Family Background, Education and Earnings in Kenya[*]

Anthony Wambugu

Working Papers in Economics no 76

June 2002

Abstract

This paper uses data collected in 2000 to first estimate OLS and ordered probit models to measure correlation between family background and workers' education. Then, human capital earnings functions are estimated to examine to what extent family background accounts for observed correlation between workers' education and earnings. Subsequently, it estimates returns to education with education treated as endogenous. Having well-educated parents is associated with greater educational attainment and earnings. Returns to education decline slightly when parents' education controls are in the earnings function. Instrumental variable estimation or self-selection correction suggests that estimates of returns to education may be larger than conventional estimates suggest.

JEL Classification: J24; J31; I21; O15

Keywords: Earnings function, return to education, instrumental variables

Correspondence to: Anthony Wambugu, Department of Economics, University of Göteborg, P.O.Box 640 405 30, Göteborg, Sweden,

E-mail: Anthony.wambugu@economics.gu.se

[*] The comments and suggestions from Arne Bigsten, Henry Ohlsson, Simon Appleton, Roger Wahlberg, Sten Dieden, Wilfered Nyangena, Anton Nivorozhkin and seminar participants are gratefully acknowledged. I thank Arne Bigsten for the opportunity to take part in the data collection exercise. I thank the AERC for financial support and the CSAE for the data.

1. Introduction

This paper studies the impact of family background on educational attainment and wages of adult workers in Kenya. It investigates whether omitted family background characteristics result in over estimated returns to education. It also examines whether treating education as endogenous in the wage function affects estimates of returns to workers' education. Although researchers of African labor markets are aware of the potential role of family background for education and labor market outcomes, lack of data on workers' family backgrounds has resulted in sparse empirical work on the issue. In particular, there is little work that considers educational attainment and earnings together.

Two strands of research guided this study. The first concerns demand for education. Strauss and Thomas (1995) and Shultz (1988) surveyed the literature and concluded that parents' education is positively associated with greater child education. The result is also reported in recent studies in Africa (e.g. Glewwe and Jacoby, 1994 in Ghana, Tansel, 1997 in Ghana and in Cote d' Ivoire, Lloyd and Blanc, 1996 in six countries, and Appleton, 2001 in Uganda). The effect might differ by gender of the child.

However, to analyse education and earnings one requires the eventual educational attainment of a worker. In a study of wage earners from a 1980 survey in Nairobi, Knight and Sabot (1990) reported binary probit estimates showing that parents' education was associated with higher probability of primary or secondary school completion. But university education was not in the analysis.[1] Also, the impact of each parents' education was not identified. Further, since education attainment is an outcome of ordered discrete decisions on whether to continue or withdraw from school (Lillard and Willis, 1994; Glick and Sahn, 2000) a different empirical model can take this into account.

[1] Castro-Leal et. al. (1999) reported that the top and bottom income quintiles received 44% and 2% of the subsidy allocated to tertiary education respectively. Access to tertiary education is therefore worth investigating.

The second strand of research guiding this study is on returns to education. A summary of estimates of returns to education world-wide (Psacharopoulos, 1994) reported that, private returns to education in Sub-Sahara Africa were 41.3% at primary level, 26.6% at secondary level, and 27.8% at higher level of education. A survey of returns to education in Sub-Sahara Africa based on the model developed by Mincer (1974) is provided by Appleton, et. al. (1996). One critique of conventional estimates of returns to education is that they may be biased because family background measures are omitted or the measures impact of on education decisions. In short, earnings and education may have common determinants.

There are three potential sources of endogenous education. First, as noted by Grilliches (1977), measurement error in education biases the estimate towards zero. Second, omitted unobservable productive worker characteristics may be positively correlated with both education and wages leading to upward biased estimates. Third, according to Card (1995, 1999) a discount rate bias may lead workers with high discount rates to acquire less education. The return to education for this group would be lower than the conventional estimate of return to education.

Researchers have addressed the problem in a variety of ways. One approach is to include measures of ability in wage functions (e.g. Grilliches, 1977; Blackburn and Neumark, 1995). They find the standard estimates are upward biased. Knight and Sabot (1990) controlled for workers' ability using test scores from Ravens Progressive Matrices test. They reported that the ability measure had little direct impact on urban wages in Kenya and Tanzania. Glewwe (1996) reported a similar result in Ghana.

Other studies (e.g. Heckman and Hotz, 1986 for non-farm men in Panama; Lam and Schoeni, 1993 for married men in Brazil; Krishnan, 1996 for young workers in urban Ethiopia; and Kingdon, 1998 in India), control for family background in wage functions. The common finding is that family background has positive impact on wages and returns to education may fall. Armitage and Sabot (1987) used the data used by Knight and Sabot (1990) and reported that the return to a year of secondary education rose with parents' education. They did not examine how parents' education affected returns to different levels of education and endogenous education was not considered. Also, the measure of parents' education used does not separate the impact of mother and father's education.

Family background measures have also been used as instruments for education. Most studies of this nature are from developed countries (e.g. Levin and Plug, 1999; Dearden, 1999; Uusitalo, 1999; Callan and Walker, 1999; and Oosterbeek and Ophem, 1999). Other studies use variation in education determinants to construct instruments for education. For example, Angrist and Krueger (1991) used quarter of year when a worker was born while Harmon and Walker (1995) used change in minimum school leaving-age. The key finding is that estimates from this procedure often exceed OLS estimates and the difference is large in some studies.

Harmon and Walker (1995) and Vella and Gregory (1996) used another approach that involves correction for self-selection bias in estimated returns to education.[2] The point is that observed educational attainment is an outcome of non-random decisions. The return to education for a given level of education is only observed for those workers that attained that level of education. Manda (1997) applied this method and reported that there was no evidence of self-selection bias. However, the data used did not have measures of family background and some of the identifying restrictions may be weak.

Other studies (e.g. Ashenfelter and Krueger, 1994) used data on twins to difference out family background fixed effects. They reported that ability bias may be small but measurement error could bias conventional estimates downwards. An alternative is to treat family fixed effects as fixed over time and use panel data methods to difference them out. Angrist and Newey (1991) used this method and reported returns to education that exceeded conventional estimates.

The empirical analysis in this study used data collected in 2000 from four urban centres in Kenya. They contain measures of family background among other variables that may have determined workers' education. It estimates the association between parents' education and two education indicators: years of education and the highest level of education attained. OLS and ordered probit models were estimated taking account of all levels of education. To study impacts of parents' education on earnings and returns to workers' education a flexible wage earnings function that allows for non-linear effect of worker education is used. This allows an assessment of how parents' education impacts on returns to different levels of education. The potential problem of endogenous education is also examined.

