ECONOMIC STUDIES

DEPARTMENT OF ECONOMICS
SCHOOL OF ECONOMICS AND COMMERCIAL LAW
GÖTEBORG UNIVERSITY
124

ESSAYS ON EARNINGS AND HUMAN CAPITAL IN KENYA

Anthony Wambugu

ISBN 91-88514-83-8
ISSN 1651-4289 print
ISSN 1651-4297 online

ESSAYS ON EARNINGS AND HUMAN CAPITAL IN KENYA

2003

Anthony Wambugu
Department of Economics
Göteborg University

Abstract

Among Sub-Sahara Africa countries, Kenya has had a rapid educational expansion. This dissertation provides empirical analyses of the impact of education on labor earnings in Kenya, based on surveys of manufacturing firms and a survey of households in the 1990s. It consists of four papers.

Paper 1 examines whether real earnings and private returns to education in manufacturing labor market changed over the 1990s. Results indicate that, real earnings standardized for differences in observed worker and firm characteristics rose over the survey period. But returns to human capital were constant. Further, the results indicate that returns to education are highest for workers in the top part of the earnings distribution, suggesting that, education worsens earnings inequality among manufacturing workers.

Paper 2 uses the 2000 wave of the manufacturing firms survey to examine whether failure to control for family background in earnings functions, or to treat education as endogenous to wage formation, results in significant bias in estimates of private returns to education. Parental education has significant impact on a worker's education, and estimates of the effect of education on wages in Kenya's manufacturing that do not control for parental education are upward biased. When education is instrumented, results suggest that, standard estimates of private returns to education may be downward biased if endogenous schooling is not modelled. But this hinges on the validity and quality instruments.

Paper 3 analyses a household survey to identify the impact of education on employment and earnings. All levels of education reduce the chances of agricultural employment, while higher education reduces the chances of entry into the informal sector also. Perhaps it is because education raises private and public sectors entry probabilities. Decomposition results indicate that, differences in individual and household characteristics explain a substantial part of the women-men gap in sector entry probabilities. Returns to primary education are highest in the informal sector while returns to secondary education are highest in the private sector. Women have higher returns to education than men, and selectivity controls in the earnings function indicate no evidence of selectivity bias except for women in the public sector.

Paper 4 also analyses a household survey to examine the impact of education on households' economic activity combinations and incomes. Results suggest that, as education increases in the household from primary to secondary level, the tendency is to diversify income-generating activities. At higher levels of education, households tend to generate income from wage employment only. Further, the impact of education on total earned income is substantial. And while the largest impact of education is on household wage income, lower education has a notable impact on household farming income. This result is important, as many primary graduates are likely to enter the farming sector.

Keywords: education, human capital, earnings, family-background, employment, household, activity farming, own business, informal sector, Kenya.

Acknowledgements

I have learned and received help from many people during my studies in Göteborg. I am immensely grateful to each and every one of you, and I wish you well. I thank Arne Bigsten, the dissertation advisor. His advice and suggestions have made the work reach this stage. He is always pleasant to listen to. I thank the Lecturers and Professors that taught me in the graduate program. I also thank Henry Ohlsson, Simon Appleton, Katarina Nordblom, Francis Teal, Sten Dieden, Wilfred Nyangena, Christer Ljungvall, Karin Kronlid, Anders Isacksson, John Obere, Germano Mwabu, Roger Wahlberg, Måns Söderbom, Joseph Wang'ombe, Abebe Shimeles, Fr. Amadeo Paolino, Björn Ohlsson, and Karl Lundvall. I am very grateful for the help received from Eva-Lena Neth Johansson, before coming to Sweden and since then. I knocked at the offices of Eva Jonason and Gunilla Leander sometime. I thank them for their help too. I have received much support from my family and I appreciate that very much. The African Economic Research Consortium fellowship for doctoral studies and stipendium from Göteborg University are acknowledged with gratitude.

Tack så mycket.
Göteborg, den 7 januari 2003
Anthony Wambugu

TABLE OF CONTENTS

Introduction

Kenya has had one of the most rapid educational expansions in Sub-Sahara Africa since 1963, the year of political independence. In the 1970s when the government could not meet the demand for secondary education, local communities pooled resources to increase the number of secondary schools. Bigsten (1984) and Knight and Sabot (1990) discuss educational expansion in Kenya and study its impact on incomes. Hughes (1991) argues that the demand for education in Kenya is closely tied with smallholders' economic strategies, where decline in land sizes pushes individuals to search for wage employment, to supplement household budgets. But to gain access into wage employment more and more education is required.

This dissertation uses micro-economic data from firms and households, collected in the 1990s, to investigate earnings determination with emphasis on the role of education. Education is one among several dimensions of human capital. In a recent book, Kooreman and Wunderlink (1997) define human capital as "all those qualities of a person, such as knowledge, health, skills and experience, that affect his or her possibilities of earning current and future money income, psychological income, and income in kind"(pp 181). This definition illustrates the multi-dimensional nature of human capital. Schultz T.W. (1960, 1961, 1975) recognized that investment in human capital is an important way to improve the welfare of people around the world and urged economists not to be hesitant in the inquiry of human capital investments and returns. A substantial amount of research has been conducted since then as the surveys by Schultz, T. P. (1988), Strauss and Thomas (1995), Appleton and McKinnon (1996), Appleton (2000) and Psacharopoulos (1994) illustrate.

Many commentators consider education to be a crucial factor in many aspects of the development process. For example, investments in education are emphasized as one way to reduce poverty in less developed countries (see World Bank, 2000). And the study of the relationship between education and incomes can improve understanding of income distribution (Bigsten, 1984). Research into the role of education has increased in recent years for at least five reasons (Knight, 1996): (i) accumulation of evidence that education represents human capital; (ii) suitable micro-economic data in less developed countries; (iii) emphasis in new growth theory on human capital and externalities; (iv) the interactions between education and other dimensions of human capital such as better health and nutrition, and (v) the strong growth of some East Asian countries is partly attributed to educational investments.

Substantial empirical micro-economic evidence has accumulated since the 1960s indicating that education and labor market earnings are positively correlated. For example, a worker in Sub-Sahara Africa earns 13 per cent more for each additional year of education compared to 7 per cent in OECD countries. Workers, who complete primary education in SSA, earn on average 41 per cent more than their counterparts with no education (Psacharopoulos, 1994). If the figures hold, then education is a profitable investment for individuals in SSA. However, there are concerns and scepticism regarding the level of economic returns to education.[1]

In estimating economic returns to education it is assumed that wage differentials do not change in response to changes in labor market conditions. The objective in "*Real Wages and Returns to Human Capital in Kenyan Manufacturing*" is to test empirically whether real

[1]Haveman and Wolfe (1984) provide a discussion of other important benefits of education

wages and returns to education changed over the 1990s across quantiles of the earnings function. The human capital earnings function (Becker and Chiswick, 1966; Mincer, 1974; Willis, 1986) is used to analyse earnings of workers in manufacturing enterprises located in four urban centres in Kenya.[2] Because of its desirable features the human capital earnings function has come to dominate research on earnings (see Chiswick, 1997). Hence it is the main tool of analysis used here. In addition to a test of whether returns to education changed over the 1990s, the study applies recently developed techniques for quantile regression analysis to examine changes in returns to education across the earnings distribution based on a one-group model (Bushnisky, 1994). The results show that real wages standardized for worker and firm characteristics changed upwards in the 1990s. The other finding is that while Mincerian returns to education vary across quantiles of the earnings distribution, they seem to be stable over the survey period.

The paper on*"Family Background, Education and Earnings in Kenya***"**, addresses another concern about estimates of returns to education. It is often argued and some empirical evidence (e.g. Lam and Schoeni, 1993) suggests that, returns to education may be subject to omitted family background bias. The question is whether failure to control for family background injects substantial bias into the estimates. New micro-data are used to examine the potential omitted family background on economic returns to different levels of education. The standard Mincerian earnings function is used and the finding is that the bias is lower than in other countries where standard estimates are on average 20 per cent higher when family background is not controlled for. The study also examines, in the context of a less developed country, whether failure to treat education as endogenous to wage formation results in significant bias in education effects. This issue has received much attention in developed countries but little evidence is available for less developed countries.[3] The earnings analysis in this part uses a two-equation model and Instrumental variable method. The results suggest there is some bias, a result that is in line with studies in developed countries that instrument for education in wage functions. But the result depends on quality and validity of instruments.

Surveys of urban enterprises can improve understanding of how urban labor markets operate. But, they also raise other questions. In Kenya, going by trends in the last decade majority of labor market entrants are not likely to obtain urban wage jobs. Instead, they enter small-scale agriculture and the informal sector. For example, between 1990 and 1999, the number of Kenyans in the informal sector increased by over 200 per cent (Government of Kenya, 2001). Returns to education will depend on the effect education has on access and incomes in more than one sector. In "*Education, Employment and Earnings in Kenya*", the importance of education on access to five employment types (public sector work, private sector work, informal sector work, agriculture, and unpaid family work) is analysed. The data used are from a survey of rural and urban households. Estimates of a five-way multinomial logit model indicate that education is highly correlated with employment type. In particular, education is essential to access wage employment. Hughes (1991) notes that such a link is likely to fuel demand for education where wage jobs are few. Decomposition of women-men differential in employment allocation probabilities suggests that a substantial part is

[2] The data come from surveys of manufacturing enterprises organized under the World Bank's Regional Program on Enterprise Development (RPED). They were collected in 1993, 1994, and 1995. To these is added data collected in 2000 from more or less the same enterprises but organized under the United Nations Industrial Development Organization (UNIDO).

[3] The data used are from the 2000 survey described in footnote 3. Some unique variables related to family background and availability of school facilities were collected in this survey wave but not in the other survey waves.

accounted for by individual and household characteristics. The study also conducts earnings function analysis for the public sector, private sector and informal sector. Studies usually consider only the first two sectors. But in Kenya majority are outside formal wage sector. The notable result is the positive income returns to primary education in the informal sector. A joint model of employment assignment and earnings determination is used in the analysis. The results suggest that selectivity bias might not be a major problem.

In "*Education and Household Earned Income in Kenya*" the focus is on the multiplicity of activities from which households generate income. It examines the relationship between education and the economic activity combinations in which a household derives income. This is important especially for economic strategies of smallholders faced with land scarcity. For this part of the study, a discrete choice multinomial logit model is employed. The data are from a survey of rural and urban households. The data show that, the practice of activity combination is common in Kenya. Even in rural areas where farming is said to be the dominant activity, there are hardly any households that are pure farmers.

Second, at low levels of education, the household tends to diversify into different economic activities. But as a household gains access to higher education, there is tendency to specialize and rely on wage employment. In all the activity combinations, wage earnings constitute a substantial share of household total earnings. The study also examines the income gains associated with different quantities of education. Earnings functins are estimated for total household earnings, farm earnings, wage earnings, and own business earnings. The results show that education has positive impact on total household earnings. In addition, an encouraging result is that primary education improves farm earnings. This suggests the type of modernization taking place in farming rewards primary educated labor. Higher levels of education are rewarded mainly in wage employment.

References

Appleton, S., Hoddinot, J., and McKinnon, J. (1996), Education and health in Sub-Sahara Africa, *Journal of International Development*, Vol. 8, 3, 307-339.

Becker, G.S. and Chiswick, B.R. (1966), Education and the distribution of earnings, *American Economic Review*, 56, 358-69.

Bushnisky, M. (1994), Changes in the U.S. wage structure 1963-1987, Application of quantile regression, *Econometrica,* 62, 405-58.

Chiswick, B. (1997), Interpreting the coefficient of schooling in the human capital earnings function, *Policy Research Working Paper 1790*, The World Bank

Haveman, R. and Wolfe, B. (1984), Schooling and economic well-being: the role of non-market effects. *Journal of Human Resources*, 19, 377-407.

Hughes, R. (1991), Examining the roots of educational demand: the case supporting rural agrarian development, *World Development*, vol. 19, No. 2/3, pp 213-223.

Knight, J. (1996), Human capital in economic development. Editorial introduction, *Oxford Bulletin of Economics and Statistics*, 58, 1

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

Kooreman, P. And Wunderlink, S. (1997), *Economics of household behavior*, St. Martins Press.

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

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

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

Schultz, T.P. (1988), Schooling Investments and Returns, in: H. Chenery and T.N. Srinivasan, *Handbook of Development Economics*, vol. 1 (Amsterdam, North Holland), 543-630.

Schultz, T.W. (1960), Capital formation by education, *The Journal of Political Economy*, Vol. 68, Issue 6, 571-583.

Schultz, T.W. (1961), Investments in Human Capital, *American Economic Review*, 51, 1-17.

Schultz, T.W. (1975), The value of the ability to deal with disequilibria, *Journal of Economic Literature*, 13, 872-876.

Willis, R.J (1986), Wage determinants: A survey and reinterpretation of human capital earnings functions, in: O.Ashenfelter and R.Layard, eds., *Handbook of Labor Economics*, Vol. I (Elservier Science).

World Bank (2000), World development report, 2000/2001: Attacking poverty, Oxford University press, New York.

Real Wages and Returns to Human Capital in Kenyan Manufacturing*

By

Anthony Wambugu
Department of Economics
Göteborg University
Box 640
SE-405 30 GÖTEBORG
Sweden

2002a

Abstract: This paper studies how real wages and private wage returns to human capital in Kenya manufacturing sector changed over the 1990s. The analysis uses employer-employee matched data from a survey of firms conducted in 1993, 1994, 1995, and 2000. Quantile earnings regressions are used to describe the conditional wage distribution over this period. Among workers in the median and in the bottom and top quantiles of the wage distribution, the wages of the more educated are higher than for the less educated. The wage premia to education for workers in the top quantile is higher than that of workers in the bottom quantile. The results suggest that education has a positive effect on manufacturing wage inequality. Unmeasured factors may complement schooling in wage determination giving rise to differences in wage premia to education across quantiles. The regression estimates also indicate that over the survey period, the real wage standardized for observable characteristics of workers and firms increased in all quantiles of the wage distribution, while the wage premia to education was stable.

JEL Classification: J3 O1
Keywords: Quantile regression, returns to schooling, Kenya

*I thank Arne Bigsten, Simon Appleton, Francis Teal, Måns Söderbom and seminar participants for comments and suggestions on earlier versions of the paper. Thanks to Arne Bigsten for the data and opportunity to join the team for the 2000 survey.

1. Introduction

Real wages and private wage returns to human capital have been a concern of development research for a long time. In particular, the analysis of investments in and returns to human capital has received much attention since the work of Schultz (1960, 1961, 1975). In recent years, education has been emphasised because investments in the poor people's human capital is considered a potential way to reduce poverty (World Bank, 2000). Changes in the structure of wages and private wage returns to human capital could also provide insights into how labor markets operate to reward skills and influence wage earnings distribution. For example, in the developed countries there were large changes in wages and returns to skill during the 1980s and 1990s. Katz and Autor (1999) survey the theory and empirical evidence. The explanations for the changes include shifts in factors that influence demand and supply for labor and changes in technology. Because demand and supply factors are likely to have changed in less developed economies, wages and returns to skills may have changed there also. Little empirical evidence is available about changes over time. This study inquires into real wages and private wage returns to human capital in Kenya over the 1990s.

During the 1990's, Kenya's economy performed poorly.[1] The growth in real GDP was less than 2.5 per cent in six out of nine years between 1991 and 1999, while the average rise in population was close to 3 per cent. This means that per capita GDP stagnated or declined. The rates of growth in agriculture and manufacturing sectors were low. For example, in five of the years manufacturing recorded rates of growth below 2 per cent while agricultural output declined in some years. In the 1990s also, formal wage employment expanded very slowly and many workers are now absorbed by the informal sector (excluding small-scale farming).[2] The sector expanded by almost 250 per cent over this period (Government of Kenya, 2001). In the early 1990's, the government instituted economic reforms including, removal of price controls, freeing the foreign exchange rate, and other trade and financial sector reforms. The reforms and the poor economic performance may have had an impact on the performance of firms and by extension, wages and employment patterns.

Returns to human capital may have changed in this period not only due to low demand for labor occassioned by poor economic performance but also due to the continued expansion in supply of educated labor. A review of several studies on changes in returns to human capital in less developed countries (Pritchett, 2001) finds that returns may increase, decrease, or remain stable over time. In Kenya, Appleton, Bigsten and Manda (1999) find that returns to education for workers in urban areas declined between 1978 and 1995 particularly for secondary graduates. But Appleton (2002) notes that it is not known how returns to education in Kenya changed over the 1990s unlike in Uganda where he finds a rise in returns.

The aim of this paper is to inquire into what happened to real wages and private wage returns to human capital for manufacturing sector workers in Kenya over the 1990s. A survey of manufacturing firms is used. It comprises four waves conducted in 1993, 1994, 1995 and 2000. Previous estimates of returns to human capital in Kenya are based on ordinary least squares earnings regressions. However, recent studies in developed countries (e.g. Bushnisky,

[1] Table A1 presents some economic and education indicators for the 1990's

[2] In Kenya it covers a wide range of activities. For example, shoe shining, road-side sellers, door-to-door traders, small-scale artisans and metal workers, and food kiosks. Many of these activities use only small amounts of physical capital.

1994, 1998; Machado and Mata, 2001) show that both the level and change over time in returns to skills and experience can differ across the earnings distribution. At a theoretical level, Card (1995) presents a model in which there is variation in returns to education across individuals. Also, a focus on the whole earnings distribution is important because changes in returns to education and experience have implications for earnings inequality. In recent years there is renewed emphasis on income inequality in development research. Therefore in addition to the standard earnings regressions the analysis uses quantile earnings regressions to obtain a broader view of the levels and changes in wages and returns to human capital.

The next section, the issues in the literature on returns to education in less developed countries are outlined. Section 3 reviews the two methods commonly used to measure returns to education. Section 4 describes the data used in the analysis and Section 5 specifies the empirical model. The estimation results are presented in Section 6 and section 7 concludes.

2. Education and Earnings: A Survey of Issues

A survey of returns to education by Psacharopoulos (1994) summarizes an aggregate pattern where returns to schooling are (i) higher in private sector employment than in public sector employment; (ii) highest at primary level and lowest at tertiary level.; (iii) higher in developing countries especially in Africa, than in developed countries; and (iv) higher for women than men. However, Bennell (1996) argued that the pattern of returns to education is unlikely in SSA because the surveyed studies are based on diverse methods, data quality, and countries that differ in size and records of economic performance. He proposed that, it is better to search for patterns in returns to schooling at country level. Besides, a survey (Appleton et al, 1996) of Mincerian returns to education for several SSA countries shows that returns are higher for higher education levels. Recent estimates of returns to education (e.g Bigsten et al, 2000; Mwabu and Shultz, 2000; and Jones, 2001) report a similar result.

A major source of skepticism about estimates of returns to schooling is that observed wage differences between workers may fully or partly reflect differences in ability and not productivity differences due to schooling. So if worker ability is omitted the estimates may be biased. Available empirical evidence does not appear to support this. A detailed study (Knight and Sabot, 1990) shows that while ability (reasoning ability test scores) had a small effect on earnings, this did not reduce the impact of human capital (cognitive skills). Other omitted variables have been considered. For example, failure to control for family background may inject upward bias into estimated returns to education (e.g Lam and Schoeni, 1993). Behrman and Birdsall (1983) and Glewwe (1996) find that school quality is correlated with earnings.

Farm production function estimates (e.g Pinckney and Kimuyu, 1995; Appleton and Balihuta, 1996) and manufacturing production function estimates (e.g. Bigsten et. al. 2000) provide more direct productivity effects of human capital. Bigsten et al, (2000) estimate earnings and production functions based on firm surveys in five African countries. They find positive wage returns to human capital and relatively higher returns to physical capital. Jones (2001) uses data from Ghana's manufacturing firms to examine the relation between education and earnings. She finds that more educated workers are more productive than less educated workers and firms pay wages that reflect productivity. This supports a key assumption of a competitive labor market, that wages reflect value of marginal product.

Another source of skepticism about standard estimates of returns to education is that they are largely based on samples of wage earners. The question is whether returns to schooling for wage earners are a good guide to returns to schooling in other forms of employment. In Uganda, Appleton (2001) finds that there is no major difference between the returns to education in farming, wage employment, and self-employment. Empirical evidence is required for other countries to see if this is also the case there. A related concern is that, wage benefits on which standard returns to education are based, exclude externalities and direct consumption benefits. Schultz (1988) notes that these benefits may be large. Although such benefits are difficult to measure, Appleton and Balihuta (1996) and Weir and Knight (2000) have done this with respect to farm production. They find that in Uganda and Ethiopia respectively, the education of neighboring farmers has positive effect on an individual farmer's output. That is having an educated neighbor promotes better farming.

With regard to changes in wage returns to education, the assumption in the standard model is that wage increment due to additional education is constant over time. There is some empirical evidence to suggest that this may not be the case. For example, Moll (1996) found that in South Africa, the return to primary education for Africans declined between 1960 and 1975, but stabilized thereafter. In contrast, the return to secondary schooling remained strong in this period. In Ghana the return to secondary and post-secondary schooling in rose in 1987-1991 (Canagarajah and Thomas,1997). On the other hand, Krishnan, Sellassie, and Dercon (1998) find that returns to education in Ethiopia's urban labor market did not respond to labor market reforms between 1990 and 1997. But in Uganda, Appleton (2002) finds a rise in returns to education over the 1990s.

In developed countries, empirical work on changes in wage structure and returns to human capital (e.g Bushnisky, 1994, 1998 and Machado and Mata, 2001) concentrates on the whole wage distribution. The results show that during the 1980s and 1990s, there were large changes in returns to education and experience for workers at different points on the wage distribution, and the return to education is not identical across the wage distribution. Little research of this nature is available from Africa with the exception of Mwabu and Schultz (1996) on South Africa and Nielsen and Rosholm (2001) on Zambia.

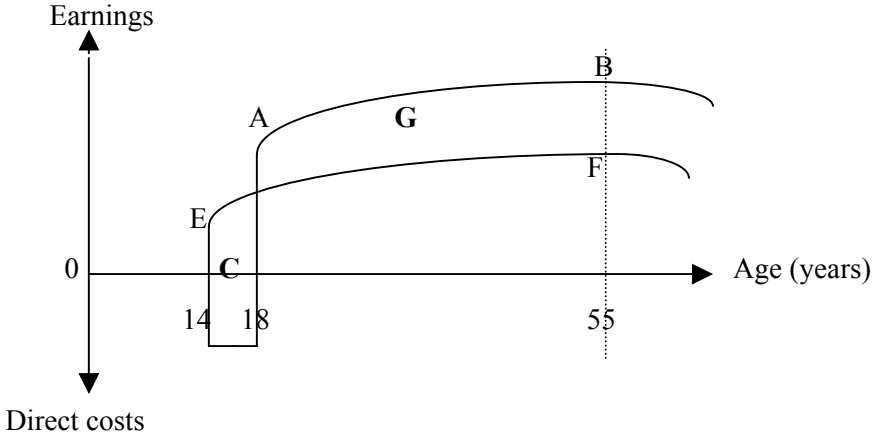
In summary, the section highlights several issues, but not all are taken up in this study. Instead, the objective of the study is a modest one. It is to examine whether real wages and private wage returns to education for workers in Kenya maufacturing enterprises are identical for low wage and high-wage workers and the changes in the 1990s. Before moving on to the empirical work the next section sets how returns to human capital are measured.

3. Measurement of Returns to Human Capital

The human capital model is the basis for the measurement of returns to schooling. The idea is that an individual spends part of his/her lifetime in school to enhance their productivity. It is assumed that schooling imparts skills that have value in the labor market. However, to acquire the skills, the individual foregoes earnings during the time spent in school. The wages paid should reflect the worker's productivity. One method to estimate returns to investments in human capital is to compare costs and benefits (Psacharopoulos, 1995). An important concept in this regard is that of the age-earnings profile. Figure 1 illustrates the age-earnings profiles

of two workers. One worker studied up to primary level and the other studied up to secondary level. Assume the primary graduate entered the labor market aged 14. The age-earnings profile may look like EF. Labor market earnings rise with age at first and then decline with age. The secondary graduate entered the labor market aged 18. The cost (*C*) of 4 years of secondary education has two components: direct cost and opportunity cost (foregone earnings). The age-earnings profile may look like AB. The earnings gain from secondary education is *G*.

Figure 1: Age-earnings profiles



The return to human capital is the discount rate that would equalize the sum of present discounted stream of schooling costs, to the sum of present discounted stream of wage benefits. In this illustration, the rate of return to secondary education would be that discount rate (r_s) that satisfies the expression (1).

$$\sum_{t=1}^{37} \frac{(w_s - w_p)}{(1+r_s)} = \sum_{t=1}^{4} \frac{(w_p + c_s)}{(1+r_s)} \qquad (1)$$

where w_s is the earnings of a secondary graduate and w_p is the earnings of a primary graduate. The left hand side represents the benefit and the right hand side represents the costs. The difference $(w_s - w_p)$ is the earnings gain labeled *G* in Figure 1, which the graduate will receive for 37 years. It comes at a cost $(w_p + c_s)$ during 4 years of secondary schooling.

The second, and more widely used method is the human capital earnings function. The simple schooling version is due to Becker and Chiswick (1966) while Mincer (1974) introduced work experience into the model. Willis (1986) provides a survey of the theory and empirical literature on the development of the human capital earnings function. The basic Mincerian human capital model relates the natural logarithm of earnings (*wage*) to years of schooling (*sch*) completed, years of labor market experience (*exp*), and years of labor market experience squared. The square term in labor market experience accounts for the curvature depicted in Figure 1. The basic earnings function is

$$\ln wage = f(sch, exp, exp^2) \qquad (2a)$$

The schooling coefficient is interpreted as an estimate of the Mincerian rate of return to schooling and assumed to be constant across different levels of schooling. To estimate returns to education at different points of the schooling distribution, the basic model is extended with years of schooling completed entered as a quadratic (see Willis, 1986 for this model and Bigsten et. al., 2000 for an application).

$$\ln wage = f(sch, sch^2, exp, exp^2) \quad (2b)$$

The return to a small increment in schooling in this model is the partial derivative with respect to schooling evaluated at a given point on the education-earnings profile.

Another flexible formulation of the earnings function is obtained if we break up the total years of schooling into years spent at each schooling level.

$$\ln wage = f(prim, sec, post, exp, exp^2) \quad (2c)$$

where *prim* is the years of primary school, *sec* is the years of secondary school, and *post* is the years of post secondary school. This will yield returns to education within a given level.

The Mincerian returns to schooling from the earnings functions above would equal private wage returns to schooling if (i) the cost of schooling is the opportunity cost of the student's time, that is, earnings foregone when attending school[3]; (ii) earnings differentials reflect productivity differentials; (iii) individuals live for ever and (iv) the increment in earnings is constant overtime. The latter assumption is relaxed later so as to test whether returns to schooling changed during the 1990's.

4. Empirical Specification

This section lays out the econometric model and estimation procedure used in this paper. The human capital earnings function described in Section 3 is the main tool of analysis. First, ordinary least squares is used to estimate semi-logarithmic earnings equations to obtain the effect of education on expected log earnings as is common in the literature. Then, to describe the entire conditional earnings distribution, the earnings equation is estimated using the quantile regression estimator introduced by Koeneker and Basset (1978) at three quartiles: lower quartile (25th percentile), median (50th percentile), and upper quartile (75th percentile). With larger data set earnings functions can be estimated at more quartiles to give a richer description of the data.

The advantages of quantile regressions include greater resistance to outliers in the dependent variable, a more detailed description of how explanatory variables correlate with the dependent variable, and it is a way to discover heteroskedasticity in data (Deaton, 1997). In the present application, quantile regressions describe how economic returns to human capital vary across quantiles of the earnings distribution. The schooling coefficient at the lower quartile shows the schooling effect for workers at the lowest 25 per cent of the wage distribution. Estimates at the median show the schooling effect for workers at the middle, and

[3] The foregone earnings for a particular level of schooling are the earnings of schooling level immediately below.

estimates at the upper quartile show schooling effect for workers in the top 25 per cent of the wage distribution.

Following Bushnisky (1994, 1998), the quantile regression model of the earnings function can be specified as follows:

$$\ln w_i = x_i^{'}\beta_\theta + u_{\theta_i} \tag{3a}$$

$$Quant_\theta(\ln w_i \mid x_i) = x_i^{'}\beta_\theta ; Quant_\theta(u_{\theta_i} \mid x_i) = 0 \tag{3b}$$

where *w* denotes real hourly wage, *x* is a vector of explanatory variables, and u_θ is a random error term. The *i = 1,……,n,* indexes individual worker and *n* is the number of workers in the sample. The parameter vector is denoted by β_θ and $Quant_\theta(lnw_i|x_i)$ is the θth conditional quantile of *lnw* given x_i. The estimation procedure is that of Bushnisky (1994, 1998). Unlike in least squares where parameter estimates minimize the sum of squared errors, quantile regression parameters minimize the absolute sum of the errors from a particular quantile of the log earnings across workers. The problem is to obtain the θth quantile regression parameters to

$$Min\left\{\sum_{i:\ln w_i \geq x_i^{'}\beta_\theta} \theta \mid \ln w_i - x_i^{'}\beta_\theta \mid + \sum_{i:\ln w_i < x_i^{'}\beta_\theta} (1-\theta) \mid \ln w_i - x_i^{'}\beta_\theta \mid\right\} \tag{4}$$

If $\theta = 0.50$, this is the median regression or least absolute deviation (LAD) estimator. Other quantile regressions are estimated by weighting the absolute sum of the errors. For example, when $\ln w_i \geq x_i^{'}\beta_\theta$ the deviation is positive and the weight used is θ. When $\ln w_i < x_i^{'}\beta_\theta$ the deviation is negative and the weight is $1-\theta$. The solution to expression (4) is obtained by setting up a linear programming problem for the full sample and then linear programming algorithms are used to obtain the solution. The paper estimates earnings functions at three quantiles simultaneously. This allows hypotheses testing of cross-quantiles restrictions. For example, are the education effects identical in the bottom and top quantiles? To avoid understating the standard errors a bootstrap method is used (Bushnisky, 1994).

The earnings function analysis is in two parts. First, Mincerian earnings functions are estimated on a sample pooled across survey waves. The specification includes dummy variables for survey waves and the coefficients on these dummies form the basis to test whether the level of real wages standardized for sample composition changed over the 1990s. Economic returns to education across quantiles are also computed and compared. Second, the standard Mincerian earnings function is extended to include controls for firm characteristics and occupation of the worker, to see the effect this has on the magnitude of estimated returns to education across quantiles and on change in real wages. Third, the standard Mincerian earnings function is allowed to differ across the survey waves by including interactions between each survey wave and the explanatory variables. This helps to answer the question of whether education effects on earnings changed over the 1990s and whether the change was identical across the earnings distribution.

5. Data and Summary Statistics

The paper analyses employer-employee matched data from surveys of enterprises in Kenya. The first three waves (1993, 1994, and 1995) were organized under the World Bank's Regional Program on Enterprise Development (RPED). Nine countries (Burundi, Camerron, Cote d' Ivoire, Ghana, Kenya, Rwanda, Tanzania, Zambia, and Zimbabwe) were covered by the RPED. The Kenya RPED survey was funded by the Swedish International Development Agency (SIDA). A joint team of the Department of Economics, Gothenburg University and Department of Economics, Nairobi University undertook the surveys. In 2000 the United Nations Industrial Development Organization (UNIDO) funded a fourth survey, that followed as closely as possible the enterprises in the RPED. It was conducted by a joint team of the Center for the Study of African Economies (CSAE), Oxford University in collaboration with the Department of Economics, Gothenburg University, University of Nairobi, and the Federation of Kenya Employers (FKE).

The survey of Kenya manufacturing cover firms located in the capital, Nairobi; Mombasa, the main sea port and two inland urban centres (Nakuru, and Eldoret). The first wave of RPED was in February-March 1993, the second in May-June 1994, and the third in August-September 1995. A detailed account of the Kenya RPED survey and some studies is in Bigsten and Kimuyu (2001). The firms are in four main sub-sectors that comprise about 73 per cent of all manufacturing employment: wood, textiles, food, and metal sub-sectors. Seventy-five per cent of the primary sample are formal firms and 25 percent are informal firms. The formal firms are a random sample from the Central Bureau of Statistics file of registered firms as it was the best available source. For the informal firms a sampling frame of firms in the four urban centres was constructed. In the RPED surveys, letters of introduction were sent to formal firms while for the informal firms a team went directly and requested for an interview. A total of 224 enterprises were interviewed in 1993, 216 in 1994, 218 in 1995, and 190 in 2000. In waves two, three, and four, some firms had to be replaced because they had closed down, declined to be interviewed, or could not be retraced.

In every firm interviewed, up to ten workers among those present on the interview date were chosen at random. A workers' questionnaire was used to collect information about individual wages, non-wage benefits, tenure in current firm, and individual characteristics such as age and education attainment among other information. Tables 1 and 2 present summary statistics of variables used according to survey wave. Gender is represented by a dummy variable for men. Most of the workers (85 per cent) are men. On average workers are aged 34 years and time spent with the firm is 9 years. The long tenure may suggest labor turnover rate is low. Both age and time in the firm are expected to be associated with higher earnings according to human capital theory. Age captures both an experience and age effect while tenure may be associated with learning and acquisition of firm-specific skills. Occupation attainment is represented by a set of dummy variables. The data show that most (58 per cent) of the workers are in production department.

