Education and the Distribution of Earnings in Kenya: Evidence from Quantile Regression Analysis

Anthony Wambugu
School of Economics
University of Nairobi
Nairobi, Kenya

Abstract

The paper used matched employer-employee data to investigate whether there are asymmetries in the effects of observed dimensions of human capital, such as education and work experience on the conditional earnings distribution of manufacturing sector workers in Kenya. Ordinary least squares method was used to estimate the conditional expected earnings, while quantile regression method was used to estimate the earnings function at different points along the conditional earnings distribution. The regression estimates show that the earnings premium associated with education and experience varies across the conditional earnings distribution. This suggests that education and experience contribute to widening the earnings dispersion in the Kenya manufacturing sector labor market.

JEL Classification: J3; O1

Keywords: Quantile regression; Earnings, Return to education; Africa; Kenya

1. Introduction

Understanding the role of human capital in explaining individual earnings differences is a key theoretical and empirical topic in labor economics (Borjas, 2007; Cahuc and Zylberberg, 2004). An large empirical literature (see Bennell, 1996; Psacharopoulos and Patrinos, 2004; Kingdon and Patrinos, 2010) estimates the mean earnings premium for education in African labour markets based on Becker (1964) and Mincer(1974) human capital theory. Although the earnings premium for education can be heterogeneous (Card, 1995) with implications for earnings inequality, this aspect has not received as much attention as the mean return to education. For example, Martins and Pereira (2004) found that earnings premium for education increase over the earnings distribution for 15 European countries. In contrast, Arulampalam, Booth and Bryan (2010) found that earnings premium for on-the-job training in 10 European countries do not vary over the earnings distribution. Such evidence is scarce for African labour markets. The purpose of this study is to investigate whether or not there are asymmetries in the earnings premium for education and work experience and other earnings determinants in Kenya. The objective is to apply quantile regression analysis (Bushnisky, 1994, 1998) to investigate whether real earnings and earnings determinants differ across the conditional earnings distribution. The paper used a unique dataset containing both firm-level and individual worker-level variables in estimating earnings equations. The paper proceeds as follows. Section 2 describes the data used. Section 3 presents the estimation strategy involving ordinary least squares and quantile regression. The descriptive statistics and estimation results are presented in Section 4. Section 5 concludes.

2. Earnings function

The earnings function is widely used to examine individual wage differences. Assume the earnings equation

$$\ln w_i = x_i^{'}\beta + \varepsilon_i \qquad (1)$$

where ln*w* denotes the natural logarithm of hourly earnings (in 1990 Kenya shillings)., *x* is a vector of covariates including workers' age in years, square of age, years of education, square of years of education, years worked in the firm, square of years worked in the firm. Dummy for male workers, and four dummy variables for survey waves (wave 1 dummy is the benchmark). Quadratic variables in education, age, and job tenure in the firm allow for non-linear effect on earnings. The β is a vector of parameters, and $\varepsilon \sim N(0,\sigma^2)$ random error terms.

The quantile regression model (Buchnisky, 1994, 1998) is expressed as

$$\ln w_i = x_i^{'}\beta_\theta + u_{\theta_i} \quad (2a)$$

$$Quant_\theta(\ln w_i \mid x_i) = x_i^{'}\beta_\theta ; Quant_\theta(u_{\theta_i} \mid x_i) = 0 \quad (2b)$$

w and x are defined as in (1) and u_θ is a random error term. The parameter vector is denoted by β_θ and $Quant_\theta(lnw_i|x_i)$ is the θth conditional quantile of *lnw* given x_i. The θth conditional quantile regression parameters are obtained by minimizing the absolute sum of the errors from a particular quantile of the log earnings across workers. That is

$$Min\left\{\sum_{i:\ln w_i \geq x_i^{'}\beta_\theta} \theta \mid \ln w_i - x_i^{'}\beta_\theta \mid + \sum_{i:\ln w_i < x_i^{'}\beta_\theta} (1-\theta) \mid \ln w_i - x_i^{'}\beta_\theta \mid\right\} \quad (3)$$

If $\theta = 0.50$, this gives the least absolute deviation (LAD) estimator. Other conditional quantile earnings functions are estimated by assigning different weights to the absolute residuals. When $\ln w_i \geq x_i^{'}\beta_\theta$ the residual is positive and the weight is θ. When $\ln w_i < x_i^{'}\beta_\theta$ the residual is negative and the weight is $1-\theta$. Linear programming methods are used to solve the minimization problem (3).

