the labour (N=1,612) with average real monthly earnings of MWK14,681 compared to MWK9,119 in 2010 who have now exited (N=870) the labour market. Between these two groups, average earnings have increased by 61%. Consequently, those that have exited the market have been replaced by higher earning individuals.

b. **Employed in both waves:** We observe that earnings are higher amongst individuals employed in both years compared to those only employed in either period. The gap in earnings between these two groups is stable, i.e. MWK9,119 versus MWK15,746 (1.73 times higher) in 2010 and MWK14,681 versus MWK23,758 (1.62 times more) in 2013. Moreover, those employed in both states also experienced an increase in earnings. Specifically, their earnings increased from MWK15,746 in 2010 to MWK23,758 in 2013, representing an increase of 51% over three years.

Considering the importance of ganyu employment in Malawi, this analysis (as in Table 1) is repeated is repeated for ganyu. The results are given in Table 2 where a similar pattern is observed as that from Table1. First, the largest increase in earnings is observed for individuals employed in both waves. Second, those employed in wave 2 only (new entrants) earn more compared to individuals found in wave 1 only.

3.2.2 Employment status and education attainment

Table 3 shows that of those without education qualification ("None"), 41% were unemployed in both waves and this forms the majority. On the other hand, 66% of those with university education were employed in both periods. New entrants into the labour market (wave 2 only) have more education compared to those that have dropped out of the labour market (wave 1 only).

3.2.3 Identifying sources of increases in real earnings

Chart 1 shows that real monthly earnings increased on average across all occupations except non-governmental organisations (NGOs) which might have been negatively affected by donor aid withdrawals that Malawi faced.[4] These results compare well with Farm Input Subsidy Programme (FISP) evaluation panel data being currently being analysed by other researchers in Malawi where large increases have been observed in real ganyu wages (nominal ganyu daily wage rates divided by maize prices) for some districts between 2012 and 2015 (see Chart 2).

Given this comparison, one can argue that the increases in the earnings may be genuine despite the facts that it may be difficult to isolate the main factors from the many drivers behind this some. Gross Domestic Product (GDP) growth may be a contributor but economy only grew at an average of about 3% per year. As for ganyu wages, it seems that with the availability of food in many parts of the Country, the ganyu workers tend to have higher bargaining power and usually ask for more wages. It is also worth noting that in real terms, minimum wages were adjusted upwards twice between 2010 and 2013; first by 50% effective 1st January 2011 and second by 34% effective 1st July 2013. These could explain the increases in earnings in the private and government sectors although this largely depends on effective implementation and monitoring.

4. ECONOMETRIC RESULTS

4.1 Earnings and education in cross section

According to the human capital theory (as discussed in Section 2.2), it is argued that investment in education improves workers' skills resulting in high productivity and, therefore, higher earnings (Mincer, 1974). We estimate earnings functions the two waves of the panel using OLS. Table 4 shows the results and our dependent variable is the log of real monthly earnings[5].

Across both waves, the results show that there are large returns to education and potential labour market experience. The strongly positive returns to education are consistent with other findings in Malawi (Chirwa & Matita (2009), Chirwa & Zgovu (2002). Similar results have been found in other African countries (e.g. in Cameroon by Ewoudou & Vencatachellum (2006), in Rwanda by Lassibille & Tan (2005), and Bennell (1996) for Sub-Saharan Africa)). This is in contradiction of the assertion by Psacharopoulos and Patrinos (2002) and Psacharopoulos (1994) that returns to education in developing countries are concave.

The negative and significant gender dummy is consistent with the general finding that females earn less than their male counterparts (Chirwa & Matita (2009). We also find that average earnings in the enterprise sectors are significantly higher lower than in the private sector at the 1 percent. Ganyu earnings are also lower at the 5% level of significance in the 2010 model only; the difference becomes insignificant in the 2013 and pooled models.

[4] The "Other" category is made up of 22 individuals only or 0.33% of all individuals. Perhaps we need to add them up to the dominant category or a category with similar characteristics.

[5] The interpretation of the coefficients is the percentage change in the monthly earnings given a unit change in an explanatory variable. For dummy variables the percentage effect of a change from the base category.

4.2 Homogeneous returns to education

We begin with the basic model assuming that returns to education are homogeneous and also ignoring and selectivity bias into the formal or informal sectors. Table 5 shows results based on OLS, fixed effects and random effects.

The fixed effects model does not work well indicating very low within variation in our variables especially after inclusion of the time dummy. Including a time dummy further reduces the variation in the data and this somehow causes the results to change dramatically[6]. Specifically, we see that the coefficient for education for OLS is almost 9 times as bigger as that of fixed effects- a sign that there is little within variation. We, therefore, concentrate on OLS and random effects. The Breusch Pagan LM test yields significant results indicating that the random effects model is more appropriate compared to OLS. The time dummy shows that average monthly earnings are higher in 2013 compared to 2010.

The random effects model differs from OLS in two aspects. First, in the random effects model, NGO employees earn significantly more than those in private sector at the 5% level of significance; the result is insignificant for OLS. Second, the coefficients for the enterprises are slightly lower in the random effects than in OLS.