Table 1: Summary Statistics of Continuous Variables for each Survey Wave and for Full Sample

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Total
Age (years)					
Mean	35	33	33	35	34
Median	33	31	32	33	32
SD	9	8.7	8.9	9.4	9
Tenure (years)					
Mean	8.1	7.3	7.6	8.5	7.9
Median	6	5	6	6	6
SD	7.1	6.9	7.1	7.6	7.2
Education (years)					
Mean	8.8	8.8	9.2	9.6	9.1
Median	9	9	9	11	9
SD	3	2.9	2.9	2.6	2.9
Employment					
Mean	176	117	146	168	152
Median	60	50	54	50	53
SD	468	261	316	297	351
Output/worker (1990 Kshs)					
Mean	393,990	634,700	584,886	526,107	531,279
Median	215,115	217,432	268,885	266,783	242,963
SD	513,815	1,507,690	1,086,705	730,484	1,029,260
Capital/worker(1990 Kshs)					
Mean	413,595	403,892	459,904	805,060	508,869
Median	206,346	204,955	229,152	378,780	236,737
SD	566,369	571,206	577,197	1,742,257	970,540
Real wage/worker(1990 Kshs)					
Mean	14, 545	14, 374	21, 824	33,559	20,603
Median	10, 790	11, 662	14, 837	20, 444	13,352
SD	12, 569	14, 284	37, 548	68, 910	39,566

Table 2: Summary Statistics of Dummy Variables for each Survey Wave and for Full Sample (per cent)

Variable	Wave 1	Wave 2	Wave 3	Wave 4	Total
Male worker	89	85	85	81	85
No education	1	0	0	0	0
Some primary	16	16	13	8	13
Full primary	45	43	44	39	43
Full secondary	38	39	39	48	41
University	1	1	3	4	2
Management worker	6	2	2	4	4
Administration worker	10	11	5	22	12
Sales worker	3	5	3	5	4
Supervisory worker	10	8	10	12	10
Technician worker	7	10	16	12	11
Production worker	63	62	62	45	58
Firm in Nairobi	68	67	66	58	65
Firm in Mombasa	13	15	15	23	17
Firm in Nakuru	10	10	10	8	10
Firm in Eldoret	9	8	9	11	9
Foods sector	18	17	18	24	19
Wood sector	14	12	12	8	12
Textile sector	7	6	8	12	8
Metal sector	25	24	22	19	23
Bakery sector	4	8	8	5	6
Furniture sector	14	17	17	11	15
Garments sector	14	14	12	14	13
Machinery sector	4	2	3	7	4

Few workers have no education; most workers have either completed primary (43 per cent) or secondary (41 per cent) education. The average years of education completed is 9. The rise in average years of education over the survey period is probably because retiring workers have less education, while new entrants have more education. The years of education completed are derived from information on the highest level of education completed and the grade attained. Given that grade repetition is possible, the total years of education a worker spent in school could be understated.

An enterprise questionnaire was used to collect firm-level information from the manager or another senior person in the firm. Among the labor information they provide is the total laborforce, percent of unionized labor, labor turnover, total labor costs, permanent labor and casual labor, and expected change in employment. Other information is on output, sales, expenses, and capital stock. The sector in which a firm operates is indicated by dummy variables. The largest poportion (23 per cent) of firms are in the metal sector and the smallest proportion (4 per cent) are in machinery sector. Firm location is indicated by a set of dummy variables. Most workers (65 per cent) are in firms located in Nairobi, which reflects concentration of manufacturing there.

Firm size is measured in number of workers employed by the firm. The average number of employees is 152 with a large dispersion since the sample includes very small and very large firms. The average size declined slightly over the survey period. Figure 2 in the appendix plots the aggregate employment figures in Table A2 which are derived from government statistics. The plots for the private manufacturing sector (private emp) and for the whole manufacturing sector (total emp) are almost horizontal which indicates the slow rise in manufacturing sector employment over the 1990s. The capital per woker in sample firms is calculated as the replacement value of plant and equipment in Kenya shillings divided by total number of workers in the firm. In 2000 the ratio was 1.8 times what it was in 1993. Output per worker is calculated as the total value of output in Kenya shillings divided by total number of workers. In 2000 it was 1.3 times larger than in 1993 on average.

Table 3 presents the distribution of constant price(1990=100) earnings in Kenya shillings according to education level. Earnings include the basic wage plus allowances (e.g. for food, transport, and housing) and production and Christmas bonuses. Three observations in Table 3 are as follows: (i) the average hourly wages increase with quantiles such that workers at the top quantile have the highest earnings in all education classes; (ii) the more educated workers receive higher earnings; and (iii) the average hourly earnings increased over the survey period for all workers (defined by education classes). Compared with the rise in average real private consumption demand of 3 per cent between 1993 and 1999 (see Table A1 in appendix) the rise in raw wage is about two times. At firm level, the average wage per worker in 2000 is at least 50 per cent higher than in 1993. This is consistent with the official statistics. Table A2 and the plot in Figure 1 in appendix show that annual wage per worker in the private manufacturing (Priv wage) and in the whole manufacturing sector (All wage) rose especially from the mid 1990s. The data suggest that wages grew and employment stagnated in the manufacturing sector. The sections that follow present results of how education and other variables highlighted in this section influence manufacturing wages.

Table 3: Average Monthly and Hourly Earnings in 1990 Kenya Shillings

Wave	N	Mean	P25	P50	P75
Below primary education					
Wave 1	186	1127.7	742.4	955.2	1256.1
		5.8	3.8	4.9	6.5
Wave 2	154	1036.2	739.2	881.3	1110.2
		5.1	3.5	4.4	5.7
Wave 3	137	1526.4	975.2	1346.1	1599.8
		8.9	5.7	7.8	9.3
Wave 4	81	1623.6	1124.6	1525.0	1844.7
		9.3	5.4	7.7	9.5
Total	558	1272.3	799.0	1038.4	1476.5
		6.9	4.1	5.5	7.9
Full primary education					
Wave 1	492	1383.4	791.9	1004.4	1438.0
		7.1	4.1	5.2	7.4
Wave 2	419	1236.2	684.9	940.1	1397.3
		6.1	3.4	4.5	7.1
Wave 3	472	1553.4	965.6	1299.2	1698.9
		9.0	5.6	7.6	9.9
Wave 4	385	2686.8	991.1	1398.1	1875.5
		14.1	4.9	7.1	9.4
Total	1768	1677.8	815.8	1162.7	1610.5
		8.9	4.2	6.0	8.7
Full secondary education					
Wave 1	417	2338.5	935.4	1484.8	2722.1
		11.9	4.8	7.4	13.9
Wave 2	379	1719.4	821.9	1181.5	1860.7
		8.8	4.0	6.1	9.7
Wave 3	419	2456.7	1109.4	1535.2	2342.8
		14.3	6.5	8.9	13.6
Wave 4	473	2925.4	1246.4	1982.2	3398.1
		17.6	6.3	10.1	17.6
Total	1688	2393.3	1025.0	1514.2	2590.6
		13.4	5.4	8.1	13.4
University education					
Wave 1	9	6488.9	3464.4	5939.0	8908.5
		29.5	17.9	30.7	43.2
Wave 2	14	5360.0	2397.3	4081.1	5821.9
		25.9	12.4	21.1	30.8
Wave 3	35	7958.1	2497.5	5436.6	11315.0
		46.3	14.5	31.6	65.8
Wave 4	43	11110.9	4320.4	6763.8	12459.5
		54.6	21.3	32.0	64.4
Total	101	8809.3	3464.4	5555.8	10731.7
		45.5	18.3	30.0	57.3

Source: computation from survey data. N is number of observations and P denotes percentile.

6. Earnings Function Analaysis

This section discusses the results of how real hourly wages changed over the 1990s based on earnings regression estimates. It also presents wage premia to education and the change in the wage structure over the 1990s at the mean and across quantiles of the earnings distribution.

6.1. Real Wages and Returns to Human capital

The Mincerian Earnings Function

The description of data in Section 5 indicates that real hourly wages changed over the survey period. However, part of the change in wages may be due to differences in characteristics of surveyed workers and firms. To obtain the change in real wages standardized for differences in observable worker characteristics, the human capital earnings function is estimated. Table 4 presents the results. The OLS estimates of the earnings function are presented in the first column. The second column presents the estimates of the earnings function controlling for firm fixed effects (FE). The next three columns present estimates of the earnings function for the first quartile (Q25), median (Q50), and third quartile (Q75) of the earnings distribution. To allow for time effects on earnings, dummy variables for survey waves are included. Also, because more than one worker is drawn in most firms the standard errors of estimated coefficients are corrected for clustering at firm level.

The time effects are significant and based on the coefficient of the dummy variable for wave four, the implied change in real hourly wage over the survey period is computed as $100*(e^b-1)$ following Halvorsen and Palmquist (1980), where b is the time dummy variable coefficient. In the OLS regression this works out to a rise of 35 per cent in earnings over a period of seven years. Controlling for human capital and firm fixed effects, the increase (42 per cent) is higher. Across quantiles the highest increase (40 per cent) is for workers at the median of the earnings distribution. Workers in the bottom quantile and those at the top quantile received wage rises similar to the mean increase. These increases are lower than the increases derived from the raw wage data, which points to the importance of controlling for sample characteristics. An F-test of the null hypothesis that wage increases across quantiles are identical has an observed F(6, 4104)-value of 2.28 with p-value = 0.03. Therefore, the null hypothesis may be rejected. This implies that wage increases over the survey period across quantiles are not identical.[4] It is highest for median worker.

The variables usually used to measure human capital (age, tenure, and education) have positive effect on earnings. Age effects are significant and the age-earnings profiles are concave. Tenure effects are small and mostly insignificant except for workers in the bottom quantile where there is positive effect. It may suggest that the internal wage structure favors workers in this quantile or their firm-specific skills are rewarded with higher wages. Education effects are significant and the education-earnings profile is convex as the derivative of log wage with respect to education evaluated at 6, 10, 14, and at the average years of education shows (Table 5). First, the pattern of Mincerian returns to education is identical across quantiles; returns rise with education. Second, returns to education for workers in the

[4] An earnings function was estimated on a sub-sample of workers excluding firms that entered the panel in 2000. The implied change in wages is similar to the one derived in the full sample.

top quantile are higher than for workers in the bottom quantile. For workers with 6 years of schooling, the return (multiplied by 100) range from 3 per cent at the first quartile to 5 per cent at the third quartile. The mean return (OLS) is 4 per cent. For workers with 10 years of schooling, returns vary from 10 per cent at the first quartile to 16 per cent at the third quartile with mean returns being 13 per cent. With 14 years of schooling returns range from 16 per cent at the first quartile to 27 per cent at the third quartile. The observed F(4, 4104)-value for an F-test of the null hypothesis that the education effects are identical across quantiles is 22.89 with p-value of 0.00. Hence the null hypothesis that the education effects are identical across quantiles of the wage distribution may be rejected. The rise in wage premia to education with quantiles suggests that education may be positively related to wage inequality in manufacturing sector.

An alternative specification (see 2c in section 3) breaks the total years of education completed into years of primary education (S_p), secondary education (S_s), and tertiary education (S_t) education. The specification is similar to that used by Moll (1996) to examine returns to education in South Africa. Here it is adapted to Kenya's education system.

$$S_p = \begin{cases} x, 0 \le x \le 7 \\ 7, x > 7 \end{cases}$$

$$S_s = \begin{cases} 0, x \le 7 \\ x - 7, 7 < x \le 13 \\ 6, x > 13 \end{cases}$$

(5)

$$S_t = \begin{cases} 0, x < 13 \\ x - 13, x \ge 13 \end{cases}$$

where x is years of schooling completed. The education coefficients in this earnings regression are interpreted as the return to education. The return can vary across the education levels but it is uniform within each level. The regression results are presented in Table 6. As in the earlier specification of the earnings function, returns to education increase with education and are highest for workers in the 75th percentile. In particular, return to primary education ranges from 1 per cent in the 25th percentile to 4 per cent in the third quartile. The mean return is 3 per cent. Secondary education returns range from 7 per cent at the lower quantile to 14 per cent at the upper quantile and the mean return is 11 per cent. Tertiary education returns are over 30 per cent across quantiles. In this specification, the observed F (6, 4103)-value for a test of the null hypothesis of equal education effects across quantiles is 28.03 with p-value of 0.00. Hence the null hypothesis may be rejected, which implies that education coefficients differ across quantiles of the wage distribution.

Table 4: Regression Estimates of the Mincerian Earnings Function. Education specified as a Quadratic Function in Years of Education: All workers

Explanatory variable	OLS	FE	Q25	Q50	Q75
Age (years)	0.06***	0.04***	0.05***	0.06***	0.05***
	(4.62)	(5.68)	(6.41)	(5.25)	(4.93)
Age squared	-0.0005***	-0.0004***	-0.0005***	-0.0006***	-0.0004***
	(3.07)	(3.51)	(4.83)	(3.55)	(2.76)
Education (years)	-0.09***	-0.07***	-0.08***	-0.10***	-0.11***
	(5.40)	(5.42)	(4.11)	(5.80)	(5.63)
Education squared	0.0112***	0.0083***	0.0088***	0.0115***	0.0134***
	(11.19)	(12.15)	(7.85)	(11.02)	(11.20)
Time with firm (years)	0.01	0.00	0.01***	0.004	0.01
	(1.60)	(0.99)	(2.60)	(0.72)	(1.33)
Time with firm squared	-0.0000	0.0000	-0.0001	0.0001	0.0000
	(0.16)	(0.22)	(0.68)	(0.60)	(0.07)
Male worker	0.02	-0.003	0.09***	0.05	0.02
	(0.51)	(0.12)	(2.61)	(1.59)	(0.50)
Wave 2	-0.10***	-0.11***	-0.09***	-0.09***	-0.09**
	(2.95)	(4.58)	(3.08)	(3.41)	(2.03)
Wave 3	0.30***	0.33***	0.37***	0.36***	0.31***
	(8.49)	(13.99)	(13.18)	(13.17)	(9.03)
Wave 4	0.30***	0.35***	0.30***	0.34***	0.31***
	(6.63)	(11.41)	(7.82)	(9.76)	(7.56)
Constant	0.20	0.68***	0.07	0.25	0.53***
	(0.88)	(4.83)	(0.50)	(1.26)	(3.28)
Adj. R^2/Pseudo R^2	0.35	0.28	0.17	0.20	0.23
Sample size			4115		

Notes: The dependent variable is the logarithm of real hourly earnings. t statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%

Table 5 Returns to Education (per cent) From Earnings Function in Table 4

Education	OLS	FE	Q25	Q50	Q75
6 years	4	3	3	4	5
10 years	13	10	10	13	16
14 years	22	16	17	22	27
Average years	11	8	8	11	13

Table 6: Regression Estimates of the Mincerian Earnings Function. Education specified as a Spline Function in Years of Education: All workers

Explanatory variable	OLS	F.E.	Q25	Q50	Q75
Primary education (years)	0.03***	0.03***	0.01	0.03***	0.04***
	(3.22)	(3.14)	(1.25)	(3.80)	(6.18)
Secondary education (years)	0.11***	0.08***	0.07***	0.10***	0.14***
	(12.49)	(14.43)	(11.33)	(16.46)	(33.08)
Tertiary education (years)	0.33***	0.25***	0.37***	0.34***	0.36***
	(10.36)	(14.49)	(9.65)	(19.20)	(11.59)
Constant	-0.10	0.44***	-0.13	-0.02	0.17
	(0.46)	(3.21)	(0.70)	(0.18)	(1.07)
Adj. R^2/Pseudo R^2	0.35	0.28	0.18	0.20	0.23
Sample size			4115		

Notes:
Regressions include controls as in Table 4 for age and its square, time in firm and its square, and dummies for waves and males. Absolute values of t-statistics within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

The Extended Earnings Function

Table 7 presents extended earnings functions with education specified as a quadratic function. It includes other wage determinants in addition to human capital variables. The difference between this earnings regression and that reported in Table 4 is that, enterprise characteristics and occupation dummies are included. The enterprise characteristics include firm size (number of workers); average labor productivity (output per worker), capital intensity (physical capital per worker), dummies for sector of business, and firm location. One potential problem is that these controls may be endogenous to wage formation. For example, higher average productivity in the firm may lead to higher wages. But higher wages could improve average productivity. Hence one has to be cautious in the interpretation of the results.

Age and education effects remain significant after controls for firm characteristics are included. Workers in large firms receive higher wages than comparable workers in small firms; the elasticity of hourly wages with respect to firm size is about 0.07 across quantiles. Workers in firms with higher average labor productivity also receive higher wages and wages in all occupations are greater than wages of production workers. Wages of workers in firms located in Mombasa, Eldoret, and Nakuru are lower than for comparable workers in firms located in Nairobi. There are significant sectoral wage differentials also, which may reflect compensating wage differentials for differences in working conditions. The implied change in real wages over the survey period across the wage distribution is about 36 per cent. The change is not substantially different from that derived from earnings regression with only controls for human capital (Table 4) despite the large number of controls added.

The returns to education calculated from the extended earnings function are presented in Table 8. It shows that the effect of education will tend to be underestimated in earnings functions that include these controls. A possible explanation is that education can influence wages by influencing the choice of occupation (see Knight and Sabot, 1990), sector or firm size a worker enters. A tendency to underestimate returns to education is also noted in farm production functions (see Appleton and Balihuta, 1996) when variable inputs whose use depends on education are included in the farm production functions that also include education. Table 9 presents estimates of returns to education for the specification of the earnings function with a schooling spline. In this specification also the estimated returns are lower than those derived from the standard human capital earnings function in Table 6. Thus, part of the return to education is a return to post-education choices.

The returns to education in Tables 5, 6, 8, and 9 may overstate or understate the actual private wage returns. First, one assumption of the underlying model is that the cost of schooling is the foregone earnings during the time spent in school. Direct private costs are assumed to be equal to zero. This may understate direct private costs especially at secondary level of education. As a result, Mincerian returns to education may overstate private wage returns to education for workers with higher levels of education. For workers with primary education, Mincerian returns may understate private wage returns. This is because foregone earnings for primary school children are overestimated. Primary school children in lower grades are too young to work and even if they work during their later grades, they do not earn a wage equivalent to the wage an adult without schooling would earn.

Table 7: Regression Estimates of the Extended Earnings Function. Education specified as a Quadratic Function in Years of Education: All workers

Explanatory variable	OLS	FE	Q25	Q50	Q75
Male worker	0.09**	0.09***	0.12***	0.12***	0.09**
	(2.50)	(3.47)	(5.97)	(4.98)	(2.38)
Age (years)	0.04***	0.04***	0.04***	0.04***	0.04***
	(4.02)	(5.85)	(4.31)	(5.36)	(4.59)
Age squared	-0.0003***	-0.0004***	-0.0004***	-0.00***	-0.00***
	(2.66)	(4.02)	(3.44)	(3.96)	(2.63)
Education (years)	-0.07***	-0.06***	-0.04***	-0.06***	-0.06***
	(4.71)	(4.91)	(3.20)	(4.11)	(3.77)
Education squared	0.0073***	0.0062***	0.0043***	0.01***	0.01***
	(7.62)	(9.45)	(5.52)	(7.79)	(6.71)
Time with firm (years)	0.00	0.00	0.00	0.00	-0.00
	(0.65)	(0.33)	(1.09)	(1.24)	(0.64)
Time with firm squared	0.0000	0.0000	0.0000	0.00	0.00
	(0.13)	(0.33)	(0.30)	(0.10)	(1.09)
Employment (logarithm)	0.07***	-0.04	0.07***	0.07***	0.08***
	(6.18)	(1.03)	(6.90)	(7.40)	(6.53)
Capital per worker (logarithm)	-0.00	-0.04	0.02*	-0.01	-0.03***
	(0.05)	(1.28)	(1.89)	(1.31)	(2.68)
Output per worker (logarithm)	0.05***	-0.00	0.04***	0.04***	0.05***
	(3.28)	(0.17)	(4.21)	(6.83)	(5.46)
Management worker	0.93***	0.96***	0.77***	1.00***	1.19***
	(10.47)	(21.16)	(12.03)	(15.74)	(14.63)
Administrative worker	0.43***	0.44***	0.29***	0.43***	0.57***
	(11.02)	(15.48)	(6.39)	(9.31)	(14.31)
Sales worker	0.31***	0.24***	0.20***	0.27***	0.42***
	(4.46)	(5.35)	(3.26)	(3.62)	(5.35)
Supervisory worker	0.37***	0.39***	0.35***	0.39***	0.41***
	(10.66)	(14.21)	(7.56)	(8.30)	(7.45)
Technician worker	0.13***	0.15***	0.08***	0.10***	0.13***
	(3.95)	(5.60)	(3.54)	(3.63)	(3.08)
Firm in Mombasa	-0.10**		-0.09***	-0.05*	-0.09***
	(2.31)		(3.77)	(1.86)	(2.58)
Firm in Nakuru	-0.43***		-0.39***	-0.39***	-0.40***
	(9.40)		(11.76)	(10.36)	(7.66)
Firm in Eldoret	-0.43***		-0.35***	-0.42***	-0.45***
	(8.65)		(11.52)	(15.24)	(11.10)
Wood sector	0.07		0.05	0.06	0.05
	(1.05)		(1.47)	(1.24)	(0.89)
Textile sector	-0.18**		-0.20***	-0.14***	-0.16***
	(2.37)		(5.61)	(3.60)	(3.75)
Metal sector	0.10		0.10***	0.10***	0.11***
	(1.59)		(3.85)	(4.88)	(3.46)
Bakery sector	-0.14		-0.14***	-0.19***	-0.18***
	(1.61)		(3.40)	(4.55)	(3.92)
Furniture sector	0.14**		0.16***	0.14***	0.11***
	(2.09)		(4.79)	(5.46)	(3.38)
Garments sector	-0.04		-0.04	-0.06**	-0.10**
	(0.62)		(1.43)	(2.47)	(2.33)
Machinery sector	0.14		0.20***	0.21***	0.10
	(1.62)		(3.69)	(6.87)	(1.60)
Wave 2	-0.07**	-0.06***	-0.08***	-0.08***	-0.07***
	(2.31)	(2.85)	(3.07)	(4.17)	(2.61)
Wave 3	0.38***	0.40***	0.40***	0.36***	0.37***
	(12.05)	(17.20)	(18.12)	(21.15)	(12.84)
Wave 4	0.31***	0.31***	0.28***	0.31***	0.29***
	(8.23)	(9.25)	(12.35)	(15.83)	(8.74)
Constant	-0.15	1.36**	-0.57***	0.03	0.31
	(0.49)	(2.35)	(2.59)	(0.12)	(1.61)
Adj. R^2/Pseudo R^2	0.52	0.39	0.30	0.33	0.36

Notes: The dependent variable is the logarithm of real hourly earnings. t statistics in parentheses* significant at 10%; ** significant at 5%; *** significant at 1%

Table 8: Returns to Education (per cent) From Extended Earnings Function in Table 7

Education	OLS	FE	Q25	Q50	Q75
6 years	2	1	1	2	3
10 years	8	6	5	7	8
14 years	13	11	8	12	14
Average years	6	5	4	6	7

Table 9: Regression Estimates of the Extended Earnings Function. Education specified as a Spline Function in Years of Education: All workers

Explanatory variable	OLS	FE	Q25	Q50	Q75
Primary education (years)	0.02***	0.02***	0.01	0.02***	0.03***
	(3.12)	(3.08)	(1.21)	(3.40)	(3.94)
Secondary education (years)	0.04***	0.04***	0.03***	0.04***	0.04***
	(6.26)	(7.14)	(5.21)	(8.06)	(6.87)
Tertiary education (years)	0.26***	0.22***	0.24***	0.24***	0.31***
	(8.55)	(13.43)	(6.14)	(9.02)	(7.15)
Constant	0.02***	0.02***	-0.61***	-0.12	0.03
	(3.12)	(3.08)	(3.50)	(0.88)	(0.19)
Adjusted R^2/Pseudo R^2	0.53	0.40	0.31	0.34	0.36
Sample size			4115		

Notes:
Regressions include controls as in Table 8 for firm size, output per worker, capital per worker, and dummies for survey waves, occupation, sector, and firm location.
Absolute values of t-statistics within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Second, education can have positive external benefits that are not captured by standard estimates of returns to education. Haveman and Wolfe (1984) discuss many such potential benefits. Further, as noted by Bigsten (1984), education gives access to the labor market. The direct wage returns may be low but access benefit could be large. The data used are for those already in the labor market and have wage income. With more detailed data it is possible to take into account those workers that are educated but have no jobs or are educated but in other income generating activities. In addition, the literature survey pointed out that estimated returns could be overstated if the worker's family background is omitted in the earnings function.

6.2. Change in Structure of Manufacturing Sector Earnings over the 1990s

Section 6.1 examines how real wages changed over the 1990s, and how wage premia to education vary across quantiles of the wage distribution. This section discusses the results of an empirical test of whether the earnings structure for workers in manufacturing enterprises shifted over the 1990's. Tables 10 and 11 present OLS and quantile regressions of the Mincerian earnings function with the same variables as in the corresponding regressions in Tables 4 and 6, and with human capital variables interacted with time dummies. This allows one to assess whether the increase in wages over the survey period is also associated with change in returns to human capital (only the coefficients of human capital variables and those of the interactions with wave four dummy are reported).

Table 10: Regression Estimates of the Extended Earnings Function with Survey Waves Interacted with Explanatory Variables. Education specified as a Quadratic Function in Years of Education: All workers

Explanatory variable	OLS	FE	Q25	Q50	Q75
Male worker	0.04	0.00	0.13	0.02	-0.02
	(0.67)	(0.04)	(1.56)	(0.39)	(0.23)
Age (years)	0.04*	0.03**	0.05***	0.04**	0.04
	(1.67)	(2.48)	(2.69)	(1.97)	(1.34)
Age squared	-0.0001	-0.0002	-0.0005*	-0.00	-0.00
	(0.41)	(0.89)	(1.78)	(0.62)	(0.15)
Education (years)	-0.06**	-0.07***	-0.06***	-0.06***	-0.04
	(2.46)	(2.85)	(4.39)	(3.33)	(1.47)
Education squared	0.01***	0.01***	0.01***	0.01***	0.01***
	(6.61)	(6.55)	(5.93)	(7.60)	(6.79)
Time in firm	0.01	-0.00	0.01	-0.00	0.00
	(0.54)	(0.25)	(0.68)	(0.22)	(0.12)
Time in firm squared	-0.00	0.00	-0.00	0.00	-0.00
	(0.55)	(0.08)	(0.69)	(0.00)	(0.63)
Wave4xMale worker	-0.09	0.00	-0.04	0.02	0.02
	(0.98)	(0.07)	(0.50)	(0.15)	(0.20)
Wave4xAge (years)	0.03	0.02	-0.01	0.03	0.02
	(1.21)	(1.05)	(0.28)	(0.95)	(0.44)
Wave4xAge squared	-0.00	-0.00	0.00	-0.00	-0.00
	(1.41)	(1.27)	(0.07)	(1.02)	(0.62)
Wave4xEducation	-0.04	0.00	-0.10	-0.05	-0.02
	(0.99)	(0.14)	(1.60)	(1.06)	(0.33)
Wave4xEducation squared	0.00	-0.00	0.01**	0.00	0.00
	(1.27)	(0.04)	(2.31)	(1.25)	(0.51)
Wave4xTime in firm	0.01	0.00	0.01	0.02	0.01
	(0.64)	(0.40)	(0.44)	(0.96)	(0.87)
Wave4xTime in firm squared	0.00	0.00	0.00	-0.00	0.00
	(0.32)	(0.56)	(0.70)	(0.49)	(0.22)
Constant	0.38	0.73***	-0.06	0.43	0.51
	(1.04)	(2.95)	(0.15)	(1.20)	(1.11)
Adj. R^2/Pseudo R^2	0.36	0.29	0.18	0.20	0.24

Notes: The dependent variable is the natural logarithm of real hourly earnings. t statistics within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Interactions of the explanatory variables with dummies for waves 2 and 3 included but are not reported to save space.

An F-test of the null hypothesis that the coefficients of the three time dummies and their interactions with human capital variables are jointly equal to zero has an observed F(27, 358)-value of 8.63 with p-value of 0.00. The F(24, 358)-value for a similar test in the specification in which education is specified as a spline function (Table 11) is 9.94 with p-value of 0.00. These F-values are significantly different from zero. Therefore in both specifications the null hypothesis may be rejected. This implies that the earnings structure as a whole shifted in the survey period. The interaction of years of education with wave four dummy in the OLS regression is statistically insignificant across quantiles in the specification with education as a quadratic. This points toward stable effect of years of education on wages over the survey period. When years of education are splined, the effect of primary education on earnings for workers in the top quantile fell, while for those with 8 to 13 years of education in the bottom and top quantiles the education effect rose.

Table 11: Regression Estimates of the Extended Earnings Function with Survey Waves Interacted with Explanatory Variables. Education is specified as a Spline Function in Years of Education: All workers

Explanatory variable	OLS	FE	Q25	Q50	Q75
Male worker	0.03	-0.01	0.16**	0.04	-0.07
	(0.52)	(0.17)	(2.05)	(0.40)	(0.52)
Age (years)	0.04*	0.03**	0.05***	0.04*	0.04
	(1.73)	(2.50)	(2.98)	(1.88)	(1.37)
Age squared	-0.0001	-0.0002	-0.0005*	-0.00	-0.00
	(0.50)	(0.94)	(1.92)	(0.68)	(0.17)
Primary education (years)	0.05***	0.03**	0.02	0.03**	0.05***
	(2.94)	(2.30)	(0.99)	(2.06)	(4.97)
Secondary education (years)	0.12***	0.09***	0.06***	0.11***	0.15***
	(9.04)	(9.59)	(4.95)	(13.59)	(14.05)
Tertiary education (years)	0.40***	0.36***	0.46***	0.46***	0.48***
	(6.47)	(7.38)	(5.79)	(3.86)	(4.37)
Time in firm (years)	0.01	-0.00	0.01	0.00	-0.00
	(0.59)	(0.23)	(0.52)	(0.19)	(0.35)
Time in firm squared	-0.0002	0.0000	-0.0003	-0.00	-0.00
	(0.54)	(0.09)	(0.51)	(0.28)	(0.13)
Wave4xmale worker	-0.08	0.01	-0.07	-0.32	0.10
	(0.86)	(0.17)	(0.75)	(0.66)	(0.74)
Wave4xAge (years)	0.04	0.02	0.00	0.00	0.01
	(1.24)	(1.15)	(0.19)	(0.01)	(0.29)
Wave4xage squared	-0.00	-0.00	-0.00	0.03	-0.00
	(1.41)	(1.31)	(0.34)	(1.57)	(0.48)
Wave4xPrimary education	-0.03	0.00	-0.01	-0.00	-0.06**
	(1.39)	(0.14)	(0.37)	(1.49)	(2.08)
Wave4xSecondary education	0.03	0.01	0.05**	-0.01	0.04**
	(1.45)	(0.63)	(2.30)	(0.44)	(2.17)
Wave4xTertiary education	-0.07	-0.11*	-0.09	0.01	-0.15
	(0.92)	(1.92)	(0.91)	(0.79)	(1.35)
Wave4xTime in firm	0.01	0.00	0.02	-0.15	0.02
	(0.55)	(0.25)	(0.98)	(1.28)	(1.20)
Wave4xTime in firm squared	0.00	0.00	0.00	0.01	-0.00
	(0.34)	(0.63)	(0.30)	(0.72)	(0.24)
Constant	0.11	0.49**	-0.09	-0.00	0.37
	(0.30)	(2.01)	(0.26)	(0.16)	(0.74)
Adj. R^2/Pseudo R^2	0.36	0.29	0.19	0.21	0.24

Notes: The dependent variable is the natural logarithm of real hourly earnings. t statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Interactions of the explanatory variables with dummies for waves 2 and 3 included but are not reported to save space.

Table 12 presents previous estimates of returns to education in Kenya that consider a time dimension. Since the studies do not estimate quantile regressions, comparison is confined the returns to education derived from ordinary least squares in this study. The study by Bigsten et. al. (2000) is included because it uses pooled manufacturing survey data for the early 1990s. Comparison across studies is not straightforward because of differences in data, time periods, specification of earnings functions, and measurement of variables. Bennel (1996) notes that cross-country differences in these aspects make comparison of returns to education across countries difficult. It can be argued that such differences are also likely to make comparison across studies in the same country difficult. Nevertheless a comparison can indicate the general trend. Bigsten (1984) estimates returns to education using the 1977/78 labor force survey and compares them with those of studies from the 1960s. He concludes that returns to primary and university education declined. But returns to secondary school

remained high. Milne and Neizert (1994) use the 1977/78 and 1986 urban labor force surveys and find that between 1978 and 1986 the return to primary education for a 30 year old declined while returns to secondary education rose. Appleton, Bigsten, and Manda (1999) use the 1978 and 1986 urban labor force surveys and data from wave three of the RPED survey. They find that returns to primary education fell from 10 per cent in 1978 to 2 percent in 1995. Returns to secondary education fell from 34 per cent to 13 per cent while returns to university education were stable.