The remainder of the study is divided into five sections. Section 2 outlines how family background might influence educational attainment and earnings. Section 3 contains the empirical models and Section 4 contains a description of the data. Section 5 presents the estimation results and section 6 summarizes and concludes the study.

2. Optimal Education and Earnings

The theory of human capital is based on the notion that an individual's incentive for investing in education is provided by the prospect of future gain in earnings. Following Grilliches (1977) suppose the following relation gives the wage:

$$w = w\,(s, A, \mu) \qquad (1)$$

where w is the wage, s is education, A is a measure of ability, and , μ is an unobserved worker specific characteristic that is independent of worker's ability. Other factors such as family background may also affect the wage.

[2] An ordered probit education model is estimated and a selection term is computed and included in wage function.

Educational investments may be looked at as resulting from optimizing behavior, where long-lived individuals invest in education to maximize present discounted value of wealth (B). Further, it is assumed that the cost of education is opportunity cost. The problem is to

$$\text{Max B(s)} = \int_0^\infty w(s, A, \mu) e^{-r(s+t)} dt \qquad (2)$$

The rate *r* is the discount rate applied to future earnings stream. Card (1995, 1999) argues that individuals with different discount rates acquire different amounts of education. Moreover, this could be a source of bias in estimated returns to education. The first order condition of (1) with respect to *s* gives the stopping rule for educational investment.

$$\frac{\partial w / \partial s}{r} = w(s, A, \mu) \qquad (3)$$

It indicates that education investments will be made until the present value of marginal benefits (left hand side) equals the marginal costs (right hand side). Assuming that parents are altruistic (see Becker and Tomes, 1976, 1979), that markets are perfect; and that education is a pure investment good, optimal education would prevail. Family background would have little impact on worker's education.

However, the assumptions may not hold and there can be room for family background to influence education. More educated parents may have better access to funds to finance education or they invest more in out-of-school human capital (e.g. better home education, health, and nutrition) and hence complement formal education. They may have better information on returns to education and hence less uncertainty in educational investments. The key point is that family background may be associated with different marginal benefits and marginal costs of educational investments and this translates to workers acquiring different quantities of education.[3]

3. Econometric Specification

Based on human capital theory and previous studies, the building blocks in this study are the wage earnings function and the educational attainment function. Begin with the wage function

3.1. Wage Earnings Model

The basic earnings function based on the model developed by Mincer (1974) has a semi-logarithmic form and can be written as

$$\ln w_i = X_i^{'}\gamma + S_i\delta + u_i ; \qquad u_i \sim \text{i.i.d.} N(0, \sigma_u^2) \qquad (4)$$

where w_i is the real hourly wage of worker *i*, S_i is worker's education in years. The regressors in vector X are worker's age and square of age, gender dummy, tenure in current firm and dummy variables for firm location in Mombasa, Nakuru and Eldoret. γ and δ are parameters to be estimated, and u_i is a random error term.

In estimating equation (4) it is assumed that (i) the effect of education on earnings is linear; (ii) education is not correlated with the error term; (iii) there is no omitted variable bias that results from not controlling for ability or family background. To address these issues, the basic specification is extended. First, controls for parents' education are added to address assumption (iii) partially. Second, to address assumption (i) the square of education is added

[3] Behrman and Kenan (1996) review analytical approaches to human capital investments and returns

(see Willis, 1986 for a discussion and Bigsten et al, 2000 for an application). Third, Instrumental Variable method and a selectivity model address assumption (ii).

3.2. Educational Attainment Model

The effects of family background on the highest level of education completed are obtained from an ordered probit model which can be described as follows: Assume there is a latent variable, S_i^* measuring the highest education level completed by the *i*th worker. Hence

$$S_i^* = H_i^{'}\beta + \varepsilon_i ,\ \varepsilon_i \sim \text{ i.i.d.N}(0, 1) , \tag{6a}$$

$$S_i = j \text{ if } \mu_{j-1} \leq S_i^* \leq \mu_j , \tag{6b}$$

where S_i denotes the highest education level reported directly by the ith worker. The thresholds parameters μ_j, j = 0,1, 2, 3 are estimated along with parameter vector β. H_i is a vector of regressors that include age and square of age that may pick up time and lifecycle effects. Distances to nearest school facilities when worker was of school age can measure school availability or direct school costs. Regional dummies may capture regional variations in education development and other region specific factors. Family background is measured by parents' education. ε is a random variable distributed as standard normal. This predicted probability of a worker having j as the highest level of education is

$$\Pr ob(S_i = j) = \Phi(\mu_j - H_i^{'}\beta) - \Phi(\mu_{j-1} - H_i^{'}\beta), \tag{6c}$$

where $\Phi(.)$ is the cumulative normal distribution function.[4] Then the impact of explanatory variables on the probabilities can be obtained. The impacts of family background on years of

[4] An alternative is an ordered logit model where the logistic is substituted for the normal distribution.

education completed are modelled using an ordinary least squares model with the same regressors as in the ordered probit.

4. Data and Descriptive Statistics

This study relies on data from a survey of about 200 manufacturing firms in four urban centers in Kenya, that is, Nairobi, Mombasa, Nakuru, and Eldoret. It is the fourth wave of a panel survey started in 1993 and was fielded in October/November 2000. The data were collected in face-to-face interviews with firm managers and a sample of up to 10 workers in every firm visited. The relevant variables for the analysis are age, education, years of tenure in current firm, wages, their parents' education and the proximity to primary and secondary schools at the time the worker was of the relevant school age. Earlier waves of the panel survey did not collect information on workers' family background. The analysis is based on a sample of 843 workers between the ages of 16 and 64 years.

Table 1 presents variable description and sample statistics. The proportion of female workers is relatively small and over 80% of the workers are employed in Nairobi and Mombasa. The average age is about 35 years while the average tenure is slightly below 9 years. The latter may suggest low turnover of workers. About three-quarters of the workers lived within 3 kilometres of a primary school when they were of primary school age. In contrast, about two-thirds had no secondary school within 3 kilometres when they were of secondary school age.

Workers' educational attainment appears high with the average being 9.6 years of education. The highest level of education completed by about half the workers is secondary education while a small proportion never completed primary education. But parents' educational attainment is low with over 70% of the parents having primary education or less. In particular, mothers have lower education than fathers. Over 40% of the former are uneducated. The data also suggest that workers attained more education than their parents.