The quantile regression approach permits the impact of earnings function covariates to be investigated across quantiles. For example, the education coefficient at the lower quartile shows the schooling effect for workers at the lowest 25% of the conditional earnings distribution. Estimates at the median show the schooling effect for workers at the middle, and estimates at the upper quartile show schooling effect for workers in the top 25% of the conditional earnings distribution. Moreover, since quantile regression estimates minimize the absolute sum of errors, they are relatively more resistant to outliers than OLS estimates (Deaton, 1997 and Buchnisky, 1998). To test cross-quantiles restrictions on covariates, simultaneous conditional quantile earnings equations are estimated. For example, is the effect of education, experience or gender on earnings identical across quartiles?

3. Data

The paper analysed four waves of a firm-level survey in Kenya collected under the World Bank's Regional Program on Enterprise Development (RPED)(1993, 1994, and 1995 waves) and under the United Nations Industrial Development Organization (UNIDO)(2000 wave). The firms are located in the capital, Nairobi; Mombasa, the main seaport and two inland urban centers (Nakuru, and Eldoret). The firms are spread across four main sub-sectors (wood, textiles, food, and metal) that comprise about 73 per cent of manufacturing employment. Formally registered firms constitute 75 per cent of the primary sample while 25 percent are informal sector firms. A total of 224 firms were interviewed in 1993, 216 in 1994, 218 in 1995, and 190 in 2000. In waves two, three, and four, some firms were replaced because they had closed down, declined to be interviewed, or could not be retraced. The survey collected firm-level information on production, investment, finance, investor confidence, labor, and infrastructure from the manager or another senior person in the firm. In addition, information on occupation wages, non-wage benefits, tenure in current firm, and individual characteristics and pay was collected from ten workers

4. Empirical results

4.1 Descriptive Statistics

Appendix Table A1 summarizes, for each wave and total sample, data on variables used in the analysis. Most of the workers (85 per cent) are men. The mean age is 34 years and the mean job tenure is 9 years and most (58 per cent) are in production department. Most workers have either completed primary (43 per cent) or secondary (41 per cent) education. The average years of education completed is 9. The rise in average years of education over the survey period is probably because retiring workers have less education, while new entrants have more education. The total years of education a worker spent in school could be understated because of grade repetition. The largest proportion (23 per cent) of workers is in the metal sector and the smallest proportion (4 per cent) is in machinery sector. Most workers (65 per cent) are in firms located in Nairobi, which reflects concentration of manufacturing there. The mean firm size (number of workers) is 152 but the standard deviation is large since the sample includes very small and very large firms. Output per worker (real value of output in Kenya shillings (1990=100/total number of workers) in 2000 is 1.3 times the value in 1993, while the average capital per worker (real replacement value of plant and equipment in Kenya shillings (1990=100)/total number of workers) is 1.8 times that in 1993.

Table 1 summarizes monthly and hourly earnings in Kenya shillings (1990=100) by education level and survey wave. Earnings include the basic wage plus allowances (e.g. cash allowance for food, transport, and housing) and production and Christmas bonuses.

Table 1: Mean and percentile real earnings in Kenya shillings by education and survey wave