4.3 Heterogeneous returns to education

The basic model, whose results are discussed in Sections 4.1 and 4.2, disregards the differences in the level of educational attainment by looking at a single overall education level- years of schooling. This homogeneous model assumes that there are no differential trends in the returns to education for different levels of education. As earlier discussed, there is little statistical evidence and causal empiricism for the homogenous model. The heterogeneous model provides the alternative and looks at the different levels of education as having separate effects on earnings. Using this model specification, we, therefore, replace years of schooling (S) with an educational dummy variable to represent the different educational categories, namely "no education", "primary education", "secondary education" and "tertiary education" in the manner discussed in Section 2.2.1.

The returns to education based on the random effects model are summarised in Chart 3. Full results are presented in Table 6. Regardless of gender, the returns to education increase with the level of education supporting a convex relationship between education and earnings. Our results confirm the finding that the returns increase with level of schooling in Sub-Saharan Africa but are contrary to the literature supporting concave rates of returns as already pointed out in Sections 4.1 and 4.2 above.

Also consistent with the international literature is the finding that female workers tend to have much higher returns on education than male workers particularly at higher levels of education. Similar findings have been established in Malawi by Chirwa (2008) and Chirwa & Matita (2009). Although more men enter the labour market, they tend to earn less than their female counterparts suggesting that female education is more effective in generating returns in Malawi. The high rates of return to higher education for females and tertiary education may just be a reflection of the short supply of workers with skills. It is worth noting that the education system in Malawi has greatly emphasised on primary education especially with the introduction of free primary education in 1994. There has been an expansion in primary education attainment and this possibly explains the low returns.

[6] We tested as to whether we need to be using a time dummy and the results support its inclusion.

4.4 Formal and informal sector segmentation

Just like in many other developing countries, the formal sector in Malawi only absorbs a small percentage of the labour force. Consequently, most people are involved in self-employment activities or in paid employment in the informal sector (Chirwa & Matita, 2009). According to results based on the Malawi labour force survey 2013, about 89% of the employed population aged between 15 and 64 years is employed in the informal sector[7]. Due to the large size of the informal sector, studying earning differentials between economic sectors is important for policy targeting. The results are presented in Table 7 based on OLS and random effects estimators.

Returns to education are positive in both sectors but with larger magnitudes in the formal sector compared to the informal sector for both OLS and random effects. The convex relationship between earnings and experience is also maintained in both sectors except that the coefficient is bigger in the informal sector than the formal sector. This implies that experience matters more in the informal sector. The time dummy coefficient is also positive in both models but larger in the informal sector-an indication that earnings grew much stronger within the informal sector.

Interestingly, a number of things are noted that are different when the sample is split between the formal and informal sectors than when it is not. First, we find lower statistically significant coefficients for gender in the formal sector compared to the formal sector. In addition, the coefficients for regions are insignificant in the formal sector. On the other hand, in the informal sector, the differences are not only significant but the coefficients are also larger. Second, within the informal sector, ganyu employers earn more compared to their counterparts in enterprises. No wonder ganyu or casual employment is the most dominant form of employment in the informal sector and in Malawi in general. Previous research in Malawi by Bose & Livingstone (1993); Chirwa & Zgovu (2002) has shown that casual employment on peak labour tasks may be well paid, above the daily minimum wage, but is short-lived. Third, sectoral analysis of earnings sheds some light on the importance of distinguishing between the different types of employment sector in estimating rates of return on education in developing countries. Studies that fail to take this into account tend to overstate the returns to education by assuming that returns are the same in both the formal and informal sectors.

4.5 Robustness checks

We conduct a number of robustness checks to see if our main results are preserved under different conditions or assumptions.

a. **Dividing the sample by gender:** We report OLS and random effects regression results in Table 8 having split the sample by the gender of earners. A version of this is already discussed in Section 4.3 but assuming heterogeneous returns to education. The table shows that the key results after dividing the sample into male and female sub-samples are preserved. The ranking of sectors is roughly the same for the full sample, although women are seen to be much worse in ganyu employment compared to men as shown by a negative and significant coefficient for females. The magnitudes of the coefficients are stable with and without split sample.

b. **Experiment with individuals employed in both waves only:** As we earlier noted in Section 3.2, these individuals experienced the largest jump in earnings over three years and also possess the highest levels of education. Again, the results from Table 8 (last two columns to the right) again do not differ much compared to those based on full sample (which includes those only employed in either waves 1 or 2).

[7] In our data set, the informal sector is about 77.54% amongst individuals with positive earnings.

c. **Alternative treatment of outliers:** Table 9 gives results based on four alternative ways of dealing with outliers as discussed in Section 3.1.2. Overall, the results do not change much if we consider both magnitude and direction although we tend to have larger estimates when N is larger.

4.6 Measurement error using panel data

In panel data, measurement error, just like non-random attrition bias, is of concern in any data set and an attempt is usually made in the literature to arrive at results that are robust to these concerns (Deaton, 1997. There are two types of measurement error, discussed below following Wooldridge (2002):

a. **Measurement error in the dependent variable:** The assumptions we make about the measurement error are important. First is the usual assumption that the measurement error has zero mean. However, if this is not the case, then the estimation of the intercept is affected. The second assumption relates to the relationship between the measurement error and the explanatory variables included in the model. If the measurement error in the dependent variable is statistically independent or uncorrelated with each explanatory variable, then the OLS estimators from equation are consistent (and possibly unbiased as well). Consequently, measurement error does not bias the coefficients but only leads to larger standard errors than when the dependent variable is not measured with error, i.e. it leads to loss of efficiency.