Table 12: Estimates of Returns to Education from Previous Studies in Kenya

Study	Data	Primary	Secondary	University
Bigsten (1984)[a]	LFS, 1977/78	5	26	11
Bigsten et al (2000)[b]	RPED, 1993-95	4	12	22*
Milne & Neizert (1994)[c]	LFS,1978	9	11	-
	LFS,1986	7	16	-
Appleton, Bigsten & Manda (1999)[d]	LFS, 1978	10	34	61
	LFS, 1986	5	16	20
	RPED, 1995	2	12	69
Manda (1997)[e]	LFS, 1978	18	56	-
	LFS, 1986	13	37	-
	RPED, 1993-95	5	13	53

Notes
RPED: Regional Program on Enterprise Development
LFS: Labor force Survey
(a) Returns to education for urban areas Table V.10, column 1. The dependent variable is log cash income. The regressors are education dummies, vocational and on-the-job training dummies, and experience in years.
(b) Part of a cross-country study of five African countries. The dependent variable is ln (monthly earnings). Regressors include schooling, schooling squared, age, age squared, tenure, tenure squared, and male dummy. Based on manufacturing workers. Evaluated at 6, 10, and 14 years
(c) The dependent variable is ln (hourly earnings). Regressors included are schooling, schooling squared, age, age squared, female dummy, location dummies, and occupation dummies. The schooling effect reported is for a worker aged 30 years.
(d) The dependent variable is ln (monthly earnings). The regressors include schooling dummies, potential experience, a second and third order polynomial in potential experience, male dummy, and location dummy variables. Only returns to schooling for manufacturing workers are shown in this Table.
(e) Dependent variable is ln(hourly earnings). Age, age squared, vocational training dummy, occupation dummies, and location dummies are included in separate earnings equations for each schooling level.
* Computed from the estimated model with schooling set to 16 years.

Taken together, the results in this study and those in earlier studies point to decline in returns to primary and secondary education. Given that output growth has been slow in recent years, one explanation may be that demand for educated labor increased at a slower rate than supply of educated labor. When wage premia to education are rigid as over the 1990s, it may be because different types of educated labor are easily substitutable or family support makes workers fail to revise reservation wages downwards or there are imperfections in the labor market for particular education class (see Bigsten, 1984). It is also possible that manufacturing workers may be cushioned from the expansion in supply of educated labor if firms pay relatively higher wages to elicit effort or reduce turnover among their workers.

The empirical evidence on changes in returns to human capital from outside Kenya is mixed and is in most cases for periods before the 1990s period. Moll (1996) used South African data and found that returns to primary education for Africans fell from 8 per cent in 1965 to 3 percent in 1975. For colored and Asians, returns fell from 16 per cent to 12 per cent. Returns to secondary education remained stable. Evidence from Ghana (see Canagarajah and Thomas, 1997) shows that the return to a year of education rose from 4 per cent in 1987 to 6 percent in 1991. Returns to post-primary levels of education increased over this period and became compressed.

The study of changes in earnings structure over time is mainly in developed countries. An exception is the study by Nielsen and Rosholm (2001) who use quantile regressions to trace the change in the public-private wage differential in Zambia between 1991 and 1996. They find that the earnings of less educated public sector workers in the bottom deciles increased relatively more than the earnings of private sector workers. But in the top deciles the wage earnings advantage of highly educated workers in public sector workers narrowed. They also report that private wage returns to education were larger in the private sector, and vary across quantiles of the wage function. A pioneering study based on U.S. data (Bushnisky, 1994) finds that returns to education in the USA were higher for workers in the top deciles of the wage distribution in the 1960's and early 1970's, but fell and converged across quantiles in the second half of the 1970's. In the 1980's, returns to education recovered and rose sharply especially for workers in the top deciles. Because labor markets differ across countries, results from one setting may not be generalized to other settings.

7. Summary and Conclusion

Data that allow matching of employees to their employers and for multiple periods in African labor markets have rarely been available. This study uses employer-employee matched data from a survey of manufacturing firms in Kenya, to study real wages and private wage returns to human capital over the 1990s. Ordinary least squares and quantile regression estimates of earnings functions indicate that real wages standardized for differences in human capital and firm characteristics increased between 1993 and 2000. The increase occurred for workers in the median and in the bottom and top quantiles of the earnings distribution. Given the slow growth in modern sector wage employment and a rapid expansion of informal sector employment (Government of Kenya, 2001), one would expect the labor supply pressure to drive down wages in a competitive labor market.

However, recent empirical evidence (e.g. Manda, Bigsten, and Mwabu, 2001 on unionization; Teal, 1996 on rent sharing and Soderbom and Teal, 2001 and Azam and Ris, 2001 on rent-sharing and efficiency wages) suggests that other models of wage determination also explain the observed variation in wages in African countries. In this paper the regression estimates indicate that workers with same observable characteristics have higher wages in larger firms than in smaller firms. Also, firms with greater average productivity pay higher wages across quantiles of the earnings distribution than low average productivity firms.

Education and experience (proxied by age of the worker) are rewarded in the manufacturing labor market, while the reward to time in firm is small. More educated workers receive higher wages in the median and in the top and bottom quantiles of the earnings distribution. Since the return to education encompasses two effects-direct effect on wages and the effect on access to wage employment, low returns at lower levels of education may reflect low effect of primary and secondary education on access to modern sector wage employment. The rapid expansion in informal sector employment may partly indicate reduced chances of securing wage employment. Workers in the top quantile of the wage distribution have higher returns to education than comparable workers in the median, who in turn have higher returns than those in the bottom quantile. Harmon, Oosterbeek, and Walker (2000) report a similar finding for the U.K, while Mwabu and Schultz (1996) find the same at higher education for

white South Africans, but the opposite for Black South Africans. Within the human capital model, a possible explanation of the greater returns to education in the top quantile than in lower quantile would be that it is due to differences among workers in unobserved productive characteristics, for example individual ability or quality of schooling. The result suggests that education in Kenya is associated with higher private wage returns to such unobserved factors.

Investigation of change over time in private wage returns to education indicates that little change occurred over the 1990s across the earnings distribution. Given the large and expanding stock of educated labor in Kenya, urgent measures are required to increase demand for educated labor in manufacturing and other economic sectors. Stagnant or declining returns to education are a concern that should be addressed if education is to play a role in poverty reduction strategies. Moreover, low returns may discourage parents from making adequate investments in the schooling of their children. The role of education in shaping income distribution also deserves further investigation.

8. 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. (2002), Education, incomes and poverty in Uganda in the 1990s. Credit Research paper No. 01/22, Center for Research in Economic Development and International Trade, University of Nottingham.

Appleton, S. and Balihuta, A. (1996), Education and agricultural productivity: Evidence from Uganda, *Journal of International Development*, 8 (3), 415-444.

Appleton, S., Hoddinot, J., and McKinnon, J. (1996), Education and health in Sub-Sahara Africa, *Journal of International Development*, Vol. 8, 3, 307-339.

Appleton, S., Bigsten, A. and Kulundu, D. (1999), Educational expansion and economic decline. Returns to education in Kenya 1978-1995, *CSAE Working paper* 99-6, Center for the Study of African Economies, Oxford University.

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

Becker, G.S. and Chiswick, B.R. (1966), Education and the distribution of earnings, *American Economic Review*, 56, 358-69.

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

Behrman, J. and Birdsall, N. (1983), The quality of schooling: Quantity alone is misleading. *American Economic Review*, 73, 928-46.

Bigsten, A. (1984), *Education and income determination in Kenya*. Gower Publishing Company Limited. Gower House, Croft Road, Aldershot, England.

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.

Bigsten, A. and Kimuyu, P. (eds., 2001), *Structure and performance of manufacturing in Kenya*. Palgrave, London.

Bushnisky, M. (1994), Changes in the U.S. wage structure 1963-1987, Application of quantile regression, *Econometrica,* 62, 405-58.

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

Canagarajah, S., and Thomas, S. (1997), *Ghana's labor market*: 1987-92, The World Bank

Card D. (1995), Earnings, schooling, and ability revisited, *Research in Labor Economics*, vol. 14, 23-48.

Deaton, A. (1997), *The Analysis of household surveys: A micro econometric approach to development Policy*. The Johns Hopkins University

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.

Government of Kenya, *Economic surveys*, various issues Ministry of planning and national development.

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

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

Haveman, R. and Wolfe, B. (1984), Schooling and economic well-being: the role of non-market effects. Journal of Human Resources, 19, 377-407.

Jones, P. (2001), Are educated workers really more productive? *Journal of Development Economics*, vol. 64, 57-79

Katz, L. and Autor, D. (1999), Changes in wage structure and earnings inequality, in: O.Ashenfelter and D. Card, eds., *Handbook of Labor Economics*, Vol. I (Elservier Science).

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

Koenker, R., and Bassett, G. (1978), Regression quantiles, *Econometrica,* 46, 1, 33-50

Krishnan, P., Selassie, G. T., and Dercon, S. (1998), The urban labor market during structural Adjustment, Ethiopia 1990-97, *CSAE working paper 98-9*, Center for the study of African economies, University of Oxford.

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

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

Manda D.K. (1997), Labor supply, returns to schooling and the effect of firm size on wages: The case of Kenya. PhD thesis *75*, Department of economics, University of Göteborg.

Manda, D. K., Bigsten, A., and Mwabu, G. (2001), Trade union membership and earnings in Kenyan manufacturing firms, *Working Papers in Economics no. 50*, Department of Economics, Göteborg University.

Milne, W. and Neizert, M. (1994), Kenya, in: Horton, S. Kanbur, R. and Mazumdar, D., eds., *Labor markets in an era of adjustment*, Vol. 2, EDI, World Bank.

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

Moll, P.G. (1996), The collapse of primary Schooling returns in South Africa, 1960-90, *Oxford Bulletin of Economics and Statistics*, 58, 185-210.

Mwabu, G.M. and Shultz, T.P. (2000), Wage premia for schooling and location, by gender and race in South Africa, *Economic Development and Cultural Change,* Vol. 48, 2, 307-334.

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

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

Pinckney, T. and Kimuyu, P. (1995), Policy and rural development: Two communities in East Africa. Department of Economics, Williams College, Williamstown, Massachusetts.

Pritchett, L. (2001), Where has all the education gone? *World Bank Economic Review*, vol. 15, 3, 367-391.

Pscharopoulos, G. (1995), The Profitability of Investment in Schooling: Concepts and Methods. *HCO Policy Working Paper* 63, Washington D.C.: The World Bank.

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

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

Schultz, T.P. (1988), Schooling Investments and Returns, in: H. Chenery and T.N. Srinivasan, *Handbook of Development Economics*, vol. 1 (Amsterdam, North Holland), 543-630.

Schultz, T.W. (1960), Capital formation by education, *The Journal of Political Economy*, Vol. 68, Issue 6, 571-583.

Schultz, T.W. (1961), Investments in Human Capital, *American Economic Review*, 51, 1-17.

Schultz, T.W. (1975), The value of the ability to deal with disequilibria, *Journal of Economic Literature*, 13, 872-876.

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

Weir, S. and Knight, J. (2000), Education externalities in rural Ethiopia: Evidence from average and stochastic frontier production functions, *CSAE WPS 2000-4*, Centre for the study of African Economies, Oxford University.

Willis, R.J (1986), Wage determinants: A survey and reinterpretation of human capital earnings functions, in: O.Ashenfelter and R.Layard, eds., *Handbook of Labor Economics*, Vol. I (Elservier Science).

World Bank (2000), *World development report, 2000/2001: Attacking poverty,* Oxford University press, New York.

APPENDIX

Table A1: Selected Economic and Education Enrollment Indicators for Kenya, 1991-1999

Indicator	1991	1992	1993	1994	1995	1996	1997	1998	1999
Economic indicators									
GDP growth (%)	2.1	0.5	0.2	3.0	4.8	4.6	2.4	1.8	1.4
Inflation rate (%)	19.6	27.3	46.0	28.8	1.6	9.0	11.2	6.6	3.5
Real private consumption	5119	5020	3953	3525	4409	4311	4739	4873	4804
Population growth (%)	3.1	3.0	2.9	2.8	2.7	2.6	2.5	2.4	-
Population size (millions)	22.7	23.4	24.0	24.8	25.2	26.3	27.1	27.9	28.7
Formal sector (000)	1442	1462	1475	1505	1557	1619	1647	1665	1674
Informal sector (000)	1063	1238	1467	1792	2241	2644	2987	3354	3739
Manufacturing growth (%)	3.8	1.2	1.8	1.9	3.9	3.7	1.9	1.4	2.4
Agricultural growth (%)	-1.1	-3.7	-4.1	2.8	4.8	4.4	1.2	1.5	1.2
Education indicators									
Primary level (millions)	5.46	5.56	5.43	5.56	5.54	5.60	5.68	5.92	5.87
Secondary level (millions)	0.61	0.63	0.52	0.62	0.63	0.66	0.69	0.70	0.64
University level (millions)	0.039	0.042	0.040	0.040	0.040	0.037	0.040	0.044	0.042
Polytechnics (millions)	0.009	0.009	0.009	0.011	0.008	0.009	0.009	0.009	0.010
Techn.Institutes (millions)	0.007	0.008	0.008	0.008	0.008	0.010	0.009	0.008	0.010

Source: Economic surveys, various issues.

Table A2: Annual Wages (1990 Kenya shillings) and Employment in Kenyan Manufacturing, 1991-99

Year	Private employment (000)	Wage per worker Private mfg (K£)	Total employment (000)	Wage per worker Whole mfg (K£)
1991	149.80	1717.36	188.9	1664.08
1992	151.00	1474.83	190.3	1422.57
1993	154.30	1155.54	193.6	1109.24
1994	158.20	1051.07	197.5	1011.55
1995	165.50	1306.47	204.8	1269.77
1996	172.30	1471.39	210.5	1421.32
1997	177.10	1587.35	214.5	1536.21
1998	180.80	1811.50	217.1	1748.54
1999	184.00	2026.14	220.1	1691.16

Note: Nominal wages deflated with CPI (1990=100) series in IMF CD-ROM data. Mfg stands for manufacturing

Figure 1: Manufacturing Sector Wage per Worker (1990 Kenya Shillings)

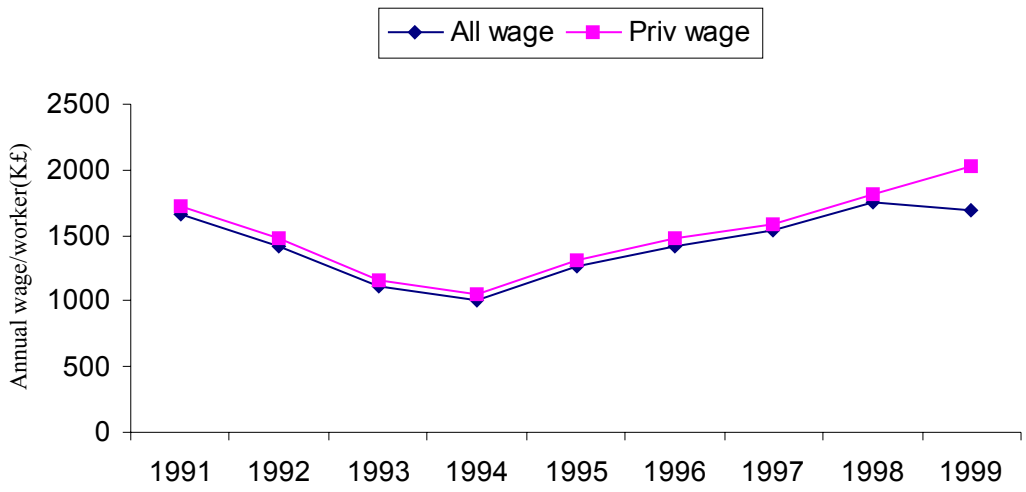
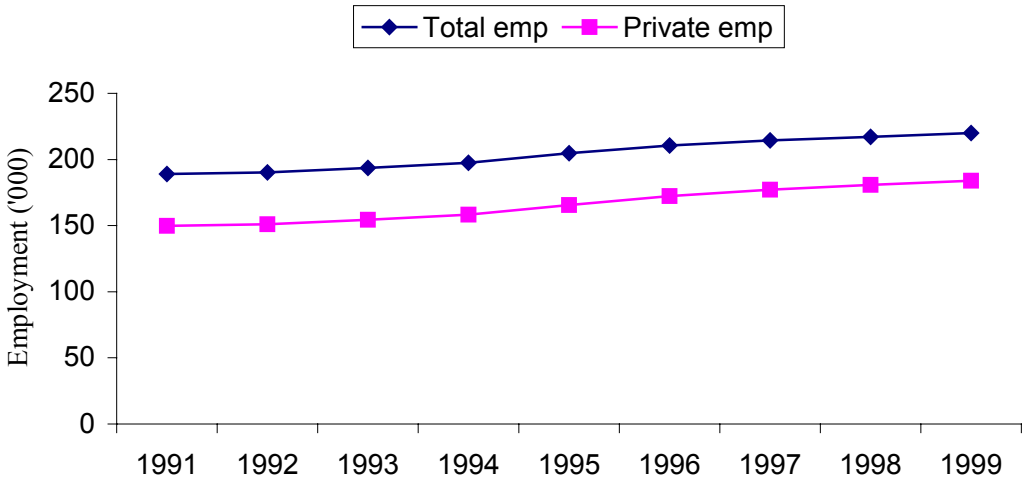


Figure 2: Employment in Kenya Manufacturing, 1991-1999



Family Background, Education and Earnings in Kenya*

By

Anthony Wambugu
Department of Economics
Göteborg University
Box 640
SE-405 30 Göteborg
Sweden

2002b

Abstract: A recent survey of manufacturing firms in Kenya is analysed to examine whether failure to take account of family background, or to treat education as endogenous in wage function analysis results in substantial bias in estimates of private wage returns to education. Regression estimates of education functions show that, workers with educated parents acquire more education than that of comparable workers with uneducated parents. Earnings function estimates indicate that parental post-primary education is positively related to workers' earnings. But omitted family background bias in economic returns to education at different levels of education is small when compared to results from other less developed countries. Regression estimates of the earnings function with education instrumented indicate that standard estimates of returns to education may be understated. But the finding is sensitive to the validity and quality of the instruments.

JEL Classification: J24; J31; I21; O15
Keywords: Earnings function, return to education, instrumental variables

*I thank Arne Bigsten, Henry Ohlsson, Simon Appleton, Roger Wahlberg, seminar participants and two anonymous people for comments and suggestions on earlier versions of the paper. I thank Arne Bigsten and the Center for the Study of African Economies, Oxford University for the opportunity to join the team that collected the data on which the paper is based.

1. Introduction

Many empirical analyses in less developed countries based on the human capital model tradition, find positive economic returns to education. Schultz (1988), Psacharopoulos (1994) and Appleton (1996, 2000) survey the literature. However, these estimates of returns to education are criticized because education is seen as a choice variable and therefore not exogenous to wage formation. For example, empirical evidence in countries of Sub-Sahara Africa (SSA), indicates that parental background can play an important role in education attainment (see Knight and Sabot, 1990 for Kenya and Tanzania; Glewwe and Jacoby, 1994 for Ghana; Tansel, 1997 for Ghana and Cote d' Ivoire; Lloyd and Blanc, 1996 for six countries, and Appleton, 2001 for Uganda).[1] Despite this finding, it is not clear in many SSA countries whether omission of family background or failure to treat education as endogenous to wage formation, results in substantial bias in standard estimates of returns to education. In part this is because the data used in many analyses lack information on family background.

One early approach in the literature is to include variables that can influence a worker's educational attainment, explicitly in the human capital earnings function used to estimate returns to education. For example, workers with greater unobserved skills may attain more education and such unobserved skills can also influence wages. Grilliches (1977) and Blackburn and Neumark (1995) discuss the subject in detail. They control for this type of bias using ability test scores and obtain estimates that are lower than standard estimates of returns to education. Knight and Sabot (1990) use data from two less developed countries (Kenya and Tanzania) to test whether human capital (cognitive skills) from schooling has independent effect on earnings or it is a signal of worker's ability. Using ability test scores in an earnings function that also controls for schooling and cognitive skills, the conclusion is that, though ability can have a role in wage formation, controlling for it does not diminish the effect of human capital on earnings. Glewwe (1996) also examines the issue in Ghana, and finds that returns to cognitive skills are lower than standard estimates of returns to schooling.

In other surveys, proxies for unobserved skills such as ability test scores may not be available. As an alternative, family background variables if available are used as controls in earnings functions. Family background can be used in its own right as a determinant of wages or as a proxy for unobserved skills. Examples from less developed countries include the studies by 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 Ethiopia's urban labor market; and Kingdon (1998) for males and females in Uttar Pradesh India. In developed countries two recent studies include Blackburn and Neumark (1995) and Dearden (1999). The common finding among the studies is that family background has positive impact on wages and returns to worker's education decline when his/her family background characteristics are included in earnings regressions.

A third approach attempts to clean out individual worker specific effects when estimating returns to education. Some studies (e.g. Ashenfelter and Krueger, 1994) use data on workers who are twin brothers. The idea is that twins share family background, while their education and earnings differ. Differencing across twins would clean out their common fixed

[1]Strauss and Thomas (1995) and Schultz (1988) survey the empirical evidence from all parts of the less developed world and come to this conclusion.

effects from the earnings function. They find that ability bias in standard estimates of returns to education is small. But measurement error could bias the estimates downwards. Others (e.g. Angrist and Newey, 1991) use panel data of workers who acquired more education while in the labor force. Assuming that individual fixed effects are time invariant, panel data methods can be used. They find that returns to education exceed standard estimates. In Kenya and other Sub-Sahara countries, panel data or data on twins are rare. So the approaches that attempt to clean out family background fixed effects have not found much application there. Moreover, the accuracy of such education changes may be a concern.

The more recent literature tries to instrument for education attainment in earnings functions. This requires variables that are strongly correlated with education but that do not directly influence earnings. For example, Angrist and Krueger (1991) use quarter of the year when a worker was born as an instrument. The idea is that due to school-entry age policy and compulsory schooling laws in the U.S.A., the quarter of birth introduces exogenous variation in education attainment. They find that standard estimates are not substantially different from instrumental variable estimates. In contrast, Uusitalo (1999) and Levin and Plug (1999) use family background variables as instruments for education. Harmon and Walker (1995), use changes in minimum school leaving-age policy in the U.K. And Card (1993) uses geographic proximity to a four-year college in the U.S.A, the motivation being that if one is close to a college, costs of attendance are relatively low and they would acquire more education.[2]

The main finding in the studies that instrument for education is that estimates of returns to education can often exceed standard estimates and the difference can be large sometimes. The analyses are however confined to labor markets in developed countries. Since education and labor markets in these countries may differ from those in less developed countries; the results may not be generalized. But the question addressed is also relevant for the latter. Yet as Schultz (2002) notes, compared with developed countries there are fewer studies that use instrumental variables to study schooling and wages in labor markets of the less developed countries.

The purpose of this paper is to use data from a recent survey of manufacturing firms in Kenya to examine schooling and wages while trying to instrument for education in the wage function. The latest survey wave has information on parental background and schooling availability at different stages of a worker's education career. Three earlier waves in the early 1990s did not collect this information.[3] Another survey with information on parental background of manufacturing workers in Kenya was implemented by Knight and Sabot in 1980, and also covered firms in Tanzania (see Knight and Sabot, 1990 for very detailed analyses of survey data). In a part of the analyses they estimate binary probit models of schooling attainment. The finding is that, workers with more educated parents were more likely to complete primary or secondary education. Armitage and Sabot (1987) use the data to examine how parental education interacts with workers' earnings. They find that the private wage return to secondary education rises monotonically with parental education. Within the human capital framework, this was interpreted to mean that family background represents out-of-school human capital investments that are complementary to formal schooling.

[2] Card (1995, 1999) surveys theory and empirical literature on this line of research

[3] A study on returns to human and physical capital in five African countries using earlier waves (Bigsten et. al., 2000) highlights and discusses the potential bias in schooling coefficients if controls for family background are left out of the earnings function.

These studies provide important results and the analysis in this paper does not diminish their importance. Rather, it complements the study of family background, education and earnings in Kenya in several ways. First, the paper estimates two schooling models. In one model the dependent variable is the total years of education completed. The model serves as the first stage to study returns to endogenous schooling. The second schooling model follows Lillard and Willis (1994) and Glick and Sahn (2000) and uses ordered probit to model the highest education level completed by a worker. They argue that schooling attainment is outcomes are discrete and ordered. The two models help to inquire into how parental background is correlated with different measures of worker's education.

The earnings function analysis first examines whether failure to control for parental background in earnings regressions results in significant bias in estimated returns to education. Rather than focus on the linear education effect only, the paper estimates a flexible earnings function to assess the relative impact of parental background on returns across levels of education. Wage regressions with measures of parental education are compared with wage regressions without parental education. Then the question of whether worker's education is endogenous to wage formation and how this influences estimates of private wage returns to education is considered. Parental education and indicator variables of schooling availability are the potential instruments to identify education attainment.

The next section reviews how family background might influence education attainment and earnings. Section 3 lays out the econometric models and section 4 describes the data. The regression estimates of education attainment functions and earnings functions are presented in section 5, and the conclusion follows in section 6.

2. Education and Earnings

The theory of human capital is based on the notion that an individual's incentive to invest in education comes from the prospect of future earnings gain due to greater productivity that comes from investments in schooling. This assumes that earnings reflect a worker's productivity. Following Grilliches (1977) the wage relation can be written as

$$w = w(s, A, \mu) \tag{1}$$

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

Individuals or their families may invest in education to maximize the present discounted value of wealth. The human capital model also assumes that the cost of education is the opportunity cost of time spent in school (foregone labor earnings). The problem that the individual faces is to maximize (2) with respect to schooling

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

where B denotes wealth, $w(.)$ is the wage relation in (1) and r is the constant discount rate applied to future earnings stream. The first order condition yields the stopping rule in expression (3) for educational investment. 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).

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

With altruistic parents (see Becker and Tomes, 1976, 1979), markets that are perfect; and education as a pure investment good, the individual would demand the optimal education. In this equilibrium, family background would have little impact on workers' education. However, the assumptions may not hold in the real world and this creates room for family background characteristics and schooling availability to influence worker's education.

For example, better-educated parents may have better access to funds to finance education especially beyond primary education. More educated parents may knowingly or unknowingly make greater out-of-school human capital investments (e.g. better home education, health, and nutrition) that complement formal education. Information problems could also be less severe for such parents. For example, they may have better information about returns to education, which reduces uncertainty in educational investments. Supply factors (e.g. distances to nearest schools) can also affect schooling availability and influence educational attainment. In summary, in a situation characterised by imperfect markets, parental background can be associated with different marginal benefits and marginal costs of educational investments. This translates into different educational attainment across individuals with different backgrounds.[4]

3. Econometric Specification

The analysis begins with a study of the impact of family background on education attainment.[5] Two schooling models are specified. In the first model education is measured as a discrete variable. One alternative to estimate the model is to fit single equation probit models. For example, Knight and Sabot (1990) estimate single equation probit models for primary school completion and for secondary school completion in Kenya's and Tanzania's urban manufacturing. In the survey used for the present paper workers reported four levels of education: below primary, full primary, secondary, and university. These are modelled jointly using ordered probit to try and improve on efficiency of the parameter estimates. The form of the ordered probit is

$$s_i^* = h_i^{'}\beta + \varepsilon_i, \quad \varepsilon_i \sim N(0,1) \tag{6a}$$

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

where s_i^* is a latent variable that measures education attainment. But this is not observed. Instead s_i, the highest education level that is reported by the *ith* worker is what is observed. The vector *h* contains a set of variables that are hypothesised to explain educational

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

[5]The sample comprises of wage earners in manufacturing firms. It can be argued that too much emphasis should not be put on the estimates of schooling models based on such a selected sample. A broader data set would be useful. However, the main merit of the data is the availability of parental background and measures of schooling availability, which allow the study of returns to endogenous schooling. But, rather than go straight to this, the paper first explores briefly schooling attainment.

attainment. The thresholds μ_j, j = 0,1, 2, 3 that determine whether a worker attains a given level of education are estimated along with the parameter vector β. The ε is a random error term, which is assumed to follow a normal distribution.[6] Estimation of this model yields coefficients to calculate the probability of a worker i having j as the highest level of education. That 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 density.

The second schooling model has total years of education completed as the dependent variable. As in other studies that instrument for education that are reviewed earlier, this model will form the first stage regression in a two-equation framework to examine whether standard estimates of returns to education differ markedly from those estimated with education instrumented. The total years of education are derived from survey information about the highest level of education completed.[7]

The latent education attainment index and the total years of education are hypothesized to depend on a number of factors. These include age (years) and age squared, to pick time and lifecycle effects on education. A dummy variable for male workers is included to capture gender differentials in education. The survey also collected information that is not usually collected in firm surveys, but which is important in modelling education attainment. One of the variables is the distance to nearest primary and secondary schools when a worker was of the relevant school age. This can proxy schooling availability and direct school costs. Dummy variables are used to indicate the various distance intervals. Another set of dummy variables indicates the region in which a worker obtained most of his/her education. This may capture regional variations in education development and other region specific factors. A set of dummy variables is used. Parental background is measured by dummy variables for parents' education. Dummy variables for the highest education level for each parent are used.

In the labor market, an individual faces a wage, which can be described by a semi-log wage function (4).

$$\ln w_i = x_i'\gamma + s_i\delta + u_i, \quad u_i \sim N(0, \sigma^2) \tag{4}$$

where w_i is real hourly wage of worker i. In the basic earnings function s_i is the years of education completed and the vector x contains experience and experience squared. Age (years) is used as a proxy for experience. In essence it represents an experience and an age effect. The earnings function also includes time in the firm (years), time in the firm squared, a dummy for male worker, and dummy variables for firm location. The aim is to estimate the parameter vectors γ and δ. u_i is a random error term.

In estimation of equation (4) the underlying model assumes that (i) the effect of education on earnings is linear; (ii) education attainment is not correlated with the error term; (iii) there is no omitted variable bias from failure to control for family background and unobserved skills. To address these issues empirically, the basic specification is extended.

[6] The alternative is to assume that the random error term follows a logistic distribution and specify the ordered logit model.
[7] Years of education completed can be modelled using ordered probit also (e.g. Harmon and Walker, 1995 and Glick and Sahn, 2000). But the sample is rather small to get enough cases into each point of the education distribution.

First, the earnings function includes indicators for family background to address assumption (iii) partially. Second, the square in years of education is included to introduce nonlinearity to the education effect. Third, education is instrumented in the wage function to address (ii). This has to borrow from the schooling demand literature. Hence the initial focus in this paper on schooling attainment models.

Education could be endogenous for three reasons. First, as noted by Grilliches (1977), measurement error in education biases the estimates towards zero. Second, omitted unobservable human capital could be positively correlated with both education and wages leading to overstated returns to education. 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.

4. The Data

The data used in this paper were collected in 2000 as part of a survey of close to 200 manufacturing firms in Kenya, a country located in East Africa. The firms surveyed are located in four urban centres (Nairobi, Mombasa, Nakuru, and Eldoret) where manufacturing and wage employment are concentrated. The United Nations Industrial Organization (UNIDO) financed the survey and data collection was in September/November. The survey used two questionnaires; an enterprise questionnaire and a worker questionnaire. Both instruments were administered in face-to-face interviews. The enterprise questionnaire was used in interviews with firm managers to record firm-level information. The workers questionnaire was used in interviews to obtain information from a sample of up to ten workers in every firm visited. In this paper, it is the information recorded in the interviews with workers that is used.

The information collected in the workers part of the survey is unique because in addition to worker characteristics, information about their parents' education was collected. Further, the survey collected retrospective information concerning geographic proximity to primary and secondary schools at the time the worker was of the relevant school age. Therefore, although the sample cannot be said to be representative of wage earners in Kenya, the availability of family background variables and indicators of schooling availability, which are absent in other surveys, justifies its use. But it is a drawback to note when interpreting the results. The sample used is limited to workers aged between 16 and 64 years. Also, it includes only workers for whom the survey recorded complete information that this paper requires to do the earnings function analysis. This resulted in a sample of 843 workers. Table 1 presents the definitions and summary statistics of the variables.

Most workers (82%) in the survey are men. The average age and time in the firm in years are 35 and 9 respectively. For 26% of the workers both parents have no education, for 31% both parents have full primary education, and for 22% both parents have at least primary education. At the time they were of primary school age, 29% of the workers had a primary school within 1 kilometre from their home, for 44% it was between 1 to 2 kilometers, 19% had one between 3 and 5 kilometres, and for 8% it was over 6 kilometres. When they were of secondary school age, 14% had a secondary school within 1 kilometre from their home, for

26% it was between 1 and 3 kilometres, for 22% it was located between 3 and 6 kilometres away, and for 38% over 6 kilometres.