Education and survey wave	N	Mean	P25	P50	P75
Below primary education					
Wave 1	186	1128 (6)	742 (4)	955 (5)	1256 (7)
Wave 2	154	1036 (5)	739 (4)	881 (4)	1110 (6)
Wave 3	137	1526 (9)	975 (6)	1346 (8)	1600 (9)
Wave 4	81	1624 (9)	1125 (5)	1525 (8)	1845 (10)
Pooled	558	1272 (7)	799 (4)	1038 (6)	1477 (8)
Earnings ratio (wave 4/wave 1)		1.44 (1.50)	1.52 (1.25)	1.60 (1.60)	1.47 (1.43)
Full primary education					
Wave 1	492	1383 (7)	792 (4)	1004 (5)	1438 (7)
Wave 2	419	1236 (6)	685 (3)	940 (5)	1397 (7)
Wave 3	472	1553 (9)	966 (6)	1299 (8)	1699 (10)
Wave 4	385	2687 (14)	991 (5)	1398 (7)	1876 (9)
Pooled	1768	1678 (9)	816 (4)	1163 (6)	1611 (9)
Pay ratio (wave 4/wave1)		1.94 (2.00)	1.25 (1.25)	1.39 (1.40)	1.30 (1.29)
Pay ratio (Primary/below prim)		1.32 (1.29)	1.02 (0.00)	1.12 (0.00)	1.09 (1.13)
Full secondary education					
Wave 1	417	2339 (12)	935 (5)	1485 (7)	2722 (14)
Wave 2	379	1719 (9)	822 (4)	1182 (6)	1861 (10)
Wave 3	419	2457 (14)	1109 (7)	1535 (9)	2343 (14)
Wave 4	473	2925 (18)	1246 (6)	1982 (10)	3398 (18)
Pooled	1688	2393 (13)	1025 (5)	1514 (8)	2591 (13)
Pay ratio (wave 4/wave1)		1.25 (1.50)	1.33 (1.20)	1.33 (1.43)	1.25(1.29)
Pay ratio (Secondary/Primary)		1.43 (1.44)	1.26(1.25)	1.30 (1.33)	1.61 (1.44)
University education					
Wave 1	9	6489 (30)	3464 (18)	5939 (31)	8909 (43)
Wave 2	14	5360 (26)	2397 (12)	4081 (21)	5822 (31)
Wave 3	35	7958 (46)	2498 (15)	5437 (32)	11315 (66)
Wave 4	43	11111 (55)	4320 (21)	6764 (32)	12460 (64)
Pooled	101	8809 (46)	3464 (18)	5556 (30)	10732 (57)
Pay ratio (wave 4/wave1)		1.71 (1.83)	1.25(1.17)	1.14 (1.03)	1.40 (1.49)
Pay ratio (University/secondary)		3.68 (3.53)	3.38 (3.60)	3.67 (3.75)	4.14 (4.38)

Source: Author's computation from survey data.

Note: N is number of observations and P denotes percentile. Hourly earnings and pay ratios are within parentheses. Monthly earnings and pay ratios are outside the parentheses.

Three points emerge from Table 1 about earnings. First, earnings in wave four were higher than in wave one for every education group. Second, for every education group, earnings are higher in the 75th percentile than at the 25th percentile. And third, the educational earnings differential between two adjacent education groups is greater at 75th percentile than at 25th percentile.

4.2 Log earnings Regression Estimates

All or part of the increase in raw wage differentials uncovered in (4.1) may reflect differences in characteristics of workers and firms. The Mincerian earnings function was estimated on pooled data controlling for characteristics of workers and firms. The results are in Table 2.

Table2. Ordinary Least Squares and Quantile Regression Estimates (Pooled data)

Explanatory variable	OLS	OLS-FE	P25	P50	P75
Age (years)	0.06***	0.04***	0.05***	0.06***	0.05***
	(4.62)	(5.68)	(6.41)	(5.25)	(4.93)
Age squared/100	-0.05***	-0.04***	-0.05***	-0.06***	-0.04***
	(3.07)	(3.51)	(4.83)	(3.55)	(2.76)
Education (years)	-0.09***	-0.07***	-0.08***	-0.10***	-0.11***
	(5.40)	(5.42)	(4.11)	(5.80)	(5.63)
Education squared/100	1.12***	0.83***	0.88***	1.15***	1.34***
	(11.19)	(12.15)	(7.85)	(11.02)	(11.20)
Tenure with firm (years)	0.01	0.001	0.01***	0.004	0.01
	(1.60)	(0.99)	(2.60)	(0.72)	(1.33)
Tenure squared/100	-0.001	0.001	-0.01	0.01	0.001
	(0.16)	(0.22)	(0.68)	(0.60)	(0.07)
Male worker	0.02	-0.003	0.09***	0.05	0.02
	(0.51)	(0.12)	(2.61)	(1.59)	(0.50)
Wave 2	-0.10***	-0.11***	-0.09***	-0.09***	-0.09**
	(2.95)	(4.58)	(3.08)	(3.41)	(2.03)
Wave 3	0.30***	0.33***	0.37***	0.36***	0.31***
	(8.49)	(13.99)	(13.18)	(13.17)	(9.03)
Wave 4	0.30***	0.35***	0.30***	0.34***	0.31***
	(6.63)	(11.41)	(7.82)	(9.76)	(7.56)
Constant	0.20	0.68***	0.07	0.25	0.53***
	(0.88)	(4.83)	(0.50)	(1.26)	(3.28)
Adjusted R^2 [Pseudo R^2]	0.35	0.28	[0.17]	[0.20]	[0.23]

