Family Background, Education and Adult Earnings in Kenya

Anthony Wambugu
School of Economics, University of Nairobi
P. O. Box 30197-00100, NAIROBI, KENYA
anthony.wambugu@uonbi.ac.ke
anthony_wambugu@yahoo.com

Abstract

This paper investigates whether the education wage premium observed among workers is attributable to better family background-proxied by parental education. The study also presents instrumental variable (IV) estimates of private returns to education. Workers born of well-educated parents tended to attain more education and enjoy a parental education wage premium independent of their own education wage premium in the labour market. Second, the positive effect of parental education on workers' education and earnings increases monotonically with parental education level; and the effects are larger for father's education compared to mother's education. Third, OLS estimates of private returns to education are biased downwards relative to IV estimates. Finally, the study found a significant negative coefficient for self-selection term, implying that unobserved factors that lead individuals to acquire more education reduce earnings relative to the mean earnings of education level attained.

JEL Classification: J24; J31; I21; O15
Keywords: Family background, earnings function, returns to education, instrumental variables

1. Introduction

The theory of human capital predicts that more educated workers would have higher labor market earnings than less educated workers (Willis, 1986). Many studies support this prediction (Psacharopoulos and Patrinos, 2004). However, there are concerns that worker's education is endogenous, or more specifically, that omitted or unobserved determinants of both education attainment and earnings may be correlated Grilliches, 1977). Another concern, especially in developing countries is whether and to what extent education promotes intergenerational mobility. For instance, strong link between parental education and their children's education attainment and earnings can promote intergenerational transmission of socioeconomic status and curtail intergenerational mobility (Becker and Tomes, 1979).

In order to provide further insight into the role of family background in Kenya, this paper investigates whether a worker's family background has a wage premium or wage penalty in Kenya's labour market. The study addresses the following questions. Does omitting family background from human capital earning functions over estimate private returns to worker's own education? Which parent's education has greater effect on worker's education and earnings outcomes? Are OLS estimates of private returns to education different from IV estimates? Due to lack of data on the family background of each worker in labour force surveys, these issues have not been adequately studied in Kenya. These issues, however, have been the focus of attention by economists elsewhere following the work of Becker and Tomes (1979) and Griliches (1977). Kenya is an appropriate setting to examine these issues. Although educational expansion has been rapid (Knight and Sabot, 1987), earnings inequality is high (World Bank, 2009).

Previous studies of demand for education in developing countries find that children of more educated parents are more likely to enrol in school (Strauss and Thomas, 1995; Shultz, 1988). However, the samples used are censored and eventual education attainment of young children is unknown. There is relatively little study of adult workers' eventual education attainment mapped on measures of access to education at the time adult worker attended school. Further, pperiodic reviews of the educational return literature suggest that an additional year of education is associated with 10% wage premium (Psacharopoulos, 1994; 2004). However, OLS estimates of Mincerian earning equations common in this literature may be biased downwards due to measurement error in education or upwards due to correlation between unobserved worker characteristics, education and wages (Grilliches, 1977). Further, high discount rate workers acquire less education and OLS estimates of educational return may be upward biased (Card, 1995, 1999)

Several approaches are used to obtain unbiased estimate of educational returns. Some studies include ability measures in earning function. While some found that OLS estimates are upward biased (e.g. Grilliches, 1977; Blackburn and Neumark, 1995) others did not find biased OLS estimates (e.g. Knight and Sabot, 1990; Glewwe, 1996). Several studies include family background characteristics in earning functions. They found significant wage premium to better family background (e.g. Heckman and Hotz, 1986; Lam and Schoeni, 1993; Krishnan, 1996; and Kingdon, 1998). Armitage and Sabot (1987) found significant effect of family background on workers' returns to education in Kenya and Tanzania. However, they only consider private return to an additional year of education. This is restrictive because the effect of family background on private returns to education can vary across education levels (primary, secondary, university). They also do not distinguish the separate effects of mother's and father's education on workers' education and earnings. This is restrictive as the effect of fathers' and mothers' education may differ (Haverman and Wolfe, 1995).