Table 2, presents the correlation between parents' and workers' education and Table 3, presents mean and median log hourly earnings by parents' education, and correlation between earnings and parents' education. The key point is that parents' education is correlated with workers' education and labor market earnings. To gauge the strength of these relationships conditional on other variables the paper turns to multivariate analysis.

5. Estimation Results

The empirical analysis has three distinct parts that closely follow the three strands of the research reviewed earlier. The section begins with educational attainment function estimates to focus on the association between parents' education and worker's education. This is followed by wage earnings function estimates with and without controls for family background. The third part presents results of an endogenous education model.

5.1. Education Attainment Function Estimates

The education function estimates are presented in Tables 4 and 5. Equations (1) and (3) in Tables 4 and 5 includes worker characteristics and parents' education. In equations (2) and (4) dummy variables for region and distances to nearest schools were added. The discussion that follows is confined to equations (2). This is because the additional variables did not reduce the size of parents' education coefficients very much. In addition, this equation gives the separate impacts of mother and father's education.

In the following discussion, the omitted dummy variables describe a worker with uneducated parents, lived less than one kilometer from nearest school facility, and attended most education in Nairobi province. Also, as pointed out frequently (e.g. Green, 1997 and Long, 1997) ordered probit estimates do not tell how changes in regressors affect predicted probabilities. For this purpose marginal effects are calculated based on equation (2) and presented in Table 6.

The results in Table 4 suggest that in the manufacturing labor market in Kenya, having an educated parent is associated with completion of more years of education. Table 6 suggests that this may be because parents' education is predicted to raise the probabilities of attaining secondary or university level of education. An F-test rejected the null hypothesis of equality of father and mother's education coefficients (see Tables 4 and 5). The differential impact may be due to a number of reasons.

First, parent's education impact may be gender-specific. For example, Glick and Sahn (2000) reported that father's education had greater impact on male children education in Conakry, Guinea. In Ghana and Cote d' Ivoire, Tansel (1997) reported that father's education had greater impact for both girls' and boys' education. However, in Ghana, parents' education impact was greater for girls than for boys and the reverse held in Cote d' Ivoire. Moreover, in Ghana mother's education had greater impact on girls' than boys' education.

Second, it may be the impact of the most educated parent that dominates. In the sample used, fathers have more education than mothers. Third, if mother's education impact is through home production, then if education increases the likelihood of mothers participating in the labor market this may reduce time allocated to home production and weaken the direct impact of mother's education.

Having the nearest primary school beyond 10 kilometres from worker's home is associated with fewer years of education. But distance beyond 6 kilometres is predicted to raise the probability of attaining only primary education. Workers within one to three kilometres from the nearest secondary school are predicted to complete fewer years of education. This may be because this variable is associated with a higher probability of ending education at primary level or below. Tansel (1997) reported a similar finding in Ghana where distance to secondary schools reduced middle school attainment.[5]

The significance of regional dummies suggests that workers who attended most education in Eastern, Western, Nyanza and Coast provinces completed fewer years of education. Table 6 suggests that the three latter variables are predicted to raise the probability of attaining primary education or less. The regional dummies may reflect income differentials unmeasured school attributes such as teacher supply and school facilities or household specific attributes.

5.2. Basic Wage Function Estimates

In equation (1a), Table 7, the return to education is 0.14. In equation (2a) the coefficients of education and its square are significant. This suggests a non-linear log wage-education relationship. The return to education ranges from 0.09 at 7 years to 0.21 at 12 years of education. An altenative specification where education is measured as 0/1 dummy variables shows that conventional returns to education (equation 1, Table 8) range from 0.03 at primary to 0.67 at university level.[6] In the three equations, age and tenure effects are significant and there is no indication of significant gender wage differential. In addition, workers in firms in Nakuru and Eldoret earn lower wages relative to those in firms in Nairobi.

5.3. Wage Function Estimates with Parents' Education

Conventional estimates of returns to education are often questioned because the wage functions used to estimate them do not control for family background. The estimates are said to be subject to family background bias (Lam and Schoeni, 1993). The objective in this part

---

[5] Distances to school facilities may be endogenous since schools may be placed where there is demand or people move to locations with better access to school facilities (Appleton, 2001 and Strauss and Thomas (1995).

[6] Calculated as [exp(b2-b1)-1]*100 divided by number of years in a level of education (see Halvorsen and Palmquist, 1980). b2 is the coefficient of dummy for higher level of education and b1 the coefficient of the dummy

is twofold. First, it is to examine the impact of parents' education on wages and second, to assess whether returns to worker's education decline when controls for parents' education are included in wage earnings function. In equations (1b) and (2b) of Table 7, having parents with post primary education is associated with higher wages relative to having uneducated parents. The earnings advantage associated with fathers' post-primary education is 27% compared to 19% for mother's post-primary education. A similar result is found in equation (2) of Table 8.

Comparing equations (1a) and (1b), controlling for parents' education reduced the return to worker's education by 7%. On the other hand, comparing equations (2a) and (2b) the drop in return to education ranges from 10% at 12 years of education to 22% at 7 years of education. In Table 8, a comparison of equations (1) and (2) suggests a similar pattern. However, as noted by Grilliches (1977) too much zeal in controlling for omitted variable bias may be counterproductive since measurement error could be exacerbated. Lam and Schoeni (1993) reported that part of the decline in returns to education after controlling for family background could be attributed to measurement error. Hence the upward bias in conventional estimates without background controls may not be large.

The evidence from other countries is mixed. In Panama Heckman and Hotz (1986) reported that, controlling for parents' education, the return to male education fell by 25% and mother's education had a larger impact on worker's wages than father's education. In Brazil, return to married males' education fell by 25% to 33% when parental background was controlled for (Lam and Schoeni, 1993). Similarly, Kingdon (1998) reported that in India, return to a year of male education fell by 16% while return to female education fell by 49% controlling for father's education in the wage equation.

---

for level immediately below it. It is assumed that primary education is 7 years, secondary education is 6 years, and university education is 3 years.

In a study of the urban labor market in Ethiopia, Krishnan (1996) reported that returns to education fell by 20% and 10% for workers in public and private sector respectively when family background is controlled for. But she demonstrated using a sector selection model that family background was important for gaining employment into these sectors. After controlling for this effect in the wage function, controlling for family background resulted in returns to education falling only by 5% and 4% in the respective sectors.

A sector allocation mechanism in which family background plays a key role may be present in Kenya. Having well educated parents' may be associated with better labor market contacts and networks that play a role in job search and allocation for young adults. For example, about 40% of workers in the sample secured their current job through friends, family and relatives. It could also reflect the ability to meet job search costs and hence to sustain longer unemployment durations until their children get a higher pay jobs or jobs with greater opportunities to acquire on the job human capital.