Note:

a) The dependent variable is the natural logarithm of hourly earnings in Kenya shillings (1990 = 100).

b) OLS standard errors are Huber-White corrected for heteroskedasticity and adjusted for intra-firm correlation in residuals due to workers clustering in firms. Quantile regression standard errors are bootstrapped.

c) T-statistics within parentheses. Significance: * significant at 10%; ** significant at 5%; *** significant at 1%

Column 1 presents OLS estimates of the effect of human capital variables, gender, and year dummies on log hourly earnings. Column 2 presents OLS estimates controlling for unobservable firm fixed effects (FE). Columns 3, 4, and 5 present quantile regression estimates. The coefficients of the education variables are significant and the sign pattern suggests a convex education-earnings profile at the three quantiles. In addition, the earnings premium to education at every education level is higher at the upper quantile than at the lower quantile. For workers with 6 years of schooling, the return (multiplied by 100) range from 3% at the first quartile to 5% at the third quartile. The mean return (OLS) is 4%. For workers with 10 years of schooling, returns vary from 10% at the first quartile to 16 per cent at the third quantile with mean returns being 13%. With 14 years of educational returns range from 16% at the first quantile to 27% at the third quantile. F-test of equality of education coefficients at the 25th, 50th, and 75th percentiles has $F_{4,\ 4104} = 22.89$(p-value= 0.00). Thus, earnings premium to education across quantiles are unlikely to be equal.

Age and age squared variables have significant effect on earnings across quantiles and the sign pattern suggests age-earnings profiles are concave. The coefficients of job tenure variable are generally insignificant except in the first quantile. This could indicate seniority wages or return to firm-specific skills (Borjas, 2007). There is no evidence of a gender earnings gap in the OLS regression. However, men in the first quartile earn $100*(e^{0.09}-1) = 9\%$(see Halvorsen and Palmquist, 1980) more than otherwise comparable women in the same quantile. This could indicate productivity differences between men and women or labor market discrimination among low wage workers.

The coefficients on year dummy variables are significant. The coefficient on year 2 dummy implies that holding other variables constant, real wage was lower than in year 1. In contrast, the coefficients of year 3 and year 4 dummies imply higher earnings than in year 1. For example, the OLS coefficient on year 4 dummy is 0.30, which implies that average percentage change in real hourly earnings over the period is 35 per cent ($100*(e^{0.30}-1)$), which works out to 5 per cent per year. When unobservable firm fixed effects are included, the absolute values of the wave dummy coefficients increase.

The implied change in real hourly earnings over the period is 42 per cent or 6 per cent per year. The quantile regression estimates indicate that changes in real hourly earnings were not uniform across the earnings distribution. The percentage increase among low wage workers (first quartile) and high wage workers (third quartile) is similar to the average increase (35 per cent), the percentage increase among middle (median) wage workers is higher (40 per cent). An F-test for equality of the wave dummy coefficients at 25th, 50th, and 75th percentiles has $F_{6, 4104} = 2.28$(p-value =0.03). So the hypothesis of equality is rejected. The earnings increases computed from the conditional earnings equations are lower than the increases derived from the raw data. This indicates that part of the observed increase in earnings is due to individual differences in human capital.

Adding control variables in the earnings function does not reduce the effect of human capital. Appendix Table A2 reports the estimates controlling for firm size (number of workers), labor productivity (output per worker), capital intensity (capital per worker), and dummies for sector, firm location and occupations. The OLS estimates are reported in column 1 and column 2. Some of the controls may be endogenous to wage formation. We lack suitable instruments. So the aim is to explore how the educational structure of earnings in the Mincerian earnings function change when the controls are included. Useful hints about other wage determination models can also be obtained.

Age and education effects remain significant even with the controls. Earnings increase with firm size. Higher average labor productivity is associated with higher earnings, consistent with efficiency wage or rent-sharing considerations in wage setting. Earnings in every occupation are higher than those of production workers and workers in firms located in Mombasa, Eldoret, and Nakuru receive lower earnings than comparable workers in firms located in Nairobi. There are sectoral wage differentials also, which may reflect unobserved factors or compensating differentials for working conditions. The implied increase in real hourly earnings over the survey period across the conditional earnings distribution is about 36 per cent. This estimate is not substantially different from that derived from earnings regression with only controls for human capital despite the large number of controls in the extended earnings function.