   In our model, one can reasonably argue that measurement error is correlated with education since people with more education tend to report their earnings more accurately. However, in the absence of additional information, it is difficult to establish if measurement error in earnings is related to any of the explanatory variables. One solution to measurement error in the dependent variable is to collect more data because more observations imply a better estimator of variance and consequently reduces the errors in inferences. This solution is beyond the researcher considering that the data is secondary data[8].

b. **Measurement error in the independent variables:** This has been considered a more serious than measurement error in the dependent variable. In panel data, the most common method of dealing with measurement error is first differencing (short and longer differencing). However, our data comes from a panel of two periods such that longer differencing is not possible. Moreover, first differencing when T=2 yields the same results as fixed effects presented in Table 5.

4.7 Age-period cohort analysis

As the name suggests, age-period-cohort (APC) captures three effects, namely age-effects to capture variation across different age-groups; period effects for variation affecting all age groups overtime and cohort effects to capture changes affecting individuals belonging to the same birth years. The general methodology of APC suffers from the identification problem arising from the

[8] Alternatively, we can use the Malawi Labour Force (2013) cross section data section for comparison purposes. Nevertheless, there is no guarantee that this is 'better' quality data.

exact linear dependence: **Cohort=Period-Age**, i.e. singular matrix with no unique solution. Researchers have developed a number of methodological solutions to this and other problems of APC:

a. The *proxy variables approach* which uses one or more proxy variables as surrogates for the age, period or cohort coefficients , e.g., O'Brien (2000);
b. The *nonlinear parametric transformation approach* which defines a nonlinear function of one of the age for the age, period or cohort variables, e.g., Fienberg & Mason (1985);Yang & Land (2006);
c. The *APC intrinsic estimator* by Fu (2000) & Yang et al (2008), e.g., Branson et al (2013).

We use the second approach owing to its simplicity and apply OLS and random effects estimation[9]. The results are presented in Tables 10 and 11. The results from both estimators show that individual characteristics are significantly related to earnings. The age effect is curvilinear and convex. As expected, education (measured by years of schooling) has a large positive effect on earnings; these results compare well with our previous findings. Individuals from the formal sector earn more than those in the informal sector. Females and individuals in the Southern region earn significantly lower than their counterparts. The period effect is significantly higher for 2013 when compared to the base category, i.e., 2010. This result is not surprising considering our earlier findings. The addition of more variables to the basic models (columns 10 and 18) reduces the age effect and it barely remains statistically significant when all controls are added. The cohort effects are concave; they are not shown in Tables 10 and 11 results but presented graphically in Chart 4.

5. CONCLUSION AND POLICY IMPLICATIONS

5.1 Conclusion

The paper sought to examine the returns to education in Malawi using the IHS3 panel data set. The random effects results based on the standard Mincerian earnings functions show that the average rate of return to years of schooling in Malawi is 10 percent for the full sample. The returns to education increase with the level of education and are also higher for females than males. Decomposition of the sample by economic sector reveals that the rate of return on education is similar in both sectors, is 10 percent in the formal sector and 9.6 percent in informal employment. The rates of return found in this study compare favourably with those observed in other studies in Malawi and other Sub-Saharan African countries including Ghana, Cameroon and Rwanda. Sectoral analysis of earnings gives us some more interesting results different than when the sample is not split. First, we observe that there is no statistically significantly difference in earnings between males and females in the formal sector. Second, we find that casual labour is more lucrative than the self-employment activities within the informal sector. These two findings highlight the importance of distinguishing between the different types of employment sectors in estimating rates of return on education in developing countries. The assumption that returns are the same in both sectors of the economy is not realistic and suggests that studies that fail to take this into account tend to overstate the returns to education. The results are robust to different model specifications. We need to further investigate the reasons behind the large increase in earnings observed between the two waves.

[9] Although our panel is short (two periods only), we present these results in order to stimulate debate on APC effects in Malawi.

5.2 Policy implications

The results have a number of implications with respect to policy. First, the positive and large returns from schooling suggest that education is a good investment. Access to education, therefore, is important for all. Second, since returns increase with the level of education, it may be important to invest more resources into higher level education. The current policy makes primary education universally free yet the returns are the lowest. The Government has recently removed subsidies that were being offered for tertiary education yet the findings from this paper show that returns to education are highest at the tertiary level. Perhaps, the Government should direct the resources 'freed' from the removal of the subsidy towards expansion of other areas within tertiary education. Third, education policy should be shaped to encourage female education considering that females with similar skills to males tend to earn more. This may be both economically efficient and equitable although currently there are fewer females entering the labour market in Malawi. This should not only target lower level education but also tertiary education where the returns are the highest.

Tables

Table 1: Mean monthly earnings by employment status and survey period

	2010				2013			
	Mean	SD	N	Percent	Mean	SD	N	Percent
Employment status								
Wave 1 only	9,119	(20,662)	870	29%	.	.	.	.
Wave 2 only	.	.	.	.	14,681	(31,995)	1,612	43%
Both waves	15,746	(50,298)	2,100	71%	23,758	(63,309)	2,100	57%
Total	**13,702**	**(43,476)**	**2,970**	**100%**	**19,826**	**(52,295)**	**3,712**	**100%**

Source: Own computation using IHPS data, earnings expressed in constant 2010 prices.

