Education and Labour Earnings Inequality in Tanzania: Evidence from Quantile Regression Analysis

Cornel Joseph,* Vincent Leyaro‡ & Michael O. Ndanshau§

Abstract

This paper uses Tanzania's 2014 Integrated Labour Force Survey data to investigate the relationship between education and labour-earning inequalities. The quantile regression method is applied to compute returns to education at different points of the earnings distribution. The estimation result reveals significant variation in the coefficients of marginal returns to education across earning distributions, and the estimated coefficients are higher at the top of earning distribution. The marginal returns to education are higher for tertiary education than primary and secondary levels across all quantiles of the earnings distribution. The results also show that OLS coefficients conceal variation in the returns to education across the earning distribution. This finding suggests that education contributes positively to the widening of earnings dispersion in Tanzania, mainly due to the strong heterogeneous effects of education on earnings. Accordingly, it is vital to have a policy in Tanzania to reduce disparities in educational attainment between the least and most educated individuals.

Keywords: labour earning inequality, education, quantile regression, Tanzania

JEL: D24 J63 L63 O490

1. Introduction

Education plays a significant role in economic transformation. Evidence shows that many developing countries have used education as a policy tool for reducing poverty, addressing inequality and promoting the standard of living (Abdullah et al., 2015; Nabassaga et al., 2020). Likewise, education can help individuals to acquire new skills, raise their productivity, and promote career change toward well-paid jobs (Baye, 2015). Given an understanding of its importance, Tanzania has implemented various educational policies, programmes and plans, including the adoption of the Universal Primary Education (UPE) programme in 1977; a launch of the Primary Education Development Plan (PEDP) for the period 2002–2006; and the implementation of Secondary Education Development Plan (SEDP) during the period between 2004 and 2009. Moreover, the country implemented a Fee-Free Basic Education Policy (FFBEP) in 2002 and 2015 for primary and secondary education, respectively. The trio of policies, programmes and plans has significantly expanded educational opportunities in Tanzania, increasing the supply of educated labour. However, like in Tanzania, income inequality has been widening in most developing countries, despite efforts to narrow educational

*Department of Geography and Economics, Mkwawa University College of Education, Iringa, Tanzania: cornell.mlacha@udsm.ac.tz (Corresponding author)
‡University of Dar es Salaam School of Economics
§University of Dar es Salaam School of Economics

 ♦ **https://doi.org/10.56279/ter.v13i2.137**

inequality. The critical questions for research and policy interest are: Can education reduce earnings inequality *while* increasing individuals' earnings? To answer this question, we need to determine the contribution of education to labour-earning distribution. Specifically, we need to know whether education affects individuals differently across the earnings distribution.

In the literature, the human capital theory associated with Mincer (1958, 1996), Schultz (1960) and Becker (1964) explains labour earning inequalities as a consequence of differing human capital stocks that determine an individual's productivity. However, from an empirical point of view, analysing labour-earning inequalities is the subject of a critical methodological debate concerning the validity of using regression methods to analyse earning differentials. Previous empirical studies in Tanzania typically relied on regression analysis based on standard linear specification, thereby focusing mainly on the mean effects of schooling (Soderbom et al., 2006; Quinn & Teal, 2008; Islam et al., 2015). While it is of interest to note that the mean effects of schooling may mask much important information in the earnings distribution, this may not be informative as to the inequality-reducing effects of education (Wang, 2013). For example, if the effects are more pronounced in the upper than in the lower tail of the earnings distribution, education increases rather than decreasing inequality. For education to necessarily promote equality, it should increase earnings more for individuals in the lower tail of the earnings distribution. If the average effects of schooling were the only information available, it is unclear whether expanded educational opportunities will increase or decrease inequality. Moreover, in a distributional setting, the literature has not investigated whether different types of education result in differing returns; or whether one type of education—vocational or academic—brings a return premium compared to the other at some point in the wage distribution. These questions are critical because a lack of information about educational tracks may lead to costly decisions for both the individual and the government (Bettinger & Baker, 2011).

This paper investigates the effects of education on labour earnings in Tanzania using the quantile regression technique. The paper specifically tries to answer the question of: How does educational composition of the workforce in Tanzania influence the distribution of labour earnings? We compare the returns to one extra year of academic education with the returns to one extra year of vocational education to investigate whether one track brings a return premium at any point in the wage distribution. Such a comparison needs to be included in the literature. The findings of the paper, which are based on the Tanzania Integrated Labour Force Survey (ILFS) dataset of 2014, add value to the existing literature on labour earnings and education in Tanzania by applying the quantile regression technique to investigate returns to education across the earnings spectrum to establish whether some workers benefit more from education, and its implication on inequality. This allows us to shed light on heterogeneous returns to different levels of education, and to answer the question of how academic and vocational education differs over the wage distribution.