The parameter estimates of the earnings function estimated to examine change in effect of human capital and other earnings determinants on earnings in different parts of the conditional earnings distribution across survey waves are presented in Table 3. In addition to education, age, tenure and their squares, gender and time dummies; the human capital variables and gender dummy are interacted with survey wave dummy variables. A joint F-test was conducted on the standard earnings function to determine whether the earnings function was stable across the survey waves. The $F_{24, 358} = 9.94$(p-value = 0.00). The null hypothesis that the estimates do not vary across waves may be rejected. Thus the earnings structure shifted over the survey period. A closer look at the log earnings regressions indicates that very few of the coefficients on interactions between the covariates and wave four dummy are significantly different from zero. With education specified as a quadratic, the interaction between the square of years of education and wave four dummy is positive and significant in the first quartile. This suggests that the non-linear impact of education on log earnings was significantly stronger in wave four than in wave one. In general, the earnings equation indicates stable education-earnings relationship over the period.

Table 3. Earnings equations including interaction variables (quadratic education function)

Explanatory variable	OLS	FE	P25	P50	P75
Male worker	0.04	0.002	0.13	0.02	-0.02
	(0.67)	(0.04)	(1.56)	(0.39)	(0.23)
Age (years)	0.04*	0.03**	0.05***	0.04**	0.04
	(1.67)	(2.48)	(2.69)	(1.97)	(1.34)
Age squared/100	-0.01	-0.02	-0.05*	-0.02	-0.005
	(0.41)	(0.89)	(1.78)	(0.62)	(0.15)
Education (years)	-0.06**	-0.07***	-0.06***	-0.06***	-0.04
	(2.46)	(2.85)	(4.39)	(3.33)	(1.47)
Education squared/100	1.04***	0.92***	0.80***	0.96***	1.08***
	(6.61)	(6.55)	(5.93)	(7.60)	(6.79)
Tenure in firm/10	0.05	-0.02	0.08	-0.02	0.01
	(0.54)	(0.25)	(0.68)	(0.22)	(0.12)
Tenure squared/100	-0.02	0.001	-0.04	0.001	-0.02
	(0.55)	(0.08)	(0.69)	(0.00)	(0.63)
Wave4*Male worker	-0.09	0.001	-0.04	0.02	0.02
	(0.98)	(0.07)	(0.50)	(0.15)	(0.20)
Wave4*Age (years)	0.03	0.02	-0.01	0.03	0.02
	(1.21)	(1.05)	(0.28)	(0.95)	(0.44)
Wave4*Age squared	-0.001	-0.001	0.001	-0.001	-0.001
	(1.41)	(1.27)	(0.07)	(1.02)	(0.62)
Wave4*Education	-0.04	0.001	-0.10	-0.05	-0.02
	(0.99)	(0.14)	(1.60)	(1.06)	(0.33)
Wave4*Education squared	0.001	-0.001	0.01**	0.001	0.001
	(1.27)	(0.04)	(2.31)	(1.25)	(0.51)
Wave4*Tenure in firm	0.01	0.001	0.01	0.02	0.01
	(0.64)	(0.40)	(0.44)	(0.96)	(0.87)
Wave4*Tenure squared	0.001	0.001	0.00	-0.001	0.001
	(0.32)	(0.56)	(0.70)	(0.49)	(0.22)
Constant	0.38	0.73***	-0.06	0.43	0.51
	(1.04)	(2.95)	(0.15)	(1.20)	(1.11)
Adjusted R^2 [Pseudo R^2]	0.36	0.29	[0.18]	[0.20]	[0.24]

Note:

a) The dependent variable is the natural logarithm of hourly earnings in Kenya shillings (1990 = 100).
b) Interactions of the explanatory variables with dummies for waves 2 and 3 included but are not reported.
c) OLS standard errors are Huber-White corrected for heteroskedasticity. Quantile regression standard errors are bootstrapped.
d) Values of t-statistics are in parentheses, * significant at 10%; ** significant at 5%; *** significant at 1%

Many studies examine wage premium to education and experience for African labor markets, but few extend to the whole wage distribution and the few use a single cross-sectional data set (e.g. Mwabu and Schultz, 1996; Girma and Keddir, 2005). Therefore, empirical evidence of changes in wage premium across the conditional earnings distribution for African labor markets is scarce. In the U.S.A. Buchnisky(1994) found that in the 1960s and early 1970s, returns to education and experience were higher for workers in top deciles of the conditional earnings distribution. But in the second half of the 1970s returns fell and became compressed across quantiles. In the 1980's, returns to education increased sharply especially for workers in top deciles. Nielsen and Rosholm (2001) used quantile regressions to examine public-private wage differential in 1991-1996 in Zambia. Results show that in the bottom decile, the earnings of less educated public sector workers rose more than those of private sector workers. But in the top deciles the earnings advantage to highly educated workers in public sector narrowed. Also, wage premia for education were higher in the private sector than in the public sector and varied across quantiles.