Table 2: Mean monthly ganyu earnings by employment status and survey period

	2010				2013			
	Mean	SD	N	Percent	Mean	SD	N	Percent
Employment status								
Wave 1 only	7,359	(12, 905)	556	40%	.	.	.	.
Wave 2 only	.	.	.	.	13,989	(28,045)	989	55%
Both waves	8,460	(15,659)	836	60%	15,764	(27,451)	806	45%
Total	**8,029**	**(14,649)**	**1,392**	**100%**	**14,813**	**(27,777)**	**1,795**	**100%**

Source: Own computation using IHPS data, earnings expressed in constant 2010 prices.

Table 3: Employment status and education attainment

Description	None	Primary	Secondary	University	Total
Unemployed in both waves	3,286	863	1,138	54	5,341
	(41.1)	*(45.5)*	*(40.5)*	*(11.9)*	*(40.6)*
Employed in wave 1 only	988	186	218	38	1,430
	(12.4)	*(9.8)*	*(7.8)*	*(8.4)*	*(10.9)*
Employed in wave 2 only	1,691	362	482	63	2,598
	(21.1)	*(19.1)*	*(17.2)*	*(13.9)*	*(19.7)*
Employed in both waves	2,037	484	969	300	3,790
	(25.5)	*(25.5*	*(34.5)*	*(65.9)*	*(28.8)*
Total	***8,002***	***1,895***	***2,807***	***455***	***13,159***
	(100.0)	***(100.0)***	***(100.0)***	***(100.0)***	***(100.0)***

Source: Own computation using IHPS data, percentages in parentheses.

Table 4: OLS results for log of real monthly earnings

Description	OLS		
	2010	**2013**	**Pooled**
Years of schooling	0.086***	0.090***	0.088***
	(0.009)	(0.008)	(0.008)
Experience	0.029***	0.049***	0.040***
	(0.005)	(0.005)	(0.004)
Experience squared	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)
Female	-0.395***	-0.395***	-0.394***
	(0.044)	(0.038)	(0.031)
Central	-0.064	-0.083	-0.049
	(0.067)	(0.087)	(0.052)
Southern	-0.267***	-0.184**	-0.210***
	(0.066)	(0.088)	(0.052)
Government	0.308***	0.215*	0.266***
	(0.106)	(0.114)	(0.080)
NGOs	0.289	0.260	0.290
	(0.359)	(0.253)	(0.246)
Other	0.166	0.19	0.246
	(0.599)	(0.227)	(0.305)
Enterprises	-0.373***	-0.345***	-0.322***
	(0.096)	(0.089)	(0.067)
Ganyu	-0.146**	0.052	-0.031
	(0.071)	(0.067)	(0.049)
Constant	8.326***	8.269***	8.251***
	(0.135)	(0.149)	(0.107)
R-squared	0.221	0.172	0.183
N	2958	3703	6661

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Table 5: OLS, Fixed effects and random effects results for log of real monthly earnings

Description	OLS			Fixed effects			Random effects		
	1	2	3	4	5	6	7	8	9
Years of schooling	0.094***	0.088***	0.087***	0.013	0.057***	0.010	0.104***	0.102***	0.100***
	(0.006)	(0.008)	(0.008)	(0.009)	(0.011)	(0.014)	(0.004)	(0.005)	(0.005)
Experience		0.040***	0.039***		0.061***	0.002		0.037***	0.036***
		(0.004)	(0.004)		(0.012)	(0.017)		(0.004)	(0.003)
Experience squared		-0.001***	-0.001***		0.000	0.000		-0.001***	-0.001***
		(0.000)	(0.000)		(0.000)	(0.000)		(0.000)	(0.000)
Female		-0.394***	-0.397***					-0.391***	-0.396***
		(0.031)	(0.031)					(0.027)	(0.027)
Central		-0.049	-0.076		0.286	0.132		-0.004	-0.011
		(0.052)	(0.051)		(0.259)	(0.276)		(0.034)	(0.034)
Southern		-0.210***	-0.231***		0.188	0.078		-0.193***	-0.200***
		(0.052)	(0.052)		(0.276)	(0.287)		(0.033)	(0.033)
Government		0.266***	0.267***		-0.028	-0.035		0.274***	0.279***
		(0.080)	(0.081)		(0.129)	(0.132)		(0.056)	(0.056)
NGOs		0.29	0.269		0.153	0.089		0.276**	0.269**
		(0.246)	(0.242)		(0.222)	(0.224)		(0.132)	(0.133)
Other		0.246	0.161		0.291**	0.240*		0.699***	0.629***
		(0.305)	(0.299)		(0.145)	(0.135)		(0.238)	(0.236)
Enterprises		-0.322***	-0.366***		-0.102	-0.146*		-0.208***	-0.239***
		(0.067)	(0.070)		(0.086)	(0.086)		(0.046)	(0.046)
Ganyu		-0.031	-0.054		0.107	0.095		-0.001	-0.025
		(0.049)	(0.049)		(0.078)	(0.077)		(0.037)	(0.037)
2013			0.320***			0.311***			0.258***
			(0.034)			(0.052)			(0.023)
Constant	8.258***	8.251***	8.156***	9.005***	7.431***	8.774***	8.256***	8.150***	8.069***
	(0.047)	(0.107)	(0.107)	(0.074)	(0.276)	(0.40)	(0.030)	(0.076)	(0.076)
R-squared	0.104	0.183	0.204	0.394	0.417	0.429	0.163	0.231	0.244
N	6661	6661	6661	6661	6661	6661	6661	6661	6661