The quantile regression technique proposed by Firpo et al. (2009) is somehow superior to the OLS technique since it allows estimating the effect of the potential determinants on all parts of the earnings distribution, and enables us to gauge the different degrees of dispersion of earnings at each educational level. Moreover, the quantile regression technique is considered a proper technique in the presence of heteroscedasticity, and the regression coefficients are not sensitive to outlying values of the dependent variable (Fournier & Koske, 2013). Not least, the use of the quantile regression technique to analyse the effect of education across quantiles of the conditional earnings distribution sheds light on whether premiums to education and other earnings determinants are identical for low- and high-earning workers, in addition to allowing the establishment of whether education ameliorates or worsens existing inequalities. The results would also help policymakers better understand the role of education in determining labour earnings in Tanzania.

The paper is organised as follows. Apart from this introductory section, section 2 briefly reviews the related literature; section 3 presents the data and empirical methods employed; section 4 presents and discusses the empirical results; and section 5 concludes the paper.

2. Review of Related Literature

The theoretical literature of this paper is based on the human capital theory underpinned by Schultz (1961), Becker (1964) and Mincer (1974). The theory assumes that investment in education is necessary to acquire skills and training, which will in turn increase individual capital (Blundell et al., 1999). This knowledge and skills will increase one's productivity, hence bringing higher labour market earnings to an individual (Tan, 2014). Consequently, the level and distribution of schooling across the population determine the distribution of earnings (Becker & Chiswick, 1966; Mincer, 1974). Therefore, the theory predicts that the supply and demand of educated people influence earnings inequality in a society.

Empirical evidence in developing countries so far shows the positive effects of education on earnings, implying that returns to schooling are convex (Schultz, 2003; Psacharopoulos & Patrinos, 2004). However, the scope of the effects varies across studies in the literature. For example, a study by Söderbom et al. (2006) that examines returns to education in Kenya and Tanzania, using data on manufacturing employees for the 1993–2001 period, established the existence of convex returns to schooling. Also, a study by Kifle (2007) that estimated the private rate of returns to education using a sample of data from formal sector employees in Eritrea found that the marginal returns to education increased with the level of education. Moreover, Sackey (2008) made a study on private returns to schooling in Ghana using the living standard survey data for 1992 and 1999, which was fitted using the Ordinary Least Squares (OLS) technique. The study found that private returns to schooling at higher levels of education increased for both female and male workers.

Kahyarara (2013) examined the extent to which levels of education of a wage employee account for wage difference in a selected sample of workers in Kenya, Tanzania, Uganda, Madagascar, Ghana, Niger, Guinea Conakry, Rwanda, Benin and Togo. The study found a positive correlation between education and wages, and the marginal return to education was more remarkable in higher levels of education. Rizwanul et al. (2015) examine the determinants of labour income in Tanzania by using the Mincerian human capital model. The study used the Tanzania National Panel Survey (NPS) data, Wave II of 2011/12. The findings showed that education and experience positively influenced earnings for both males and females.

Twumasi-Baffour (2013) used quantile and OLS regression methods to examine the role of education in determining earnings in Ghana and Tanzania by using all three rounds of the Urban Worker Surveys of 2004–2006 for both countries. The quantile regressions for both Tanzania and Ghana suggested that primary and secondary levels of education were inequality-reducing among workers in Tanzania, but not in Ghana. Moreover, the study found that tertiary education widens earnings inequality both in Tanzania and Ghana. On the other hand, using three rounds of the Urban Worker Survey of Ghana for 2004–2006, Twumasi-Baffour (2015) examined the role of education in earnings determination in Ghana. The OLS and QR techniques were applied, and the findings showed that all levels of education were associated with earning premiums across quantiles with large returns to higher levels of education. Likewise, Twumasi-Baffour (2016) used the QR technique to investigate the effect of education on the earnings distribution of urban workers in the labour market in Ghana over the period 1998/99–2005/6. The findings showed that in 1998/99, except for secondary education, premiums to post-secondary and university education relative to primary university were highest at the second quantile (median) of the conditional earnings distribution. Whilst the returns to post-secondary and university education were lowest at the top quartile of the earnings distribution, secondary education had the lowest returns at the bottom quartile. However, using the 2005/06 sample, the results revealed a consistent pattern with higher premiums to all levels of education at the top quantile (75th) of the earnings distribution.