5. Summary and Conclusion

This paper examined changes in labor market earnings and earnings premia to human capital across the earnings distribution, during a period of economic liberalization in Kenya.The manufacturing sector was one of the key sectors affected by the reforms. Using four waves of an enterprise survey in Kenya the analysis indicates that, controlling for human capital, real earnings rose by 5 to 6 per cent per year across the conditional earnings distribution. The result is surprising. In the 1990s Kenya experienced slow wage employment growth coupled with rapid expansion in informal sector employment. Increased labor supply would be expected to put downward pressure on wages. Further analysis indicates that, controlling for unobserved firm fixed effects, firm characteristics, and workers' occupation, does not eliminate the real earnings increase. Firm size, labor productivity, occupation, and sector were important earnings determinants. The firm size and productivity effects may reflect rent sharing and/or efficiency wages (e.g. Teal, 1996, Soderbom and Teal, 2001, and Azam and Ris, 2001) or union wage premium (e.g. Manda, Bigsten, and Mwabu, 2005).
Returns to education ranged from 4% at 6 years of education to 22% at 14 years of education. Given that education has direct effect and an indirect effect on earnings (Bigsten, 1984), the greater returns at higher education level may indicate greater impact of higher levels of education in accessing wage jobs. Further, returns to education are 2% to 10% higher at the 75^{th} percentile than at the 25^{th} percentile. This may indicate there are unobserved productive characteristics (e.g. innate ability or quality of education) that compliment educational attainment (Harmon, Oosterbeek, and Walker, 2000; Mwabu and Schultz, 1996). With few exceptions the interaction effect between education and time dummies was not significant. It implies that the return to education did not change significantly over the sample period-a period that was characterised by structural reforms. This contrasts with Uganda where Appleton (2002) found that returns to education rose significantly in the 1990s. During this period, Uganda's growth performance was better than Kenya's performance.
It is likely that a growing economy increased opportunities to realize and increase returns to education in Uganda, while the scope for increased returns to skills was limited in Kenya's economy.

References

Arulampalam, W., Booth, A.L. and Bryan, M.L.(2010), Are there asymmetries in the effects of training on the conditional male wage distribution? J*ournal of Population Economics* 23(1): 251-272

Appleton, S. (2002), Education, incomes and poverty in Uganda in the 1990s. *CREDIT Research paper No. 01/22*, Center for Research in Economic Development and International Trade, University of Nottingham.

Azam, J. and Ris, C. (2001), "Rent-sharing, hold-up, and manufacturing sector wages in Cote d' Ivoire", Paper prepared for the UNIDO/CSAE conference on "New Industrial Realities and Firm Behavior in Africa", 20-22 September 2001 at St. Edmund Hall, Oxford University.

Becker, G.S. (1964), Human capital, Columbia University Press for the National Bureau of Economic Research, New York.

Bennel, P. (1996), Rates of return to education: Does the conventional pattern prevail in Sub-Saharan Africa? *World Development*, 24(1): 183-199.

Borjas, *G. (*2007), *Labor Economics (*4th edition), McGraw-Hill

Bushnisky, M. (1994), Changes in the U.S. wage structure 1963-1987, Application of Quantile regression, *Econometrica*, 62(2): 405-458.

Bushnisky, M. (1998), "Recent advances in quantile regression models: A practical guideline for empirical research", *Journal of Human Resources*, XXXIII(I): 88-126.

Cahuc, P. and Zylberberg, A. (2004), *Labor Economics*, MIT Press.

Colclough, C., Kingdon, G. and Patrinos, H. (2010), The Changing Pattern of Wage Returns to Education and its Implications, *Development Policy Review*, 28(6):733-747

Card, D. *(*1995*). Earnings, schooling, and ability revisited*. In: Polachek, S. (Ed.), Research in Labor Economics. JAI Press, Greenwich, CT, pp. 23–48.

Deaton, A. (1997), *The Analysis of household surveys: A Micro-econometric Approach to Development Policy*, Washington, D.C., Published for The World Bank, The Johns Hopkins University Press.