Other studies use family background characteristics as instruments for education (e.g. Levin and Plug, 1999; Dearden, 1999; Uusitalo, 1999; Callan and Walker, 1999; and Oosterbeek and Ophem, 1999). Other instruments used include quarter of birth (Angrist and Krueger, 1991) and change in compulsory schooling law (Harmon and Walker, 1995). Ashenfelter and Krueger (1994) and Angrist and Newey (1991) use twins data and individual panel data. Bedi and Gaston (1997); Vella and Gregory, 1996; and Manda, (1997) treat the endogeneity issue as an education self-selection problem. In general instrumental variables estimates exceed conventional estimates by orders of magnitude in some studies and there is some evidence of self-selection.

The present study contributes to the human capital literature in several ways. First, it analyses a unique urban survey containing workers' family background and demand side variables like distances to nearest schools when the worker was of school going age. Typical labour force surveys in Kenya do not contain such information. Secondly, a survey of the literature by Haverman and Wolfe (1995) found that mother's education is more closely correlated with her children's education attainment compared to that of the father. Is this the case in Kenya? This study separated mothers and fathers to conduct a richer analysis of parental education effects on worker's education attainment and earnings. Thirdly, the study estimates a flexible earning function to examine the effect of family background on returns to different education levels. Finally, the paper not only focuses on the optimality of worker's education attainment, but also its deviation from predicted education due to self-selection.

The remainder of the paper is organized as follows. Section 2 discusses how family background might influence educational attainment and earnings. Section 3 presents the empirical models on which the analysis is based and Section 4 describes the data used. Section 5 presents the econometric estimates and Section 6 concludes the paper.

2. Optimal Education and Earnings

Human capital theory postulates that an individual's incentive to invest in education comes from expected net earnings. The wage relation is expressed as (Grilliches, 1977):

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

where w is wage, s is education, A is ability, and, μ is an unobserved worker attribute, independent of ability. Let us assume that individuals choose quantity of education to maximize present discounted value of wealth (V) and that cost of education is opportunity cost. Formally, the problem is

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

where r is the discount rate. The stopping rule for educational investment is:

$$(\partial w / \partial s) / r = w(s, A, \mu) \quad (3)$$

This means that the individual invests in education until the present value of marginal benefit, $(\partial w / \partial s) / r$ equals the marginal cost, ($w(s, A, \mu)$. With altruistic parents, perfect markets and education as a pure investment good, optimal quantity of education would be obtained (see Becker and Tomes, 1976, 1979; Behrman and Kennan, 1996). Individual's family background would have little impact on worker's education.

However, if the assumptions do not hold, marginal benefits and marginal costs of educational investments can differ by individual's family background. Specifically, parents with more education may be less credit constrained in financing education or invest more in other forms of human capital (e.g. health, and nutrition) to complement formal education. They may also acquire better information on education returns and hence face less uncertainty in educational investments. Finally, the opportunity cost of schooling in terms of foregone earnings maybe higher for less educated parents. Consequently, family-background specific differences in workers' education attainment may arise. The schooling demand function can be expressed as follows:

$$S = f(C, F, D) \quad (4)$$

where S denotes eventual quantity of schooling attained by a worker, C denotes individual worker characteristics, F represents family background and D denotes availability of schooling facilities at the time the worker was attending school.

3. Econometric Strategy

3.1. Earnings Function

We use the semi-logarithmic Mincerian earning function.

$$\ln w_i = X_i'\gamma + S_i\delta + u_i; \qquad u_i \sim i.i.d.N(0, \sigma^2) \quad (5)$$

where w_i is 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 to capture age specific differences in earnings; male dummy to capture gender specific differences in earnings; tenure in current firm; and location dummy variables to capture location specific differences in earnings. The unknown parameters are denoted by γ and δ and u_i is a random error term.

The first step is to estimate the earning function allowing for non-linearity in the effect of education by including the square of years of education completed (Willis, 1986; Bigsten et al, 2000). The second step is to add family background variables-parents' education in the earning function. Finally, education may be correlated with the error term in the earning function; therefore we use instrumental variable methods.

3.2. Educational Attainment Model

The effects of family background on the highest level of education completed by a worker are obtained from ordered probit model. Let us assume there is a latent variable, S_i^* measuring the highest education level completed by the *i*th worker. Hence

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

$$S_i = j \text{ if } \mu_{j-1} \le S_i^* \le \mu_j, \quad (6b)$$

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 to pick time and lifecycle effects. Distance to nearest school facilities when worker was of school going age measure school availability or direct costs of schooling. Regional dummies capture regional variations in education development and other region specific factors. Family background is measured by parents' education and ε is a random variable distributed as standard normal.