On their part, Leyaro et al. (2014) investigated the determinants of earnings of urban workers in Tanzania by using two datasets: the Integrated Labour Force Survey (ILFS) for 2000/01 and 2006; and the Urban Household Worker Survey (UHWS) for 2004, 2005 and 2006. The findings showed that returns to education increased with the level and years of education. Based on QR, the result suggested the existence of differential returns to education across the earnings distribution: primary and secondary educations were inequality-reducing, implying they were more beneficial to those with lower earnings, whereas tertiary education was inequality-increasing. Moreover, Kavuma (2015) examined the private marginal returns to education between wage employees and the self-employed in Uganda using the Mincerian framework with pooled regression models. The study used the data of two waves of a UNHS panel data (2005/06 and 2009/10). The result revealed the existence of a convexity between returns and levels of education attained. In addition, using the QR technique to investigate the heterogeneous returns to education, Kavuma (2015) found that returns to education decreased with quantile for all employment types examined.

Generally, the above literature survey reveal that the most recent empirical studies have been on the causal average effects of education on earnings. Consequently, more is needed to know about how education affects earnings distribution, particularly so in the case of Tanzania.

3. Data and Methodology

3.1 Data Sources

This study is based on the 2014 ILFS data for Tanzania collected by the Tanzania National Bureau of Statistics (NBS). The critical information collected by the survey is of two types: household and personal characteristics, and employment-related information. That information has important implications for earning determination. During the survey, individuals were required to report earnings from paid employment and self-employment (such as business and agriculture); and most individuals even reported weekly or monthly earnings. Individuals and households that reported earnings in ways other than paid or self-employment were dropped from the analysis. For those with weekly earnings, these were converted into monthly earnings to have a standard measure for all individuals. Therefore, this study is based on monthly earnings, considering hours worked. In terms of hours worked, individuals reported the number of hours they worked in the previous week, and the number of hours they usually worked. As a result of data limitations, we could not control for the quality of education in the analysis. Instead, the analysis is based on the assumption that the quality of schooling is the same across individuals and at all levels of education. The sample used in this study includes only individuals in the working age-group of 15–64 years. The sample size used consists of 11,724 individuals.

3.2 Methodology

3.2.1 Model Specification

The traditional human capital model is the standard theoretical framework for analysing the relationship between education and labour market earnings. This model implies that income distribution (or earnings) is determined by both the level of education (and schooling) and experiences in the labour market. Specifically, the model reads as:

$$ln\, E = \beta_0 + \beta_1 S + \beta_2 EX + \beta_3 EX^2 + \mu_i \qquad (1)$$

Where $ln\, E$ is the log monthly earnings, S is the number of years of schooling of an individual, and EX, EX^2 are potential years of experience (age-school-age started school) and its square, respectively. Experience square captures the declining effects of experience as individuals' age increases; and μ_i is a well-behaved stochastic error term.

Moreover, previous studies[1] have adopted the extended version of the Mincerian wage equation specified as:

[1] Among others, see Soderbom et al. (2006); Kahyarara, (2013); Kavuma et al. (2015); Falco et al. (2014)

$$ln\ E = \beta_0 + \beta_1 S + \beta_2 EX + \beta_3 EX^2 + Z_i\beta + \mu_i \qquad (2)$$

Where Z_i is a vector of control variables, including sex (takes the value of 1 if it is a male, and 0 otherwise); training (takes the value of 1 if an individual attended any training for at least a month, and 0 otherwise); sectors of employment dummies (for whether individual work for the public or private sectors, self-employment with or without employees, agricultural self-employment); log of weekly working hours; a dummy for marital status that takes a value of 1 if the labour is married, and 0 otherwise; locality (three dummies indicating whether the respondent works in Dar es Salaam, other urban, rural); and union (if an individual is a member of a trade union or not).

The S in equation (1) is at the centre of the analysis. The coefficient on years of schooling (β_1) represents the average private rate of return to one additional year of schooling (marginal returns to education), regardless of the level of education. Precisely, the coefficient of S (β_1) should capture the percentage change in earnings given a one-unit increase in the years of education.

It should be noted that the rate of return to an additional year of education (β_1) in (1) and (2) is constant across all levels of education. Noteworthy, however, is that available evidence from different parts of the world suggests that different school years impart different skills to workers and bring other returns (Schultz & Mwabu, 1998b; Nasir, 2002). Therefore, it is misleading to maintain constant rates of return (CCR) for all years of education. On this account, the model has been recast by converting the continuous years of schooling into a series of dummy variables, and by including additional variables in the estimation model. By this approach, the slope of the earnings function changes with different levels of education if there are significant differences in returns to education for each level.