In this sample, 7% of the workers have less than primary education, 38% attained full primary education, 51% attained secondary education, and 4% have university education. There is indication that workers' attained more education than their parents. For instance, while half of the workers completed secondary education, over 70% of the parents have only primary education or less. Education attainment is particularly low among the mothers; 45% of the workers had mothers who had no education compared to 28% with fathers without education. Rapid education expansion is a key feature of Kenya's education system (see Bigsten, 1984 and Knight and Sabot, 1990 for discussion).

Table 1: Summary Statistics and Definitions of variables

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

Table 1 contd….

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, presents the correlation coefficients between parents' and workers' education and earnings. The point is that parents' education is correlated with workers' education and earnings. To gauge the strength of these relationships conditional on other variables the paper turns to multivariate analysis.

Table 2: Correlation Coefficients: Parental Education and Workers' Education and Earnings Outcomes

	Workers Education and Earnings					
	Years of Education	Below primary	Primary	Secondary	University	Correlation with earnings
Father's education						
None	-0.33*	0.23*	0.16*	-0.25*	-0.08*	-0.03
Primary	0.04	-0.12*	0.06*	0.03	-0.07*	-0.11*
Post-primary	0.30*	-0.11*	-0.24*	0.23*	0.17*	0.17*
Mother's education						
None	-0.31*	0.22*	0.17*	-0.26*	-0.07*	-0.01
Primary	0.14*	-0.15*	-0.03	0.13*	-0.03	-0.09*
Post-primary	0.25*	-0.10*	-0.20*	0.20*	0.15*	0.14*

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

5. Estimation Results

This section presents the regression results in three parts. It begins with educational attainment function estimates to focus on the association between parental education and worker's education.[8] A comparative analysis of earnings functions with and without controls for parental background follows in the second part. The third part presents results of returns to endogenous education.

5.1. Education Attainment Function Estimates

Table 3 presents maximum likelihood estimates of the ordered probit regression of levels of education and ordinary least squares regression estimates of years of education. The wage relation is set aside first. To begin with, the education attainment models include only workers' characteristics and parents' education as explanatory variables. The estimates are labeled OLS (1) and OPROBIT (1). The models are then extended to include dummy variables for the province a worker obtained most education and for distances to nearest primary and secondary schools. The sizes of parents' education coefficients in the two specifications do not differ substantially. In this section, results from the full models labeled OLS (2) and OPROBIT (2) are discussed.

In interpreting the estimates of the ordered probit regression, the signs permit only the prediction of the direction of change in probabilities of extreme outcomes. In this case these are less than primary education and university education. For example, the estimates on parental education are positive, which indicates that a worker with educated parents had greater probability of being a university graduate and less likely to have less than primary education, than a comparable worker with uneducated parents. The direction of change in the probabilities of the schooling outcomes in between cannot be inferred. Therefore partial effects are calculated following Long (1997).

[8]At this point the schooling model is not linked with the earnings function. It is estimated independently as in the schooling demand literature.

Table 3: OLS and Ordered Probit Estimates of Schooling Attainment Functions: All workers

Explanatory Variables	OLS (1)		OLS (2)		OPROBIT (1)		OPROBIT (2)	
Age (years)	0.32*	(5.43)	0.29*	(4.90)	0.16*	(5.61)	0.16*	(5.17)
Age squared	-0.005*	(6.04)	-0.004*	(5.48)	-0.002*	(6.40)	-0.002*	(5.92)
Male worker	-0.14	(0.67)	0.03	(0.16)	-0.18	(1.65)	-0.08	(0.71)
Father's education								
Primary	0.79*	(3.53)	0.76*	(3.49)	0.30*	(2.64)	0.31*	(2.72)
Post-primary	1.59*	(4.78)	1.50*	(4.49)	0.75*	(4.15)	0.73*	(3.93)
Mother's education								
Primary	0.51*	(2.52)	0.51*	(2.54)	0.27*	(2.55)	0.27*	(2.47)
Post-primary	1.09*	(3.18)	0.99*	(2.94)	0.62*	(3.15)	0.60*	(2.97)
Nearest primary school								
1-3 kms			-0.05	(0.29)			-0.12	(1.27)
3-6 kms			0.00	(0.02)			-0.14	(0.98)
6-10 kms			-0.54	(1.29)			-0.42**	(2.04)
Over 10 kms			-1.16**	(2.25)			-0.44***	(1.83)
Nearest secondary school								
1-3 kms			-0.52**	(2.10)			-0.26**	(2.09)
3-6 kms			-0.23	(0.90)			-0.11	(0.85)
6-10 kms			-0.08	(0.25)			-0.01	(0.05)
Over 10 kms			0.27	(0.96)			0.12	(0.81)
Province of education								
Central			-0.35	(0.97)			-0.16	(0.84)
Eastern			-0.67***	(1.78)			-0.29	(1.42)
Western			-0.94*	(2.71)			-0.51*	(2.63)
Rift Valley			-0.08	(0.22)			-0.10	(0.46)
Nyanza			-1.43*	(3.86)			-0.75*	(3.66)
Coast			-0.81***	(1.83)			-0.47**	(1.97)
Threshold 1					1.28		0.57	
Threshold 2					2.82		2.17	
Threshold 3					5.03		4.45	
χ^2(D.F)[a]						65.39(4)		57.37 (4)
χ^2(D.F)[b]								20.64 (8)
Pseudo R^2					0.11		0.14	
Log-likelihood					-763.96		-742.08	
Constant	3.72*	(3.46)	4.99*	(4.30)				
F		20.73*		17.39				
(D.F)[a]		(4, 176)		(4, 176)				
F				2.46**				
(D.F)[b]				(8, 176)				
Adjusted R^2	0.20		0.23					
Sample size				843				

Notes:
(1) The dependent variable for OLS is years of education completed and for OPROBIT is highest education level attained.
(2) Numbers in ( ) are absolute values of t-statistics (for OLS) and z-statistics (for OPROBIT) based on standard errors robust to heteroskedasticity.
(3) Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively.
(4) (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.

Because of the non-linear nature of the ordered probit model, marginal effects on probability of a level of education when a given explanatory variable changes are a function of both the parameter estimates and the explanatory variables. The marginal effects evaluated at the

sample mean of explanatory variables are presented in Table 4. The change in probability of a worker attaining education outcome j when a continuous variable h_k changes is (Long, 1997)

$$\frac{\partial \Pr(s=j)}{\partial h_k} = \beta_k \left[\phi(\mu_{j-1} - \bar{h}\beta) - \phi(\mu_j - \bar{h}\beta) \right], \quad j = 0,1,2,3 \tag{7a}$$

where $\phi(.)$ is the normal density. For categorical variables one can calculate the difference in probability of education outcome j as variable h_k changes from h_0 to h_1, is calculated as

$$\Delta \Pr(s=j) = \Pr(s=j \mid h, h_k = h_1) - \Pr(s=j \mid h, h_k = h_0), \quad j = 0,1,2,3 \tag{7b}$$

For example, in the case of parental post-primary education, the change in probability of a worker attaining a given level of education is the difference in the probability when a parent has post-primary education, compared to a parent with no education. In the sample used, most parents have no education and hence this group would be suitable as a reference category.

Table 4: Change in Education Attainment Probabilities in the Ordered Probit Model

Variables	Below primary	Full Primary	Full secondary	University
Age (years)	0.02*	0.03*	-0.04*	-0.01*
Male worker	0.01	0.03	-0.03	0.00
Father's education				
Primary	-0.03*	-0.09*	0.11*	0.01**
Post-primary	-0.05*	-0.22*	0.22*	0.05**
Mother's education				
Primary	-0.02*	-0.08*	0.09*	0.01*
Post-primary	-0.04*	-0.18*	0.18*	0.04**
Nearest primary school				
1-3 kms	0.01	0.04	-0.04	0.00
3-6 kms	0.01	0.04	-0.05	-0.01
6-10 kms	0.05	0.12*	-0.16**	-0.01*
Over 10 kms	0.06	0.12*	-0.16***	-0.01*
Nearest secondary school				
1-3 kms	0.03***	0.08**	-0.09**	-0.01**
3-6 kms	0.01	0.03	-0.04	0.00
6-10 kms	0.00	0.00	0.00	0.00
Over 10 kms	-0.01	-0.04	0.04	0.01
Province of education				
Central	0.02	0.05	-0.06	-0.01
Eastern	0.03	0.09	-0.11	-0.01***
Western	0.06**	0.14*	-0.19*	-0.02*
Rift Valley	0.01	0.03	-0.03	0.00
Nyanza	0.10*	0.19*	-0.27*	-0.02*
Coast	0.06	0.13*	-0.17**	-0.01*

Notes: Derived from OPROBIT (2) in Table 3. For dummy variables the reported effect is for a discrete change from 0 to 1. Significance at 1%, 5%, and 10% level is indicated by *, **, and *** respectively.

The mean predicted probability that a worker has less than full primary education is 4%, that of having full primary education is 40%, and those of having secondary and university education respectively are 54% and 2%. Table 4 shows that, parental education increases the probability that a worker will attain at least secondary education and is also associated with

more years of education. To test whether mother's education coefficients are identical to father's education coefficients, an F-test (years of education model) and a Chi-square test (levels of education model) were carried out. The observed F-value and Chi-square-value reported at the bottom of Table 4 are statistically significant. Therefore, the null hypothesis may be rejected in both models. This implies that father's and mother's education coefficients are not identical.[9]

The differential impact may be due to a number of reasons. First, parent's education impact may be gender-specific as recent studies suggest (e.g. Glick and Sahn, 2000 for Conakry, Guinea; Tansel, 1997 for Ghana and Cote d' Ivoire. Second, it could be that it is 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 comes through home production activities, then if education increases the likelihood of mothers labor market participation, less time may be allocated to home production activities and weaken mother's education effect.

Distances to school facilities have significant effect on education attainment at some levels. A distance greater than 10 kilometres to the nearest primary school increases the probability that a worker attained only primary education or less, and hence fewer years completed. Those within 1 to 3 kilometres from the nearest secondary school had greater probability of ending education at primary level or below. Tansel (1997) reports a similar finding in Ghana. There, distance to secondary schools reduces middle school attainment.[10] To test if distances to schools have identical effects, the F-test (years of schooling model) has an observed F (8, 176)-value of 2.46 with p-value of 0.01. The corresponding test in the levels of education model has an observed chi-square (8)-value of 20.64 with p-value of 0.00. The null hypothesis may be rejected in both models. This implies that distances to nearest schools have joint and significant effect on education attainment.

The province in which a worker received most education has significant effect on some levels of education attainment. Attendance of most education in Western, Nyanza and Coast provinces raised probability of having primary education or below, and reduced years of education completed. Such regional differences in education may reflect income differentials, unmeasured school characteristics such as teacher supply and school facilities or household specific characteristics.

5.3. The Earnings Functions with and Without Parents' Education

Estimates of returns to education are questioned because the wage functions used do not usually control for family background. The estimates are said to be subject to omitted family background bias (Lam and Schoeni, 1993). This section uses empirical evidence from earnings functions to answer the following questions. First, ccontrolling for own education does parental education significantly influence earnings? Which level of parental education has the largest impact on earnings? Second, do the private wage returns to education decline when parental education is controlled for in earnings function? Which level of worker's education has the most sensitive return to these controls? The first two columns of Table 5

[9] Schooling models with mother's and father education pooled were also estimated. The results showed strong impact of parental education on education attainment.

[10] 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).

present the earnings functions with and without parental education controls. The issue of endogenous schooling in the wage function is set aside first.

Table 5: Ordinary Least Squares and Instrumental Variable Estimates of the Earnings Function: All Workers.

Explanatory Variables	OLS (1)	OLS (2)	IV (1)	IV (2)	OLS (3)	OLS (4)
Education (years)	-0.08**	-0.08**	-0.52**	-0.38	-0.07**	-0.08**
	(2.47)	(2.19)	(2.16)	(1.36)	(2.20)	(2.14)
Education squared	0.0122***	0.0113***	0.0396***	0.0340**	0.0114***	0.0108***
	(6.63)	(5.79)	(3.09)	(2.21)	(6.17)	(5.59)
Age (years)	0.04**	0.05***	0.00	-0.01	0.04**	0.05***
	(2.25)	(2.83)	(0.11)	(0.34)	(2.26)	(2.80)
Age squared	-0.0003	-0.00*	0.0001	0.0003	-0.0003	-0.0004
	(1.27)	(1.70)	(0.51)	(1.18)	(1.24)	(1.60)
Time in firm (years)	0.02**	0.0244**	0.06***	0.06***	0.03**	0.03**
	(2.11)	(2.20)	(3.71)	(3.43)	(2.26)	(2.31)
Time in firm squared	-0.0003	-0.00	-0.00**	-0.00**	-0.00	-0.00
	(0.83)	(0.92)	(2.42)	(2.39)	(1.06)	(1.09)
Male worker	-0.02	0.05	0.05	0.07	0.02	0.08
	(0.25)	(0.88)	(0.76)	(0.94)	(0.27)	(1.34)
Father, primary education		0.07				0.08
		(1.28)				(1.30)
Father, post-primary education		0.22***				0.22***
		(3.00)				(2.94)
Mother, primary education		-0.01				0.00
		(0.23)				(0.07)
Mother, post-primary education		0.17**				0.16*
		(1.97)				(1.89)
Province of education dummies				Yes	Yes	Yes
Distance dummies, primary school						Yes
Distance dummies, secondary school						Yes
Constant	0.77**	0.38	2.74**	2.05	0.89***	0.51
	(2.26)	(1.13)	(2.35)	(1.59)	(2.68)	(1.53)
R-squared	0.43	0.45	0.21	0.22	0.44	0.47
F (D.F.)-value of excluded instruments			8.11 (8.22)	7.46 (7.22)		
P-value			0.00	0.00		
Partial R^2 of excluded instruments			0.15	0.10		
χ^2(D.F.)-value, distances to schools			5.20 (8)	6.43 (8)		
P-value			0.74	0.60		
χ^2(D.F.)-value, parental education			2.69 (4)	2.36(4)		
P-value			0.61	0.67		
χ^2(D.F.)-value, province of education			12.60 (6)			
P-value			0.05			
χ^2(D.F.)-value, cver-identification test			19.33 (16)	6.84(10)		
P-value			0.25	0.74		
χ^2(D.F.)-value, Hausman test, OLS vs. IV			18.99 (9)	7.52 (14)		
P-value			0.03	0.91		
Sample size			843			

Notes: The dependent variable is logarithm of hourly earnings. Robust z-statistics in parentheses. * Significant at 10%; ** significant at 5%; *** significant at 1%. Ins indicates the variable is one of the excluded instrument.

The answer to the question of whether parental education affects wages in the manufacturing labor market is yes. Regression estimates of the earnings function which controls for parental education-OLS (2) show that, even controlling for own education, post-primary education of a worker's parents has significant effect on earnings. But parental primary education does not seem to significantly influence earnings. This is in contrast with the result in the schooling models, where both parental primary education and parental post-primary education has significant impact on education attainment. The coefficient of father's post-primary education in the earnings function is 0.22, implying that a worker whose father has post-primary

education receives a wage that is 27% higher (exp(0.22)-1)*100 than that of a worker with comparable characteristics except the father has no formal education. The corresponding earnings advantage associated with mothers' post-primary education is 19%.

The answer to the question of whether education effects fall when parental education controls are included is yes. Evaluated at 6 years, 10 years, and 14 years of education, the returns (multiplied by 100) to education are 7%, 16%, and 26% in earnings regression without parental background controls-OLS(1). With controls for parental background-OLS (2), the corresponding returns to education are 6%, 15%, and 24%. While there is a drop in returns to education at the three points when controls for parental education are included, the highest decline (14%) is at 6 years of education. This suggests that the phenomena represented by parental education operates more strongly at this level given the earnings function estimated.

The evidence on the effect of family background on earnings from other less developed 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) found 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) addresses the issue of family background bias with data from the urban labor market in Ethiopia. She estimated earnings functions with and without controls for parental background and found that returns to education fell by 20% and 10% for workers in public and private sector respectively. But using selectivity-corrected earnings regressions she demonstrated an important point. The importance of family background in determining earnings was because family background influenced allocation into employment sectors. Once the selectivity effect is controlled for, the impact of family background on earnings is weakened. Returns to education fall by only 5% in public sector and 4% in the private sector upon controlling for selectivity effects. The return to family background in this labor market comes through the impact of family background on entry effect into sectors. It is likely that a selectivity mechanism is in operation in the Kenyan labor market with family background determining sectors or industries in which individuals work given that formal sector jobs are not generated at a rate to match the growth in supply of educated labor.

The impact of parental education could reflect returns to labor market networks and connections. A hint of this comes from responses to a question in the survey that asks workers to say how they came to know about the their current job. Around 40% of the workers say that they got information from friends, family members and relatives. This may reflect the role that networks and contacts play in this labor market. More educated parents may have better labor market contacts and networks that assist in job search and allocation into sectors. More educated parents may have relatives who are educated and working in urban centres, who provide information about job opportunities in urban areas, especially low skilled jobs that are not widely publicized.[11] The parental education effect on wages may also reflect their capacity to meet job search costs and hence sustain longer unemployment durations. In the

[11] Bigsten (1996) discusses the role of information in circular migration decisions of smallholders in Kenya

end, the child-worker may secure a higher pay job or a job with greater opportunities to acquire human capital.

The positive effect of parental education on earnings could also represent home environment. Families make many decisions and take many actions regarding investments in home learning, health and nutrition. For example Knight and Sabot (1990) found that, in Kenya better-educated parents were more likely to send their children to public secondary schools, which were the schools that produced graduates with relatively higher cognitive skills. It is possible that the decisions and actions of families complement formal schooling, and result in skills that are rewarded in the labor market. If such complementarity is stronger for better-educated parents, this may become visible as differences in wages.

Another interpretation is suggested by Grilliches (1977). He cautions against too much zeal in controlling for parental background, noting that if workers' education is measured with error and family background is correlated with actual education, this error could be exacerbated forcing returns to education down. For example, Lam and Schoeni (1993) find that in Brazil, part of the reduction in returns to education that would otherwise be attributed to family background is due to measurement error in education. It is possible that in the survey used some workers reported their education level and grade incorrectly. If this is serious, both the reduction in returns to education and the direct impact of parental education calculated from OLS (2) may be due to measurement error.

5.4. Earnings Function with Endogenous Education

Up to this point the education relation (part 5.2) and the wage relation (part 5.3) are studied separately. Provided one can find variable(s) that induced variation in workers' education attainment but have no direct impact on earnings, education can be instrumented. In this paper parental education, distances to nearest schools, and the province in which the worker received most education could have the two properties. Because the results hinge on the instruments validity test and over-identification tests are reported.

In the initial instrumental variable earnings regression-IV (1), parents' education, distances to nearest schools, and province of education are treated as excluded (from the earnings function) instruments. The return to education at 6 years of education is –4%, which is lower than the OLS estimate of 7%. At 10 and 12 years of education the returns are 28% and 44% respectively compared to 16% and 26% in the OLS earnings regression. This finding is similar to the findings in the literature cited in the introduction. Instrumental variable estimates exceed standard estimates. But before one can make a firm conclusion, the quality and validity of instruments have to be examined.

The quality of instruments is assessed using tests proposed by Bound et. al. (1995). Good quality instruments should have significant effect in the reduced form schooling model.[12] The F-test of excluded instruments rejects the null hypothesis that the instruments set has no joint significance in the reduced-form schooling model.[13] This suggests that they have joint significant impact on workers' education. Bound et al also suggest that addition of

[12] The reduced form schooling model includes all the variables in the schooling equation discussed in section 5.2 and variables in the wage function. The estimates are not reported.
[13] All test-statistics for the various tests and the p-values are reported at the bottom of Table 5.

instruments in the reduced form equation should improve the explanatory power of that model. The partial R-squared suggests that the instrument set explains 15 % of the variation in years of education. The test of over-identification does not reject the null hypothesis that the instruments are valid and the Hausman test suggests that OLS and IV (1) are not identical. This suggests that in this specification of the earnings function it is necessary to instrument.

However, this result depends critically on the instruments. Therefore sensitivity analysis is essential. First, tests of exogeneity of parental education, distances to schools, and province of education separately are computed. The test for province of education rejects the null. This suggests that province of education may not be excluded from the earnings function. Therefore, the earnings function is estimated with province of education included among the earnings determinants. This earnings regression has the label IV (2). Now the Hausman test statistic is not significant. It implies that the results with province of education removed from the set of excluded instruments-IV (2) and the standard earnings function with province of education included as a wage determinant-OLS (3), are not statistically different. It suggests that in this specification of the earnings function it was not necessary to instrument for education.

The second type of sensitivity analysis was to estimate an earnings function specification with all variables (excluded and included instruments) as wage determinants. The results have the label OLS (4) and show that the province where a worker obtained most education has direct effect on earnings. An F-test of the hypothesis that the province dummies have identical effects has an observed F (6, 176)-value of 3.29 with a p-value of 0.004. The null hypothesis may be rejected. A similar test for parental education has an observed F (4, 176)-value of 6.19 with p-value of 0.00. Hence the null hypothesis may be rejected also. It is only distances to school facilities for which the null hypothesis of no joint effect on earnings could not be rejected since the observed F (8, 176)-value is 0.79 with p-value of 0.61.[14] This suggests that it is only distances to nearest schools, which are not, omitted variables in the earnings function. But distances to schools have very low explanatory power in the schooling models, which would make them very weak instruments. The partial R-squared that relates to distances to nearest schools is 0.0195. Bound et. al. caution against use of weak instruments.

The results of the sensitivity analysis have one message. It is that while the joint test of over-identification for the full set of instruments suggests that they are valid, that is, can be excluded from the earnings function, the sensitivity analysis casts doubt over the result given that parental education and province of education have direct impact on earnings. As Deaton (1997) notes, getting good instruments is a problem and it is also hard to make the case that instrumental variable estimates are preferable to OLS estimates.[15]

[14] Tests for mother and father's education separately show that each has an effect on earnings. Similarly, separate tests for distances to primary schools and distances to secondary schools show that neither set has direct effect on earnings.

[15] Harmon and Walker (1995) and Vella and Gregory (1996) treat education as a self-selection problem. They add a selection term to the earnings function. Manda (1997) uses this approach for a sample of male workers in Kenyan enterprises and found no evidence of self-selection bias. We derived a selection term from the ordered probit model and included it in an earnings function. The estimated coefficient on this term was –0.33 and significant. However, Carneiro, Heckman, and Vytlacil (2001) point out that the view underlying the self-selection model is different from the view underlying instrumental variables model. Following the literature on family background and earnings we focus on instrumenting education in the wage function.

6. Summary and Conclusion

Returns to education are an important input into development policy. But it is often argued that estimates derived from earnings functions that fail to control for family background or for endogenous schooling may be misleading. Family background and measures of schooling availability at the time adult workers went to school are usually missing in labor market surveys. This paper uses data collected in 2000 in a survey of manufacturing firms in Kenya, that has this information, to study schooling and wages. To begin with, regression estimates of education functions show that parental education is an important determinat of the quantity and level of education attained. The estimates suggest that father's and mother's education has statistically separate effects on education as found in other African countries. The results illustrate that the higher the level of education individuals acquire the greater the effect on their children's education.

Second, wage function estimates show that workers with fathers that have post-primary education receive wages that are 27% higher than those of comparable workers whose parents are uneducated. Mother's post primary education is associated with 19% higher wages. This may reflect sample selectivity into private sector wage employment as Krishnan (1996) finds in Ethiopia. Third, controlling for parental education in the wage function reduces the level of returns to worker's education. The return to education at 6 years falls by 14%, that to 10 years of education falls by 6%, and that to 14 years falls by 8%. These percentage declines are relatively lower when compared to the average of about 20% in studies reviewed for a number of other less developed countries. A part of the return to worker's own education is a return to parents' education. But parental education controls do not render the worker's education effect insignificant.

Third, wage function estimates with education instrumented show that standard estimates of returns to education may be understated. But while treating education as endogenous to wage formation is a sound strategy to obtain consistent estimates on which to derive returns to education, getting good quality and valid instruments is difficult. The use of parental education to instrument education may be inapropriate, which would seem to support the conclusion of the review by Card (1999) of studies that instrument for education.

7. References

Appleton, S. (2000), Education and health at household level in Sub-Sahara Africa. *CID Working Paper No. 33*, Havard University.

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. (1984), *Education and income determination in Kenya*. Gower Publishing Company Limited. Gower House, Croft Road, Aldershot, England.

Bigsten, A. (1996), The circular migration of small-holders in Kenya, *Journal of African Economies*, vol. 5, Issue 1, pp 1-20.

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.

Card, D. (1993), Using geographic variation in college proximity to estimate the return to schooling. *National Bureau of Economics Research Working Paper No. 4483*. Cambridge Massachussetts, U.S.A.

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.

Carneiro, Heckman, and Vytlacil (2001), Estimating the return to education when it varies across individuals, unpublished manuscript, U of Chicago, April, 2001.

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

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

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.

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

Schultz, 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.

Schultz, T.P. (2002), Why governments should invest more to educate girls, *World Development*, vol. 30, no. 2, pp 207-225.

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.

Education, Employment and Earnings in Kenya*

By

Anthony Wambugu
Department of Economics
Göteborg University
Box 640
SE-405 30 Göteborg
Sweden

2002c

Summary: This paper uses a survey of rural and urban households in Kenya to examine the relation between education, employment, and earnings. For both women and men, the strongest impact of education is to raise the probability of entry into the public and private sector wage employment. Men with more than primary education have lower probability of entry into the informal sector, while those with less than full secondary education reduce their chance of agricultural employment. Women in households that own land are less likely to work off the farm. Decomposition of the difference in employment allocation probabilities between women and men shows that, a substantial part of the differential can be explained by differences in their characteristics. The highest monetary reward for primary education is in the informal sector, while returns to secondary education are highest in the private wage sector. Women have higher return to education than men, which could reflect greater effect of education on access to employment for women. Correcting the earnings function estimates for sample selection bias does not alter the standard estimates except for women in the public sector.

* The suggestions and comments from Arne Bigsten and Henry Ohlsson are gratefully acknowledged

I. INTRODUCTION

Investments in education by individuals and households aim to improve access to employment and earnings.[1] In the 1990s, a feature of Kenya's employment structure is the enormous expansion of the informal sector compared to the near stagnation in formal sector employment.[2] A main purpose of this study is to examine how education influences employment type and whether and how other individual and household characteristics have a role. This is important to know because different jobs have different opportunities for skill acquisition and earnings. Also, as Demery and Grootaert (1990) note, such analysis can give indications of heterogeneity or discrimination in the labor market. A second purpose is to compare private returns to education across three employment types. The level of returns to education can indicate the potential for education to improve economic welfare and to give incentives to individuals and households for educational investments.

Previous analyses of employment and earnings determination in Sub-Sahara Africa countries offer four views or approaches to the labor market. One approach is to augment the human capital earnings function of the type developed by Becker and Chiswick (1966) and Mincer (1974) with variables that indicate the employment type a worker is in. For example, in part of a study on education and income determination in Kenya, Bigsten (1984) finds that type of employment influences earnings in rural and urban areas. But he points out that employment type may be simultaneously determined with earnings. Knight and Sabot (1990) also use this approach to examine the effect of the public-private sector wage differences on return to secondary education in Kenya and Tanzania. They find that return to secondary education can be distorted if the sectoral dimension is ignored. In this approach, wages in the identified sectors differ in levels while rewards to worker characteristics are identical.

A second approach is to assume that workers face two labor market choices. The authors estimate human capital earnings functions to assess the public-private or the formal-informal wage differentials.[3] Examples include the studies by Lindauer and Sabot (1983) on Tanzania, and Nielsen and Rosholm (2001) on Zambia. Public sector workers are reported to enjoy a wage premium over private sector workers. In Kenya, Appleton, Bigsten and Manda (1999) examine returns to education in urban wage and self-employment sectors. A major assumption of the studies in this approach is that workers are randomly assigned into sectors.

The third approach takes into account the possibility that earnings and employment type are simultaneously determined. The studies by Van der Gaag and Vijverberg (1988) on Cote d'Ivoire and Van der Gaag, Stelcner, and Vijverberg (1989) on Cote d' Ivore and Peru that used a switching regression model fall into this category. They find that taking account of sector allocation could reverse the direction of the wage differential. In urban Zambia Andersson (1993) finds that public and private wage structures differ and that the formal-informal distinction was important for men, but not for women. In part of his study on Ghana, Teal (2001) considers earnings determination in the public and private sectors with correction for the non-random assignment into the wage sector. He found no evidence of sample

[1] Individuals can also acquire education if they consider it to have intrinsic value.

[2] Informal sector activities cover a wide range, from the street seller with only a small bag to small-scale metal workers with equipment. In 1991 there were around one million workers in the informal sector; which increased to around four million workers in 2000 (Government of Kenya, 1995, 2001).

[3] Other examples of binary events in the labor market include; employed or not employed, working or not working.

selectivity bias. In other studies (e.g Glewwe, 1996 on Ghana and Nielsen and Nielsen, 2001 on Zambia) sector allocation is treated as two binary events (labor market particpation and sector allocation).

A fourth strand of research assumes that workers face multiple choices of sectors or employment types in the labor market. For example, Thomas and Vallee (1996) examine earnings in the informal, formal, and regulated sectors within manufacturing in the Cameroon. Glick and Sahn (1997) examine gender and schooling impacts on employment and earnings in the private sector, the public sector, and in self-employment businesses in Conakry, the capital of Guinea. Krishnan, Sellassie, and Dercon (1997) study the Ethiopian urban labor market on the basis of five employment states: the private sector, public sector, and self-employment, unemployment, and non-participation. A study on Cote d' Ivoire (Vijverberg, 1993) estimate wage and off-farm earnings functions for men and women corrected for sample selectivity based on four employment states: farm self-employed, off-farm self-employment, wage employment, and non participation. The public-private distinction was not considered. Appleton, Collier, and Horsnell (1990) consider the private and public sector distinction in Cote d'Ivoire with non-participation as a base choice. They also consider a three-sector model of the wage sector, with the private sector split into union and non-union segments. In Kenya, Kabubo (2000) estimates earnings functions for the public and private sector with selection based on four employment states. The distinction between informal sector employment and other employment types and the determination of informal sector earnings is not considered. Given the growing role of the informal sector in Kenya's employment structure, it is important to analyze available data to highlight the characteristics and incomes of those that work in this sector.

The above studies advance understanding of labor markets in Africa. In particular, given that many people work outside the wage sector, it is important to use available surveys to look into other sectors as well. This paper complements research into education, employment and earnings in Kenya within the fourth strand of the literature in three directions. First, it models the full range of employment types: public sector employment, private sector employment, informal sector employment, agricultural employment, and unpaid family work. The aim is to identify the impact education and other individual characteristics have on the probability of falling in one of the five employment types. Second, the paper investigates the source of male-female differential in access to employment, through decomposition of the entry probability differentials. Third, private returns to education across three sectors-the public sector; the private sector, and the informal sector are compared for men and women and checked for selectivity bias.

The paper is organized as follows. A description of the data used in the analysis is in Section 2. Section 3 identifies the determinants of employment type. Earnings function analysis for different sectors with and without sample selectivity correction is in Section 4 and Section 5 concludes the paper.

II. DATA

The data are from a survey of urban and rural households in Kenya, with information on 59,183 individuals in 10, 857 households. The Central Bureau of Statistics (CBS) conducted it in August to September 1994. The information includes a range of variables useful for

analysis of different socio-economic issues. The sample used for the analysis in this paper is restricted to workers aged 15-65 at the time of the survey. Only the main occupation is considered and only workers who were not studying and had full information on crucial variables for earnings function analysis e.g. education and information to generate household-level variables, are included in the analytic sample. Of the 22, 579 workers that meet the criteria, 66% work in agriculture; 11% in private formal sector; 8% in public sector; 9% in the informal sector, and 6% are unpaid family workers.

Table 1 presents the summary statistics for the workers by employment sector for the variables used in the analyses. The average age of men in the sample is 35 years while that of women is 33 yeas. The unpaid family workers are the youngest; men in this sector are younger (24 years) than women (28 years). In the other sectors men are older than women; in the public sector men are 37 years old on average while women are 5 years younger. The age gap in the private sector is identical to that in the public sector. However the private sector workers are younger than public sector workers; on average men are aged 34 years and women are 29 years. In both the agriculture sector and the informal sector men are aged 35 years on average while women are 34 years and 33 years respectively. Most (70 per cent) of the workers in this sample are married. But the proportion of married men in unpaid family work is relatively small (19 per cent). Also, relatively fewer (44 per cent) of women in private sector are married. The data also indicate that most men (69 per cent) are household heads compared to 19 per cent of women. In most households there are small children (0-6 years), and school age children (7-14 years).

Turning to education, 27 per cent of men and 41 per cent of women have no education. While there is no major gap in primary education between men and women, the proportion of women who have completed secondary education is about half (7 per cent) that of men. The respective proportions of men and women with no education are lowest in the public sector (6 per cent and 7 per cent) and highest in unpaid family work and agriculture (39 per cent to 45 per cent). Almost 65 per cent of both men and women in unpaid family work have less than full primary education. In agriculture, 73 per cent of the women and 69 per cent of the men fall in this education class. Labor demand in the public sector is the most skill-intensive; Sixty-three per cent of the men and 69 per cent of the women completed at least some secondary education. The corresponding proportions are 35 per cent and 31 per cent in the private sector and 31 per cent and 22 per cent in the informal sector.