The predicted probability of *j* as the highest level of education is

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

where $\Phi(.)$ is the normal CDF. The model is estimated by maximum likelihood method. The marginal effects of the ordered probit model are computed following Greene (1997). The effect of family background on completed years of education is estimated using a linear model with same regressors as those in the ordered probit. The linear model is expressed as follows:

$$S_i = H_i'\psi + u_i \quad (7)$$

where S denotes number of years of schooling completed and the vector H contains the same explanatory variables as in equation (6a), u is the error term, and ψ is the vector of parameters estimated. The model is estimated by ordinary least squares method.

4. Data and Descriptive Statistics

This study analyses data from a unique firm survey of 200 manufacturing firms in four urban centers (Nairobi, Mombasa, Nakuru, and Eldoret) in Kenya. It was based on a sampling frame compiled by the World Bank for a panel survey in 1993-1995. The data used in this study were collected in October/November 2000 in face-to-face interviews with firm managers and a sample of up to 10 workers in every firm visited. It is the only wave that contains key education demand and supply variables required for the present study. 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 and school availability indicators. The analysis is based on a sample of 843 workers aged 16-64 years.

Table 1 provides description of each variable and reports sample statistics of the variables used in the study. Majority of workers were male and over 80% of the workers were in Nairobi and Mombasa, the centres of manufacturing in Kenya. The average age is about 35 years and average tenure is close to 9 years, suggesting low worker turnover. 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 lived more than 3 kilometres from the nearest secondary school when they were of secondary school age.

The average years of education completed is 9.6 years and about half of the sample completed secondary education. Only a tiny proportion never completed primary education. Education attainment among the workers' parents was low. Over 70% had attained primary education or less. Mothers had

lower education than fathers with over 40% of the mothers having no formal education. Clearly, workers attained more education than their parents, an indication of educational expansion in Kenya.

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. Results

The econometric results are presented in three parts. Examination of the effect of parental education on eventual education attainment of their children is based on educational attainment function estimates. To examine the effect of parental education on worker's earnings and returns education, we use earning function estimates with and without controls for family background. Finally, to gauge whether or not workers' education is endogenous the joint estimates of the education attainment function and the earning function are used.

5.1. Education Attainment Function Estimates

The education function estimates are reported in Tables 4 and 5. Equations (1) and (3) include worker characteristics and parents' education. In equations (2) and (4) dummy variables for region and distances to nearest schools were added. The following discussion is confined to equation (2) 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.

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. In Table 5, the ordered probit estimates do not show how changes in regressors affect predicted probabilities (Green, 1997 and Long, 1997). Hence marginal effects are calculated based on equation (2) and reported in Table 6.

Parental education significantly increases the eventual years of education a worker completed (Table 4). Table 6 shows that having an educated parent raises the probabilities of completing secondary education and university education. F-test rejected the null hypothesis of equality of father and mother's education coefficients (Tables 4 and 5). Differences in the effect of mother and father schooling on workers' education attainment may due to a number of reasons. First, parent's education impact may be gender-specific (see Glick and Sahn, 2000; Tansel, 1997). Second, the impact of the most educated parent may dominate. 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.

The coefficients of dummy variables for Eastern, Western, Nyanza and Coast regions are negative and statistically significant. Thus workers who attended most education in these regions completed fewer

years of education than those in Nairobi. Table 6 indicates that most workers were likely to have completed only primary education or lower. Such regional differences in education attainment may reflect income differentials, unmeasured school attributes such as teacher supply and facilities or household specific attributes.

5.2. Basic Wage Function Estimates

OLS estimates of the earning function are reported in Table 7 under equation (1a). The private return to education is about 0.14. The non-linear effect of education on earnings is captured under equation (2a). There the coefficients of years of education and education squared are statistically significant. Private returns are low for workers with few years of education and high for those with more years of education. The marginal 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 level to 0.67 at university level.[1] Age and tenure effects are significant in estimated equations. Earnings increase with age but at a decreasing rate; longer tenured workers earn more than short tenured ones. Coefficients on some loocation dummy variables are significant. Specifically, workers in firms located in Nakuru and Eldoret earn lower wages relative to workers in firms located in Nairobi (capital city).