The recast model has converted straight years of schooling variables *(S)* into dummy variables representing the different levels of education:

$$lnE_i = \gamma + \alpha_1 D_{Pr} + \alpha_2 D_{Sec} + \alpha_3 D_{Ad} + \alpha_4 D_{Te} + \beta_1 EX + \beta_2 EX^2 + Z_i\beta + \mu_i \qquad (3)$$

Where, *DPr* is the dummy for primary school education; D_{Sec} is the dummy for lower secondary school education; D_{Ad} is the dummy for upper secondary education; and D_{Te} is a dummy for tertiary education. Other variables are as already defined.

Equation (3) is commonly estimated using the OLS technique that focuses on the conditional mean effect on the regressand (Gujarati & Porter, 2009). However, for this paper, the OLS technique is used only for comparative purposes since it solely focuses on the mean, which produces under- or over-estimated results (Binder & Coad, 2010). Another weakness of using the OLS is the possibility of a correlation between the regressand, or one of the endogenous variables, and the model's error term, which could produce misleading results.

3.2.2 Estimation Strategy

We estimate the association of the concentration of human capital with individual earnings using the quantile regression technique. We utilise these regression models because they produce different effects along the distribution of the dependent variable, instead of estimating the model with mean effects (Buchinsky, 1998). Quantile regression gives a more comprehensive depiction of the impact of the independent variables on the explained variable than that of OLS. While this estimation technique does not account for the endogeneity problem present in the study, it does address the heterogeneity of the variable. Through this method, it is possible to observe the differences in the impact of education on the model across different quantiles of the earnings distribution. However, just like OLS, quantile regression would produce misleading results because, as previously mentioned, it does not address the self-selection bias of the variable.

Based on Koenker and Bassett (1978), quantile regression estimation is characterised by the minimisation of an equation that reads as:

$$Min_{\beta \varepsilon R^k} \sum\nolimits_{i:\varepsilon(y_t \geq x_t\beta)} \theta|y_t - x_t\beta| + \sum\nolimits_{i:\varepsilon(y_t < x_t\beta)} (1-\theta)|y_t - x_t\beta| \qquad (4)$$

Where it is the dependent variable, it is k by one vector of explanatory variables, β is a vector of coefficients, and i is the quantile to be estimated.

Following Bushnisky (1998), the quantile regression model of the earnings function is specified as follows:

$$ln\, w_i = x'_i\beta + \mu\theta_i \qquad (5)$$

$$Quant_\theta(ln\, w_i/x_i) = x'_i\beta_\theta; Quant_\theta(\mu\theta_i|x_i) = 0 \qquad (6)$$

Where lnE denotes monthly earnings, x is a vector of independent variables, and $u\theta$ is a random error term. The *i=1,.........., n,* is an index for individual workers, and n is the number of workers in the sample.

The vector of parameters denoted by β_θ and $Quant_\theta(ln\, w_i/x_i)$ is the θ^{th} conditional quantile of lnE given x_i. Since the quantile regression estimates minimise the absolute sum of the errors from a particular quantile of earnings across individuals, the problem is to obtain parameter estimates of the θ^{th} quantile regression in equation (6), which reads as:

$$Min\left\{\sum_{i:lnw_i \geq x'_i\beta_\theta} \theta|lnw_i - x'_i\beta_\theta| + \sum_{i:lnw_i < x'_i\beta_\theta} (1-\theta)|lnw_i - x'_i\beta_\theta\right\} \ldots\ldots\ldots..(7)$$

The median regression is when $\theta = 0.50$. Other quantile regressions are estimated by weighting the absolute sum of the errors. On the one hand, if laws $x'_i\beta_\theta$, then the deviation is positive, and θ is the weight used. On the other hand, when $lnw_i \geq x'_i\beta_\theta$, the deviation is negative, and the weight used is $1-\theta$.

The quantile regression method was used to estimate earning functions at three different percentiles of earnings distribution: the first quantile, the median and the third quantile of log monthly earnings. The OLS technique is nonetheless used, and the results are compared with those obtained by the quantile technique. Noteworthy is that, the OLS captures the effect of education and other covariates of an individual on the average earnings. At the same time, QR studies the determinants of earnings at other distribution points, for example, the bottom or top quartile. Estimating the model at different quantiles will enable us trace the entire conditional distribution of earnings given a set of regressors. After that, comparisons of the estimated returns (premiums) across the whole earnings distribution could help analyse the extent to which education raises or reduces existing inequalities. Another advantage of employing the quantile regression estimation method is that the coefficients are not sensitive to outlying values of the dependent variable (Twumasi-Baffour, 2013).