The measure of earnings used in this paper is the annual earnings from main occupation in Kenya shillings. For the self-employed the survey reports net income. On average, the earnings of men in the public sector are one and quarter times the earnings of women in the same sector. In the private sector, men earned on average one and three quarters times the average earnings of women, while in the informal sector men earned more than twice what women earned. The table also shows that relatively larger proportions of persons in unpaid work and agriculture are in households that received some transfer income in cash or in kind. The other group that has this feature is women in the informal sector. Most (70 per cent) of the households have some land. For unpaid family workers the proportion of those that come from households with land is much lower. It ranges from 36 per cent to 38 per cent. Also, compared to men, lower proportions of women in formal wage employment and in the informal sector are from households with land. Land is potentially important since lack of

land in the Kenyan context may push individuals to search for other types of employment. But it could also be a springboard to other types of employment such as informal sector businesses. For example, land may be used as collateral to secure loans to start own business.

Around 45 per cent of both men and women are absolutely poor; the respective proportions are largest in unpaid family work (58 per cent and 46 per cent) and agriculture (54 per cent and 47 per cent). The least poor persons are in public sector where the respective proportions of men and women are 24 per cent and 21 per cent. Moderate proportions (29 per cent to 38 per cent) of workers in private and informal sectors are absolutely poor.

The differentials in individual and household characteristics may explain the differentials in access to employment and also earnings. There may be differentials in returns to these characteristics also, which could further affect differentials in sectoral allocation and earnings. In the next section, the roles of the factors above in sectoral allocation are examined. In a subsequent section, earnings will be examined.

Table 1: Summary Statistics of Variables used in the paper by Sector and Gender

	Unpaid workers		Agriculture		Public sector		Private sector		Informal sector		Total	
	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women	Men	Women
Age (years)	24	28	35	34	37	32	34	29	35	33	35	33
(std dev)	(10)	(10)	(14)	(13)	(8)	(8)	(10)	(10)	(10)	(10)	(12)	(12)
Married	19	67	65	71	88	66	75	44	85	64	70	69
Household head	16	4	63	18	88	34	75	25	85	31	69	19
No education	44	42	39	45	6	7	14	23	15	35	27	41
Some primary	20	21	30	28	16	15	31	29	32	27	29	27
Full primary	9	15	14	14	15	9	20	17	21	16	16	14
Some secondary	8	11	8	7	18	18	15	15	12	11	11	9
Full secondary	17	10	7	5	31	31	16	11	16	10	13	7
Post-secondary	1	1	1	1	10	17	3	4	2	1	3	2
University	1	1	0	0	4	3	1	1	1	0	1	0
Absolutely poor*	58	46	54	47	24	21	38	36	29	36	45	44
Small child present	52	74	72	78	74	77	71	76	78	75	72	77
School child present	68	62	73	71	61	64	54	57	57	67	66	69
Prime age adult present	96	91	88	86	81	82	84	86	87	86	87	86
Household has land	36	38	80	83	62	54	61	44	60	50	71	74
Household has transfer income	45	36	45	46	28	34	34	33	29	45	39	44
Annual income (mean)	-	-	-	-	46,908	37,672	35,812	20,428	90,941	40,225	52292	33181
Number of observations	330	961	5583	9286	1284	513	2036	664	1047	875	10280	12299

Source: computed from survey data. *The poverty line is defined as Kenya shillings 987.27 in rural areas and Kenya shillings 1489.63 for urban areas. Poverty status defined following Central Bureau of Statistics poverty line based on the 1994 survey.

III. DETERMINANTS OF JOB ATTAINMENT

Job attainment can be considered an outcome of interaction between supply and demand factors. On the demand side, labor demand is a function of marginal productivity. According to the economic theory of human capital, individuals can improve their productivity through skills acquisition (through education and experience). Unobserved skills also play a role in productivity improvement. On the supply side the labor supply decision is an outcome of worker's utility maximization; with income, market commodities, and leisure as arguments in the utility function. Factors that influence expected earnings and reservation earnings would have bearing on labor supply. For the self-employed, human capital and other assets are important in starting own businesses; they may determine the relative costs and returns from setting up own business.

Because the study examines more than two employment types, a binary event model is not enough. Therefore, instead of a single equation logit model, a multinomial logistic model is used to identify the individual and household characteristics of those in various employment types. That is, the characteristics, which correlate with the probability of individual i being in, sector m. This model also improves on efficiency of parameter estimates. The form of the multinomial logistic model is

$$(1)\ \Pr ob(y_i = employment_m) = p_{im} = \frac{e^{\beta_m' x_i}}{\sum_{j=1}^{J} e^{\beta_j' x_i}}, i = 1,....,N; m = 1,.....,J$$

where J indexes employment sectors, with N the sample size, x_i a vector of regressors that influence labor demand and supply decisions, and assumed to be exogenous. This equation for the probability of an outcome can be derived either as a probability model or as a discrete choice problem among alternatives (see Long, 1997). Given the central place occupied by agriculture and informal sectors in Kenya's employment structure, the equation models entry into five employment types: unemployed (unpaid workers), agricultural sector, public sector, private sector, and informal sector. This results in a five-way multinomial logit model. It can be assumed that individuals take into account that their individual, household and regional characteristics, present both constraints and opportunities to enter various employment types.

Therefore, the explanatory variables in the logit model are grouped into three categories: individual characteristics (education, age, gender, marital status, household headship); household characteristics (presence of children and prime age adults); and assets (land ownership, and non-labor income). The variables could influence the employment sector by influencing expected earnings and reservation earnings. For example, if skill requirements differ across sectors, the expected wages can also differ. Reservation earnings can also differ across sectors. For instance, informal sector workers can have autonomy and flexibility that leads to lower relative reservation earnings. Presence of children may raise demand for home produced services and make individuals to enter an employment type that is compatible with this. Household headship may signal that the individual takes the economic responsibility for the household. While it is possible that some of the explanatory variables may be endogenous to employment type allocation, there are no good instruments for them. The results are to be interpreted in this background.

The first task of the paper is to assess the gender differential in access to various employment types. This is done using a simulation exercise based on a decomposition method (Borooah, 2001) for multinomial logit model. It allows the study to answer the question of how much of the observed differences in mean employment probabilities between women and men can be attributed to differences in observed characteristics and how much is left unexplained. The first step in the decomposition is to estimate the multinomial logit employment model (1). The parameter estimates of this model are used to predict, for each person in the sample, the probability of being in each of the five employment types. The mean values of the predicted probabilities for men and women and for all persons are shown in Table 2, panel A. For example, the mean probability of being in agriculture for men and women were, 54.3 per cent and 75.5 per cent respectively. The mean predicted probabilities are equivalent to the sample proportions of men and women in each employment type. We will have use for the mean probabilities shortly.

Table 2: Predicted individual probabilities of employment types for men and women

	Unpaid workers	Agriculture sector	Public sector	Private sector	Informal sector
A. Mean probabilities (S_m)					
All persons	5.7	65.9	8.0	12.0	8.5
Men	3.2	54.3	12.5	19.8	10.2
Women	7.8	75.5	4.2	5.4	7.1
B. Gender probabilities (P_m)					
Men	4.7	59.6	8.9	18.1	8.7
Women	6.2	72.0	6.8	6.2	8.8

Source: calculated from multinomial logit coefficients

In the second step, the multinomial logit estimates are used to simulate two hypothetical scenarios that help to uncover gender differentials in access to employment types. First, suppose all persons in the sample are men. Predict p_{im}^{M}, the probability of person *i* being in employment type *m (m = 0, 1, 2, 3,4)*. Next, suppose all persons are women and predict p_{im}^{W}, the probability of person *i* being in employment type *m*. The probabilities predicted under the two hypothetical situations are referred to as gender probabilities. The estimates of respective mean predicted gender probabilities, p_m^M and p_m^W are shown in Table 2, panel B. The difference between the mean values of the gender probabilities of men and women, can be attributed to unequal returns to characteristics (coefficients) of persons who have comparable observed characteristics except gender. It shows that, if the characteristics of all persons in the sample were evaluated using male coefficients, 5 per cent of the sample would be unpaid family workers; 60 per cent would be in agriculture; 9 per cent would be in the public sector; 18 per cent would be in the private sector; and 9 per cent would be in the informal sector.

The gender differential in entry probabilities to different sectors is the ratio, $\lambda_m^W = p_m^W / p_m^M$, of being in sector *m* if the person is a woman compared to a man. If the ratio is one, the two probabilities are equal and there is no differential. The mean probabilities (or sample proportions), s_m^W and s_m^M, in Table 2, panel A, of women and men respectively being

in different sectors, deviate from the estimated gender probabilities (p_m^M and p_m^W). This shows that men and women differ not only in gender but also in their characteristics.

The overall differential in probabilities of being in sector *m* is given by $\mu_m^W = s_m^W / s_m^M$. If the ratio is equal to one, there is no overall differential in employment type allocation probabilities between women and men of being in sector *m*. A measure of the differential in probabilities of women and men being in a sector *m*, which is due to differences in their characteristics, is given by $\delta_m^W = \mu_m^W / \lambda_m^W$. Table 3 shows the estimates of gender, characteristics, and overall differential in probabilities of being in the five sectors.

Table 3: Estimates of Gender, Attributes and Overall Differential in Employment Allocation Probabilities of Women Relative to Men

Sector	Gender differential 1-λ	Attributes differential 1-δ	Overall differential 1-μ
Unpaid workers	-32	-85	-144
Agriculture	-21	-15	-39
Public sector	34	55	66
Private sector	66	21	73
Informal sector	-1	33	32

Source: Estimation results

A negative sign indicates differential against men to be in the particular sector. The overall differential is not exactly the sum of the component differentials. The difference is the interaction between the component differentials. Table 3 shows that, with respect to unpaid work, gender and characteristics differentials reinforced each other to yield a large overall differential for women to be in unpaid work. The characteristics differential explains a large part of it. Similarly, the overall differential in agriculture is tilted towards women.

In the other sectors, the overall differential is in favor of men. For example, in the public sector, women faced a gender disadvantage of 34 per cent, and a characteristics disadvantage of 55 per cent. This leads to an overall disadvantage of 66 per cent. In contrast, women faced in the private sector, much larger gender disadvantage (66 per cent) while the characteristics disadvantage was smaller (21 per cent). The gender advantage for women to be in the informal sector is very small (1 per cent) and it is outweighed by the characteristics disadvantage; this implied an overall disadvantage to be in the informal sector.

Tables 4 and 5 present multinomial logit regression estimates of the type of employment for men and women respectively. Wald tests of the hypothesis that all coefficients except intercepts associated with each of ten pairs of employment categories are zero, that is, to test if categories can be combined were computed. The chi-square statistic for every pair was significant at 0.001 level. Hence the hypothesis may be rejected for men and for women; which suggests that the five-way split is an appropriate view of employment structure. A property of the multinomial logit model is that of Independence of Irrelevant Alternatives (IIA). Hausman tests of the null hypothesis that employment types are independent are conducted and the data do not reject it. The test statistics are placed at the bottom of the Tables.

The other issue is that to obtain unique probabilities the coefficients of one of the outcomes in the multinomial logit model should be normalized to zero. In this paper the

unpaid family workers employment category is omitted. The coefficients of the other categories are interpreted as the effect of the associated explanatory variable on the log odds of the particular employment type relative to unpaid family work. Any sector can be omitted to serve as the base category; while coefficient estimates have different interpretation depending on base category, the probabilities remain the same. An alternative way to interpret the coefficients of the multinomial logit is to calculate the relative odds ratios (exp(b)) where b is the estimate associated with a given covariate. A third way to interpret the coefficients is to calculate partial changes in probabilities (see Long, 1997). But because the model is non-linear, the coefficients cannot be interpreted as representing the partial change in predicted probability of an individual to be in a given employment type, from change in a covariate.

Table 4: Multinomial Logit Estimates of Employment Type Allocation for Men
(Unpaid Family Workers is the omitted category)

Explanatory variables	Agriculture sector	Public sector	Private sector	Informal sector
Age (years)	0.10***	0.52***	0.28***	0.25***
	(2.89)	(11.36)	(7.41)	(5.81)
Age squared	-0.0012**	-0.0062***	-0.0036***	-0.0032***
	(2.54)	(10.64)	(7.30)	(5.79)
Some primary	0.55**	1.93***	1.40***	1.34***
	(2.34)	(6.79)	(5.71)	(5.26)
Full primary	0.86***	2.89***	1.84***	1.80***
	(2.90)	(8.53)	(6.05)	(5.72)
Some secondary	0.64**	3.17***	1.65***	1.33***
	(2.04)	(9.04)	(5.16)	(3.98)
Full secondary	-0.06	3.23***	1.07***	0.95***
	(0.20)	(9.85)	(3.57)	(3.05)
Post-secondary	-0.09	4.28***	1.80***	1.26**
	(0.15)	(7.09)	(3.04)	(2.01)
University	-0.14	3.91***	1.25**	0.70
	(0.22)	(6.57)	(2.08)	(1.06)
Household head	1.84***	2.00***	1.80***	2.28***
	(7.13)	(7.27)	(6.94)	(8.03)
Married	0.34	0.81***	0.64**	0.95***
	(1.35)	(3.00)	(2.50)	(3.52)
Small child present	0.17	-0.17	-0.06	0.11
	(1.00)	(0.91)	(0.33)	(0.59)
School child present	-0.10	-0.20	-0.39**	-0.27
	(0.51)	(1.01)	(2.00)	(1.37)
Prime age adult present	0.15	0.12	-0.07	0.22
	(0.46)	(0.34)	(0.21)	(0.63)
Received transfer income	-0.09	-0.28	-0.15	-0.37**
	(0.52)	(1.61)	(0.89)	(2.09)
Household has land	0.77***	-0.21	-0.28	-0.29
	(3.25)	(0.83)	(1.18)	(1.20)
Constant	-6.31***	-13.96***	-5.82***	-6.82***
	(6.27)	(14.52)	(7.46)	(8.01)
Wald χ^2 (D.F)	2728.17 (112)			
Pseudo R^2	0.21			
Log-Likelihood	-10130.65			
Sample size			10280	

Notes: Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors adjusted for clustering on household. The sample evidence is for the null hypothesis of IIA. $\chi^2(84)$=-19.46; $\chi^2(84)$= 10.697 $\chi^2(84)$= 61.222; and $\chi^2(84)$= -3.201 with agriculture, public sector, private sector, and informal sector omitted. Regressions include fourteen dummy variables for regions and area of residence.

Table 5: Multinomial Logit Estimates of Employment Type Allocation for Women
(Unpaid Family Workers is the omitted category)

Explanatory variables	Agriculture sector	Public sector	Private sector	Informal sector
Age (years)	0.04*	0.46***	0.14***	0.24***
	(1.65)	(9.89)	(4.45)	(7.78)
Age squared	-0.0002	-0.0056***	-0.0017***	-0.0030***
	(0.58)	(8.53)	(3.89)	(7.11)
Some primary	0.52***	2.08***	0.98***	0.53***
	(3.91)	(7.85)	(5.34)	(3.32)
Full primary	0.21	2.01***	0.65***	0.37**
	(1.46)	(6.54)	(3.14)	(2.02)
Some secondary	-0.00	2.87***	0.80***	0.13
	(0.03)	(10.10)	(3.79)	(0.65)
Full secondary	0.09	4.01***	0.88***	0.43**
	(0.46)	(13.57)	(3.78)	(1.99)
Post-secondary	0.21	5.40***	1.91***	0.39
	(0.60)	(13.85)	(4.96)	(0.92)
University	0.31	4.62***	1.86***	-0.36
	(0.55)	(7.45)	(3.49)	(0.44)
Household head	1.55***	2.35***	1.66***	2.15***
	(8.21)	(10.64)	(7.65)	(10.62)
Married	0.13	-0.31*	-1.08***	-0.21
	(1.13)	(1.89)	(7.69)	(1.51)
Small child present	0.13	-0.01	0.06	0.09
	(1.13)	(0.05)	(0.40)	(0.63)
School child present	0.04	-0.19	-0.17	0.02
	(0.37)	(1.31)	(1.37)	(0.20)
Prime age adult present	0.14	0.19	-0.21	0.20
	(1.04)	(1.02)	(1.12)	(1.20)
Received transfer income	0.02	-0.08	-0.16	0.23**
	(0.17)	(0.56)	(1.25)	(1.99)
Household has land	0.97***	-0.11	-0.46***	-0.27*
	(8.17)	(0.66)	(3.12)	(1.83)
Constant	-5.06***	-13.66***	-3.51***	-5.69***
	(7.87)	(14.85)	(5.95)	(10.03)
Wald χ^2 (D.F)	3029.87 (112)			
Pseudo R^2	0.25			
Log-Likelihood	-8236.32			
Sample size			12299	

Notes: Robust z statistics in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors adjusted for clustering on household. Regression included fourteen regional dummies. The sample evidence is for the null hypothesis of IIA. $\chi^2(84)$= -72.39; $\chi^2(84)$= -8.72; $\chi^2(84)$= 20.80; and $\chi^2(84)$= 11.517 with agriculture, public sector, private sector, and informal sector omitted. Regressions include fourteen dummy variables for regions and area of residence.

The partial change in probability of a given employment type when a continuous variable x_k, changes is calculated as

$$(2) \qquad \frac{\partial \Pr(y = m \mid x)}{\partial x_k} = P(y = m \mid x)\left[\beta_{km} - \sum_{j=1}^{J} \beta_{kj} \Pr(y = j \mid x)\right]$$

and the difference in predicted probability from a discrete shift in a dummy variable from x_s to x_e is computed as

$$(3) \qquad \Delta \Pr(y = m \mid x) = \Pr(y = m \mid x, x_k = x_e) - \Pr(y = m \mid x, x_k = x_s)$$

Tables 6 reports the partial changes or differences in probabilities for men and women respectively, from changes or shifts in included regressors.

Table 4: Multinomial Logit, changes in probabilities of employment type: Men and Women

	Men				Women			
Explanatory variables	Agriculture sector	Public sector	Private sector	Informal sector	Agriculture sector	Public sector	Private sector	Informal sector
Age (years)	-0.052*	0.028***	0.019***	0.007***	-0.018*	0.007***	0.003***	0.011***
Some primary	-0.227**	0.099***	0.095***	0.042***	-0.029***	0.039***	0.017***	-0.001***
Full primary	-0.296***	0.195***	0.077***	0.035***	-0.063	0.058***	0.014***	0.005**
Some secondary	-0.336**	0.312***	0.046***	-0.012***	-0.176	0.167***	0.025***	-0.006
Full secondary	-0.422	0.447***	-0.004***	-0.014***	-0.339	0.361***	0.011***	-0.010**
Post-secondary	-0.519	0.639***	-0.049***	-0.062***	-0.619	0.688***	0.015***	-0.043
University	-0.487	0.650***	-0.082**	-0.072**	-0.465	0.510***	0.038***	-0.048
Household head	-0.010***	0.012***	-0.013***	0.043	-0.011***	0.017***	0.003***	0.044***
Married	-0.095	0.025***	0.029**	0.047***	0.074	-0.006*	-0.050***	-0.016
Small child present	0.050	-0.022	-0.031	0.003***	0.011	-0.002	-0.002	-0.002
School child present	0.053	-0.001	-0.044**	-0.009	0.012	-0.004	-0.007	0.000
Prime age adult present	0.028	0.001	-0.041	0.012	0.015	0.001	-0.013	0.004
Received transfer income	0.036	-0.011	-0.001	-0.025**	-0.004	-0.002	-0.006	0.013**
Household has land	0.252***	-0.044	-0.134	-0.071	0.196***	-0.017	-0.057***	-0.079*

Note:: * significant at 10%; ** significant at 5%; *** significant at 1%.

Additional schooling for men up to some secondary education reduces the chance of agricultural employment. For example, if a man has some primary education the chance of agricultural employment falls by 23 per cent and by 34 per cent for a man with some secondary education. On the other hand, for a woman with some primary education the chance of agriculture work falls by only 3 per cent. For men full secondary education and above has insignificant effect on the chances of working in agriculture. But for women the insignificant effect set in at full primary education level. More schooling for men raises the probability of working in the public sector. The probability of working in the public sector for a man with some primary education rises by 10 per cent, while with university education the chance rises by 65 per cent. The corresponding rises in probability for women are 4 per cent and 51 per cent. Except at post-secondary level of education, the partial effects of women education on probability of public sector employment are lower than those men education.

A man with less than full secondary education increases his chance of being in private sector employment. However, it seems that additional schooling is associated with lower partial effects. For example, a man that has full primary education raises the chance of private sector employment by 8 per cent compared to 5 per cent for a man with some secondary education. Men with full secondary education and above are less likely to be in private sector employment and the absolute partial effect rises with education. For instance, a secondary graduate reduces his chance of private sector work by 0.4 per cent, while a university graduate reduces the chance by 8 per cent. For women, some secondary education or below raises probability of private sector employment as in the case of men, but the partial effects are not large. However, in contrast to men, full secondary education and above for women raises the probability of private sector work. A woman with full secondary education raises the probability by 1 per cent and by 4 per cent if she has university education.

The informal sector seems to attract men with full primary education or less. However, the partial effect (3.5 per cent) for a man with full primary education is lower than for a man with less than full primary education (4.2 per cent). Secondary education and above

reduces the chances of informal sector work. If a man has some secondary education, the probability of informal sector work falls by 1 per cent while with university education the probability falls by 7 per cent. For women, less than full primary education significantly reduces the chance of informal sector work. But a woman with full primary education raises her chance of entry into the informal sector. On the other hand a woman with more than primary education reduces her chance of informal sector work. However, the partial effects are small and only the effect for full secondary education is statistically significant.

Some demographic factors are also important covariates of employment type allocation. The partial effect of age shows that, the older a man or a woman is the more likely he or she is to be in non-agricultural employment. For a man with average characteristics, age reduces the chance of agricultural employment by 5 per cent, while for a woman it is 2 per cent. For men, age raises the probability of public sector employment by 3 per cent compared to 2 per cent for private sector work and 1 per cent for informal sector work. The partial effects of age on the probabilities of non-agricultural employment for women are smaller than those of men. The largest effect (1 per cent) is for public sector and informal sector work.

A man that is a household head increases the probability of public sector employment compared to a man that is not a household head, and reduces the chance of agricultural employment. He is also likely to be in the informal sector although this is not significant. In addition, household headship is associated with lower chance of private sector employment. For a woman household head the chances of working outside agriculture are higher. The most noticeable result is that household headship seems to push women into the informal sector. This may suggest that with smaller chances of entering the other sectors this sector provides an outlet for those with the economic responsibility to provide at least subsistence.

A married man is more likely to work outside agriculture. The chance of a married man to be in public employment is 3 per cent higher than for a comparable but unmarried man. A married man is also likely to be in private sector or informal sector. The largest effect (5 per cent) is in informal sector work. On the other hand, a married woman is more likely to work in agriculture although the effect is not significant. However, she is significantly less likely to be either in public sector or private sector employment. In rural areas such a situation is consistent with labor allocation pattern where the husband works off-farm and the wife takes care of the shamba (farm). Pooling of resources and acquisition of new networks through marriage may aid entry into informal sector employment by the man. Networks can also be important in wage sector employment. The negative effect of marriage on wage sector employment for women may also reflect employer preferences that may not be in favor of women workers.

Turning to land and non-labor income, men in households that received some transfer income were less likely to be in the informal sector. In contrast, women from such households were more likely to be in the informal sector. Land ownership raised the chance of men and women being in agriculture by 25 per cent and 20 per cent respectively. For women, land ownership is also associated with lower chances of non-agricultural employment, especially private sector and informal sector work.

IV. EARNINGS FUNCTION ANALYSIS

The Standard Earnings Function

The analysis begins with estimation of the standard human capital relationship. The dependent variable is the natural logarithm of annual income. The incomes are in Kenya shillings. The explanatory variables are education and age (experience). Because the survey records the highest level of education, dummy variables were used for education. This allows the return to education to vary across education levels. Age is used as a proxy for work experience. It captures both an age effect and an experience effect. Table 7 presents the OLS estimates of the earnings functions for the three sectors separately for men and women and Table 8 presents F-tests of the hypothesis that all coefficients in the earnings functions are equal for pairs of regressions. The null hypothesis of equality may be rejected in most pairs. Private returns to education conditional on employment in one of the three sectors are computed. These are Mincerian returns to education since only the earnings foregone when a worker was in school are factored into the calculations. Therefore private wage returns at education levels with high direct private costs may be exaggerated.

Table 7: Separate Ordinary Least Squares Regression estimates of the Earnings Functions by Sector: Men and Women

	Men			Women		
Explanatory variables	Public sector	Private sector	Informal sector	Public sector	Private sector	Informal sector
Some primary	0.15	0.16*	0.45***	0.26	0.12	0.61***
	(1.15)	(1.88)	(2.77)	(0.89)	(1.03)	(4.57)
Full primary	0.33**	0.44***	0.53***	0.79**	0.47***	1.01***
	(2.53)	(4.95)	(3.10)	(2.50)	(3.72)	(6.74)
Some secondary	0.58***	0.54***	0.58***	0.96***	0.55***	1.07***
	(4.76)	(5.72)	(2.94)	(3.45)	(4.32)	(6.16)
Full secondary	0.68***	0.93***	0.83***	1.11***	1.04***	1.26***
	(5.87)	(10.17)	(4.55)	(4.08)	(7.62)	(6.66)
Post-secondary	0.90***	1.20***	0.77***	1.46***	1.29***	0.84*
	(7.11)	(9.56)	(2.71)	(5.39)	(4.81)	(1.76)
University	1.42***	2.00***	2.87***	2.11***	2.65***	2.55***
	(10.74)	(6.44)	(4.99)	(6.97)	(6.49)	(5.11)
Age (years)	0.07***	0.12***	0.09***	0.08*	0.12***	0.08***
	(3.00)	(7.61)	(3.10)	(1.76)	(6.25)	(2.82)
Age squared	-0.00**	-0.00***	-0.00**	-0.00	-0.00***	-0.00**
	(2.44)	(6.60)	(2.31)	(1.11)	(5.13)	(1.99)
Constant	8.44***	7.20***	7.83***	7.48***	6.83***	7.25***
	(19.50)	(26.47)	(15.94)	(9.80)	(22.53)	(14.45)
Sample size	1284	2036	1047	513	664	875
Adj R^2	0.15	0.18	0.07	0.23	0.26	0.12
Average RORE						
Full primary	0.06	0.08	0.10	0.17	0.09	0.25
Full secondary	0.10	0.16	0.09	0.09	0.19	0.07
University	0.37	0.64	2.23	0.57	1.33	0.88

Notes: The dependent variable is log annual income. Robust t-values within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors adjusted for clustering on household .RORE stands for rate of return to education. It is assumed full primary=7 years; full secondary=4 years; and university=3 years. Regressions include fourteen dummy variables for regions and area of residence.

For men and women the returns to university education are the highest in all sectors. For men, returns to education in the private sector are larger than in the public sector. In the informal sector, the returns to primary and secondary education are identical. While primary education for women in public and informal sectors seems to have larger returns than secondary

education, in the private sector the returns to women education are convex. As in the summary of returns to education by Psacharopoulos (1994) returns to women education are higher than returns to male education in the formal sectors.

Table 8: Tests of Equality in Earnings Structures

Pair	F(D.F.) for all coefficients	F(D.F) for education variables
All men vs. all women	11.74 (9, 4910)	3.59 (6, 4910)
Men vs. women public sector	3.59 (9, 1557)	2.07 (6, 1557)
Men vs. women private sector	7.16 (9, 2288)	0.55 (6, 2288)*
Men vs. women informal sector	6.39(9, 1643)	1.65 (6, 1643)*
Men public vs. men private	9.10 (9, 3065)	2.65(6, 3065)
Men public vs. men informal	1.49 (9, 2234)*	1.79 (6, 2234)
Men private vs. men informal	8 (9, 2842)	2.42(6, 2842)
Women public vs. women private	3.98 (9, 1036)	1.39 (6, 1036)*
Women public vs. women informal	0.65 (9, 1285)*	0.87 (6, 1285)*
Women private vs. women informal	3.79 (9, 1379)	2.34 (6,1379)

* indicates insignificant difference. D.F. is degrees of freedom.

The Selectivity-corrected Earnings functions

The returns to education derived from the earnings function estimates of the previous section were obtained using ordinary least squares (OLS) and samples of workers in the specific sectors. It is assumed that workers are randomly assigned to employment types. If this is not the case, there is potential for sample selectivity bias in the coefficients of the earnings function. The coefficients, not only reflect the effect of the explanatory variables on earnings, but also the effect of the explanatory variables on the probability of entry into those employment types. There may be unobserved factors that affect both the type of employment and earnings, such that the random errors in the earnings function are correlated with the random errors in the process that determines employment type. The question is: Does failure to control for the potential bias affect estimates of returns to education significantly?

To check for the presence of sample selectivity bias, a two-step approach that Lee (1983) suggested is used (see also Lee and Trost, 1984). The first step is to estimate the multinomial logit model of job attainment by maximum likelihood and to use the results to predict the employment type of each individual in the sample. The second step is to calculate additional variables called the inverse Mill's ratios (or selectivity-correction terms) for each individual and to include them in the earnings regression. This would correct for sample selectivity bias. The standard earnings function estimates can be compared with the selectivity bias corrected earnings function estimates to assess whether and how selection into employment types would affect the returns to characteristics, for example education.

The extended earnings function with selection variable is

$$(4) \qquad \ln w_{ij} = \alpha_{0j} + \sum \alpha_k s_{ik} + \alpha'_{2j} X_i + \theta_{ij} \lambda_{ij} + u_{ij}, j = 1,2,.....M ; u_i \sim N(0, \sigma^2)$$

where w_{ij} is earnings of individual i in sector j, with s_{ik} the education level, X_i a vector of regressors, λ_{ij} the selectivity-correction term, and u_{ij} a random variable.[1] The coefficient on λ_{ij} is the covariance between the error term in the earnings equation and the error term in the multinomial logit job attainment equation. It measures the impact of non-random assignment

[1] $\lambda_j = \phi(H_j)/\Phi(H)$, where $H_j = \Phi^{-1}(P_j)$. ϕ is the standard normal density and Φ is the normal cdf

into employment types, and the sign indicates the nature of selectivity bias. If it is negative, it means that the unobserved factors that increase the probability of a worker being in the particular employment type also lead to lower earnings. If significant, sector allocation is non-random and including the selectivity term in the earnings equation corrects for the potential sample selectivity bias.

The two-step approach requires variable(s) to identify the employment allocation process. Schultz (1990) suggests that household assets and non-labor income can serve this purpose. In this paper, dummy variables for land ownership and for receipt of transfer income are used. The two variables are expected to influence workers' reservation wages. In addition, dummy variables for presence of children of various ages are included. These may proxy the costs of taking care of these household members. The use of demographic variables to achieve identification can be found in previous studies (e.g. Glick and Sahn, 1997; Appleton et al, 1999; Krishnan et al, 1997; and Kabubo, 2000). Tables 9 and 10 present OLS estimates of the extended earnings function with and without correction for selectivity bias for men and women respectively.

Table 9: Ordinary Least Squares and Selectivity-bias corrected Regression Estimates of the Earnings Function by sector: Men

Explanatory variable	Public sector		Private sector		Informal sector	
	OLS	Selection	OLS	Selection	OLS	Selection
Some primary	0.16	0.10	0.14	0.14	0.38**	0.37**
	(1.27)	(0.60)	(1.62)	(1.52)	(2.17)	(2.04)
Full primary	0.35***	0.25	0.37***	0.38***	0.42**	0.41**
	(2.67)	(1.12)	(4.28)	(3.97)	(2.30)	(2.17)
Some secondary	0.59***	0.46*	0.43***	0.43***	0.41**	0.41**
	(4.80)	(1.79)	(4.57)	(4.23)	(2.01)	(2.01)
Full secondary	0.69***	0.52	0.75***	0.76***	0.60***	0.60***
	(5.80)	(1.61)	(8.25)	(8.14)	(3.10)	(3.09)
Post-secondary	0.93***	0.72*	1.02***	1.02***	0.59**	0.60**
	(7.30)	(1.82)	(8.36)	(8.30)	(2.11)	(2.14)
University	1.39***	1.18***	1.72***	1.72***	2.27***	2.29***
	(10.30)	(2.97)	(5.86)	(5.86)	(4.19)	(4.22)
Age (years)	0.05**	0.02	0.08***	0.08***	0.05*	0.05
	(2.00)	(0.54)	(4.98)	(4.50)	(1.67)	(1.58)
Age squared	-0.00	-0.00	-0.00***	-0.00***	-0.00	-0.00
	(1.50)	(0.34)	(4.47)	(3.98)	(1.14)	(1.05)
Married	0.18*	0.16*	0.26***	0.26***	0.36**	0.34**
	(1.84)	(1.73)	(3.95)	(3.94)	(2.36)	(2.00)
Inverse Mills Ratio		-0.15		0.03		-0.06
		(0.58)		(0.18)		(0.20)
Constant	8.95***	9.73***	8.27***	8.22***	8.87***	8.99***
	(20.50)	(6.64)	(29.53)	(20.20)	(15.86)	(10.95)
Sample size	1284	1284	2036	2036	1047	1047
Adj R^2	0.18	0.18	0.24	0.24	0.13	0.13
Average RORE						
Full primary	0.06	0.04	0.06	0.07	0.07	0.07
Full secondary	0.10	0.08	0.12	0.12	0.05	0.05
University	0.34	0.31	0.55	0.54	1.44	1.47

Notes: The dependent variable is log annual income. Robust t-values within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors adjusted for clustering on household. RORE stands for rate of return to education. It is assumed full primary=7 years; full secondary=4 years; and university=3 years. Regressions include fourteen dummy variables for regions and area of residence.