5.3. Wage Function Estimates with Parents' Education

As noted by Lam and Schoeni (1993), OLS estimates of private wage returns to education are questioned on account of omitting family background from earning functions. Earning function estimates including controls for parents' education are reported in Table 7 under equations (1b) and (2b). A worker whose father has post-primary education earned 27% more compared to a worker whose father has no schooling. The corresponding earnings advantage is 19% for mother's post-primary education. A similar result is reported in Table 8 under equation (2).

In Tables 7 and 8, the effect of worker's education on earnings is generally lower by around 7-22% once controls for parents' education are included in the earning function. 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. Krishnan (1996) found that educational returns fell by 20% and 10% for workers in public and private sector respectively when family background is controlled for. But once selection effects into various employment sectors were taken into account, the effect of family background on educational returns fell by 5% and 4% in the respective sectors.

5.4. Wage Function Estimates with Endogenous Education

So far we have modeled the education relation and the wage relation separately as is common in the literature. However, as discussed earlier, OLS estimates of private wage returns to education may be biased and inconsistent if worker's education attainment is endogenous. To investigate this, the education attainment function and the wage equation are estimated using instrumental variable (IV) method. Distances to nearest schools, and regions where worker attended most schooling are identifying instruments.

[1]This is [exp(b2-b1)-1]*100 divided by years in a level of education (Halvorsen and Palmquist, 1980). b2 and b1 are coefficients of dummies for higher and immediate lower levels. Primary=7, secondary= 6, tertiary= 3.

The instrumental variable estimation results are reported in Table 7 under equations (1c) and (2c). The IV estimate of the private return to education in the basic wage function (equation 1c) is 0.24 compared to the OLS estimate of 0.14 in equation (1a). Hausman test rejects the null hypothesis of equality of OLS and IV estimates (see Table 7). In the alternative specification with square of education (equation 2c) the IV estimate of the return to education at 7 years of education is 0.04, which is lower than the OLS estimate of 0.09 in equation (2a). But the IV estimates of the return to education at 10 and 12 years are double the OLS estimates. Hausman test rejects the null hypothesis of equality of OLS and IV estimates in this case as well. Previous studies reviewed earlier also reported IV estimates twice or three times the OLS estimates in other countries.

There are concerns about IV estimation in general (Bound et al 1995; Murray, 2006). Although IV estimates differ from OLS estimates they may not be preferable if they are not robust to concerns about instruments validity. In the present study, instruments would be invalid if they are weakly correlated with workers' education and not excludable from wage equation. Using various tests the quality of instruments was assessed (Table 7). First, in the reduced-form education attainment equation the null hypothesis of equal coefficient estimates on instruments was rejected. This implies that the IVs are strongly correlated with workers' years of education. Second, over identification (OID) test (Deaton, 1997) was used. Residuals from the IV wage equation are regressed on all instruments. 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.

Educational investment decision can also be treated as a sel-selection problem. OLS would underestimate (overestimate) returns to education for those with comparative advantage in more(less) schooling intensive occupations (Bedi and Gaston,1997). Using reduced-form ordered probit model estimates where the dependent variable is the highest education level a self-selection term (lambda) is calculated from the residualsand included as an additional regressor in the earning function. The results are reported in Table 8 under Equation (3). First, the coefficient for lamda is negative and statistically significant. Statistical significance of this coefficients indicates that education is endogenous and the negative sign suggests that education over-achievers have characteristics that reduce their earnings below the average of their education group. It implies that OLS estimates are downward biased (Harmon and Walker, 1995). Table 8 shows that estimated education coefficients controlling for self-selection are at least twice the OLS estimates.

6. Summary and Conclusion

This paper investigates the role of family background in determining a worker's eventual educational attainment and adult earnings in Kenya's urban labor market. If individuals face same costs and benefits from educational investments, family background should not have any role in determining educational attainment. In addition, educational investment decisions may not be independent of the earnings determination process. Empirical evidence on the issue is thin due to lack of data on family background of adult workers in labour force surveys. Without such evidence, it is difficult to debate the role of education expansion in promoting equality of opportunity and intergenerational mobility.