However, we must cautiously interpret the marginal returns to education from QR estimates since they do not control for endogeneity problems (Schultz & Mwabu, 1998a; Twumasi-Baffour, 2016). According to Mwabu and Schultz (1996), the errors in the quantiles may be heteroscedastic because of ability and education, or other covariates may not be independent, thus making the quantile regression variances biased. Therefore, the quantile regression understates standard errors (Twumasi-Baffour, 2015). We, therefore, utilise bootstrap estimates of the asymptotic variances of the quantile coefficients within 100 repetitions.

The primary variable used in the analysis are individual labour monthly earnings as the dependent variable, potential years of working experience (age-school-age started school), a series of dummies capturing education levels attained by individuals (primary, secondary and tertiary), and training (takes the value 1 if an individual attended any training for at least a month, and 0 otherwise) as a variable of the interest to capture the returns to education. We also include variables to capture socio-economic characteristics of labour, such as sex (takes the value of 1 if it is a male, and 0 otherwise); sectors of employment dummies (whether an individual works for the public or private sector, self-employment with or without employees, and agricultural self-employment); log of weekly working hours; a dummy for marital status (takes a value of 1 if the labour is married, and 0 otherwise); dummy variables for locality (Dar es Salaam, other urban, rural); and union (if an individual is a member of a trade union or not).

4. Empirical Results

Table 1 presents the overall OLS results in column 1; followed by quantile regression (QR) estimates of the earnings function for the 25th, 50th, and 75th quantiles, respectively. The results are as discussed below.

4.1 Education

Like previous studies (Mwabu & Schultz, 1996; Twumasi-Baffour, 2013; Leyaro et al., 2014; Twumasi-Baffour, 2016), the marginal returns to education increased considerably over the quantiles of the conditional distribution of earnings (Table 1).

Table 1: The OLS and Quantile Regression Estimates

Variables	OLS	1st Quantile (25th)	2nd Quantile (50th)	3rd Quantile (75th)
The levels of education-reference category are primary education				
Lower Secondary	0.398***	0.374***	0.349***	0.389***
	(0.022)	(0.026)	(0.024)	(0.029)
Upper secondary	0.690***	0.598***	0.614***	0.741***
	(0.066)	(0.052)	(0.061)	(0.090)
Tertiary	1.076***	0.986***	0.999***	1.136***
	(0.039)	(0.041)	(0.039)	(0.045)
Tvet	0.090***	0.096***	0.090***	0.127***
	(0.020)	(0.023)	(0.023)	(0.022)
Exper	0.026***	0.026***	0.028***	0.032***
	(0.004)	(0.004)	(0.004)	(0.004)
Expersq	-0.043***	-0.047***	-0.046***	-0.053***
	(0.007)	(0.007)	(0.007)	(0.008)
Sex	0.373***	0.368***	0.347***	0.355***
	(0.017)	(0.019)	(0.020)	(0.020)
Married	0.124***	0.090***	0.122***	0.121***
	(0.023)	(0.026)	(0.021)	(0.022)
Union	0.416***	0.482***	0.414***	0.367***
	(0.029)	(0.035)	(0.033)	(0.040)
Logwwh	0.353***	0.365***	0.326***	0.269***
	(0.025)	(0.028)	(0.028)	(0.034)
Casual	-0.172***	-0.110***	-0.180***	-0.215***
	(0.020)	(0.028)	(0.024)	(0.028)
Youth	-0.104***	-0.227***	-0.081**	0.012
	(0.033)	(0.040)	(0.038)	(0.034)
The regional dummies-reference category is Dar es Salaam				
Other urban	-0.322***	-0.295***	-0.291***	-0.298***
	(0.018)	(0.023)	(0.020)	(0.022)
Rural	-0.600***	-0.643***	-0.589***	-0.508***
	(0.030)	(0.044)	(0.035)	(0.039)
Status in the employment-reference category is agriculture				
Public	0.391***	0.519***	0.445***	0.324***
	(0.043)	(0.059)	(0.046)	(0.048)
Private	0.662***	0.992***	0.640***	0.357***
	(0.034)	(0.044)	(0.037)	(0.039)
Self-employed with	0.861***	0.969***	0.815***	0.785***
	(0.044)	(0.052)	(0.046)	(0.061)
Self-employed without	0.273***	0.410***	0.257***	0.193***
	(0.032)	(0.043)	(0.038)	(0.035)
Constant	9.653***	8.969***	9.798***	10.562***
	(0.108)	(0.120)	(0.110)	(0.143)
Observations	11,724	11,724	11,724	11,724
R-squared	0.401			

Notes: Dependent variable is the log of monthly labour earnings. For OLS regressions, robust standard errors are in parentheses, and for quantile regressions, bootstrapped standard errors using 100 replications are in parentheses, *** p<0.01, ** p<0.05, * p<0.1.