The favourable impact of family background on wages may also suggest that a better home learning environment and investments in health and nutrition enhance the human capital acquired from a given quantity of education. This may subsequently command higher wages in the labor market. Attendance at better quality schools may have a similar effect. For example, Knight and Sabot (1990) found that, graduates of public secondary schools in Kenya had relatively higher cognitive achievement. More importantly, parents' education was positively associated with probability of attending such schools.

5.4. Wage Function Estimates with Endogenous Education

The first two parts of the analysis model the education relation and the wage relation separately as is common in the literature. To incoporate endogenous education a reduced-form education equation is added to the wage equation to form a two equation model. Parents' education, distances to nearest schools, and regions are the identifying instruments.

In Table 7 equations (1c) and (2c) show results of instrumental variable estimation. The return to education in the basic wage function (equation 1c) is 0.24 compared to the conventional OLS estimate of 0.14 in equation (1a). The Hausman test rejects the null hypothesis of equality of OLS and IV estimates. In the alternative specification with square of education the IV return to education (equation 2c) at 7 years of education is 0.04, which appears lower than the OLS estimate of 0.09 in equation (2a). But the IV returns to education at 10 and 12 years are double the OLS estimates. The Hausman test rejects the null hypothesis of equality of OLS and IV estimates in this model. Studies cited in the introduction also report estimates that are some times twice or three times the OLS estimates.

There are concerns regarding estimation by IV in that though IV estimates differ from OLS estimates it is not clear that they are preferable. In this application, the instruments may be poor, that is, they may be weakly correlated with workers' education. Second, the instruments may not be valid, that is, they may not be legitimately excluded from the wage equation. And third, even if they are valid, the instruments may not be indicating access to education. They may be picking other effects.

The quality of instruments was assessed using tests proposed by Bound et. al. (1995). A test on excluded instruments in the reduced-form education equation rejected the null hypothesis of equal coefficient estimates on instruments. This suggests that they have joint significant impact on workers' education. Bound et al also suggest that addition of instruments in the reduced form equation should improve explanatory power of the model. Regressing education and its square on instruments yielded an R^2 of about 0.19 in each case.

To assess validity of instruments, an over identification (OID) test (see Deaton, 1997) was used. Residuals from the IV wage equation are regressed on all instruments used in the reduced form education equation. The R^2 from this regression is multiplied by the sample size to yield a chi-squared distributed test statistic with degrees of freedom equal to the number of over-identifying instruments. The null hypothesis of valid instruments is not rejected.

The results from the ordered probit model of highest level of education attained (equation 2, Table 5) can be used to control for self-selection when estimating returns to education for these levels. Equation (3), Table 8 shows selectivity corrected earnings function estimates. The estimated returns to education are at least twice the OLS estimates and the self-selection term has negative and significant coefficient. Harmon and Walker (1995) and Vella and Gregory (1996) reported a similar finding while Manda (1997) reports selection coefficients with mixed signs and significance. The significant negative selection term implies that OLS leads to estimates that are downward biased (Harmon and Walker, 1995).

6. Summary and Conclusion

Family background can have important impact on educational investment decisions. For this reason it is argued that returns to education based on earnings functions that omit family background measures may be too high. In addition education attainment is a choice variable that may not be independent of the earnings determination process. Empirical evidence is thin due to lack of data on family background of adult workers. This paper used a sample of workers in Kenya manufacturing firms from a 2000 survey to examine these relationships.

From estimation of OLS and ordered probit models it turns out that parents' education is positively associated with worker's education, measured in years completed or highest level completed. Unlike previous studies the paper estimates separate impacts of each parents' education. It found that father's education had greater impact. But since this may reflect any number of effects one must be careful in concluding that the education of one or the other parent has larger impact. The conclusion to draw is that if potential parents are given more education, this is likely to lead to more education for their children.

The second finding is that comparing wage functions with and without controls for parents' education, returns to workers' education in the former are lower. This would suggest that a portion of the conventional estimate of returns to workers' education can be due to the impact of parents' education on wages. The results suggested that having parents who completed post-primary education is associated with higher earnings. However, any bias in conventional estimates from omitting family background measures is not very large and appears to be similar to studies in other countries.

A third finding of the study is that estimation of a two-equation model of workers' education attainment and earnings, suggests that conventional estimates of returns to education for Kenya manufacturing workers may be downward biased. Estimates from Instrumental Variable estimation or selectivity correction methods yield estimates that are substantially higher than the conventional estimates. Accounting for the education attainment decision when estimating returns to education appears to be important. In this context conventional estimates may be a lower bound on returns to education.

In closing, family background is important in predicting workers' education. In turn, workers' education has positive wage returns that rise with the quantity of education. To the extent that differential educational attainment by family background reflects liquidity constraints or other constraints, the results are supportive of public policy towards greater access to education for children of less educated parents. With more data, it would be desirable to study whether the relationships estimated in this study vary across gender and across time periods.

7. References

Appleton, S. (2001), What can we expect from universal primary education? In *Uganda's Recovery. The Role of Farms, Firms and Government*. Edited by Ritva Reinikka and Paul Collier. The World Bank, Washington D.C.

Appleton, S., Hoddinot, J, and McKinnon, J., (1996), Education and health in sub-Sahara Africa, *Journal of International Development*. Vol. 8, No. 3, 307-339.

Armitage, J. and Sabot, R. (1987), Social economic background and the returns to education in two Low-income countries, *Economica* 54, number 213, Feb 1987, 103-108.

Ashenfelter, O. and Krueger, A. (1994), Estimates of the economic returns to education from a new sample of twins, *American Economic Review* 84, 1157-1173.

Angrist, J. and Krueger, A. (1991), Does compulsory school attendance affect education and Earnings, *Quarterly Journal of Economics*, 106, 979-1014.

Angrist, J. and Newey, W. (1991), Over-identification tests in earnings with fixed effects, *Journal of Business and Economic Statistics*, 9(3), 317-323.

Becker, G, and Tomes, N. (1976), Child endowments and the quantity and quality of children, *Journal of Political Economy*, Vol.84, Issue 4, part 2, 143-162.

Becker, G, and Tomes, N. (1979), An equilibrium theory of the distribution of income and intergenerational correlation, *Journal of Political Economy*, Vol.87, Issue 6,1153-1189.

Behrman, J. and Kenan, W.(1996), Human capital formation, returns and policies: Analytical approaches and research questions, *Journal of International Development*: Vol. 8, 3, 341-373.