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Table 6: OLS and random effects results for log of earnings using education categories

Description	OLS			Random effects		
	Male	**Female**	**All**	**Male**	**Female**	**All**
Primary education	0.186***	0.132	0.162***	0.147***	0.162**	0.151***
	(0.066)	(0.105)	(0.060)	(0.047)	(0.071)	(0.039)
Secondary education	0.579***	0.810***	0.650***	0.583***	0.851***	0.667***
	(0.073)	(0.112)	(0.064)	(0.048)	(0.085)	(0.042)
Tertiary education	1.949***	2.163***	2.053***	1.885***	2.197***	2.007***
	(0.142)	(0.178)	(0.133)	(0.095)	(0.125)	(0.077)
Experience	0.052***	0.025***	0.040***	0.046***	0.024***	0.035***
	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)
Experience squared	-0.001***	-0.000***	-0.001***	-0.001***	-0.000***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female			-0.394***			-0.394***
			(0.031)			(0.026)
Central	-0.098	-0.196***	-0.113**	-0.038	-0.138***	-0.056*
	(0.060)	(0.062)	(0.051)	(0.042)	(0.053)	(0.033)
Southern	-0.229***	-0.365***	-0.270***	-0.191***	-0.319***	-0.233***
	(0.060)	(0.065)	(0.050)	(0.041)	(0.053)	(0.032)
Government	0.215**	0.246*	0.226***	0.240***	0.152	0.226***
	(0.084)	(0.135)	(0.074)	(0.062)	(0.107)	(0.054)
NGOs	0.01	0.347	0.133	0.131	0.247	0.203
	(0.193)	(0.426)	(0.232)	(0.139)	(0.250)	(0.126)
Other	0.078	0.118	0.137	0.612**	0.35	0.539***
	(0.330)	(0.310)	(0.248)	(0.303)	(0.256)	(0.195)
Enterprises	-0.095	-0.693***	-0.330***	0.034	-0.574***	-0.198***
	(0.083)	(0.111)	(0.071)	(0.054)	(0.082)	(0.045)
Ganyu	-0.028	-0.018	0.014	0.009	0.003	0.042
	(0.058)	(0.085)	(0.050)	(0.042)	(0.074)	(0.037)
2013	0.285***	0.368***	0.309***	0.238***	0.279***	0.249***
	(0.039)	(0.047)	(0.034)	(0.028)	(0.036)	(0.022)
Constant	8.499***	8.521***	8.636***	8.515***	8.502***	8.647***
	(0.099)	(0.121)	(0.086)	(0.072)	(0.105)	(0.058)
R-squared	0.188	0.234	0.229	0.239	0.289	0.279
N	3944	2717	6661	3944	2717	6661

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Table 7: Regression results on log of monthly earnings by economic sector

Description	Formal		Informal	
	OLS	Random	OLS	Random
Years of schooling	0.095***	0.100***	0.081***	0.096***
	(0.012)	(0.007)	(0.008)	(0.006)
Experience	0.033***	0.029***	0.042***	0.040***
	(0.009)	(0.006)	(0.004)	(0.004)
Experience squared	-0.001***	-0.000***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.142*	-0.123**	-0.453***	-0.466***
	(0.085)	(0.063)	(0.032)	(0.030)
Central	-0.02	0.012	-0.098*	-0.018
	(0.121)	(0.068)	(0.053)	(0.039)
Southern	-0.073	-0.047	-0.282***	-0.247***
	(0.123)	(0.067)	(0.054)	(0.037)
Government	0.230***	0.252***		
	(0.085)	(0.063)		
NGOs	0.226	0.265*		
	(0.229)	(0.142)		
Other	0.127	0.586***		
	(0.280)	(0.203)		
Ganyu			0.313***	0.210***
			(0.057)	(0.037)
2013	0.189**	0.134***	0.352***	0.302***
	(0.074)	(0.042)	(0.041)	(0.027)
Constant	8.041***	8.034***	7.850***	7.857***
	(0.220)	(0.135)	(0.116)	(0.086)
R-squared	0.223	0.281	0.152	0.170
N	1732	1732	4929	4929

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference category for region and year are Northern region and 2010. Base categories for employment are Private sector in formal sector and Enterprises in the informal sector.

Table 8: Regressions of log monthly earnings by gender and employees in both waves