Table 10: Ordinary Least Squares and Selectivity-bias Corrected Regression Estimates of the Earnings functions by Sector: Women

Explanatory variables	Public sector		Private sector		Informal sector	
	OLS	Selection	OLS	Selection	OLS	Selection
Some primary	0.16	-0.21	0.05	0.04	0.37***	0.34***
	(0.58)	(0.70)	(0.39)	(0.37)	(2.86)	(2.65)
Full primary	0.61*	0.22	0.36***	0.35***	0.68***	0.66***
	(1.96)	(0.66)	(2.66)	(2.62)	(4.50)	(4.38)
Some secondary	0.80***	0.05	0.42***	0.42***	0.70***	0.66***
	(2.84)	(0.16)	(3.29)	(3.20)	(3.99)	(3.69)
Full secondary	1.02***	-0.02	0.84***	0.84***	0.83***	0.79***
	(3.73)	(0.07)	(6.23)	(6.23)	(4.54)	(4.16)
Post-secondary	1.28***	-0.12	0.89***	0.89***	0.42	0.29
	(4.73)	(0.27)	(3.41)	(3.42)	(0.92)	(0.61)
University	1.85***	0.65	2.29***	2.27***	2.06***	1.90***
	(6.21)	(1.58)	(6.01)	(5.99)	(6.37)	(5.30)
Age (years)	0.09**	-0.02	0.12***	0.12***	0.09***	0.11***
	(2.09)	(0.39)	(6.03)	(5.66)	(3.54)	(3.72)
Age squared	-0.00	0.00	-0.00***	-0.00***	-0.00***	-0.00***
	(1.44)	(0.67)	(5.18)	(4.88)	(2.68)	(2.99)
Married	0.05	0.21**	0.06	0.07	-0.13	-0.17*
	(0.57)	(2.03)	(0.75)	(0.63)	(1.42)	(1.77)
Inverse Mills' Ratio		-0.72***		-0.02		0.21
		(3.96)		(0.12)		(1.04)
Constant	7.74***	11.78***	7.34***	7.37***	7.92***	7.42***
	(10.01)	(9.02)	(23.03)	(16.43)	(16.87)	(11.40)
Sample size	513	513	664	664	875	875
Adj R^2	0.30	0.32	0.33	0.33	0.22	0.22
Average RORE						
Full primary	0.12	0.04	0.06	0.06	0.14	0.13
Full secondary	0.13	-0.01	0.15	0.16	0.04	0.03
University	0.43	0.32	1.09	1.06	0.81	0.68

Notes: The dependent variable is log annual income. Robust t-values within parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors adjusted for clustering on household. RORE stands for rate of return to education. It is assumed full primary=7 years; full secondary=4 years; and university=3 years. Regressions include fourteen dummy variables for regions and area of residence.

Results for the male sample indicate that the selection term effect is negative and insignificant for the public and informal sectors and positive and insignificant in the private sector. Hence, the earnings of a man with average characteristics in either sector would not differ significantly from the earnings of a man randomly selected into the sectors. The same result turns up for women, except in the public sector where the selection term effect is negative and significant. This implies that the earnings of a woman with average characteristics in the public sector are lower than the earnings of woman that would be drawn randomly into the sector.

Controlling for sample selection lowers the size of parameter estimates of education dummies for the public sector workers. While the estimates at higher education levels remain significant, those at lower levels become insignificant. For women, the inclusion of the selection term not only lowers the size of education coefficients, but also makes them insignificant. In the private sector, the education coefficients are almost the same as in the uncorrected earnings function for men and for women. Similarly in the informal sector no change is noticeable at primary and secondary levels. The number of cases with university education in the informal is small to make a firm statement.

When the OLS estimates in Tables 9 and 10 are compared with those in Table 7, it turns out that controlling for regional dummies and marital status in the former regressions makes the education coefficients to stay the same or to decline. One exception is the return to

secondary education for women; it increases. Another observation is that even after extending the earnings function, returns to university education are the highest in all the sectors. Second, in public and private sector, most estimates indicate that returns to women education exceed returns to male education. The relatively high return to education especially in the formal wage sectors may reflect scarcity rent for those that manage to access the sectors. In the informal sector, return to primary education for a women is substantial and for both men and women returns to secondary school in this sector are low. A study on Uganda (Appleton, 2002) uses data for 1992 and 2000 to estimate returns to education. The 1992 survey is close to the 1994 survey in Kenya used here. The Mincerian estimates from the Uganda study are 7 per cent, 8 per cent and 18 per cent for primary education, secondary education, and university education respectively. It seems that returns to education were higher in Kenya around that time. But the estimates for 2000 indicate that returns to education in Uganda have increased to 17 per cent, 11 per cent, and 23 per cent respectively.

A comparison across sectors suggests that for both men and women the returns to full secondary and university education are highest in private sector. The men with university education in the informal sector also earn substantial returns. But results from the sector allocation model suggested that high level of education discourages men and women from entering the informal sector. For women, primary education is rewarded most in the informal sector, while for men the returns appear uniform across sectors. On the other hand, for both men and women, secondary education is rewarded most in the private wage sector. In general reward to university education is high in all sectors.

The results suggest that for men, the monetary reward to education in the private sector exceeds that in the public sector. In contrast, Glick and Sahn (1997) find that low educated men received higher returns to education in the public sector while the private sector rewards higher education more in Conakry, Guinea. In Ethiopia's urban labor market, Krishnan, Selassie, and Dercon (1997) find higher returns to education in the public sector than in private sector for men in 1994. But the returns to women's education are higher in the private sector at all schooling levels. In Conakry, Guinea, returns to women's' schooling in the public sector are substantial and are almost double those of men in the same sector.

The highest return to primary education is in the informal sector, while the return to secondary education in this sector is the lowest across the three sectors. Nielsen and Nielsen (2001) also find that in urban areas of Zambia, the return to primary education in the informal sector exceeds the return in the formal sector. The higher returns to education in informal sector may indicate that the activities in the sector are well suited for the realization of the productive effect of primary education. In Kenya, Neizert (1996) considers the earnings of a sample of 188 workers in Nairobi's micro-enterprises and corrects the estimates for potential bias that might emanate from the type of contract a worker holds. Education is measured in years of secondary education. For a regular worker, education raises earnings by 9 per cent. Appleton, Bigsten and Manda (1999) estimate returns to education for 254 urban self-employed persons in 1978 and 629 persons in 1986. The earnings function pooled men and women. The returns to primary education rose from 9 per cent to 12 per cent. But returns to secondary education fell from 40 per cent to 15 per cent.

An issue with respect to informal sector incomes is that the estimated functions do not control for physical capital. Information on capital use in the informal sector businesses

may not be available in household surveys like the one used in this paper. However, many informal sector businesses may have only small amounts of capital. Hence the earnings are mostly returns to labor. Also, capital expenditure maybe correlated with education (Glick and Sahn, 1997) if more educated informal sector workers spend more on capital than their less educated counterparts. Since the aim is to estimate the total return to education, capital and other inputs correlated with education may be excluded to avoid underestimating the education effect.

A comparison of the return to primary education in the informal sector estimated in this paper (recall 1994 data is used) with that in 1986 cited above suggests hat for men it is likely to have declined while for women the estimate is about the same. On the other hand, the return to secondary education is much lower suggesting that they have declined overtime. The low returns to secondary education may be one explanation of the negative effect of secondary education on entry into the informal sector. On the other hand, the relatively higher returns to primary education may explain the positive effect of primary education on the entry into the informal sector (see Table 7). No firm conclusion with regard to the return to university education can be made because the number of cases with this level of education and in the informal sector is very small.

V. SUMMARY AND CONCLUSION

This study investigates the importance of worker's human capital among other factors on an individual's type of employment and earnings. It uses a national survey of households in Kenya. The results from five-way multinomial logit models suggest that, education is important in allocating workers among unpaid family work, agricultural employment, the public sector, the private sector, and the informal sector. In particular, education discourages entry into agricultural employment, and at higher levels, also discourages entry into the informal sector. This may be because education gives access to better opportunities in wage employment that are relatively secure and have stable income. Education is relatively more strongly correlated with public sector employment than with private sector employment. This may indicate that the public sector hiring criteria puts emphasis on formal education much more than the private sector does.

Decompositions of the differential in the average probabilities of employment allocation between women and men indicate that, a substantial part is explained by differences in human capital and household characteristics. These characteristics, among them human capital, seem to keep women in agriculture and unpaid family work, while men seem to dominate the public and the private wage sectors. Thus women are less likely than men to be in the wage sectors.

The impact of education on incomes is positive in the three sectors for which data on income are available, that is, the public sector, the private sector, and the informal sector. Returns to secondary education are highest in private wage sector and women have higher returns to education than men. On the other hand, while returns to secondary education in the informal sector are low, returns to primary education especially for women are substantial. The return to education from the earnings function comprises two components. There is the effect of education on access to employment and the return within employment. In the wage sector, the return to education for women may be positive and higher than that of men. But,

unless women can access wage employment the returns to education may not be realized. Education of women raises the probability of entry into wage sectors. Hence investments in education can help them access the wage sector. In the informal sector, primary education for women raises the probability of entry and has substantial returns. Investment in primary education is likely to put more women in the sector.

Other factors also affect the probability of employment allocation and earnings. Experience and lifecycle effects captured by age effects are important. Similarly, there are important household characteristics and regional factors. While land reduces chances of off-farm work for both men and women, the effect is significant only for women. The significant effect of other factors suggests that education alone would not be enough to influence employment patterns. Nevertheless, it is a key factor.

VI. REFRENCES

Andersson, P. (1993), Labor market structure in a controlled economy: The case of Zambia. *Ekonomiska Studier* no. 37, Department of Economics, University of Gothenburg.

Appleton, S., Collier, Paul, and Horsnell, P. (1990), Gender, education, and employment in Cote d'Ivoire. Social dimensions working paper no. 8, The World Bank.

Appleton, S., Hoddinott, J. and Krishnan, P. (1999), The gender wage gap in three African countries. *Economic Development and Cultural Change*, Vol. 47. 2

Appleton, S., Bigsten, A. and Kulundu, D. (1999), Educational expansion and economic decline. Returns to education in Kenya 1978-1995, *CSAE Working paper* 99-6, Center for the Study of African Economies, Oxford University.

Becker, G.S. and Chiswick, B.R. (1966), Education and the distribution of earnings, *American Economic Review*, 56, 358-69.

Bigsten, A. (1984), *Education and income determination in Kenya*. Gower Publishing Company Limited. Gower House, Croft Road, Aldershot, England

Borooah V. (2001), *Logit and Probit: Ordered and Multinomial Models*, Sage publications, Thousand Oaks, California

Demery, L. and Grootaert, C. (1993), "Priority analysis" in Demery, L.; Ferroni, M.; Grootaert, C; with Wong-Valle, J. eds (1993), *Understanding the social effects of policy reform*. The World Bank, Washington D.C.

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

Glick, P. and Sahn, D. (1997), Gender and education impacts on employment and earnings in West Africa: Evidence from Guinea, *Economic Development and Cultural Change*, Vol. 45.

Government of Kenya (2001), Economic survey 2001

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

Kabubo, J. (2000), Male-Female wage differences and labor market outcomes in Kenya. *African Journal of Economic Policy*, vol.7, Number 2.

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

Krishnan, P., Sellasie, T., and Dercon, S. (1998), The urban labor market during structural adjustment: Ethiopia 1990-1997. *CSAE WPS/98-99*. Centre for the Study of African Economies, Oxford University.

Lee L. (1983), Generalized econometric models with selectivity" *Econometrica,* Volume 51, Issue 2 (March, 1983), 507-512

Lee, L. and Trost, R. (1984), Technical training and earnings: A polychotomous choice model with selectivity. *The Review of Economics and Statistics*, Volume 66, Issue 1 (Feb., 1984), 151-156.

Lindauer, D. and Sabot, R. (1983), The public-private wage differential in a poor urban economy. *Journal of Development Economics* 12, no. 3, pp137-152.

Long, J. (1997), *Regression models for categorical and limited dependent variables*, Sage Publication Thousand Oaks, California.

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

Neizert, M. (1996), Human capital in the hot sun. *Economic Development and Cultural Change.*

Nielsen, S., and Westergård-Nielsen, N. (2001), Returns to schooling in less developed countries: New evidence from Zambia. *Economic Development and Cultural Change*, Jan 2001; Vol. 49, Iss. 2

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

Psacharopoulos, G. (1994), Returns to investments in education: A global update. *World Development*, 22(9): 1325-1343

Schultz, T.P. (1990), Testing the neoclassical model of labor supply and fertility, *Journal of Human Resources*, 25: 599-634

Teal, F. (2002), Education, incomes, poverty and inequality in Ghana in the 1990s *CSAE WPS/2001-21*,Center for the study of African Economies, Oxford

Thomas, M. and Vallee, L. (1996), Labor market segmentation in Cameroonian manufacturing, *Journal of Development Studies*, vol. 32, No. 6, pp 876-898.

Van der Gaag, J. and Vijverberg, W. (1988), A switching regression model for wage determinants in the public and private sectors of a developing country. *The Review of Economics and Statistics*, Volume 70, Issue 2 (May, 1988), 244-252.

Van der Gaag, J., Stelcner, M. and Vijverberg, W. (1989), Wage differentials and moonlighting by civil servants: Evidence from Cote d' Ivoire and Peru, *The World Bank Economic Review* vol. 3, no. 1, pp 67-95.

Vijverberg, W. (1993), Educational investments and returns for women and men in Cote d'Ivore. *The Journal of Human Resources*, Issue 4, Special Issue: Symposium on Investments in Women's Human Capital and Development (Autumn, 1993), 933-974

Education and Household Earned Income in Kenya*

By

Anthony Wambugu
Department of Economics
Göteborg University
Box 640
SE-405 30 Göteborg
Sweden

2002d

Abstract: A survey of rural and urban households in Kenya is analyzed to explore the impact of education on household economic activity combinations and earnings. It considers three productive activities; farming, wage work, and own family business. The results suggest that, as education increases in the household from primary to secondary level, the tendency is to generate income from more than one economic activity. At higher levels of education, households tend to generate income from wage employment alone. The results also suggest that education raises total household income in rural and urban areas. Income function estimates for separate activities suggest that the wage income gain from education is substantial. And there is farm income and own business income gain from education also. This is encouraging because many Kenyans depend on incomes from farming and self-employment.

* I thank Arne Bigsten for comments and suggestions and for access to the data.

I. Introduction

Education is emphasized as a key element in the process of development.[1] Recent discussions on how to reduce poverty in less developed countries also emphasize the role of education (e.g World Bank, 2000). In Kenya, Mwabu et. al. (2000) and Geda et. al (2001) find that education is the major factor that determines whether a household is poor or not. However, the processes that underlie the aggregate finding require investigation. For households, education can improve the chances to access high return economic activities or to raise earnings within economic activities. This paper uses a survey of households in rural and urban Kenya to explore empirical evidence on the impact of education on household economic activity combinations and the impact of education on household income from farming, wage work, and own business.

A starting point to consider how education correlates with household activity mix and earnings is the dual economy model of the development process. The model identified two sectors; a rural subsistence sector and an urban industrial sector. The rural subsistence sector was assumed to hold surplus labor such that economic transformation would be accompanied by rural to urban labor migration, without changing the marginal product in the rural subsistence sector or urban wages.[2] However, the nature of labor transfer, and hence economic change, may take other forms. First, rural to urban labor migration may be circular (temporary) (see Bigsten, 1996 for study on Kenya). Second, expansion of off-farm economic activities permit households to allocate labor to more than one activity within the rural economy, without entering rural to urban migration. Reardon (1997) emphasizes the importance of off-farm activities in absorbing surplus labor in rural areas. Third, as noted by Ellis (1998) holding of multiple economic activities is not confined to rural households. Households in the urban economy may also allocate labor to multiple activities. For instance, Bigsten and Kayiizi-Mugerwa (1996) find that households in Kampala, the capital of Uganda diversified income sources during the country's economic decline in the 1970s and 1980s. The question is whether education is correlated with household activity combination behavior.

Kenya is an appropriate place to examine the impact of education on household activity combinations and earnings. First, many households engage in more than one economic activity. For example, Bigsten (1985) showed that smallholder farmers in Central Kenya derived income from farm and off-farm work and education was one factor in changing the structure of household incomes. For the same region, Julin (1993) reported that smallholder households divided hours of work between farm and off-farm work, and education significantly influenced hours of work to each of the activities. Second, education expansion has been rapid (Bigsten, 1984; Knight and Sabot, 1990).

Four paths of research into the impact of education on earnings can be identified. The most common path focuses on wage earnings mostly in urban labor markets; overviews of rates of return are provided by Psacharopoulos (1994) for all world regions and Appleton, Hoddinot, and McKinnon, (1996) for SSA countries. The main finding is that more educated workers receive higher wages. Recent studies on SSA countries (e.g. Bigsten et. al.2000; Mwabu and Shultz, 2000; Nielsen and Nielsen, 2001; and Jones, 2001) find this also.

[1]Education is one dimension of human capital. Others include better health and nutrition. Increased attention has been given to investments in human capital following the work of Schultz T.W. (1961). Schultz T.P., (1988), Strauss and Thomas (1995) and Appleton (2000) survey the literature.

[2] Harris and Todaro (1970) show that the constant urban wage implied in the traditional dual economy model might not hold. Ranis and Fei (1961) introduced the idea that growth may take place in the agricultural sector, while Corden and Findlay (1975) extend the model to include capital mobility induced by differential returns to capital.

A second path of research focuses on the effect of education on farm earnings. Comparatively less empirical evidence from this path is available for SSA agriculture.[3] Two surveys of developing countries' studies (Lockheed et.al., 1980 and Phillips, 1994) conclude that the evidence appears to support the Schultz hypothesis (Schultz, 1975) that education enhances farm efficiency in states of disequilibria, i.e. where farmers are taking up new methods, crops and inputs. In Masaka district in Uganda, Bigsten and Kayiizi-Mugerwa (1995) find insignificant effect of education on farm output. They argue that in a declining economy, the state Uganda was prior to the survey, there may be no productivity gain from education.

Other studies on African countries from 1980 onwards have mixed results. In Burkina Faso, the education of household members other than the head increases farm technical efficiency, while household head's education improves allocative efficiency (Ram and Singh, 1988). In Ugandan households, Appleton and Balihuta (1996) find that primary education raises crop production. Besides, the education of neighboring farmers has favorable impact on an individual farmer's output. Education gains also vary across the country's regions. Moock (1981) studies male farm managers in Western Kenya. He finds that less than four years of education had little impact on maize output. Similarly, Bigsten (1984) does not find significant impact of education on crop and livestock production in Kenya. But Pinckney and Kimuyu (1995) find that primary education raises farm out significantly among households in two coffee growing communities, one in Murang'a district in Kenya and the other in the Kilimanjaro region of Tanzania. However, secondary education has insignificant effect. A recent study (Weir, 1999) finds that in rural Ethiopia, at least four years of education are required to realize farm productivity gains from education. Like in Uganda, the education of neighboring farmers has favorable impact on individual farmer's output.

The third research path examines whether education has income gain in family businesses. There is little empirical evidence on this activity and results also are mixed. For instance, Vijverberg (1991) finds weak positive impact of education on the profit of non-agricultural family enterprises in the Cote d' Ivoire. But in Ghana, Vijverberg (1995) finds that the education of an entrepreneur has a small positive impact on business income. Also, the rate of return was close to the rate of return to education for wage employees. Furthermore, the education of other family members has positive impact on business income.[4]

A fourth and recent path of research considers the impact of education in more than one activity to obtain a holistic view of income gains from education. For example, Jolliffe (1998) uses a survey of farm households in Ghana and finds that, while cognitive skills (a proxy for human capital) raises total household income and off-farm income, it does not raise farm incomes. Appleton (2001) uses a national survey of households in Uganda and finds that education has comparable income returns on the farm, wage employment, and self-employment. In addition, primary education reduces the probability of farming, but increases probability of self-employment. On the other hand, secondary education has strong positive impact on probability of receiving wage earnings.[5]

[3]In the USA, early studies estimated farm technical and allocative effects of education. Welch (1970) noted that education allows workers to produce more for given inputs (worker effect or technical efficiency) and education assists workers select and reallocate inputs (input selection and input allocative effects. Ram (1980) argued that education has allocative effect because it reduces the marginal cost of acquiring and using production information while raising the marginal benefit of information. Thus labor that makes use of information more intensely would have higher returns from education

[4] Moock et. al. (1990) finds positive education impacts in Peruvian self-employment enterprises.

[5]Taylor and Yunez-Naude (2000, 2001) find positive impact of household education on earnings in Mexico. Fafchamps and Quisumbing (1999) find that, in Pakistan, households with better-educated males have higher off-farm earnings and also reallocate labor from farm to off-farm work.

From the foregoing, analysis of the impact of education focuses on earnings from a single economic activity. And not many examine the impact of education on how households combine income-generating activities alongside the productive effects. This study uses data from a survey of rural and urban households in Kenya. The aim is to address three questions: First, how is household education related to the probability of a household generating income from given mix of economic activities? Second, what is the total income gain to household education? Third, what is the income gain to household education in farming, wage employment, and own business?

A short review of the relationship between education and household income is presented in Section II. The structure of economic activities and household earnings are explored in section III to indicate the importance of various economic activities. The sample households are also characterized in terms of demographics, education, assets, and gender. The impact of education on households' mix of activities is explored in Section IV. Section V explores the impact of education on total income, farm income, own business income, and wage income. Section VI concludes.

II. Education and Household Income

Households may generate earnings from farming, own business, and wage employment. Total income (y) from the three activities is represented by equation (1) while the labor constraint is represented by equation (2). Equation (2) states that labor allocation across activities should not exceed total household labor (l). In the short-run, land, capital and allocation across activities may be taken as fixed.

$$y = \sum y_i(e, x_i, l_i, d_i), \quad i = f, o, w \tag{1}$$

$$\sum l_i = l, \quad i = f, o, w \tag{2}$$

f, o, w stands for farming, own business, and wages respectively. Income from each activity is a function of education (e), non-labor inputs and household characteristics (x), labor input (l), and other factors (d) such as prices.[6]

The household is assumed to maximize total income (y) subject to resource constraint (2) with respect to labor input to obtain optimal labor allocation. This requires that labor be applied to each activity until the values of marginal products are equal across activities.[7] The optimal labor input can be expressed as:

$$l_i^* = l(e, z, d), \quad i = f, o, w \tag{3}$$

where z consists of x_o, x_f, *and* x_w. Substituting the optimal labor inputs into the earnings equations for each activity gives reduced form earnings functions represented by

$$y_i^* = y(e, z, d), \quad i = f, o, w \tag{6}$$

The total household earned income is the sum of income from farming, own business and wages. The general form of the total household income is

[6]The contribution of education in one activity is assumed not to diminish the potential to contribute to another activity (see Jolliffe, 1998, Taylor and Yaude-Nunez, 2001, Janvry and Sadoulet, 2001).

[7]It is assumed that markets function well, there is no risk, and family labor and hired labor are perfect substitutes. However, information problems and institutional barriers can create a wedge between values of marginal products across activities.

$$y = y(e, z, d) \quad (7)$$

To study the impact of education on household activity combinations, it can be assumed that the combination that has the highest utility is selected. It is assumed that the way a household combines income-generating activities is determined by education, labor force, and other household characteristics. The characteristics may help to break constraints or may push households to make particular activity combinations.

IV. Data and Sample Characteristics

The data used in the paper are from a survey of urban and rural households (National Household Welfare Monitoring and Evaluation Survey [WMSII]) in Kenya. The survey instrument, which was a structured questionnaire, was administered between June and September 1994. This is a nationally representative survey that fielded detailed questions on household members' and household socio-economic characteristics, household expenditures, incomes and economic activities. For this study, data were drawn from three files: two household-level files and one individual-level file. In each file, additional variables were generated. The number of households in the sample for analysis is 9,183 and the characteristics are in Tables 1 and 2.

Table 1: Summary statistics (dummy variables) (percent)

Variable	Total sample	Rural sample	Urban sample
Male head	76	76	80
land ownership	70	79	15
Head's Education			
No education	38	42	10
Some primary	25	26	19
Full primary	13	13	16
Some secondary	10	9	19
Full secondary	11	8	25
Post-secondary	2	2	7
University	1	0.5	4
Most educated member			
No education	41	42	31
Some primary	24	25	17
Full primary	15	15	14
Some secondary	9	8	14
Full secondary	10	8	16
Post-secondary	2	1	5
University	1	0.4	2
Region/Area of residence			
Nairobi	3	-	19
Central	16	17	15
Coast	9	7	17
Eastern	17	17	14
North eastern	7	8	2
Nyanza	16	16	16
Rift Valley	24	26	16
Western	8	9	-
Rural residence	86		

Source: Calculations based on survey.

Male-headed households constitute over three-quarters of the households in urban and rural areas.[8] The average age of household members between 15 and 65 years and not actively studying is 34 years; it is higher in rural households (35 years on average), than in urban households (31 years). On average the household labor force consists of two adults; at least half

[8]The household head in the survey is the person whose authority in key decisions is acknowledged by household members.

are women. Seventy-nine per cent of rural households own land compared to 15 per cent of urban households. The median land size, measured in acres, is two acres in rural areas while in urban area land sizes are quite small.

With regard to education; thirty-eight per cent of household heads have no education; the proportion is higher in rural areas (42 per cent), than in urban areas (10 per cent). One-quarter of household heads have not completed primary education, with a higher proportion in rural areas (26 per cent) than urban areas (19 per cent). Eleven per cent of household heads have full secondary education; the proportion in urban households (25 per cent) is three times that for rural households. At higher education (post-secondary and university), only 3 per cent of household heads have reached this level; the proportion for urban households is 11 per cent compared to 2 per cent for rural households. In sixty-five per cent of the households, the highest education level among other members of the labor force is below full primary; the proportion ranges from 48 percent in urban areas to 67 percent in rural areas. Urban and rural areas seem to be at par with regard to full primary education; but a larger percentage of urban households have a worker with secondary education. To summarize, the statistics indicate higher education attainment in urban households than in rural households.

Total earned income is divided into three components: wages, farm income, and own business income. To obtain farm earnings, the first step is to calculate value of marketed output as the sum of value of crops, livestock and livestock products marketed in the previous twelve months. Then, because many households consume part of own production, value of crops, livestock and livestock products consumed at home is added to the value of marketed output, to arrive at value of total farm output. From the value of total farm output, value of livestock purchased, income from sale of land, and land rental income, and expenditures on seed, fertilizer, pesticides, leasing land, hired labor, and hired animal labor are subtracted. Summary statistics for household incomes are reported in Table 2.

Table 2: Summary statistics (continuous variables). Household income (Kenya shillings '000)

	Total			Rural			Urban		
Variable	Mean	Med.	SD	Mean	Med.	SD	Mean	Med.	SD
Labor force	2.28	2.00	1.18	2.31	2.00	1.19	2.10	2.00	1.12
Share of women(%)	57	50		58	50		51	50	
Average age	34.27	32.50	8.83	34.79	33.00	9.00	31.19	30	7.00
Land holding(acres)	3.99	2.00	33.48	4.53	2.00	35.95	0.76	0	8.71
Total household income	75296	40000	185659	66543	37080	131965	128047	59000	366177
Wages income	44978	18320	164839	35252	15200	110977	103599	43200	335801
Net farm income	20518	6680	50899	23245	8653	52909	4085	0.00	31926
Net business income	9799	0.00	39085	80445	0.00	31189	20373	0.00	68880
farm income, log	7.06	8.81	4.16	7.93	9.07	3.56	1.83	0.00	3.61
wage income, log	9.60	9.82	1.77	9.42	9.63	1.78	10.69	10.67	1.24
business income, log	3.06	0.00	4.52	2.98	0.00	4.43	3.55	0.00	4.96
household income, log	10.52	10.60	1.20	10.44	10.52	1.20	10.99	10.99	1.11
Wage income share	58	56		54	52		82	100	
Farm income share	32	21		37	29		4	0	
Business income share	10	0		9	0		14	0	

Note: summary statistics are unconditional, i.e. for all households with and without earnings from an economic activities

On average farm earnings constituted 37 per cent of total earnings in rural areas compared to 4 per cent in urban areas. Wage employment is the most important source of earnings. In urban areas, 80 per cent of earnings are generated from this source on average. Even in rural areas, where farming is assumed to be the main activity, wage earnings constitute at least 50 per cent of total household earnings. This seems to be in line with past work on Kenya. For example, Bigsten (1985) and Bevan et. al. (1989) reported that in Central Kenya wage income constituted 28 per cent and 37 per cent of total household earnings in 1974/75 and 1982 respectively. Wage

income remains a major source of income and appears to have become even more important for rural households. The measure of own business earnings used in this study is arrived at after expenditures on materials, labor and leasing of equipment and land are subtracted from the gross own business earnings. This source of income seems to play a much smaller role in income generation; rural households derived 9 per cent of total earnings from own business while in urban areas it was higher (14 per cent).

Table 3 shows the distribution of households over seven activity combinations. The categories are based on gross earnings in farming, own business and wage work. Forty-nine per cent of the sample households combined farming with wage work; the proportion is higher in rural areas (55 per cent), than in urban areas (13 per cent). A quarter of sample households engaged in all the three activities; the proportion in rural areas (28 per cent) is almost four times that in urban areas. Seventeen per cent of the households generated wage income alone; the proportion of households in this category is larger in urban areas (53 per cent) than in rural areas (12 per cent). Another difference is that while a quarter of urban households derived some earnings from wages and own business, less than 5 per cent of rural households had this combination. The message in Table 3 is that activity combination is widespread.

Table 3: Percentage distribution of household economic activity combinations

Economic activity combinations	Total	Rural	Urban
Wage work	18(17)	13(12)	52(53)
Farming	1(1)	2(1)	0(0)
Own business	0(0)	0(0)	0(0)
Wage work and farming	49(49)	54(55)	13(13)
Wage work and own-business	7(7)	4(3)	26(26)
Farming and own business	0(0)	0(0)	0(0)
Wage work, farming, and own business	23(25)	26(28)	8(8)

Note: Figures in ( ) are for the analytic sample while figures outside the ( ) are for data as they are in the data file. Column figures may not add to 100 because a small percentage of households have no earnings in listed activities.

Household earned income is summarized in Table 4 by activity combination. Three observations are as follows. (i) Wage income constitutes the largest share of total earned income in households that combine activities. (ii) In terms of income, the most remunerative activity combination in urban areas is own business and wage employment. In rural areas it is a combination of the three activities. (iii) Households that engaged in wage employment alone in rural areas have the lowest level of total income. (iv) the large standard deviations of some income sources indicate the huge variation across households even within the same activity combination.