Source: Calculation based on ILFS data 2014.

Evidence of the convex relationship between labour earnings and education levels is found at all quantiles of the earnings distribution. Moreover, the results suggest that lower-secondary, upper-secondary or tertiary education is associated with labour earnings premium across all quantiles relative to primary education. This variation in the rates of return across quantiles can be interpreted as the composition effect of a change in the educational composition of the workforce. The highest premiums to all levels of education are most prominent at the top quantile (75^{th}) of the conditional earnings distribution. This suggests that education reduces earnings inequality over time, consistent with increasing education levels in the population. Thus, individuals with more abilities earn more from additional investment in education than those with lower abilities. The F-tests show that the coefficients of education dummies at different quantiles are significantly different at the 1 percent significance level.

4.2 Potential Experiences and Training

Potential experiences and training influenced the log of monthly earnings positively across quantiles. Table 1 indicates that potential labour market experiences and training positively correlated with a log of monthly earnings, and the magnitudes of the effect increases as the quantile increases. The returns to experience and training are more prominent at the upper end of the earnings distribution. This signifies that training and potential explanations are more prominent in highly-paid jobs than lower-paid occupations.

4.3 Sex and Marital Status

The QR estimated coefficients for male (sex) shows a positive effect on the log of monthly earning, and the magnitudes of the effect increase as the quantile increases, and the effect is premium at the 75^{th} quantile (Table 1). This confirmed the existence of discrimination where women earn less than men in the labour market. Moreover, the QR results indicate that the marital status (married) variable positively affects the log monthly earnings. The effect increases as the quantile increases, and the most significant coefficient is at the 75^{th} quantile.

4.4 Locality

Table 1 shows further that, relative to living in Dar es Salaam, living in rural and other urban areas has a negative impact on the log of monthly earnings, and the magnitudes of the effect decrease as the quantile increases. Thus the effect is much strong on workers in low-paid occupations.

4.4 Union Membership

The results in Table 1 further indicate that union is associated with earnings premium at all quantiles. Union memberships positively impact labour earnings; however, earnings benefits are disproportionately skewed toward lower-wage earners. This implies that union membership has higher labour-earning benefits at the lower tail of the earning distribution; thus, a bargaining power is much stronger for the workers on low-paid jobs. The results are consistent with the findings of Schultz and Mwabu (1998) in South Africa.

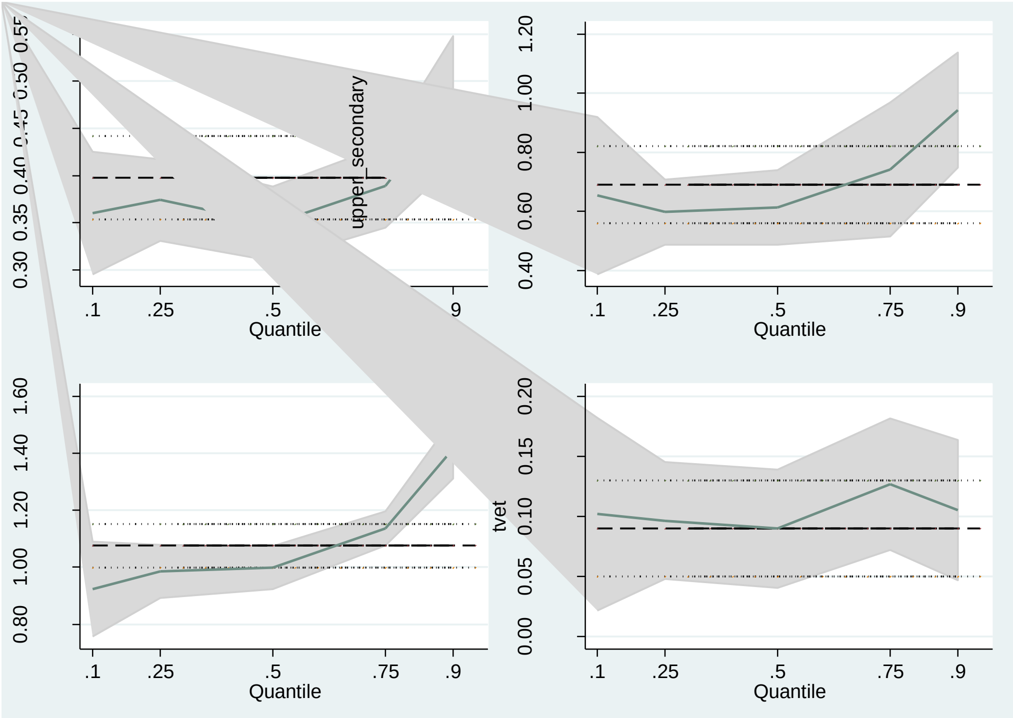
4.5 Weekly Working Hours
An important determinant of earnings inequality among the working population is the number of hours worked (generally captured by the number of hours worked per week in all jobs). The quantile regression result shows that the reward for working more is highest for workers at the lower end of the earnings distribution (Table 1). This could be due to differences in how time spent at work is recorded, such as lower-income earners may be more likely to benefit from overtime pay. In contrast, extra hours by middle and high-income earners may be compensated as a part of the basic remuneration package.