Halvorsen, R. and Palmquist, R. (1980), The interpretation of dummy variables in semi logarithmic equations, *American Economic Review*, 70(3): 474-475.

Harmon, C., Oosterbeek, H., and Walker, I. (2000), The returns to schooling. A review of
evidence, issues and deficiencies in the literature, Center for the Economics of Schooling, London School of Economics and Political Science.

Machado, J. and Mata, J. (2001), "Earnings functions in Portugal 1982-1994: Evidence from Quantile regressions", *Empirical Economics*, 26(1): 115-134.

Manda, D. K., Bigsten, A., and Mwabu, G. (2005), Trade union membership and earnings in Kenyan manufacturing firms, *.Applied Economics*, 37(15): 1693-1704

Martins, P. and Pereira, P.(2004), Does education reduce wage inequality? Quantile regression evidence from 16 countries, *Labour Economics*, 11(3): 355-371

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

Mwabu, G.M. and Schultz, T.P. (1996), Educational returns across quantiles of the wage function: Alternative explanations for returns to schooling by race in South Africa, *The American Economic Review*, 86(2): 335-339.

Nielsen, H.S. and Rosholm, M. (2001), The public–private wage gap in Zambia in the 1990s: A quantile regression approach, *Empirical Economics*, 26(1): 169-182.

Pscharopoulos, G. and Patrinos, H.(2004), Returns to Investment in Education. A Further Update, Eeducation Economics, 12(2): 111-134

Söderbom, M. and Teal, F. (2001), "Firm size and human capital as determinants of productivity and earnings". *CSAE WPS 2001-9*, Centre for the study of African Economies, Oxford University.

Teal, F. (1996), The size and sources of economic rents in a developing country Manufacturing labour market, *Economic Journal* 106 (437):963–976.

Appendix

Table A1: Summary Statistics of Variables for each Survey Wave and for Full Sample

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Total
Male worker	89	85	85	81	85
Education					
No education	1	0	0	0	0
Some primary	16	16	13	8	13
Full primary	45	43	44	39	43
Full secondary	38	39	39	48	41
University	1	1	3	4	2
Occupation					
Management worker	6	2	2	4	4
Administration worker	10	11	5	22	12
Sales worker	3	5	3	5	4
Supervisory worker	10	8	10	12	10
Technician worker	7	10	16	12	11
Production worker	63	62	62	45	58
Firm in Nairobi	68	67	66	58	65
Firm in Mombasa	13	15	15	23	17
Firm in Nakuru	10	10	10	8	10
Firm in Eldoret	9	8	9	11	9
Sector					
Foods sector	18	17	18	24	19
Wood sector	14	12	12	8	12
Textile sector	7	6	8	12	8
Metal sector	25	24	22	19	23
Bakery sector	4	8	8	5	6
Furniture sector	14	17	17	11	15
Garments sector	14	14	12	14	13
Machinery sector	4	2	3	7	4
Age (years)					
Mean(Median)	35(33)	33(31)	33(32)	35(33)	34(32)
Standard deviation	9	8.7	8.9	9.4	9
Tenure (years)					
Mean(Median)	8.1(6)	7.3(5)	7.6(6)	8.5(6)	7.9(6)
Standard deviation	7.1	6.9	7.1	7.6	7.2
Education (years)					
Mean(Median)	8.8(9)	8.8(9)	9.2(9)	9.6(11)	9.1(9)
Standard deviation	3	2.9	2.9	2.6	2.9
Employment					
Mean(Median)	176(60)	117(50)	146(54)	168(50)	152(53)
Standard deviation	468	261	316	297	351
Output/worker					
Mean(Median)	393990(215115)	634700(217432)	584886(268885)	526107(266783)	531279(242963)
Standard deviation	513815	1507690	1086705	730484	1029260
Capital/worker					
Mean(Median)	413595(206346)	403892(204955)	459904(229152)	805060(378780)	508869(236737)
Standard deviation	566369	571206	577197	1742257	970540
Wage/worker					
Mean(Median)	14545(10790)	14374(11662)	21824(14837)	33559(20444)	20603(13352)
Standard deviation	12569	14284	37548	68910	39566

Table A2. Extended earnings equations with quadratic education function (n=3862)