Description	OLS		Random effects		Employed in both waves	
	Male	Female	Male	Female	OLS	Random
Years of schooling	0.090***	0.080***	0.099***	0.100***	0.099***	0.110***
	(0.008)	(0.009)	(0.006)	(0.007)	(0.008)	(0.005)
Experience	0.050***	0.025***	0.047***	0.024***	0.038***	0.034***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Experience squared	-0.001***	-0.000***	-0.001***	-0.000***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female					-0.372***	-0.369***
					-0.041	-0.037
Central	-0.065	-0.148**	0.001	-0.077	-0.025	0.039
	(0.058)	(0.065)	(0.043)	(0.054)	(0.058)	(0.045)
Southern	-0.192***	-0.327***	-0.163***	-0.278***	-0.171***	-0.136***
	(0.061)	(0.066)	(0.041)	(0.053)	(0.060)	(0.043)
Government	0.238***	0.302**	0.265***	0.256**	0.187*	0.201***
	(0.090)	(0.141)	(0.065)	(0.112)	(0.095)	(0.062)
NGOs	0.074	0.624	0.13	0.471*	0.271	0.265*
	(0.193)	(0.444)	(0.141)	(0.275)	(0.254)	(0.147)
Other	-0.033	0.333	0.598*	0.614*	0.009	0.347
	(0.354)	(0.414)	(0.349)	(0.337)	(0.291)	(0.225)
Enterprises	-0.100	-0.810***	0.016	-0.672***	-0.274***	-0.145***
	(0.082)	(0.107)	(0.054)	(0.088)	(0.084)	(0.054)
Ganyu	-0.068	-0.169*	-0.031	-0.140*	-0.053	-0.040
	(0.055)	(0.087)	(0.043)	(0.075)	(0.053)	(0.044)
2013	0.290***	0.378***	0.244***	0.288***	0.310***	0.259***
	(0.039)	(0.047)	(0.029)	(0.036)	(0.042)	(0.028)
Constant	7.969***	8.179***	7.901***	8.027***	8.054***	7.993***
	(0.127)	(0.131)	(0.094)	(0.127)	(0.126)	(0.098)
R-squared	0.169	0.200	0.209	0.243	0.229	0.272
N	3944	2717	3944	2717	4195	4195

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Table 9: Random effect results based on alternative treatment of outliers

Description	Random effects				
	Millionaires	**100th percentile**	**Robust regression**	**Extreme residuals**	**Outliers included**
Years of schooling	0.101***	0.090***	0.090***	0.100***	0.105***
	(0.005)	(0.004)	(0.004)	(0.005)	(0.005)
Experience	0.037***	0.033***	0.032***	0.036***	0.038***
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)
Experience squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Female	-0.392***	-0.377***	-0.373***	-0.396***	-0.394***
	(0.027)	(0.026)	(0.026)	(0.027)	(0.028)
Central	-0.009	-0.038	-0.033	-0.011	-0.015
	(0.034)	(0.032)	(0.032)	(0.034)	(0.035)
Southern	-0.192***	-0.218***	-0.217***	-0.200***	-0.202***
	(0.033)	(0.032)	(0.032)	(0.033)	(0.034)
Government	0.258***	0.288***	0.295***	0.279***	0.255***
	(0.056)	(0.051)	(0.051)	(0.056)	(0.059)
NGOs	0.276**	0.156	0.162	0.269**	0.246*
	(0.134)	(0.112)	(0.112)	(0.133)	(0.133)
Other	0.630***	0.604**	0.608***	0.629***	0.605**
	(0.237)	(0.236)	(0.233)	(0.236)	(0.236)
Enterprises	-0.238***	-0.268***	-0.242***	-0.239***	-0.236***
	(0.046)	(0.044)	(0.043)	(0.046)	(0.047)
Ganyu	-0.024	-0.036	-0.027	-0.025	-0.023
	(0.037)	(0.036)	(0.035)	(0.037)	(0.038)
2013	0.250***	0.248***	0.250***	0.258***	0.261***
	(0.023)	(0.022)	(0.022)	(0.023)	(0.023)
Constant	8.050***	8.180***	8.173***	8.069***	8.014***
	(0.077)	(0.072)	(0.071)	(0.076)	(0.079)
R-squared	0.237	0.229	0.234	0.244	0.245
N	6664	6608	6597	6661	6678

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Table 10: OLS results for Age-Period-Cohort effects on log of monthly earnings

Description	OLS estimation							
	10	11	12	13	14	15	16	17
Age	0.007***	0.083***	0.065***	0.062***	0.057***	0.057***	0.056***	0.064*
	(0.001)	(0.008)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)	(0.035)
Age squared		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of schooling			0.090***	0.083***	0.076***	0.076***	0.075***	0.075***
			(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Female				-0.407***	-0.375***	-0.375***	-0.377***	-0.379***
				(0.031)	(0.032)	(0.032)	(0.031)	(0.032)
Central				-0.06	-0.067	-0.067	-0.094*	-0.095*
				(0.052)	(0.052)	(0.052)	(0.051)	(0.050)
Southern				-0.237***	-0.250***	-0.250***	-0.273***	-0.280***
				(0.053)	(0.054)	(0.054)	(0.053)	(0.052)
Formal sector					0.219***	0.219***	0.245***	0.249***
					(0.046)	(0.046)	(0.046)	(0.046)
2013							0.311***	0.309***
							(0.034)	(0.038)
Constant	8.726***	7.459***	7.112***	7.542***	7.627***	7.627***	7.510***	7.269***
	(0.046)	(0.126)	(0.119)	(0.139)	(0.139)	(0.139)	(0.140)	(0.425)
Cohort effects	N	N	N	N	N	N	N	Y
R-squared	0.005	0.028	0.121	0.163	0.169	0.1686	0.189	0.198
N	6661	6661	6661	6661	6661	6661	6661	6661