The results show that workers whose parents had primary or post-primary education attained more education than those workers whose parents had no education. In addition, workers whose parents had post-primary education attained more education than those whose parents had completed only primary education. There also exists parent-specific differences in the estimated effect of parental education on workers' education attainment. In particular, the effect of father's education exceeds that of mother's education. However, we 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.

Earning function estimates show that workers whose parents had attained post-primary education receive higher pay than those whose parents had no education. Furthermore, the results suggest that part of the worker's own education wage premium is attributed to having better educated parents. However, workers own education still has a high pay-off. This is reinforced by instrumental variable estimates which suggest that conventional estimates of educational returns may be downward biased. In this context conventional estimates may be a lower bound on returns to education.

Having well educated parents may be associated with better labor market contacts and networks that lead to productive job search. For example, about 40% of workers in the sample secured current job through friends, family and relatives. It could also reflect ability to finance longer job search duration leading to higher pay jobs or jobs with greater opportunities to acquire on the job human capital. The parental education wage premium may also reflect better home learning environment and investments in health and nutrition that enhance the human capital acquired from a given quantity of education. This may command higher wages in the labor market. Parental education may be associated with greater probability of attending high quality schools. Workers who attended better quality schools may have greater skills and subsequent earnings.

Taken together, the results show that family background continues to have a key role in predicting workers' education and labour market earnings among industrial workers in Kenya. To the extent that differential educational attainment by family background reflects liquidity constraints or other constraints, the results are supportive of policy interventions toward greater education attainment for children of less educated parents. Such interventions have the potential to reduce intergenerational transmission of low socio-economic status from parents to workers through educational investments. This would in turn increase intergenerational social mobility.

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.

Armitage, J. and Sabot, R. (1987), Social economic background and the returns to education in two Low-income countries, *Economica* 54 (213): 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*, 84(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*, 87 (6), 1153-1189.

Bedi, A.S. and Gaston, N. (1999), Using variation in schooling availability to estimate educational returns for Honduras. *Economics of Education Review* 18, 107–116

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

Bound, J., Jaeger, D. and Baker, R.(1995), Problems with instrumental variables estimation when the correlation between the instruments and endogenous explanatory variables is weak, *Journal of the American Statistical Association*, 90: 443-50.

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*: 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, 543-550

Card, D. (1995), Earnings, education and ability revisited, in: S.Polachek (ed.), *Research in Labor Economics*, 14: 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, 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

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.

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: 63-87

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

Grilliches, Z. (1977), Estimating the returns to education: some econometrics problems

Econometrica, 45(1), 1-22.

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

Haverman, R., and B. Wolfe. (1995), The determinants of children attainments: A review of methods and findings. *Journal of Economic Literature* 23: 1829–78.

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,* 23, 507-542.

Kingdon, G. (1998), Does the labour market explain lower female education in India?, *Journal of Development Studies* 35(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), 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 (4), 710-40.

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

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

Manda, D. K. (1997), Labour Supply, Returns of Education, and the Effect of Firm Size on Wages: The case of Kenya. *Doctoral thesis 75*, University of Gothenburg.
Murray, M.P. (2006), Avoiding Invalid Instruments and Coping with Weak Instruments, *Journal of Economic Perspectives*, 20(4): 111–132

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

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

Psacharopoulos, G. and Patrinos, H. A. (2004), Returns to Investment in Education: A Further Update, *Education Economics* 12 (2): 111-34.

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). Schooling Attainment, Parental Education, and Gender in Cote d'Ivoire and Ghana, *Economic Development and Cultural Change*, 45(4): 825-56

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.

World Bank (2009), Kenya Poverty and Inequality Assessment. Poverty Reduction and Economic Management Unit Africa Region, Washington D.C., World Bank.

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	0.14

	of secondary school age was 6-10 kilometer, =0 otherwise	
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

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

TABLE 2: CORRELATIONS BETWEEN PARENTS' 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]
μ_1	1.28		0.57		1.29		0.59	
μ_2	2.82		2.17		2.83		2.19	
μ_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 education												
Primary					0.06	[0.99]	0.07	[1.26]				
Post-primary					0.24*	[3.05]	0.22*	[2.98]				
Mother 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