Table 5 displays the distribution of households over activity combinations by poverty status and location. Households are categorized into poor and non-poor status based on the poverty line calculated by the Central Bureau of Statistics form the 1994 survey. In rural areas, 30 to 55 per cent of households in each combination are absolutely poor. The highest concentration of poor households is in wage work, farming, farming and wage work. In urban areas, 23 to 38 per cent of households in each combination are poor. Although activity combination is not concentrated among poor households, in rural areas at least 30 per cent of households in each activity combination are poor. Chi-square tests of independence reject the null hypothesis that activity combination and poverty status are independent (bottom of Table)

Table 4: Distribution of incomes by activity combinations

	Total sample			Rural areas			Urban areas		
	Mean	Med	SD	Mean	Med	SD	Mean	Med	SD
Farm income	38111	14840	91389	30583	14840	39249	455940	455940	604723
Wage income	59983	24400	186627	30439	11080	82848	99332	42000	263488
Farming / wage emp									
Wage income	32334	13500	108752	31231	12800	110053	60788	43200	60541
Net farm income	29166	11400	62639	29609	11800	63109	17726	5400	47719
Household income	61500	34790	126587	60840	33896	127951	78516	56640	82517
Wage share	0.52	0.53	0.30	0.51	0.51	0.30	0.81	0.90	0.21
Farm share	0.48	0.47	0.30	0.49	0.49	0.30	0.19	0.10	0.21
Own bus/ wage emp									
Wage income	81583	36500	243164	38750	22000	63857	114081	48180	313967
Net own bus. income	47996	28800	93427	27465	18000	35507	63573	36000	117691
Household income	129579	69100	305546	66215	41760	81944	177654	92260	392345
Wage share	0.56	0.51	0.21	0.56	0.52	0.21	0.57	0.51	0.20
Own busi. share	0.44	0.49	0.21	0.44	0.48	0.21	0.43	0.49	0.20
Farming/Bus./ wages									
Wage income	52223	23700	209895	46719	22760	128060	169542	43100	790536
Net farm income	22291	10550	40343	22691	10700	40843	13765	8045	26320
Net busi. Income	26185	12000	53533	25278	12000	53007	45519	24000	60854
Household income	100699	59520	233211	94688	58629	156631	228826	87315	824536
Wage share	0.45	0.45	0.23	0.44	0.44	0.23	0.54	0.51	0.22
Farm share	0.27	0.20	0.23	0.28	0.21	0.23	0.12	0.08	0.13
Own busi share	0.28	0.26	0.19	0.28	0.26	0.19	0.34	0.36	0.21

Source: calculated from WMSII data

Table 5: Economic Activity Distribution by poverty status and location

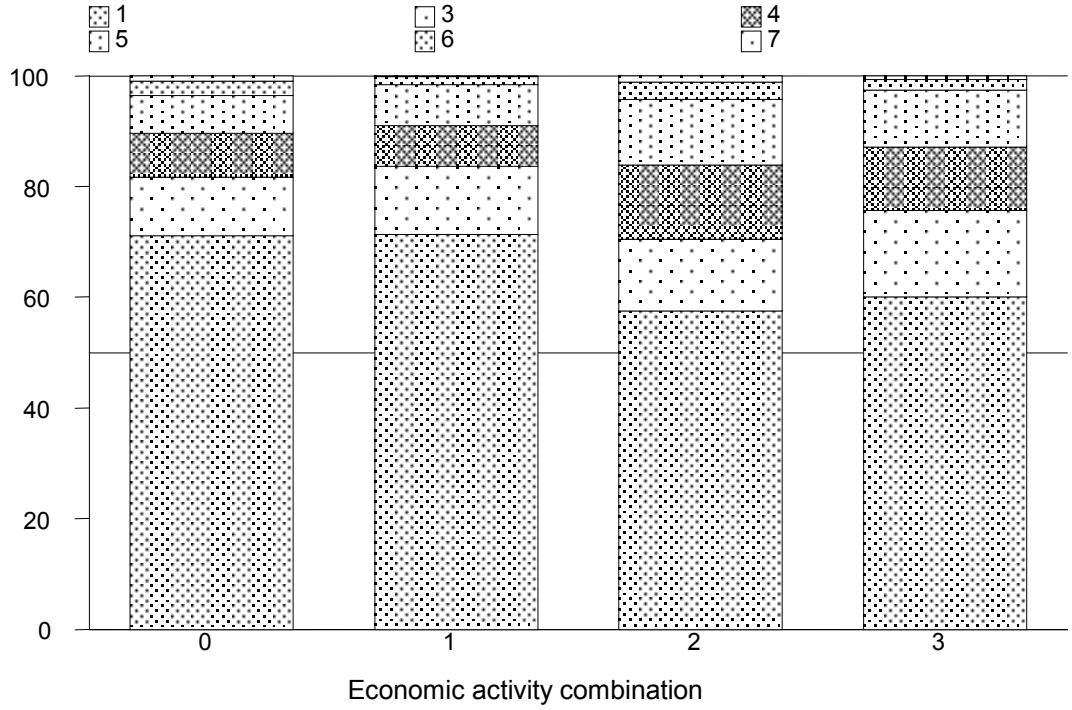
	Rural		Urban		Total	
Economic activity combination	Not poor	Poor	Not poor	Poor	Not poor	Poor
Wage work	45.36	54.64	74.38	25.62	57.80	42.20
Farming	50.45	49.55	-	-	51.33	48.67
Own business	-	-	-	-	-	-
Farming and wage work	57.14	42.86	76.92	23.08	57.87	42.13
Own business and wage work	64.37	35.63	74.71	25.29	70.25	29.75
Own business and farming	64.29	35.71			64.29	35.71
Farming, wage work and business	68.38	31.62	61.54	38.46	68.07	31.93

Pearson chi2(6) = 101.2268 for total sample; chi2(5) = 164.0398 for rural sample and chi2(5) = 10.2790 for urban sample. Absolute poverty line: Kenya shillings 978.27 in rural areas and 1,489.63 in urban areas per month per adult equivalent. The poverty line the one calculated by the Central Bureau of Statistics based on the survey data.

Activity combination seems to be a common practice, and incomes and poverty status of households in different activity combinations vary. How is education correlated with activity combinations and incomes? In the rural sample (Figure 1), at least 70 per cent of the households in each activity combination have primary education or less; with particular concentration in wage work or farming and wage work. However, at least 10 per cent of households in each activity combination have secondary education or above. Similarly in urban areas(figure 2), low education households are in all activity combinations; but they do not cluster in particular activity combination as in rural areas; at least 30 per cent of the households in each activity combination have secondary education or above. Chi-square tests reject the null hypothesis of independence between household head's education and activity combinations.[9]

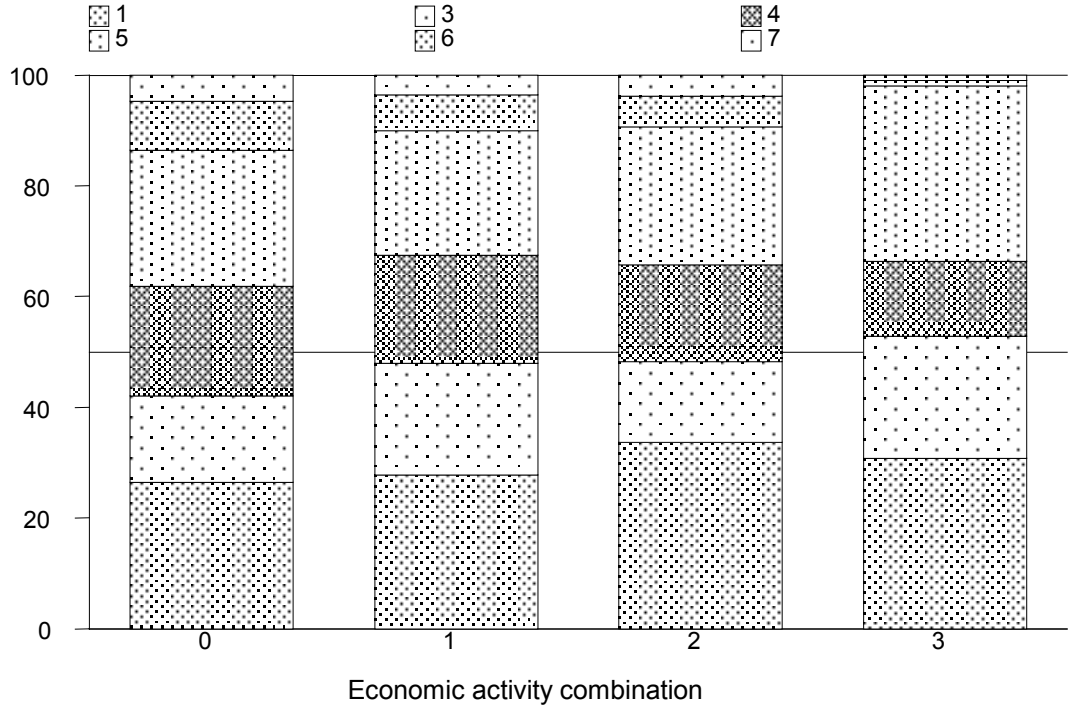
[9] $\chi^2(18) = 517.44$ for total sample, $\chi^2(18) = 255.50$ for rural sample, and $\chi^2(18) = 31.43$ for urban sample

Figure 1: Percent of rural households by head's education and activity combination



Notes. Education levels: 1=below primary education, 3=complete primary education, 4=incomplete secondary education, 5=complete primary education, 6=post-secondary education, 7=university degree. Economic activities: 0=wage employment, 1=farming and wage employment, 2=own business and wage employment, 3=farming, own business, and wage employment.

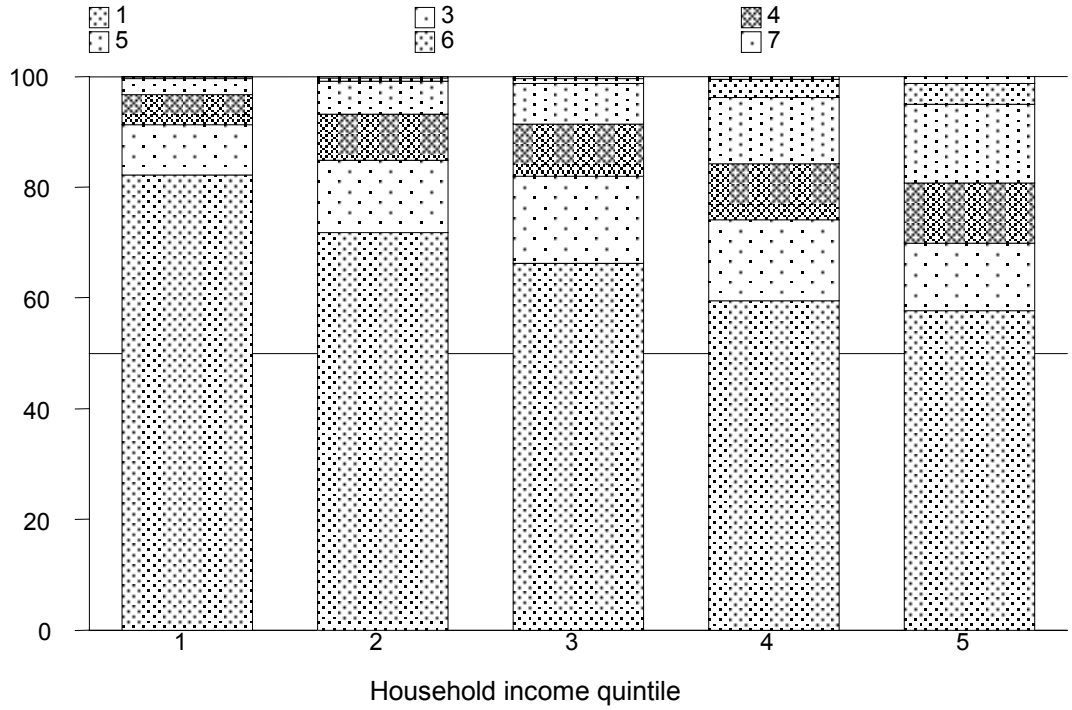
Figure 2: Percent of urban households by head's education and activity combination



The relation between education and total household earnings is depicted in Figures 3 and 4 for rural and urban areas respectively. In rural areas, around 80 per cent of households in the lowest income quintile have less than primary education; the proportion declines to around 60 per cent in the top income quintile. Conversely, the proportion with secondary education and above rises from less than 10 per cent in the lowest quintile to around 20 per cent in the top quintile. Like in

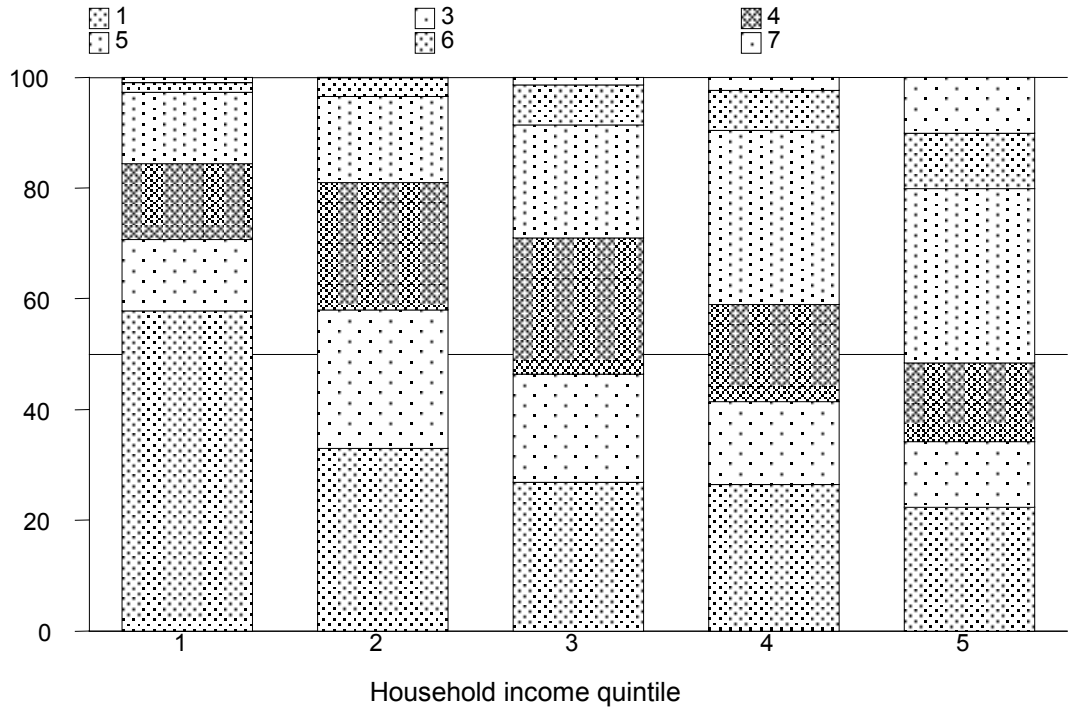
rural sample, majority (60 per cent) of urban sample in the lowest income quintile has below primary education; it declines to 20 per cent in the top quintile. In the lowest quintile, around 15 per cent have secondary education or above; increasing to over 50 per cent in the top income quintile. Chi-square tests reject the null hypothesis that education and earnings independent.[10]

Figure 3: Percent of rural households by head's education and income quintile



Notes. Education levels: 1=below primary education, 3=complete primary education, 4=incomplete secondary education, 5=complete primary education, 6=post-secondary education, 7=university degree.

Figure 4: Percent of urban households by head's education and income quintile



[10] $\chi^2(24) = 880.6$ for total sample, $\chi^2(24) = 550.4$ for rural sample, and $\chi^2(18) = 183.63$ for urban sample

To summarize, education and household activity combination may be related. Education and total earned income may also be related. Some households with high education are in lower income quintiles and some households with low education are in upper income quintiles. Similarly, every activity combination has households with low and high education. This means that the impact of education is not be uniform, and other factors may also predict earnings and activity combinations. In subsequent sections, the study turns to multivariate econometric analysis to explore the impact of education on activity combinations and household incomes.

V. Econometric specifications

The previous section explores the relationship between education and household activity combinations with the help of Table 3 and Figures 1 and 2. To explore the relation further, a reduced form multinomial logit model is specified. This helps to explore how education is correlated with the probability of household *i* holding activity combination *m, p_{im}* controlling for a number of other covariates.[11] The form of multinomial logit is

$$\Pr ob(y_i = combination_m) = P_{im} = \frac{e^{\beta_m' x_i}}{\sum_{j=1}^{J} e^{\beta_j' x_i}}, i = 1,...., N; m = 1,....., J \qquad (8)$$

where *J* indexes activity combinations, *N* is sample size, x_i is vector of explanatory variables, assumed to be exogenous.[12] β are coefficients associated with a given activity combination.

Three sets of explanatory variables are included in the multinomial logit: human capital (education); other characteristics of the labor force (size, age, gender); and geographic characteristics (region and area of residence). To measure the impact of education, six dummy variables for the highest level of household head's education, and six dummy variables for the highest education level among other labor force members are included. The second set of dummy variables is included because even where the head is not educated there may be another adult who is educated. The average age of labor force and average age squared is included as proxies for accumulated experience and household's position along the lifecycle.

Because households may be constrained in activity combinations by quantity of labor, labor force, measured as number of household members 15 to 65 years old and not in school is included. A larger labor force can widen the scope for engaging in multiple activities through reduction in indivisibilities problem that arises due to inadequate labor (Bigsten, 1996). Gender differences and household responsibilities are allowed for by adding a dummy for male-headed households. The share of women in the labor force is also included to take account of the importance of women labor in the households. The two gender variables can also reflect differences in preferences towards activity combinations. Because opportunities to combine activities are likely to differ across regions, six dummy variables for province of residence are included. The variables can control for regional differences in supply of income earning opportunities, output and inputs prices, state of infrastructure and availability of public services (e.g. health care and school facilities).

[11]The binary event of participation in a specific income activity can be modelled with a single equation probit (e.g. Bigsten and Kayiizi-Mugerwa, 1995, Appleton, 2001 and Lazlo, 2001).

[12]In the long run, fertility related variables and location are likely to be endogenous. Coulombe and McKay (1996) argue that they may be treated as exogenous in the short run. Land is excluded since it may be endogenous (Appleton, 2001)

Turning to earnings determination, a total household earnings function is specified to study the impact of education on total household earnings from wages, farming, and own business. The form of the model is:

$$y_i = w_i.\alpha + \varepsilon_i, \quad \varepsilon_i \sim N(0,\sigma^2), \tag{9}$$

where the dependent variable y_i is natural logarithm of total annual household earnings for household *i (i = 1,.....,N)*, and it is assumed to be a linear function of w_i a vector of covariates. α is the vector of coefficients associated with the covariates, and ε_i is a random error term. Separate activity-specific earnings functions are also specified to investigate the effect of education on earnings from the three economic activities from which households generate income, that is, wage work, farming, and own business. Because not every household in the sample generated income from each activity, the dependent variables have zeros recorded for these households.[13] Consistent parameter estimates can be obtained using tobit. The tobit model has been used recently in studies by Janvry and Sadoulet (2001) and Adams (2001) to analyze the determinants of several sources of earned and unearned household incomes in two less developed countries. The model is used in this paper also. Long (1997) notes that tobit avoids wastage of information.[14] In the present application it is particularly suitable for own business and farm incomes where many households have zero income. The form of the Tobit model is

$$y_{ik} = \begin{cases} y_{ik}^*, & if \quad y_{ik}^* > 0 \\ 0, & otherwise \end{cases} \tag{10}$$

where

$$y_{ik}^* = x_i.\beta + u_i, \quad u_i \sim N(0,\sigma^2)$$

with y_{ik} the log of observed annual earnings[15] of household *i* in activity *k*, y_{ik}^* the household's (latent) earnings, x_i a vector of explanatory variable, and u_i a random error term.

Among the covariates in the earnings functions are education and a quadratic in average age of the laborforce. Education is entered in form of six dummy variables for highest education of household head and six for the other most educated labor force member. Age is included to capture experience and lifecycle effects on earnings. A dummy for male-headed household and the share of women in labor force are included as proxies for gender differences in labor earnings. Also, the earnings functions include the natural log of labor force size. A larger labor force may be associated with higher earnings, if for example some household workers gain access to well-paid jobs. Because land is a key asset for Kenyan households and a factor of production in farming, it is also included in the income functions. Around twenty-one percent of the households in rural areas and 75 per cent in urban areas have no measured land. Therefore, a dummy variable for land ownership is included in the earnings function. Earnings are likely to differ across regions of the country, and therefore to control for this, dummy variables for province of residence are included in the earnings functions.

[13]In particular, 1 per cent, 24 per cent, and 68 per cent of households in the sample had zero wages, farm and own business earnings respectively in the 12 months preceding the survey.

[14]Julin (1993) and Fafchamps and Quisumbing (1999) are examples of studies that estimate Tobit for hours worked on farm and off-farm by households in Central Kenya and rural Pakistan respectively. Flood and Urban (1998) compare tobit with other models using time use data from Sweden.

[15]The constant one is added to farm, wages, and own business incomes to allow log transformation

VI. Activity Combinations and Education

This section presents estimation results of the activity combination model (8), estimated on the full, rural and urban samples respectively. The objective is to identify the activity combinations that households with different levels of education engage in. Because households' activity combinations within a cluster may be related, clustering effects are allowed for. Also, because estimation requires adequate sample size in each category, the model is estimated with four most frequent activity combinations: wage employment; farming and wage employment; wage employment and own business; and wage employment, farming and own business.[16] The maximum likelihood estimates of the multinomial logit model are reported in Table 6.

One estimation issue is that in order to estimate unique probabilities, estimates of one activity combination have to be normalized to zero. The coefficients of other activity categories are interpreted with reference to the normalized category. Also, the estimates do not show the change in probability of a household holding an activity combination, when explanatory variables change. For example, a positive coefficient would not imply that a rise in the associated variable increases p_{im}, the probability of household i falling in category m. The probability of another activity category may rise by relatively larger amount such that, p_{im} falls.

The multinomial logit puts restrictions on agents' choices; the independence of irrelevant alternatives assumption (IIA). This study on the covariates of household activity combinations is exploratory in nature. Nevertheless, the Hausman specification test of the IIA assumption was implemented. The result suggests that the null hypothesis of independent activity combinations may not be rejected (see Table 7). The test compares maximum likelihood estimates based on full sample with maximum likelihood estimates in which one combination m is dropped, while households that actually had the combination m are dropped. Under the null hypothesis the two sets of estimates should be close.

For interpretation, households with wage income only are the omitted category. The coefficients of other categories are interpreted relative to this category. The sign of a coefficient shows how risk-ratio, $p(y_i=m)/p(y_i=0)$, the ratio of probability of a household engaging in activity combination m, relative to wage employment changes when a covariate changes. Because the direction of change in $p(y_i=m)$, for a change in an explanatory variable is not clear from the sign of associated coefficient, Table 8 reports the changes in probability computed according to Long (1997). Because the change depends on values of all explanatory variables and coefficients of each category, the partial effect and the associated coefficient can have different signs. For continuous variables the partial effect is computed as

$$\frac{\partial \Pr(y=m \mid x)}{\partial x_k} = P(y=m \mid x)\left[\beta_{km} - \sum_{j=1}^{J} \beta_{kj} \Pr(y=j \mid x)\right] \tag{11}$$

For dummy variables, the difference in probability when x_k jumps from x_s to x_e is given by

$$\Delta \Pr(y=m \mid x) = \Pr(y=m \mid x, x_k = x_e) - \Pr(y=m \mid x, x_k = x_s) \tag{12}$$

[16] Sample sizes of own business only, farming only, and farming and business only households are small (1% or less of sample).

Table 6: Multinomial Logit Estimates of Household Activity Combinations: Rural and Urban Areas

Explanatory variables	Total sample			Rural sample			Urban sample		
	Farming and wage	Own business and wage	All three activities	Farming and wage	Own business and wage	All three activities	Farming and wage	Own business and wage	All three activities
Age (years)	0.13***	0.12***	0.14***	0.13***	0.07	0.13***	0.03	0.13**	0.30**
	(5.30)	(2.63)	(5.01)	(4.85)	(1.09)	(4.39)	(0.34)	(2.04)	(2.05)
Age squared	-0.0012***	-0.0017***	-0.0014***	-0.0012***	-0.0013	-0.0013***	-0.0001	-0.0017*	-0.0040*
	(3.74)	(2.75)	(3.83)	(3.50)	(1.47)	(3.39)	(0.11)	(1.87)	(1.81)
Male head	0.13	0.02	0.23*	0.21	0.14	0.30**	0.11	0.16	0.65
	(1.05)	(0.13)	(1.78)	(1.61)	(0.63)	(2.16)	(0.32)	(0.52)	(1.47)
Education of head's head									
Some primary	0.56***	0.63***	0.71***	0.65***	0.67***	0.79***	-0.23	0.24	0.04
	(4.62)	(3.44)	(5.33)	(4.90)	(2.81)	(5.52)	(0.55)	(0.79)	(0.07)
Full primary	0.40**	0.24	0.70***	0.38**	0.35	0.69***	0.27	-0.21	0.07
	(2.48)	(1.11)	(3.95)	(2.13)	(1.22)	(3.58)	(0.63)	(0.63)	(0.13)
Some secondary	0.05	0.30	0.49**	0.09	0.55*	0.58**	-0.08	-0.21	-0.54
	(0.31)	(1.39)	(2.56)	(0.42)	(1.67)	(2.57)	(0.18)	(0.68)	(0.91)
Full secondary	0.16	0.23	0.56***	0.22	0.48	0.60***	-0.10	-0.21	-0.00
	(0.96)	(1.07)	(3.07)	(1.10)	(1.59)	(2.84)	(0.23)	(0.69)	(0.01)
Post secondary	-0.52**	0.05	-0.29	-0.57**	0.24	-0.19	-0.33	-0.39	-2.36*
	(2.04)	(0.17)	(1.07)	(2.03)	(0.51)	(0.61)	(0.55)	(1.05)	(1.96)
University	-1.14***	0.38	-0.07	-1.61***	0.26	0.06	-0.25	0.06	-1.87
	(2.69)	(1.00)	(0.19)	(3.12)	(0.33)	(0.11)	(0.40)	(0.12)	(1.40)
Most educated other member									
Some primary	0.94***	0.49***	1.10***	1.03***	0.61**	1.17***	0.21	0.18	0.66
	(7.41)	(2.71)	(7.57)	(7.10)	(2.36)	(7.21)	(0.61)	(0.74)	(1.36)
Full primary	0.97***	0.48**	1.05***	1.19***	0.65**	1.23***	-0.56	0.16	0.65
	(6.67)	(2.28)	(6.40)	(7.07)	(2.24)	(6.61)	(1.50)	(0.54)	(1.20)
Some secondary	0.83***	0.43*	0.88***	1.07***	0.69*	1.09***	-0.32	0.10	0.52
	(5.20)	(1.82)	(5.07)	(5.36)	(1.81)	(5.14)	(0.84)	(0.33)	(0.98)
Full secondary	0.93***	0.68***	0.94***	1.04***	0.73**	1.01***	-0.19	0.46	0.74
	(5.56)	(2.88)	(5.09)	(5.61)	(2.12)	(4.90)	(0.41)	(1.46)	(1.49)
Post secondary	0.73**	-0.48	0.89***	1.20**	1.11	1.46***	-0.16	-1.04*	0.07
	(2.39)	(1.10)	(2.72)	(2.48)	(1.44)	(2.85)	(0.27)	(1.96)	(0.08)
University	0.45	-0.29	0.10	0.68	-0.40	0.08	-1.63	-0.47	0.80
	(1.07)	(0.48)	(0.19)	(1.32)	(0.31)	(0.15)	(1.41)	(0.70)	(0.57)
Share of women	0.93***	0.15	0.87***	1.12***	0.08	0.99***	-0.11	0.38	1.38**
	(5.49)	(0.64)	(4.65)	(5.96)	(0.21)	(4.73)	(0.25)	(1.02)	(2.23)
Log Labor force	0.35***	-0.32*	0.65***	0.38***	-0.35	0.70***	0.75**	-0.27	0.33
	(3.25)	(1.76)	(5.28)	(3.37)	(1.40)	(5.35)	(2.18)	(1.02)	(0.74)
Region/area of residence									
Central	0.16	0.51*	1.79***				-0.11	0.44	1.87***
	(0.42)	(1.77)	(3.39)				(0.17)	(1.19)	(2.80)
Coast	-0.35	0.45*	1.01*	-0.36	0.13	-0.67**	-1.43*	0.34	1.04
	(0.96)	(1.84)	(1.92)	(1.19)	(0.36)	(1.98)	(1.75)	(1.21)	(1.51)
Eastern	0.19	-0.83***	1.41***	-0.10	-2.11***	-0.52*	0.80*	-0.19	2.03***
	(0.52)	(2.66)	(2.68)	(0.35)	(4.59)	(1.71)	(1.71)	(0.49)	(3.38)
North eastern	1.22***	0.16	1.43**	1.11***	-0.51	-0.39	0.64	0.56	2.70***
	(3.01)	(0.40)	(2.25)	(3.20)	(0.92)	(0.81)	(0.87)	(0.89)	(2.78)
Nyanza	0.85**	0.97***	2.96***	0.94***	0.76*	1.46***	0.53	0.76**	2.34***
	(2.28)	(3.02)	(5.73)	(2.71)	(1.80)	(4.10)	(1.15)	(1.97)	(3.92)
Rift Valley	0.68*	-0.02	1.40***	0.55**	-0.25	-0.34	0.96**	-0.41	1.03
	(1.92)	(0.09)	(2.73)	(2.01)	(0.70)	(1.18)	(2.42)	(1.22)	(1.57)
Western	2.45***	0.86	3.64***	2.32***	0.44	1.89***			
	(5.29)	(1.28)	(6.13)	(6.03)	(0.64)	(4.78)			
Rural area	2.77***	-0.37*	2.77***						
	(13.13)	(1.83)	(11.62)						
Constant	-6.19***	-3.30***	-8.78***	-3.54***	-2.34**	-4.37***	-2.79*	-3.35***	-10.53***
	(10.73)	(4.19)	(12.10)	(6.05)	(2.06)	(6.69)	(1.90)	(2.95)	(4.09)
Sample size	9055	9055	9055	7751	7751	7751	1304	1304	1304
Pseudo R^2	0.17			0.09			0.08		
Log-Likelihood	-8885.60			-7405.97			-1390.78		

Note: * Significant at 10%; ** significant at 5%; *** significant at 1%

Table 7: Hausman test Statistics of IIA assumption for multinomial logit models. Degress of freedom within parentheses

	Total			Rural			Urban		
	Chi-sq	p-value	Evidence	Chi-sq	p-value	Evidence	Chi-sq	p-value	evidence
(a)	10.77 (50)	1	For Ho	3.47(46)	1	For Ho	-0.95(46)	1	For Ho
(b)	1.56(50)	1	For Ho	0.22(46)	1	For Ho	0.24(46)	1	For Ho
(c)	19.53(50)	1	For Ho	1.98(46)	1	For Ho	0.15(46)	1	For Ho

The omitted combination in (a) farming and wage; (b) business and wage; (c) wage, business and farming

Table 8 shows that, for households with incomplete primary education the largest inducement is to combine all the three activities. With only some primary education, a household is less likely to rely on wage income only. This may reflect that this level of education is unlikely to lead to well paid wage jobs. When education increases to full primary, household head's education reduces the probability of farming and wage work by 4 per cent and raises the probability to combine all three activities by 7.6 per cent. Full primary education of the most educated worker continues to induce the household to combine wage and farming as well as the three activities.

If household head has some secondary education, the probability to engage in the three activities goes up by 10.5 percent. It seems that secondary education of head is crucial to activity combination. Partial or full secondary education of other labor force also raises the probability of engagement in all activities, but it has relatively larger impacts on engagement in farming and wage activity. Where head has tertiary education the probability of household to combine farming and wage work is reduced. Conversely, the probability to engage in wage work is increased.[17]

To summarize, the relation between household activity combinations and education differs across education levels, and household members. The household head's education is important. At low education of household head the tendency is to combine farming, wage employment and own business. But at higher levels of education the households concentrate on wage employment. The education of other household labor force also induces households to combine three activities. But the impact of their education leans more towards farming and wage employment. This may indicate that as the household gains access to education, households spread income generation into more than one income source. Once they gain access to tertiary education, wage employment becomes the main income generation activity.

There are other factors that are important correlates of activity combinations. The older the labor force the more likely rural households are to engage in farming and wage activities or in the three activities. In urban households age of laborforce increases probability to engage in own business and wage activity or in the three activities. This may suggest that diversification of economic activities takes time. Older households may have developed networks that help them enter new income generating activities. They may also have larger savings and can therefore break liquidity constraints, such as may be present in setting up own business. Younger households seem to be mainly drawn to wage activity.

Male-headed households in rural areas are more likely to combine the three activities. This may signal scope to leave the primary income earning activity to other household members, while he supplies labor to other activities. The share of women in the labor force raises probability to engage in farming and wage activity in rural areas. It also raises probability to enage in all three activities for rural and urban households. The larger the labor force, the more unlikely it is that households in rural areas will engage in farming and wage activity. Instead, the more likely it is that households will engage in all three activities. Larger labor force may help dilute the indivisibilities problem since more workers are available to take advantage of income earning opportunities that may become available. For the urban households, larger labor force raised probability of engaging in wage and farming activity.

[17] Education may be correlated with family background, quality of schooling, or unobserved productive abilities of household labor force. The estimated education impacts may be overstated.

Table 8: Multinomial Logit Estimates, Change in Predicted Probabilities: Rural and Urban

	Rural			Urban		
Explanatory variables	Farming and Wage	Own business and Wage	Farming, wage, own business	Farming and Wage	Own business and Wage	Farming,business and wage
Age (years)	0.007***	-0.001	0.005***	-0.003	0.021**	0.013**
Male head	-0.003	-0.002	0.025**	0.003	0.020	0.026
Head's education						
Some primary	0.005***	0.001***	0.044***	-0.028	0.055	-0.001
Full primary	-0.040**	-0.002	0.076***	0.035	-0.049	0.005
Some secondary	-0.093	0.008*	0.105**	0.001	-0.031	-0.021
Full secondary	-0.065	0.004	0.086	-0.003	-0.038	0.003
Post-secondary	-0.114**	0.021	0.054	-0.015	-0.048	-0.055*
University	-0.382***	0.035	0.272	-0.019	0.032	-0.048
Other most educated						
Some primary	0.026***	-0.009**	0.053***	0.011	0.017	0.035
Full primary	0.046***	-0.010**	0.033***	-0.053	0.033	0.041
Some secondary	0.042***	-0.007*	0.025***	-0.034	0.019	0.031
Full secondary	0.050***	-0.005**	0.014***	-0.035	0.087	0.038
Post-secondary	-0.012**	-0.003	0.078***	0.006	-0.156*	0.017
University	0.139	-0.014	-0.091	-0.087	-0.077	0.081
Share of women	0.097***	-0.022	0.009***	-0.030	0.059	0.065
Labor force	-0.024**'	-0.019	0.080***	0.079**	-0.080	0.016

Notes: corresponding coefficients * significant at 10%, ** significant at 5%, and *** significant at 1%

V. Household Income and Education

This section first presents estimates the total household income function. The estimates of separate income functions for farm income, wage income, and own business income follows. Table 9 presents OLS and median estimates of total household income function. The two sets of estimates are similar and the discussion centers on median estimates since it is resistant to outliers in the dependent variable (Deaton, 1997). The incremental total income gains from education are displayed Table 10.