Bound, J., Jaeger, D. and Baker, R.(1995), Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variables is weak, *Journal of the American Statistical Association*, 90, 443-450.

Bigsten, A., P.Collier, S.Dercon, B. Gauthier, M. Fafchamps, J.W.Gunning, A.Issacksson, A.Oduro, R.Oostendorp, C. Pattillo, M. Soderbom, F.Teal, A.Zeufack, and S.Appleton (2000), Rates of return to human and physical capital in Africa's manufacturing sectors, *Economic Development and Cultural Change*: Volume 48, 4, 801-827.

Blackburn, M. and Neumark, D. (1995), Are OLS estimates of the returns to education biased downwards? Another look, *Review of Economics and Statistics* 77, 217-229.

Callan, T. and Walker, H.(1999), The economic return to education in Ireland, *Labor Economics,* 6(1999), 543-550

Card, D. (1995), Earnings, education and ability revisited, in: S.Polachek (ed.), *Research in Labor Economics*, vol. 14 (JAI Press: Greenwich, Connecticut), 23-48.

Card, D.(1999), The causal effect of education on earnings in: O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, vol. 3 (Elsevier Science), 1801-1863.

Dearden, L. (1999), The effects of families and abilities on men's education and earnings in Britain, *Labor Economics* 6(1999), 551-567.

Deaton, A.(1997), *The Analysis of Household Surveys: A Micro econometric Approach to Development Policy*. The John Hopkins University Press, Baltimore, Maryland,U.S.A

F. Castro-Leal, J.Dayton, L.Demery, K.Mehra (1999), Public social spending in Africa. Do the poor benefit?, *World Bank Research Observer*, vol.14, no. 1, 49-72.

Glewwe, P. (1996), The relevance of standard estimates of rates of return to schooling for education policy: A critical assessment, *Journal of Development Economics* 51, 267-290.

Glewwe, P., and Jacoby, H. (1994), Student achievement and education choice in low-income countries. Evidence from Ghana, *Journal of Human Resources*, 29, 841-64.

Glick, P. and Sahn, D.(2000), Education of girls and boys in a West African country: the effects of parental education, income, and household structure, *Economics of Education Review* 19, (2000), 63-87

Green, W.H.(1997), *Econometric Analysis*. Third Edition. Prentice-Hall International Inc.

Grilliches, Z. (1977), Estimating the returns to education: some econometrics problems *Econometrica,* vol. 45, No.1, 1-22.

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

Harmon, C. and Walker, I. (1995), Estimates of the economic return to education for the United Kingdom, *American Economic Review* 85, 1278-1286

Heckman, J., and Hotz, V. (1986), An investigation of the labor market earnings of Panamian males: Evaluating the sources of inequality, *Journal of Human Resources*, vol. 23, 507-542.

Kingdon, G. (1998), Does the labour market explain lower female education in India?, *The Journal of Development Studies* (Oct 1998) Vol. 35, Iss. 1, 39-65.

Knight, J. and Sabot, R. (1990), *Education, Productivity and Inequality: The East African Natural Experiment.* Oxford University Press, Oxford.

Krishnan, P. (1996), Family background, education, and employment in urban Ethiopia, *Oxford Bulletin of Economics and Statistics*, 58, 1 (1996), 167-183.

Lam, D. and Schoeni, R. (1993), Effects of family background on earnings and returns to education: Evidence from Brazil, *Journal of Political Economy*, 101, number 4, 710-40.

Levin, J. and Plug, E.(1999), Instrumenting education and the returns to education in the Netherlands, *Labor Economics*, 6(1999) 521-534.

Lillard, L., and Willis, R. (1994), Intergenerational educational mobility: Effects of family and State in Malaysia, *Journal of Human Resources.* XXIX. 4, 1126-1165.

Lloyd, C., and Blanc, A. (1996), Children's education in sub-Saharan Africa: The role of mothers and others, *Population and Development Review* 22(2): 265-98.

Long, S. (1997), *Regression models for categorical and limited dependent variable models.* Thousands oaks California, Sage.

Manda, D.(1997), Labor supply, returns to education and the effect of firm size on wages. The case of Kenya. PhD thesis, Göteborg University.

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

Oosterbeek, H. and Ophem, H.(2000), Education choices: preferences, discount rates, and rates of returns, *Empirical Economics* (2000) 25, 15-34.

Pscharopoulos, G. (1994), Returns to investment in Education. A Global Update, *World Development* 22 (9), 1325-1344.

Shultz, T.P. (1988), Education investments and returns, in: H. Chenery and T.N. Srinivasan, eds., *Handbook of Development Economics*, vol. 1.(North-Holland, Amsterdam), 543-630.

Strauss, J. and Thomas, D. (1995), Human resources: Empirical modelling of household and family decisions, in: Behrman, J. and Srinivasan, T.N., eds., *Handbook of Development Economics*, vol. III (Elsevier Science B.V), 1883-2005.

Tansel, A. (1997), Education attainment, parental education, and gender in Cot d'Ivoire and Ghana, *Economic Development and Cultural Change*, Vol. 45, Issue 4, 825-856.

Uusitalo, R. (1999), Returns to education in Finland, *Labor Economics*, 6(4), 569-580.

Vella, F. and Gregory, R. (1996), Selection bias and human capital investment: estimating the rate of return to education for young males, *Labor Economics* 3, 197-219.

Willis, R. (1986), Wage determinants: A survey and reinterpretation of human capital earnings functions, in: O.Ashenfelter and R.Layard, eds., *Handbook of Labor Economics*, (North Holland, Amsterdam), 525-602.