4.6 Type of Employment
The impact of employment types on labour earnings inequality is assessed by quantile regression specification, with a dummy variable for being self-employed (with and without employees); and working in the public sector and working in the private sector (working in agriculture as the reference category). The quartile regression results provide robust evidence that public, private and non-agricultural self-employment employees earn more relative to those in the agriculture sector (Table 1). The difference in earnings is huge for workers at the bottom of the earnings distribution. The magnitude of this earnings gap at the 75th quantile is relatively small (relative to the gap at the 25th quantile) in all employment types.

4.7 Comparisons of OLS and Quantile Regression Results
Column 1 in Table 1 shows further that education levels, as in the OLS estimates, increase earnings along the conditional earnings distribution, and the magnitudes of coefficients of education levels for QR estimates are more significant than the OLS estimates at the highest quantile (75^{th}). Furthermore, potential experiences and training positively influenced the log of monthly earnings in OLS and quantile regression (QR) estimates. However, the coefficients of training and potential labour market experiences for QR estimates are larger than those of OLS estimates. Similar to the OLS estimates, the QR estimated coefficients for male (sex) shows that they have a positive effect on the log of monthly earning, and the magnitude of the OLS estimate is larger than the QR estimates.

Likewise, both OLS and QR results in Table 1 indicate that the marital status (married) variable positively affects the log of monthly earnings. The OLS estimated coefficient (0.124) is larger than QR estimates across quantiles. Concerning locality, the results indicates that both OLS and QR estimates are lower for rural and other urban as compared to Dar es Salaam. Moreover, the magnitude in other urban coefficients is larger in OLS estimates, while that of rural ones is much larger in lower quantiles. Also, the union membership coefficient is positive and statistically significant in the OLS and QR estimates, but the QR estimate is higher than OLS estimates in the lower quantile. Similarly, like union membership, the magnitude of the weekly working hours coefficients is much higher in the lower quantile than in the OLS estimate. Nonetheless, both the OLS and quantile regression results showed that the log of earnings is higher for employees in public, private and non-agricultural self-employment earnings compared to those in the agriculture sector; and the magnitudes of coefficients are larger in lower quantiles than in the OLS estimates (Tables 1).

Finally, we show a comparison of OLS and quantile regression estimates for each level of education graphically in Figure 1. The figure confirms that both the mean and median regressions are different. The quantile regression estimates of each educational level lie outside the confidence intervals of the OLS regression. Quantile regression methods capture a large disparity along the wage distribution. In this manner, these are helpful over the OLS regression, which assumes identical returns to education in the same education group.



Note: Graphs made using the 'grqreg' Stata module (Azevedo, 2004). In each figure, the dashed (horizontal) line and the continuous line show the OLS and quantile regression estimates, respectively. The two dotted lines and the shaded region around the continuous line depict 95% confidence intervals for the two estimates.

Figure 1: Comparison of OLS and Quantile Regression Estimates of Education Levels

5. Conclusions

A quantile regression technique was applied to analyse the effects of education on earnings along different percentiles of the earnings distribution in Tanzania. The estimation results have shown much heterogeneity in returns to education since the marginal returns to education increased considerably over the quantiles of the conditional distribution of labour earnings. The finding suggests that educational improvements are essential to increase workers' earnings. Furthermore, the findings revealed that all levels of education increased with earnings at the mean and along the conditional earnings distribution. Consequently, this was evidenced by a convex relationship between earnings and education at all quantiles of the earnings distribution, where the highest premium to all levels of education was at the top quantile of the conditional earnings distribution.

The policy implication from the finding is that a policy promoting access to higher education for all can lower inequality. Thus, education policies to promote equality and support disadvantaged students in achieving better academic outcomes should be essential for promoting equality in society. Given the increase in returns to education at every level, it is also essential to reduce disparities in the levels of education attained between the least and the most educated. Investing in preventing dropouts and raising educational attainment could mitigate the labour-earning inequality in education.

References

Abdullah, A., Doucouliagos, H. & Manning, E. (2015). Does Education Reduce Income Inequality? a Meta-Regression Analysis. *Journal of Economic Surveys*, 29(2): 301–316.