The results suggest that the total income return rises with education of household head. The increment to total earnings from having a head with full primary education over having a head without education is 48 per cent in rural areas and 38 per cent in urban areas.[18] Where the head has full secondary education, the gain to total earnings above that to full primary education is 39 per cent in rural areas and 35 per cent in urban households. The income advantage to households with heads educated up to university over those educated to full secondary education is 68 percent and 80 per cent in rural and urban areas respectively.

Education of the most educated other labor force member also raises total household earnings also. The income gain is lower than that due to head's education. The rural income regressions indicate that the total earnings of a household where the highest education among other labor force members is full primary education, is 15 per cent above that of a household without other members who are educated. If it is full secondary education, the income gain is 28 per cent over full primary education. Income gains to university education are negligible.

[18] Earnings gain is $100[\exp(\alpha j-\alpha i)$ αj and αi are adjacent education dummy variable coefficients.

Table 9: Ordinary Least Squares and Median Regression Estimates of Total Household Earnings Function: Rural and Urban Areas (Absolute Values of t-values Within Parentheses)

Explanatory variables	Total		Rural		Urban	
	OLS	LAD	OLS	LAD	OLS	LAD
Age (years)	0.08***	0.07***	0.07***	0.06***	0.10***	0.09***
	(8.59)	(6.15)	(7.29)	(7.16)	(4.61)	(3.03)
Age squared	-0.0009***	-0.0009***	-0.0008***	-0.0007***	-0.0010***	-0.0009**
	(7.57)	(5.23)	(6.56)	(6.12)	(3.61)	(2.06)
Head's education						
some primary	0.28***	0.25***	0.27***	0.23***	0.33**	0.27
	(6.92)	(7.44)	(6.37)	(9.64)	(2.49)	(1.45)
full primary	0.43***	0.41***	0.43***	0.39***	0.40***	0.32*
	(9.28)	(9.99)	(8.56)	(9.32)	(3.13)	(1.92)
some secondary	0.50***	0.52***	0.51***	0.51***	0.47***	0.42**
	(9.98)	(12.30)	(9.24)	(9.48)	(3.57)	(2.29)
full secondary	0.67***	0.71***	0.67***	0.72***	0.65***	0.62***
	(13.95)	(17.74)	(12.54)	(10.90)	(5.01)	(3.58)
Post-secondary	0.86***	0.86***	0.96***	0.98***	0.76***	0.75***
	(12.35)	(13.62)	(10.88)	(10.53)	(5.36)	(3.46)
University	1.33***	1.34***	1.31***	1.24***	1.31***	1.21***
	(9.14)	(12.83)	(8.33)	(8.84)	(6.24)	(5.85)
Other most educated						
Some primary	0.22***	0.19***	0.25***	0.20***	-0.07	-0.05
	(5.14)	(4.22)	(5.31)	(5.63)	(0.71)	(0.40)
Full primary	0.20***	0.13***	0.22***	0.14***	0.00	-0.02
	(3.91)	(3.14)	(3.90)	(3.53)	(0.01)	(0.18)
Some secondary	0.32***	0.25***	0.36***	0.32***	-0.02	-0.03
	(5.99)	(4.87)	(6.07)	(6.25)	(0.14)	(0.28)
Full secondary	0.47***	0.41***	0.47***	0.39***	0.29**	0.27
	(8.53)	(7.91)	(7.83)	(8.07)	(2.16)	(1.64)
Post-secondary	0.66***	0.59***	0.86***	0.73***	0.20	0.23
	(7.68)	(9.31)	(8.19)	(8.64)	(1.47)	(1.27)
University	0.52**	0.41***	0.16	-0.06	0.84**	0.76
	(2.42)	(3.07)	(0.69)	(0.16)	(2.53)	(1.41)
Male head	0.22***	0.17***	0.22***	0.20***	0.26**	0.06
	(5.31)	(5.75)	(5.23)	(5.68)	(2.11)	(0.30)
Share of women	0.01	0.03	0.00	0.04	0.10	0.02
	(0.18)	(0.53)	(0.04)	(0.56)	(0.70)	(0.12)
Log labor force	0.42***	0.42***	0.39***	0.41***	0.67***	0.65***
	(11.18)	(9.70)	(9.71)	(8.96)	(6.77)	(6.18)
Has land	-0.13**	-0.13**	-0.14**	-0.15***	-0.01	0.03
	(1.97)	(2.12)	(2.01)	(3.66)	(0.06)	(0.28)
Log land size	0.26***	0.24***	0.27***	0.26***	0.09	0.08
	(10.15)	(10.24)	(10.23)	(9.59)	(1.40)	(1.04)
Region/area of residence						
Rural residence	-0.48***	-0.36***				
	(7.78)	(6.50)				
Central	0.03	0.08			0.03	-0.01
	(0.32)	(0.81)			(0.21)	(0.15)
Coast	-0.12	-0.02	-0.19**	-0.20***	-0.03	0.12
	(1.07)	(0.30)	(2.15)	(3.20)	(0.26)	(1.26)
Eastern	-0.40***	-0.26***	-0.47***	-0.42***	-0.14	0.01
	(3.52)	(2.92)	(6.66)	(9.37)	(0.84)	(0.11)
North-eastern	0.50***	0.47***	0.48***	0.38***	0.22	0.21
	(3.48)	(4.29)	(3.97)	(5.06)	(0.85)	(0.68)
Nyanza	-0.03	0.02	-0.06	-0.07*	-0.10	-0.03
	(0.29)	(0.23)	(0.99)	(1.74)	(0.83)	(0.28)
Rift Valley	-0.00	0.04	-0.03	-0.06*	-0.11	-0.03
	(0.04)	(0.39)	(0.51)	(1.69)	(1.02)	(0.32)
Western	-0.25**	-0.24*	-0.28***	-0.36***		
	(2.13)	(1.89)	(4.35)	(6.25)		
Constant	8.35***	8.36***	8.04***	8.31***	7.80***	8.16***
	(43.18)	(39.48)	(38.13)	(51.08)	(18.77)	(14.24)
Sample size	9183	9183	7876	7876	1307	1307
R^2	0.23	0.13	0.21	0.12	0.27	0.15

Dependent variable is log of total household earnings. * Significant at 10%; ** significant at 5%; *** significant at 1%

Table 10: Impact of Education on Annual Incomes (per cent above level below)

	Rural		Urban	
	Household Head	Other member	Household Head	Other member
Total earned income				
Full primary	48	15	38	(-2)
Full secondary	39	28	35	(33)
University	68	(-36)	80	(63)
Farm income				
Full primary	(23)	38	(31)	(-11)
Full secondary	8	4	(-17)	(27)
University	(-58)	(-47)	(-56)	(-49)
Business income				
Full primary	90	(32)	(-24)	(97)
Full secondary	23	(-8)	(25)	60
University	464	(-70)	(1)	(-69)
Wage income				
Full primary	76	45	77	4
Full secondary	42	26	22	32
University	90	(-21)	70	112

Source: Total income returns based on median regression. Activity-specific returns based on marginal effects of Tobit regressions

Among the other variables that predict total household earnings, the mean age and mean age squared have significant positive and negative coefficients respectively. Household earnings rise with age until 44 years in rural areas and 50 years in urban areas. The coefficient on age reflects both an experience effect and an age effect. The coefficient of the dummy variable for male-headed households is significant in rural areas, implying that a woman-headed household earned 22 per cent below man-headed households.

The earnings differential between male-headed households and female-headed households may reflect differences in unobserved factors or in earnings opportunities or in access to physical assets. Household labor force has significant effect in rural and urban samples. The region of residence is also correlated with total earnings in rural areas. With reference to rural central province, households in rural areas of other provinces had lower household income. Research on this province (e.g. Bigsten, 1985 and Bevan et. al., 1989) shows that households in the province already had diversified income sources in the 1960s and 1970s.

Tables 11 to 13 present Tobit regression estimates of activity-specific household income functions. The income gains from education in different activities are displayed in Table 10. They are calculated from the marginal effects conditional on a household having income above zero in a given activity. Table 11 presents Tobit regression estimates of farm income function. A household head with some primary education in rural areas raises farm income. The marginal effect implies that such a household earned 27 per cent more than a household with uneducated head. Beyond some primary education, secondary education of household head significantly raises farm income by 8 per cent above primary education. If household head has more than secondary education, the impact on farm income is negative. Other education in the household also has an impact on farm earnings. In a household where the highest education among other labor force members is some primary education, farm earnings are 30 per cent higher than in a household with no other educated member. Full primary education adds 6 per cent to earnings above some primary education and secondary education adds 3 per cent.

Table 11: Tobit Regression Estimates of Household Farm Earnings Function: Rural and Urban Areas (Absolute Values of t-values Within Parentheses)

Explanatory variables	Total	Rural	Urban
Age (years)	0.12***	0.14***	-0.17
	(3.87)	(4.53)	(0.54)
Age squared	-0.0012***	-0.0015***	0.0027
	(3.08)	(3.73)	(0.61)
head's education			
Some primary	0.17	0.26*	-1.43
	(1.30)	(1.96)	(0.92)
Full primary	0.29	0.23	1.25
	(1.64)	(1.32)	(0.82)
Some secondary	0.09	0.06	0.38
	(0.43)	(0.30)	(0.22)
Full secondary	0.33*	0.31*	0.38
	(1.66)	(1.72)	(0.22)
Post-secondary	-0.58	-0.30	-2.70
	(1.47)	(0.82)	(1.01)
University	-0.88	-0.66	-2.78
	(1.33)	(0.83)	(1.17)
Other most educated			
Some primary	0.23	0.29*	1.19
	(1.50)	(1.93)	(0.91)
Full primary	0.21	0.35**	-0.57
	(1.18)	(2.00)	(0.37)
Some secondary	0.24	0.36*	0.48
	(1.15)	(1.77)	(0.31)
Full secondary	0.32	0.39**	0.55
	(1.51)	(2.00)	(0.30)
Post-secondary	0.49	0.65*	0.95
	(1.26)	(1.89)	(0.35)
University	-0.58	-0.30	-2.88
	(0.95)	(0.53)	(0.65)
Male head	0.48***	0.47***	0.31
	(3.43)	(3.41)	(0.21)
Share of women	1.22***	1.32***	0.39
	(5.33)	(5.77)	(0.23)
Log Labor force	0.93***	0.89***	1.49
	(5.98)	(5.93)	(1.11)
Household has land	5.17***	4.50***	15.48***
	(15.54)	(12.90)	(10.18)
Log land size	0.44***	0.47***	0.61
	(6.27)	(7.01)	(0.86)
Region/area of residence			
Rural area	4.38***		
	(9.73)		
Central	0.43		-1.36
	(0.56)		(0.51)
Coast	-1.12	-1.48***	-3.40
	(1.38)	(3.85)	(1.27)
Eastern	0.16	-0.38	-2.15
	(0.20)	(1.43)	(1.08)
North-eastern	4.95***	4.26***	4.87*
	(5.75)	(8.42)	(1.80)
Nyanza	0.48	0.08	-2.87*
	(0.62)	(0.33)	(1.76)
Rift Valley	1.01	0.42*	3.12
	(1.30)	(1.69)	(1.60)
Western	0.81	0.36*	
	(1.04)	(1.69)	
Constant	-6.83***	-1.90**	-8.47
	(7.32)	(2.45)	(1.52)
Uncensored obs.	6974	6699	275
Left censored obs.	2209	1177	1032

Dependent variable is log of annual farm earnings. * Significant at 10%; ** significant at 5%; *** significant at 1%

The results contrast with those of Moock (1981) who found that farmers with 1 to 3 years of education had lower maize yields than uneducated ones in Western Kenya, and above this level, education added insignificant yields. Bigsten (1984) also found little effect of education on crop and livestock production. But Pinckney and Kimuyu (1995) find positive impact of primary education in a study site in Central Kenya. One explanation for the different results is that they cover different periods. In these periods, the relative importance of various factors on farm output can change. With land scarcity other factors such as education may become important in increasing efficiency (technical and allocative). Second, the productive value of education requires a dynamic environment (Schultz, 1975). If the changes require no education to implement, the productive value of education may not be significant.

Third, the studies cited use either logs of value of farm output or log of actual output as the dependent variable. In this paper, log of farm income is used. According to Welch (1970), a net income function captures both worker and allocation effects of education while the former capture worker effect alone. Phillips (1994) notes that the model used can explain differences in estimated educational impacts. Assuming 4 years of primary education, the return to household head's primary education is 7 per cent and 8 per cent for the most educated household member. Again though studies differ in many dimensions, this can be compared with the developing country average of 7.1 per cent in Lockheed et. al. (1980) and 6.1 per cent in Phillips (1994). Appleton (2000) estimated a return to 4 years of education of 10 per cent in Sub-Sahara Africa.

Table 12 presents Tobit estimates for own business earnings. Some primary education of the head and other members is associated with higher business income in rural areas. For example, households with a member who has some primary education earned 49 per cent more business earnings than one where no other member was educated. Where household head had full primary education, the increment to business earnings was 90 per cent above a household with uneducated head. This works out to 13 per cent assuming 7 years of primary education. The earnings gain from secondary completion above full primary is 23 per cent in urban areas. Secondary education of the most educated household member had an increment of 60 per cent. Assuming 4 years of secondary education, the return to household head education is 6 per cent in rural areas. In urban areas return to secondary education of other member is 15 per cent. Where head has university education, increment to earnings is large.

Some education levels, especially higher levels, have negative returns in own business and farming. This suggests that labor with higher education is reallocated to wage sector where returns are higher. Such negative returns to some levels of education are found in previous studies of the determinants of income in other parts of the less developed world (e.g. Janvry and Sadoulet, 2001, Appleton, 2001, and Estudilo and Otsuka, 1999). It is argued that in the case of rural areas it may reflect the limited opportunities for higher educated labor. In addition, only monetary gains are considered here. There may be other private gains.

The Tobit estimates for wage earnings function are presented in Table 13. Because very few cases had no wage earnings in urban areas, OLS estimates are presented for that sample. Consider wage income returns to education starting with household head's education. In rural areas, full primary education is associated with 76 per cent higher wage income above the uneducated. The earnings gain associated with full secondary education is 42 per cent above primary education while university education adds 90 per cent above full secondary education. For other labor force members, full primary education has an earnings gain of 45 per cent above earnings in households without another educated member. Secondary completion is associated with 26 per cent more earnings above primary completion.

Table 12: Tobit Regression Estimates of Household Own Business Earnings Function: Rural and Urban Areas (Absolute Values of t-values Within Parentheses)

Explanatory variables	Total	Rural	Urban
Age (years)	0.29**	0.16	1.03***
	(2.51)	(1.31)	(2.76)
Age squared	-0.0042***	-0.0025	-0.0140**
	(2.79)	(1.62)	(2.56)
head's education			
Some primary	1.68***	1.46***	1.86
	(3.60)	(3.05)	(1.10)
Full primary	2.02***	2.18***	-0.96
	(3.50)	(3.61)	(0.55)
Some secondary	2.66***	3.23***	-1.15
	(4.42)	(4.98)	(0.67)
Full secondary	2.60***	2.84***	-0.20
	(4.08)	(4.07)	(0.12)
Post-secondary	1.12	2.49**	-2.83
	(1.06)	(2.01)	(1.35)
University	3.65**	7.43***	-0.16
	(2.32)	(4.00)	(0.06)
Other most educated			
Some primary	1.38***	1.39***	1.24
	(2.71)	(2.59)	(0.84)
Full primary	1.04*	0.99	2.21
	(1.71)	(1.52)	(1.37)
Some secondary	0.95	1.01	1.41
	(1.45)	(1.43)	(0.87)
Full secondary	1.13	0.70	3.64**
	(1.61)	(0.94)	(1.98)
Post-secondary	-0.41	2.73*	-5.52*
	(0.30)	(1.91)	(1.90)
University	-2.80	-4.01	-0.02
	(1.25)	(1.54)	(0.01)
Male head	0.62	0.69	1.64
	(1.29)	(1.38)	(0.95)
Share of women	-0.01	-0.28	3.25
	(0.02)	(0.34)	(1.50)
Log Labor force	1.28***	1.68***	-1.95
	(2.62)	(3.30)	(1.33)
Household has land	0.90	0.54	3.70*
	(1.19)	(0.66)	(1.83)
Log land size	0.01	0.05	-1.31
	(0.04)	(0.16)	(1.07)
Region /area of residence			
Rural area	-0.95		
	(1.05)		
Central	4.29***		3.79*
	(3.02)		(1.80)
Coast	2.85**	-1.70	3.66**
	(2.05)	(1.59)	(2.08)
Eastern	0.50	-3.67***	-0.09
	(0.35)	(3.79)	(0.05)
North-eastern	-2.77	-7.85***	5.33
	(1.39)	(4.46)	(1.43)
Nyanza	6.76***	2.90***	5.01**
	(4.84)	(3.11)	(2.33)
Rift Valley	-0.95	-4.93***	-3.23
	(0.68)	(5.60)	(1.53)
Western	1.53	-2.56**	
	(0.97)	(2.44)	
Constant	-14.27***	-8.56***	-26.05***
	(5.81)	(3.41)	(3.87)
Uncensored obs.	2941	2492	449
Left censored obs.	6642	5384	858

Dependent variable is log of annual own business earnings. * Significant at 10%; ** significant at 5%; *** significant at 1%

Table 13: Tobit Regression Estimates of Household Wage Earnings Function: Rural and Urban Areas (Absolute Values of t-values Within Parentheses)

Explanatory variables	Total	Rural	Urban
Age (years)	0.12***	0.11***	0.16***
	(6.97)	(5.82)	(2.87)
Age squared	-0.0015***	-0.0014***	-0.0020**
	(6.58)	(5.64)	(2.32)
head's education			
Some primary	0.37***	0.35***	0.45***
	(6.89)	(6.11)	(2.71)
Full primary	0.59***	0.57***	0.57***
	(9.83)	(8.71)	(3.43)
Some secondary	0.72***	0.71***	0.66***
	(11.23)	(9.95)	(3.87)
Full secondary	0.89***	0.92***	0.77***
	(12.83)	(11.75)	(4.15)
Post-secondary	1.21***	1.33***	1.02***
	(14.99)	(12.74)	(6.12)
University	1.45***	1.56***	1.30***
	(7.17)	(8.17)	(4.16)
Other most educated			
Some primary	0.29***	0.32***	-0.08
	(4.96)	(5.15)	(0.67)
Full primary	0.34***	0.37***	0.04
	(5.25)	(5.27)	(0.31)
Some secondary	0.48***	0.56***	0.01
	(7.09)	(7.33)	(0.06)
Full secondary	0.58***	0.60***	0.32***
	(7.57)	(6.88)	(2.62)
Post-secondary	0.75***	1.03***	0.19
	(5.53)	(6.70)	(0.86)
University	0.74***	0.36	1.07***
	(2.93)	(1.23)	(3.03)
Male head	0.17***	0.20***	0.10
	(3.03)	(3.41)	(0.62)
Share of women	-0.17**	-0.19*	-0.03
	(1.97)	(1.87)	(0.15)
Log Labor force	0.44***	0.42***	0.68***
	(7.01)	(6.03)	(5.82)
Household has land	-0.31***	-0.35***	-0.26
	(3.01)	(2.98)	(1.18)
Log land size	0.14***	0.15***	0.01
	(3.49)	(3.57)	(0.07)
Region/area of residence			
Rural area	-0.73***		
	(9.04)		
Central	0.11		0.02
	(0.91)		(0.11)
Coast	0.06	-0.02	-0.10
	(0.44)	(0.16)	(0.74)
Eastern	-0.38***	-0.55***	-0.06
	(2.83)	(5.96)	(0.35)
North-eastern	-1.16***	-1.32***	-0.37
	(3.12)	(3.44)	(0.99)
Nyanza	0.00	-0.09	-0.14
	(0.02)	(1.00)	(0.98)
Rift Valley	0.07	-0.04	-0.01
	(0.59)	(0.56)	(0.07)
Western	-0.34*	-0.45***	
	(1.87)	(3.25)	
Constant	7.33***	6.90***	6.58***
	(23.06)	(19.81)	(7.39)
R-squared			0.23
Uncensored obs.	9055	7751	1304
Left censored obs.	128	125	3

Dependent variable is log of annual wage earnings. * Significant at 10%; ** significant at 5%; *** significant at 1%

In the urban areas, full primary education of household head is associated with 77 per cent higher earnings, while secondary completion adds 22 per cent above primary education. Where the head has university education, wage are 70 per cent above secondary education level. Secondary education of other members of the labor force adds 32 per cent to wage earnings above primary education, while if there is a member with university education, wage earnings are 112 per cent higher compared to a household with no other education.

Apart from education attainment in the household, there are other factors that are significantly correlated with incomes in the different economic activities. Age effects are positive and show a concave pattern; earnings in every activity increase with average age of the labor force at a decreasing rate. This suggests that experience and lifecycle effects are important in income generation. In rural areas, male-headed households realized relatively greater farm and wage earnings than woman-headed households. The share of women in the labor force significantly raises farm earnings but reduces wage earnings.

Gender effects on farm income might reflect greater propensity to adopt yield-enhancing technology. For example, in a study of maize farming in western Kenya, Ongaro (1988) finds that the introduction of new weeding technology raised output in female-headed households by greater amount than in male-headed households. The gender effects may also reflect bargaining power in the household or other unobserved factors associated with woman-headed households that make them earn less farm and wage earnings. In the case of wage earnings the result may indicate that women labor has fewer wage earning opportunities than male labor. For urban households, the share of women in household labor is associated with higher own business earnings. This may suggest that in urban areas it is women who are mainly involved in family businesses.

The size of labor force has positive and significant impact on earnings except in urban own business. The larger the labor force, the larger would be the farm and wage earnings. Additional labor force can be allocated to wage employment more easily, or if production in a given activity like farming is constrained by labor, additional members can lend a hand. Land ownership and land size are associated with higher farm earnings. On the other hand, land ownership is associated with lower wages in rural areas, while land size is associated with higher wage earnings. The land ownership dummy may be a proxy for some unobserved or unmeasured characteristics.

V. Conclusion

The purpose of this paper is to examine empirical evidence on how household education affects economic activity combinations and earnings of households in Kenya. The data, drawn from a national survey of rural and urban households indicate that, many households generate earnings from more than one economic activity. In rural areas the most common activity combination is farming and wage employment while in urban areas it is wage employment. This practice opens a channel of economic transformation, which is distinct from rural-urban labor migration, which analyses of structural change based on dual economy model emphasize.

Multinomial logit regression estimates of household activity combination suggest that education below secondary level induces households to engage in more than one income generating activity. Secondary education has strong effect on the probability of combining wage work, farming and own business. Tertiary education (post-secondary and university) of household head reduces the probability of mixing farming and wage work in rural areas, in

favor of wage work alone. In urban areas, education seems less important for activity combination. Household with high education levels derive earnings mainly from wage employment. It may suggest that wage employment that pays is accessible only to those with high levels of education. As noted by Hughes (1991) the demand for education in Kenya cannot be detached from household income strategies. With land scarcity, and the need for cash to carry out farming activities, households invested in education and with the increased supply of educated labor come a rise in hiring standards. More education is required to access wage employment.

OLS and Median regression estimates of total household income indicate that more educated households have higher total earnings. The increment to total earnings if household head has full primary education is 48 per cent and 38 per cent in rural and urban areas respectively. For a household head with full secondary education, the respective incremental return above primary education is 39 per cent and 35 per cent. The income advantage to households with heads educated to more advanced level is maintained through university level. While education of the most educated labor force member apart from the household head has total income returns in rural areas, these are relatively small compared to those of household head's education.

Activity-specific tobit regression estimates show that income returns differ across education levels and across activities. Some primary education in the households would raise farming income. Similarly, full primary education of the most educated labor force member, has an incremental farm income return of 38 per cent over a household with no other educated labor force member. Incremental farm income returns to secondary education are small; four to eight per cent above primary education. When it comes to full secondary education of the household head and of other labor force, returns are found in wage employment. The incremental wage income return to household head's secondary education is 22 per cent in urban households and 42 per cent in rural households. For the other most educated household member the return is 26 per cent in rural area and 32 per cent in urban areas. In own family business, education of household head seems to be the most important in rural areas but not in urban areas.

The results suggest why poverty status and education are correlated in Kenya. First education and earnings are correlated, and second education is correlated with economic activity combinations. The relatively higher returns to education in wage work points toward the central role that wage opportunities play in realization of returns to education. But the dominant sector in Kenya is small-scale farming. The results show that both land and education (at least primary education) are important for generation of farm incomes.

VI. References

Adams, R.H. (2002), Nonfarm income, inequality, and land in rural Egypt. *Economic Development and Cultural Change*, Vol. 50, Issue. 2, pp. 339-364.

Appleton, S. (2000), Education and health at household level in Sub-Sahara Africa. *CID Working Paper No. 33*, Havard University.

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

Appleton, S. and Balihuta, A. (1996), Education and agricultural productivity: Evidence from Uganda, *Journal of International Development*, 8 (3), 415-444.

Appleton, S., Hoddinot, J., and McKinnon, J. (1996), Education and health in Sub-Sahara Africa, *Journal of International Development*, Vol. 8, 3, 307-339.

Bevan, D., P. Collier, and J.W.Gunning, with A.Bigsten and P.Horsnell. 1989. *Peasants and Governments*: Oxford: Oxford University Press.

Bigsten, A. (1984), *Education and income determination in Kenya*. Gower Publishing Company Limited. Gower House, Croft Road, Aldershot, England.

Bigsten, A. (1996), The Circular Migration of Smallholders in Kenya. *Journal of African Economies*, 1996, vol. 5, issue 1, pages 1-20

Bigsten, A. (1985), "What do Smallholders do for a Living? Some Evidence from Kenya", in M. Lundahl (ed.), *The Primary Sector in Economic Development*, Croom Helm,London, 1985

Bigsten, A. and Kayiizi-Mugerwa, S.(1995), Rural sector responses to economic crisis in Uganda. *Journal of International Development*: Vol. 7, No. 2, 181-209.

Bigsten, A. and Kayiizi-Mugerwa, S. (1996), Adaptation and distress in the urban economy: A study of Kampala households. *World Development*, Vol. 20, pp 1423-1441

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.

Corden, W. and Findlay, R. (1975), Urban unemployment, intersectoral capital mobility and development policy. *Economica*, pp 59-78.

Coulombe, H. and Mackay, A. (1996), Modelling the determinants of poverty in Mauritania. *World Development*, 24(6), 1015-1031.

Deaton, A. (1997), *The Analysis of household surveys: A micro econometric approach to development Policy*. The Johns Hopkins University

Dercon, S. and Krishnan, P. (1996), Income portfolios in rural Ethiopia and Tanzania: Choices and constraints. *Journal of Development Studies*, vol. 32, No.6, pp 850-875.

Ellis, F. (1998), Household strategies and rural livelihood diversification. *Journal of Development Studies,* vol.35, no.1, pp 1-38.

Estudillo, J. and Otsuka, K. (1999), Green revolution, human capital, and off-farm employment: Changing sources of income among farm households in central Luzon, 1966-1994. *Economic Development and Cultural Change*. vol. 47, Issue 3, pp 497-524.

Fafchamps, M. and Quisumbing, A. (1999), Human capital, productivity, and labor allocation in rural Pakistan. *Journal of Human Resources*, vol. 34, Issue 2, pp 369-406.

Flood, L. and Urban G. (1998), Regression Analysis and Time Use Data A Comparison of Microeconometric Approaches with Data from the Swedish Time Use Survey (HUS) , *Swedish Working Papers in Economics* No.5

Geda, A., Jong, N., Mwabu, G., and Kimenyi, M. (2001), Determinants of poverty in Kenya: A household –level analysis. *ISS Working paper*, Institute of Social Studies, The Hague

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

Harris, J. and Todaro, M. (1970), Migration, unemployment and development: A two sector analysis, *American Economic Review*, vol. 60, no. 1, pp 126-42

Janvry, A. and Sadoulet, E. (2001), Income strategies among rural households in Mexico: The role of off-farm activities. *World Development*, vol. 29, no. 3, pp 467-480.

Jolliffe, D. (1998), Skills, schooling, and household income in Ghana. *The World Bank Research Observer*, vol. 12, no. 1: 81-104

Jones, P. (2001), Are educated workers really more productive? *Journal of Development Economics*, vol. 64, 57-79

Julin, E. (1993), Structural change in rural Kenya, Ekonomiska Studier 41, Department of Economics, Göteborg University.

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

Laszlo, S. (2001), *Self-employment earnings: An empirical analysis of non-farm activities in rural Peru*. Unprocessed, Department of Economics, University of Toronto.

Lockheed, M., Jamison, D., and Lau, L. (1980), Farmer education and farm efficiency: A Survey. *Economic Development and Cultural Change*, 29(1), 37-76.

Moock, P., Musgrove, P., and Stelcner, M. (1990), Education and earnings in Peru's informal nonfarm enterprises. *LSMS Working Paper Number 64*, World Bank, Washington.

Mwabu, G., Masai, W., Gesami, R., Kirimi, J., Ndeng'e, G., Kiriti, T., Munene, F., Chemengich, M., and Mariara, J. (2000), *Poverty in Kenya: Profiles and determinants*. Draft Final Report to AERC.

Mwabu, G.M. and Schultz, T.P. (2000), Wage premia for schooling and location, by gender and race in South Africa, *Economic Development and Cultural Change,* Vol. 48, 2, 307-334.

Nielsen, S. and Nielsen. W. (2001), Returns to Schooling in less developed countries: Evidence from Zambia; *Economic Development and Cultural Change*, Chicago; Jan 2001; Vol. 49, Issue. 2; 365-395

Ongaro, W. (1988), Adoption of new technology: A case study of maize production in Western Kenya. *Economiska Studier no. 22*, Department of Economics, University of Gothenburg.

Phillip, J. (1994), Farmer education and farmer efficiency: A meta-analysis. *Economic Development and Cultural Change*, Vol. 43, Issue 1, pp. 149-166

Pinckney, T. and Kimuyu, P. (1995), Policy and rural development: Two communities in East Africa. Department of Economics, Williams College, Williamstown, Massachusetts.

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

Ranis, G. and Fei, J. (1961), A theory of Economic development, *American Economic Review*, Vol. 51, No. 4, pp. 533-565.

Ram, R. (1980), Role of education in production: A slightly new approach. *Quarterly Journal of Economics*, vol. 95, Issue. 2, 365-373

Ram, R. and Singh, R. (1988), Farm households in Burkina Faso: Some evidence on allocative and direct return to schooling and male-female labor productivity differentials. *World Development*, vol. 16, No. 3, pp 419-424.

Reardon, T. (1997), Using evidence of household income diversification to inform study of rural nonfarm labor market in Africa. *World Development*, vol. 25, no. 5, pp 735-747.

Schultz, T.P. (1988), Schooling Investments and Returns, in: H. Chenery and T.N. Srinivasan, *Handbook of Development Economics*, vol. 1 (Amsterdam, North Holland), 543-630.

Schultz, T.W. (1961), Investments in Human Capital, *American Economic Review*, vol. 51, Issue. 1, pp 1-17.

Schultz, T.W. (1975), The value of the ability to deal with disequilibria, *Journal of Economic Literature*, 13, 872-876.

Taylor, J. E. and Yunez-Naude A. (2000), The returns from schooling in a diversified rural economy. *American Journal of Agricultural Economics*, 82: 287-297.

Vijverberg, W.P. (1991), Profits from self-employment. The case of Cote d' Ivoire. *World Development*, vol. 19, No. 6, pp. 683-696

Vijverberg, W.P. (1995), Returns to schooling in non-farm self-employment: An econometric case study of Ghana. *World Development*, vol. 23, No. 7, pp 1215-1227.

Willis, R.J (1986), Wage determinants: A survey and reinterpretation of human capital earnings functions, in: O.Ashenfelter and R.Layard, eds., *Handbook of Labor Economics*, Vol. I (Elservier Science).

World Bank (2000), *World development report, 2000/2001: Attacking poverty,* Oxford University press, New York.

Weir, S. (1999), The effect of education on farmer productivity in Ethiopia, *Working Paper 99-7*, Centre for the Study of African Economies

Welch, F.(1970), Education in production. *Journal of Political Economy*, vol. 78, Issue. 1, pp 35-59.

Yunez-Naude, A. and Taylor, J.E. (2001), The determinants of nonfarm activities and incomes of rural households in Mexico with emphasis on education. *World Development* vol. 29, No. 3, pp 561-572.