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

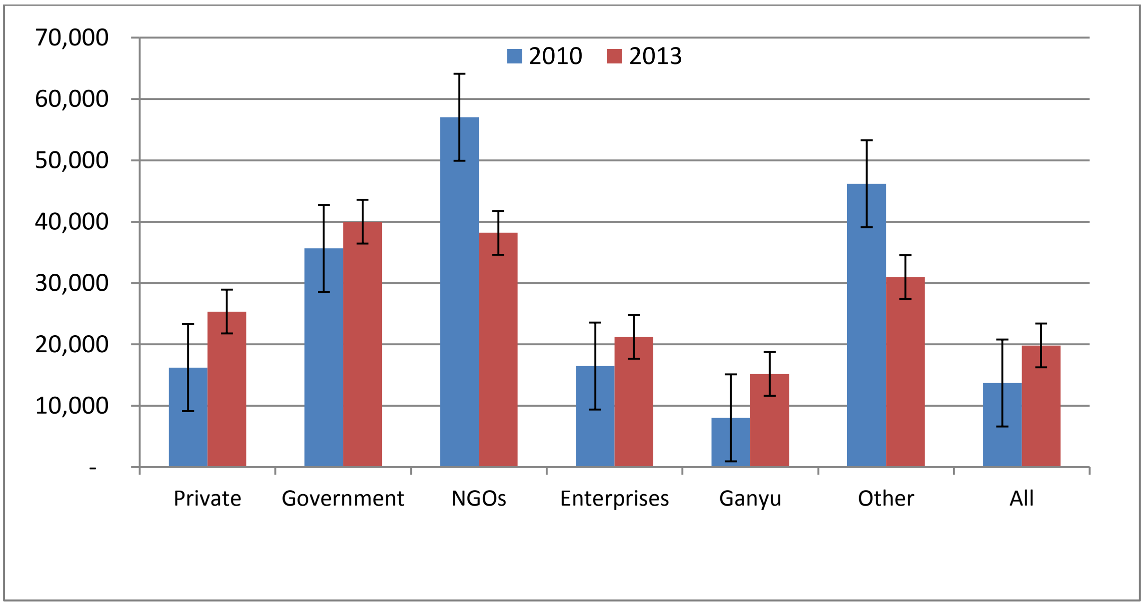
Table 11: Random effect results for Age-Period-Cohort effects on log of monthly earnings

Description	Random effects							
	18	19	20	21	22	23	24	25
Age	0.010***	0.087***	0.062***	0.059***	0.059***	0.056***	0.055***	0.052*
	(0.001)	(0.007)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.028)
Age squared		-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001***	-0.001*
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Years of schooling			0.099***	0.095***	0.094***	0.088***	0.087***	0.088***
			(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Female				-0.409***	-0.401***	-0.377***	-0.382***	-0.379***
				(0.027)	(0.027)	(0.027)	(0.027)	(0.027)
Central					-0.018	-0.023	-0.03	-0.03
					(0.034)	(0.034)	(0.034)	(0.034)
Southern					-0.213***	-0.220***	-0.228***	-0.233***
					(0.033)	(0.033)	(0.033)	(0.033)
Formal sector						0.179***	0.205***	0.206***
						(0.034)	(0.034)	(0.034)
2013							0.263***	0.271***
							(0.022)	(0.028)
Constant	8.740***	7.482***	7.122***	7.381***	7.480***	7.549***	7.430***	7.403***
	(0.042)	(0.108)	(0.103)	(0.104)	(0.106)	(0.108)	(0.108)	(0.346)
Cohort effects	N	N	N	N	N	N	N	Y
R-squared	0.011	0.033	0.180	0.211	0.218	0.2218	0.235	0.241
N	6661	6661	6661	6661	6661	6661	6661	6661

Notes: *, **, *** denote significance at 10%, 5% and 1% levels; reference categories for employment, region and year are Private sector, Northern region and 2010, respectively.