Baye, F. M. (2015). *Impact of Education on Inequality Across the Wage Distribution Profile in Cameroon:* 2005–10 (No. 2015/014). WIDER Working Paper.

Becker, G. S. (1964). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education.* New York: Columbia University Press.

Bettinger, E. & Baker, R. (2011). The Effects of Student Coaching in College: An Evaluation of a Randomized Experiment in Student Mentoring. NBER Working Paper No. 16881. *National Bureau of Economic Research*.

Buchinsky, M. (1998). Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research. *Journal of Human Resources*: 88–126.

Falco, P., Kerr, A., Pierella, P., Paci, P. & Rijkers, B. (2014).Working Toward Better Pay: Earnings Dynamics in Ghana and Tanzania. World Bank Publications.

Fournier, J. M. & Koske, I. (2013). The Determinants of Earnings Inequality. *OECD Journal: Economic Studies*, 2012(1): 7–36.

Kahyarara, G. (2013). Education and Wage in East and West Africa. *Revue d'Economie Théoriqueet Appliquée ISSN*, 1840, 7277.

Kavuma, S. N., Morrissey, O. & Upward, R. (2015). Private Returns to Education for Wage-Employees and the Self-Employed in Uganda. WIDER Working Paper, (No. 2015/021).

Kifle, T. (2007). The Private Rate of Return to Schooling: Evidence from Eritrea. *Essays in Education*, 21: 77–99.

Koenker, R. & Bassett Jr, G. (1978). Regression Quantiles. *Econometrica:* 46: 33–50.

Leyaro, V., Twumasi-Baffour, P., Morrissey, O. & Owens, T. (2014). Determinants of Urban Labour Earnings in Tanzania. CREDIT Research Paper, 2000/01–06 (No. 14/03).

Mincer, J. (1974). Schooling, Experience, and Earnings. Human Behavior & Social Institutions No. 2.

Mwabu, G. & Schultz, T. P. (1996). Education Returns Across Quantiles of the Wage Function: Alternative Explanations for Returns to Education by Race in South Africa. The *American Economic Review*, 86(2): 335–339.

Nabassaga, T., Chuku, C. A., Mukasa, A. N. & Amusa, H. A. (2020). *How Does Educational Inequality Affect Income Inequality in Africa?* African Development Bank.

Nasir, Z. M. (2002). Returns to Human Capital in Pakistan: A Gender Disaggregated Analysis. The *Pakistan Development Review*: 1–28.

Psacharopoulos, G. & Patrinos, H. A. (2004). Returns to Investment in Education: A Further Update. *Education Economics*, 12(2): 111–134.

Quinn, S. & Teal, F. (2008). *Private Sector Development and Income Dynamics: A Panel Study of the Tanzanian Labour Market*. Centre for the Study of African Economies, No. 2008–09: University of Oxford.

Rizwanul, I., Abel, K. & Joseph, N. (2015). Real Wages and Labour Productivity in Tanzania: How Do they Link? *Journal of African Studies and Development*, 7(3): 81–98.

Sackey, H. A. (2008). *Private Returns to Education in Ghana: Implications for Investments in Schooling and Migration*. African Economic Research Consortium, Research Paper No.174.

Schultz, T. P. (2003). Human Capital, Schooling and Health. *Economics & Human Biology*, 1(2): 207–221.

Schultz, T. P. & Mwabu, G. (1998a). Labor Unions and the Distribution of Wages and Employment in South Africa. *ILR Review*, 51(4): 680–703.

Schultz, T. P. & Mwabu, G. (1998b). *Wage Premia for Education and Location by Gender and Race in South Africa*. Center Discussion Paper, No. 785.

Schultz, T. W. (1961). Investment in Human Capital. *American Economic Review*, 51(1): 1–17.

Soderbom, M., Teal, F., Wambugu, A. & Kahyarara, G. (2006). The Dynamics of Returns to Education in Kenyan and Tanzanian Manufacturing. *Oxford Bulletin of Economics and Statistics*, 68(3): 261–288.

Twumasi-Baffour, P. (2013). Determinants of Urban Worker Earnings in Ghana and Tanzania: The Role of Education. CREDIT Research Paper, No 13/01.

Twumasi-Baffour, P. (2015). Determinants of Urban Worker Earnings in Ghana: The Role of Education. *Modern Economy*, 6: 1240–1252.

Twumasi-Baffour, P.(2016). Education and Earnings Inequality in Ghana. *Modern Economy, 7:* 456–469.

Wang, L. (2013). How Does Education Affect the Earnings Distribution in Urban China? *Oxford Bulletin of Economics and Statistics*, 75(3): 435–454.