TABLE 1: DEFINITION OF VARIABLES AND SAMPLE STATISTICS

Variable	Description	Mean (std)
Male	Dummy variable = 1 if worker is male, =0 otherwise	0.82
Age	Age of worker at survey	34.88(9.48)
Tenure	Number of years in current firm	8.62 (7.68)
Worker's education		
Number of years	Number of school years completed by worker	9.67 (2.52)
Below primary	Dummy variable =1 if worker's highest education level is below primary level, =0 otherwise.	0.07
Primary	Dummy variable =1 if worker's highest education level is primary level, =0 otherwise.	0.38
Secondary	Dummy variable =1 if worker's highest education level is Secondary level, =0 otherwise.	0.51
University	Dummy variable =1 if worker's highest education level is university, =0 otherwise.	0.04
Distance to primary school		
Below 1 km	Dummy variable=1 if nearest primary school when worker was of primary school age was less than one kilometer, =0 otherwise	0.29
1-3 kms	Dummy variable=1 if nearest primary school when worker was of primary school age was 1-3 kilometers, =0 otherwise	0.44
3-6 kms	Dummy variable=1 if nearest primary school when worker was of primary school age was 3-6 kilometers, =0 otherwise	0.19
6-10 kms	Dummy variable=1 if nearest primary school when worker was of primary school age was 6-10 kilometer, =0 otherwise	0.05
Above 10 kms	Dummy variable=1 if nearest primary school when worker was of primary school age was more than ten kilometer, =0 otherwise	0.03
Distance to secondary school		
Below 1 km	Dummy variable=1 if nearest secondary school when worker was of secondary school age was less than one kilometer, =0 otherwise	0.14
1-3 kms	Dummy variable=1 if nearest secondary school when worker was of secondary school age was 1-3 kilometers, =0 otherwise	0.26
3-6 kms	Dummy variable=1 if nearest secondary school when worker was of secondary school age was 3-6 kilometers, =0 otherwise	0.22
6-10 kms	Dummy variable=1 if nearest secondary school when worker was of secondary school age was 6-10 kilometer, =0 otherwise	0.14
Above 10 kms	Dummy variable=1 if nearest secondary school when worker was of secondary school age was more than ten kilometer, =0 otherwise	0.24
Province of education		
Nairobi City	Dummy variable=1 if worker received most education in Nairobi province, =0 otherwise	0.09
Central	Dummy variable=1 if worker received most education in Central province, =0 otherwise	0.13
Eastern	Dummy variable=1 if worker received most education in Eastern province, =0 otherwise	0.18
Western	Dummy variable=1 if worker received most education in Western province, =0 otherwise	0.19
Rift Valley	Dummy variable=1 if worker received most education in Rift Valley province, =0 otherwise	0.10
Nyanza	Dummy variable=1 if worker received most education in Nyanza province, =0 otherwise	0.19
Coast	Dummy variable=1 if worker received most education in Coast province, =0 otherwise	0.12
Father's education		
uneducated	Dummy variable =1 if worker's father has no formal education, =0 otherwise	0.28
Primary	Dummy variable =1 if highest education level of worker's father is primary education, =0 otherwise	0.49
Post-primary	Dummy variable =1 if highest education level of worker's father is post-primary education, =0 otherwise	0.23

Mother's education		
Uneducated	Dummy variable =1 if worker's mother has no formal education, =0 otherwise	0.45
Primary	Dummy variable =1 if highest education level of worker's mother is primary education, =0 otherwise	0.41
Post-primary	Dummy variable =1 if highest education level of worker's mother is post-primary education, =0 otherwise	0.14
Both parents education		
None/none	Dummy variable =1 if both parents have no education, =0 otherwise	0.26
None/primary	Dummy variable =1 if one parent has no education and the other has primary, =0 otherwise	0.18
Primary/primary	Dummy variable =1 if both parents have primary education, =0 otherwise	0.31
None/post primary	Dummy variable =1 if one parent has no education and the other has post-primary, =0 otherwise	0.02
Primary/post primary	Dummy variable =1 if one parent has primary education and the other has post-primary, =0 otherwise	0.10
Post primary/post primary	Dummy variable =1 if both parents have post-primary education, =0 otherwise	0.12
Hourly wage	Constant price hourly earnings in Kenya shillings	13.08 (14.41)
Log wage	Natural logarithm of real hourly earnings	2.24(0.74)
Nairobi	Dummy variable =1 if worker works in a firm located in Nairobi, =0 otherwise	0.57
Mombasa	Dummy variable =1 if worker works in a firm located in Mombasa, =0 otherwise	0.24
Nakuru	Dummy variable =1 if worker works in a firm located in Nakuru, =0 otherwise	0.08
Eldoret	Dummy variable =1 if worker works in a firm located in Eldoret, =0 otherwise	0.11
Number of observations		843

For dichotomous (0/1) variables the mean is the proportion of sample with the identified characteristic

TABLE 2: CORRELATIONS BETWEEN PARENTS' EDUCATION AND WORKERS' EDUCATION

	Years of Education	Below primary	Primary	Secondary	University
Father's education					
None	-0.33*	0.23*	0.16*	-0.25*	-0.08*
Primary	0.04	-0.12*	0.06*	0.03	-0.07*
Post-primary	0.30*	-0.11*	-0.24*	0.23*	0.17*
Mother's education					
None	-0.31*	0.22*	0.17*	-0.26*	-0.07*
Primary	0.14*	-0.15*	-0.03	0.13*	-0.03
Post-primary	0.25*	-0.10*	-0.20*	0.20*	0.15*

Significance at 10% significance level or better is indicated by "*"

TABLE 3: WORKERS' LOGARITHM OF EARNINGS BY PARENTS' EDUCATION

	N	Mean	Std dev.	Median	Correlation
Father's education					
None	236	2.20	0.72	2.11	-0.03
Primary	415	2.16	0.71	2.05	-0.11*
Post-primary	192	2.48	0.79	2.43	0.17*
Mother's education					
None	382	2.24	0.68	2.13	-0.01
Primary	346	2.17	0.75	2.04	-0.09*
Post-primary	115	2.50	0.86	2.46	0.14*

Source: Computed from sample data. Significance at 10% significance level or better is indicated by "*"

TABLE 4: OLS EDUCATIONAL ATTAINMENT FUNCTION ESTIMATES

Variables	Equation (1)		Equation (2)		Equation (3)		Equation (4)	
Age	0.32*	[5.43]	0.29*	[4.90]	0.32*	[5.49]	0.30*	[4.99]
Age^2	-0.005*	[6.04]	-0.004*	[5.48]	-0.005*	[6.09]	-0.004*	[5.55]
Male	-0.14	[0.67]	0.03	[0.16]	-0.17	[0.80]	0.01	[0.05]
Father's education								
Primary	0.79*	[3.53]	0.76*	[3.49]				
Post-primary	1.59*	[4.78]	1.50*	[4.49]				
Mother's education								
Primary	0.51*	[2.52]	0.51*	[2.54]				
Post-primary	1.09*	[3.18]	0.99*	[2.94]				
Parents' education								
None/primary					0.74*	[2.93]	0.63*	[2.57]
Primary/primary					1.34*	[5.93]	1.30*	[5.90]
None/post primary					2.10*	[3.69]	1.97*	[3.29]
Primary/post primary					1.99*	[5.82]	1.88*	[5.73]
Post primary/post primary					2.67*	[8.51]	2.45*	[7.75]
Nearest primary school								
1-3 kms			-0.05	[0.29]			-0.06	[0.32]
3-6 kms			0.00	[0.02]			0.00	[0.02]
6-10 kms			-0.54	[1.29]			-0.51	[1.20]
Over 10 kms			-1.16**	[2.25]			-1.16**	[2.28]
Nearest secondary school								
1-3 kms			-0.52**	[2.10]			-0.50**	[2.04]
3-6 kms			-0.23	[0.90]			-0.22	[0.86]
6-10 kms			-0.08	[0.25]			-0.09	[0.30]
Over 10 kms			0.27	[0.96]			0.29	[1.03]
Province								
Central			-0.35	[0.97]			-0.36	[1.00]
Eastern			-0.67***	[1.78]			-0.70***	[1.87]
Western			-0.94*	[2.71]			-0.95*	[2.76]
Rift Valley			-0.08	[0.22]			-0.09	[0.23]
Nyanza			-1.43*	[3.86]			-1.44*	[3.90]
Coast			-0.81***	[1.83]			-0.83***	[1.87]
Constant	3.72*	[3.46]	4.99*	[4.30]	3.67*	[3.43]	4.93*	[4.27]
F		20.73*		17.39				
(D.F)[a]		(4, 176)		(4, 176)				
F				2.46**				2.43**
(D.F)[b]				(8, 176)				(8, 176)
Adjusted R^2	0.20		0.23		0.20		0.23	
Number of observations	843		843		843		843	