Charts

Chart 1: Histogram of real monthly earnings by survey year and occupation

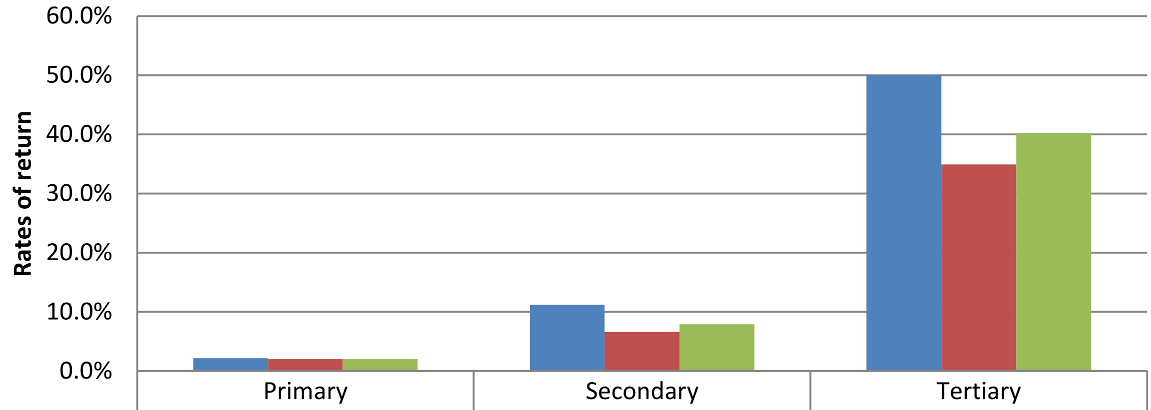


Source: Generated using IHPS data, bars represent standard errors

Chart 2: FISP evaluation panel data on real ganyu wages

Source: Farm Input Subsidy Programme (FISP) evaluation panel data

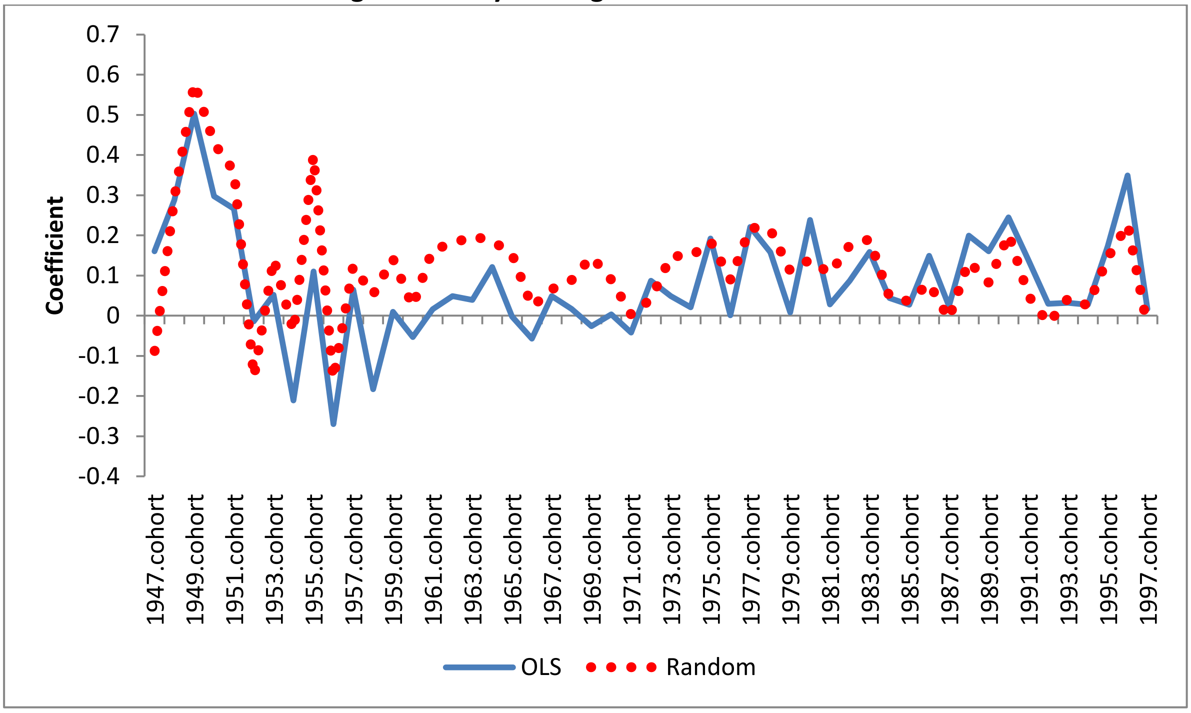
Chart 3: Rates of return on Education in Malawi, by gender



	Primary	Secondary	Tertiary
Female	2.2%	11.2%	50.0%
Male	2.0%	6.6%	34.9%
All	2.0%	7.9%	40.2%

Source: Generated using IHPS data. The omitted category is no education.

Chart 4: Cohort effects for log of monthly earnings



Source: Generated using IHPS data

References

Appleton, S., Bigsten, A. & Manda, D.K. (1999). Educational Expansion and Economic Decline: Returns to Education in Kenya, 1978-1995. *Working Paper WPS/99-6,* Centre for the Study of African Economies, University of Oxford.

Chirwa, E. & Matita, M. (2009). The Rate of Return on Education in Malawi. University of Malawi, Chancellor College, *Department of Economics Working Paper No. 2009/01*.

Chirwa, E.W. & Zgovu, E.K. (2002). *Does the Return to Schooling Depend on the Type of Employment? Evidence from the Rural Labour Market in Malawi*.

Da Maia, C. (2012). *Understanding Poverty and Inequality in Mozambique. The Role of Education and Labour Market Status.* Unpublished Doctoral of Philosophy (Economics) Dissertation, Stellenbosch University, South Africa.

Burger, R. & Yu, D. (2007). "Wage trends in post - Apartheid South Africa: Constructing an earnings series from household survey data". Working Paper 07/117, Development Policy Research Unit, University of Cape Town.

Kahyarara, G. & Teal, F. (2008).The Returns to Vocational Training and Academic Education: Evidence from Tanzania. *World Development*, Elsevier, vol. 36(11), 2223-2242.

Mincer, J. (1974). *Schooling, Experience and Earnings*. New York: National Bureau of Economic Research.

National Statistical Office (2014). *Integrated Household panel Survey 2010-2013: Household Socio-Economic Characteristics Report.* Zomba, Malawi.

National Statistical Office (2012). *Integrated Household Survey 2010-2011: Household Socio-Economic Report.* Zomba, Malawi.

Psacharopoulos, G. (1994) Returns to Investment in Education: A Global Update, World *Development*, 22 (9), 1325-1343.

Wittenberg, M. (2014). *Wages and wage inequality in South Africa 1994-2011: The evidence from household survey data.* SALDRU working papers 135, Southern Africa Labour and Development Research Unit, University of Cape Town.

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.

Yang, Y., W. J. Fu, & K. C. Land. 2004. *''A Methodological Comparison of Age-Period- Cohort Models: The Intrinsic Estimator and Conventional Generalized Linear Models.''* Sociological Methodology 34:75-110.

Yang, Yang & Kenneth C. Land. 2006. *''A Mixed Models Approach to the Age-Period-Cohort Analysis of Repeated Cross-Section Surveys: Trends in Verbal Test Scores.''* Sociological Methodology 36:75-97.