Explanatory variable	OLS	FE	P25	P50	P75
Male worker	0.09**	0.09***	0.12***	0.12***	0.09**
	(2.50)	(3.47)	(5.97)	(4.98)	(2.38)
Age (years)	0.04***	0.04***	0.04***	0.04***	0.04***
	(4.02)	(5.85)	(4.31)	(5.36)	(4.59)
Age squared/100	-0.03***	-0.04***	-0.04***	-0.04***	-0.03***
	(2.66)	(4.02)	(3.44)	(3.96)	(2.63)
Education (years)	-0.07***	-0.06***	-0.04***	-0.06***	-0.06***
	(4.71)	(4.91)	(3.20)	(4.11)	(3.77)
Education squared/100	0.73***	0.62***	0.43***	0.64***	0.72***
	(7.62)	(9.45)	(5.52)	(7.79)	(6.71)
Tenure with firm (years)/10	0.03	0.01	0.04	0.04	-0.03
	(0.65)	(0.33)	(1.09)	(1.24)	(0.64)
Tenure squared/100	0.02	0.04	0.04	0.01	0.20
	(0.13)	(0.33)	(0.30)	(0.10)	(1.09)
Employment (logarithm)	0.07***	-0.04	0.07***	0.07***	0.08***
	(6.18)	(1.03)	(6.90)	(7.40)	(6.53)
Capital per worker (logarithm)	-0.001	-0.04	0.02*	-0.01	-0.03***
	(0.05)	(1.28)	(1.89)	(1.31)	(2.68)
Output per worker (logarithm)	0.05***	-0.00	0.04***	0.04***	0.05***
	(3.28)	(0.17)	(4.21)	(6.83)	(5.46)
Management worker	0.93***	0.96***	0.77***	1.00***	1.19***
	(10.47)	(21.16)	(12.03)	(15.74)	(14.63)
Administrative worker	0.43***	0.44***	0.29***	0.43***	0.57***
	(11.02)	(15.48)	(6.39)	(9.31)	(14.31)
Sales worker	0.31***	0.24***	0.20***	0.27***	0.42***
	(4.46)	(5.35)	(3.26)	(3.62)	(5.35)
Supervisory worker	0.37***	0.39***	0.35***	0.39***	0.41***
	(10.66)	(14.21)	(7.56)	(8.30)	(7.45)
Technician worker	0.13***	0.15***	0.08***	0.10***	0.13***
	(3.95)	(5.60)	(3.54)	(3.63)	(3.08)
Firm in Mombasa	-0.10**		-0.09***	-0.05*	-0.09***
	(2.31)		(3.77)	(1.86)	(2.58)
Firm in Nakuru	-0.43***		-0.39***	-0.39***	-0.40***
	(9.40)		(11.76)	(10.36)	(7.66)
Firm in Eldoret	-0.43***		-0.35***	-0.42***	-0.45***
	(8.65)		(11.52)	(15.24)	(11.10)
Wood sector	0.07		0.05	0.06	0.05
	(1.05)		(1.47)	(1.24)	(0.89)
Textile sector	-0.18**		-0.20***	-0.14***	-0.16***
	(2.37)		(5.61)	(3.60)	(3.75)
Metal sector	0.10		0.10***	0.10***	0.11***
	(1.59)		(3.85)	(4.88)	(3.46)
Bakery sector	-0.14		-0.14***	-0.19***	-0.18***
	(1.61)		(3.40)	(4.55)	(3.92)
Furniture sector	0.14**		0.16***	0.14***	0.11***
	(2.09)		(4.79)	(5.46)	(3.38)
Garments sector	-0.04		-0.04	-0.06**	-0.10**
	(0.62)		(1.43)	(2.47)	(2.33)
Machinery sector	0.14		0.20***	0.21***	0.10
	(1.62)		(3.69)	(6.87)	(1.60)
Wave 2	-0.07**	-0.06***	-0.08***	-0.08***	-0.07***
	(2.31)	(2.85)	(3.07)	(4.17)	(2.61)
Wave 3	0.38***	0.40***	0.40***	0.36***	0.37***
	(12.05)	(17.20)	(18.12)	(21.15)	(12.84)
Wave 4	0.31***	0.31***	0.28***	0.31***	0.29***
	(8.23)	(9.25)	(12.35)	(15.83)	(8.74)
Constant	-0.15	1.36**	-0.57***	0.03	0.31
	(0.49)	(2.35)	(2.59)	(0.12)	(1.61)
Adjusted R^2 [Pseudo R^2]	0.52	0.39	[0.30]	[0.33]	[0.36]