Notes: The dependent variable is years of education completed. The numbers in [] are absolute values of t-statistics based on standard errors robust to heteroskedasticity. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively. (a) Test indicates that the null hypothesis of equal coefficients of father and mother's education maybe rejected. (b) Test indicates that the null hypothesis of equal coefficient estimates of distances to primary and secondary school facilities maybe rejected.

TABLE 5: ORDERED PROBIT EDUCATIONAL ATTAINMENT FUNCTION ESTIMATES

Variables	Equation (1)		Equation (2)		Equation (3)		Equation (4)	
Age	0.16*	[5.61]	0.16*	[5.17]	0.16*	[5.62]	0.16*	[5.22]
Age^2	-0.002*	[6.40]	-0.002*	[5.92]	-0.002*	[6.39]	-0.002*	[5.96]
Male	-0.18	[1.65]	-0.08	[0.71]	-0.19***	[1.72]	-0.09	[0.77]
Father's education								
Primary	0.30*	[2.64]	0.31*	[2.72]				
Post-primary	0.75*	[4.15]	0.73*	[3.93]				
Mother's education								
Primary	0.27*	[2.55]	0.27*	[2.47]				
Post-primary	0.62*	[3.15]	0.60*	[2.97]				
Parents' education								
None/primary					0.27**	[2.14]	0.23***	[1.83]
Primary/primary					0.58*	[4.91]	0.59*	[4.89]
None/post primary					0.91*	[2.98]	0.86*	[2.60]
Primary/post primary					1.00*	[5.58]	0.97*	[5.44]
Post primary/post primary					1.35*	[7.51]	1.29*	[7.03]
Nearest primary school								
1-3 kms			-0.12	[1.27]			-0.13	[1.31]
3-6 kms			-0.14	[0.98]			-0.14	[0.98]
6-10 kms			-0.42**	[2.04]			-0.41**	[1.98]
Over 10 kms			-0.44***	[1.83]			-0.45***	[1.86]
Nearest secondary school								
1-3 kms			-0.26**	[2.09]			-0.25**	[2.05]
3-6 kms			-0.11	[0.85]			-0.10	[0.79]
6-10 kms			-0.01	[0.05]			-0.01	[0.05]
Over 10 kms			0.12	[0.81]			0.13	[0.88]
Province								
Central			-0.16	[0.84]			-0.16	[0.84]
Eastern			-0.29	[1.42]			-0.31	[1.49]
Western			-0.51*	[2.63]			-0.51*	[2.67]
Rift Valley			-0.10	[0.46]			-0.09	[0.45]
Nyanza			-0.75*	[3.66]			-0.76*	[3.71]
Coast			-0.47**	[1.97]			-0.47**	[1.98]
Threshold 1	1.28		0.57		1.29		0.59	
Threshold 2	2.82		2.17		2.83		2.19	
Threshold 3	5.03		4.45		5.04		4.47	
χ^2(D.F)[a]		65.39(4)		57.37 (4)				
χ^2(D.F)[b]				20.64 (8)				20.57 (8)
Pseudo R^2	0.11		0.14			0.11		0.14
Log-likelihood	-763.96		-742.08			-763.605		-741.645
Number of observations	843		843			843		843

Notes: The dependent variable is highest level of education completed. The numbers in [] are absolute values of z-statistics based on standard errors robust to heteroskedasticity. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively. (a) Test indicates that the null hypothesis of equal coefficients of father and mother's education maybe rejected. (b) Test indicates that the null hypothesis of equal coefficient estimates of distances to primary and secondary school facilities maybe rejected.

TABLE 6: ORDERED PROBIT MARGINAL EFFECTS ON PREDICTED PROBABILITIES

Variables	Primary dropout		Primary graduate		Secondary graduate		University graduate	
Age	-0.01*	[4.67]	-0.05*	[4.84]	0.06*	[5.04]	0.01*	[3.38]
Age^2	0.0002*	[5.10]	0.001*	[5.46]	-0.001*	[5.75]	-0.0001*	[3.53]
Male	0.01	[0.74]	0.03	[0.71]	-0.03	[0.72]	0.00	[0.66]
Father's education								
Primary	-0.03*	[2.63]	-0.09*	[2.70]	0.11*	[2.78]	0.01**	[2.13]
Post-primary	-0.05*	[4.49]	-0.22*	[4.14]	0.22*	[5.16]	0.05**	[2.18]
Mother's education								
Primary	-0.02*	[2.36]	-0.08*	[2.48]	0.09*	[2.44]	0.01*	[2.40]
Post-primary	-0.04*	[3.58]	-0.18*	[3.13]	0.18*	[3.51]	0.04**	[2.16]
Nearest primary school								
1-3 kms	0.01	[1.23]	0.04	[1.27]	-0.04	[1.25]	0.00	[1.29]
3-6 kms	0.01	[0.91]	0.04	[1.01]	-0.05	[0.97]	-0.01	[1.07]
6-10 kms	0.05	[1.51]	0.12*	[2.47]	-0.16**	[2.03]	-0.01*	[2.69]
Over 10 kms	0.06	[1.35]	0.12*	[2.27]	-0.16***	[1.84]	-0.01*	[2.45]
Nearest secondary school								
1-3 kms	0.03***	[1.84]	0.08**	[2.13]	-0.09**	[2.08]	-0.01**	[2.12]
3-6 kms	0.01	[0.81]	0.03	[0.86]	-0.04	[0.85]	0.00	[0.91]
6-10 kms	0.00	[0.05]	0.00	[0.05]	0.00	[0.05]	0.00	[0.05]
Over 10 kms	-0.01	[0.85]	-0.04	[0.81]	0.04	[0.82]	0.01	[0.73]
Province								
Central	0.02	[0.75]	0.05	[0.88]	-0.06	[0.83]	-0.01	[0.96]
Eastern	0.03	[1.23]	0.09	[1.50]	-0.11	[1.40]	-0.01***	[1.70]
Western	0.06**	[2.04]	0.14*	[3.00]	-0.19*	[2.65]	-0.02*	[2.72]
Rift Valley	0.01	[0.43]	0.03	[0.47]	-0.03	[0.45]	0.00	[0.50]
Nyanza	0.10*	[2.57]	0.19*	[4.80]	-0.27*	[3.83]	-0.02*	[3.43]
Coast	0.06	[1.48]	0.13*	[2.34]	-0.17**	[1.96]	-0.01*	[2.54]

Notes: For dummy variables the reported effect is for a discrete change from 0 to 1. The numbers in [] are the absolute values of z-statistics. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively. Derived from equation (2) of Table 5.

TABLE 7: OLS AND INSTRUMENTAL VARIABLES EARNINGS FUNCTION ESTIMATES

Variables	Equation (1a)		Equation (2a)		Equation (1b)		Equation (2b)		Equation (1c)		Equation (2c)	
Age	0.05*	[2.85]	0.05*	[3.15]	0.06*	[3.64]	0.06*	[3.87]	0.02	[0.85]	0.03	[1.57]
Age^2	-0.0003	[1.52]	-0.0004**	[1.99]	-0.0005**	[2.15]	-0.0005*	[2.54]	0.0001	[0.52]	-0.0002	[0.78]
Male	-0.03	[0.41]	-0.01	[0.21]	0.05	[0.80]	0.05	[0.93]	0.03	[0.50]	0.06	[0.90]
Education	0.14*	[12.4]	-0.08*	[2.36]	0.13*	[10.46]	-0.08**	[2.09]	0.24*	11.22	-0.50**	[2.17]
$Education^2$			0.012*	[6.58]			0.011*	[5.76]			0.0388*	[3.12]
Tenure	0.01*	[2.85]	0.02*	[3.75]	0.01*	[2.91]	0.02*	[3.69]	0.02*	[4.06]	0.03*	[5.20]
Mombasa	-0.04	[0.62]	-0.06	[0.86]	-0.03	[0.50]	-0.05	[0.72]	-0.07	[1.14]	-0.11***	[1.71]
Nakuru	-0.72*	[8.05]	-0.70*	[8.10]	-0.69*	[7.83]	-0.67*	[7.99]	-0.72*	[6.51]	-0.62*	[6.42]
Eldoret	-0.52*	[5.26]	-0.50*	[5.29]	-0.52*	[5.36]	-0.50*	[5.41]	-0.52*	[6.05]	-0.43*	[5.37]
Father's education												
Primary					0.06	[0.99]	0.07	[1.26]				
Post-primary					0.24*	[3.05]	0.22*	[2.98]				
Mother's education												
Primary					-0.03	[0.51]	-0.01	[0.24]				
Post-primary					0.17***	[1.92]	0.17**	[1.97]				
Constant	-0.29	[0.89]	0.66**	[2.04]	-0.62***	[1.92]	0.26	[0.83]	-0.88**	[2.32]	2.33**	[2.21]
Return												
Education =7			0.09				0.07				0.04	
Education=10			0.16				0.14				0.28	
Education =12			0.21				0.19				0.44	
Adjusted R^2			0.43				0.45		0.30		0.22	
Partial R^2										0.19		0.19 (0.18)
F (D.F)[a]										8.78 (18,176)		8.78(9.36)
χ^2(D.F)[b]										34.58 (17)		19.43 (16)
χ^2(D.F)[c]										32.20 (9)		19.28 (9)
No. of observations	843		843		843		843		843		843	

Notes: The dependent variable is the logarithm of hourly earnings. The numbers in [] are absolute values of t-statistics based on standard errors robust to heteroskedasticity. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively. (a) test indicates that the null hypothesis of equal coefficient of excluded instruments maybe rejected at 1% significance level. (b) Test of over-identifying restrictions indicates that the null hypothesis of valid instruments may not be rejected at 0.0001significance level. The critical value is 35.72. (c) Hausman test indicates that the null hypothesis of no difference between OLS and IV coefficients may be rejected at 5% significance level. The critical value is 16.91.

TABLE 8: SELECTIVITY-CORRECTED EARNINGS FUNCTION ESTIMATES

	Equation (1)		Equation (2)		Equation (3)	
Age	0.05*	[3.16]	0.06*	[3.90]	0.02	[1.30]
Age^2	-0.0004**	[2.00]	-0.0005*	[2.56]	0.0001	[0.43]
Male	0.01	[0.15]	0.08	[1.31]	0.09	[1.51]
Primary	0.17**	[2.08]	0.16***	[1.90]	0.59*	[5.95]
Secondary	0.67*	[7.30]	0.60*	[6.03]	1.48*	[9.55]
Tertiary	1.77*	[12.58]	1.63*	[10.74]	3.08*	[11.82]
Tenure	0.01*	[3.25]	0.01*	[3.22]	0.01*	[3.10]
Mombasa	-0.04	[0.61]	-0.03	[0.50]	-0.01	[0.14]
Nakuru	-0.70*	[7.75]	-0.67*	[7.74]	-0.69*	[8.05]
Eldoret	-0.49*	[5.28]	-0.49*	[5.47]	-0.49*	[5.35]
Father's education						
Primary			0.10***	[1.67]		
Post-primary			0.26*	[3.42]		
Mother's education						
Primary			-0.01	[0.19]		
Post-primary			0.17**	[1.98]		
Lambda					-0.33*	[5.46]
Constant	0.54***	[1.80]	0.10	[0.32]	0.21	[0.69]
Return to primary	0.03		0.02		0.11	
Return to secondary	0.11		0.09		0.24	
Return to tertiary	0.67		0.60		1.32	
Adjusted R^2	0.40		0.42		0.42	
Number of obs	843		843		843	

Notes: The dependent variable is the logarithm of hourly earnings. The numbers in [] are absolute values of t-statistics based on standard errors robust to heteroskedasticity. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively. (a) Test indicates that the null hypothesis of equal coefficient of excluded instruments maybe rejected at 1% significance level