STELLENBOSCH UNIVERSITY
DEPARTMENT OF ECONOMICS

Education and Labour Market Outcomes in Mozambique

Carlos da Maia

2011

Abstract

In Mozambique many people work outside the wage sector. A large part of the working-age population is engaged in self-employment activities and unpaid family work. A direct implication is that different labour market segments might also differ with respect to the determinants of entry and earnings. Using *Inquérito Integrado à Forca de Trabalho*, the only available Labour Force Survey in Mozambique, and employing a multinomial logit model, we find that schooling increases the chances of wage employment and lowers the chances of self-employment and unpaid family work. Using Ordinary Least Squares (OLS) Mincerian type earnings functions we show that schooling increases earnings prospects across all sectors for which reliable earnings data are available. We checked and corrected our OLS estimates for potential selectivity bias using a modified version of Dubin and McFadden's (1984) model developed by Bourguignon, Fourier, & Gurgand (2004; 2007).

JEL classification: J31, I21

Keywords: Earnings functions, Analysis of education, Mozambique.

1. INTRODUCTION

"Prospects for reducing poverty, narrowing extreme inequalities and improving public health are heavily influenced by what happens in education. Progress towards the equalisation of opportunity in education is one of the most important conditions for overcoming social injustice and reducing social disparities in any country" (UNESCO, 2009, p. 24).

An important linkage of education and poverty alleviation is through the labour markets. Schooling is intrinsically linked to participation in the labour force. It shares a positive association with the probability of employment. More educated people are more likely to participate and to get the best paid jobs available (Bhorat & McCord, 2003, p. 135). Education is also associated with better command of earnings. Human capital theorists advocate schooling and post-schooling investment as sources of improvement of lifetime earnings (Mincer, 1974).

In developing countries such as Mozambique many people work outside the wage sector. A large part of the working-age population is engaged in self-employment activities and unpaid family work, notably subsistence agriculture and sales in informal markets. A direct implication is that different labour market segments might also differ with respect to the determinants of entry and earnings. For instance, Glick & Sahn (1997), employing multinomial logit models for Guinean data, find that the determinants of entry into various sectors of the labour market are quite different. For both men and women, more education reduces the likelihood of being self-employed while it strongly increases the likelihood of working in the public sector. In Ghana, Glewwe (1991) finds that education is positively associated with wage employment, and among the wage employed, the better educated are more likely to be in the public sector than in the private sector.

The data set we use in this study contains information on the different segments of the labour market the working-age population is engaged in. A plausible view of the labour market in Mozambique would allow for four distinct employment segment alternatives, apart from the decision not to participate or being unemployed. These include wage employment in the private sector, wage employment in the public sector, self-employment and unpaid family work.

Thus, the objectives of this study are twofold. First, we want to identify the impact schooling and other individual characteristics have on the individual probability of falling into each one of the six employment segments mentioned above. Given the multitude of "options" open to the working-age population, in this part of the study we employ a multinomial logit model.

The second objective is to link education and earnings. We use Mincerian type Ordinary Least Squares (OLS) regressions to estimate the relationship between investment in schooling and post-schooling, on the one hand, and the distribution of earnings, on the other, in the three employment sectors for which reliable earnings data are available: the wage public sector, the wage private sector, and self-employment. However, since workers might not have been randomly assigned to each sector, there is potential for sample selection bias in our OLS earnings functions. To correct for this we use a modified version of Dubin & McFadden's (1984) model. This procedure, as we will see below, is based on a multinomial logit model.

In the literature no prior studies exist employing similar methodologies to analyse such issues in Mozambique. Therefore, this study is important in its own right. In addition, such analysis, notably the first issue we tackle, gives an indication of the degree of heterogeneity found in the Mozambican labour market (Demery & Grootaert, 1993). For instance, if schooling affects the chances of entry into the public sector differently from entry into private and/or self-employment, then the questions raised here are important from a policy perspective. Many developing countries, including Mozambique, under the influence of the Bretton Woods institutions – International Monetary Fund and the World Bank – had undergone a large scale economic liberalisation, expansion of the private sector, and reduction of the civil service (Glick & Sahn, 1997, p. 794). Further, as argued by Wambugu (2002, p. 2), the level of returns to schooling in each labour market sector will indicate the education potential for improving economic welfare and will serve as an incentive for poor individuals and households to invest in education.

The paper is structured as follows. Section 2 describes the data set and reports descriptive statistics of the variables of interest. Section 3 links education to employment outcomes. Earnings functions analysis without and with sample selectivity correction are presented in Section 4. The concluding remarks are presented in Section 5.

2. DATA

The data set is sourced from *Inquérito Integrado à Forca de Trabalho*, the only available Labour Force Survey in Mozambique. This survey was conducted by the National Statistics Institute (INE) in conjunction with the Ministry of Labour over the years 2004 and 2005. It surveyed a stratified random sample of 17,151 households designed to be representative at the national and provincial levels and by place of residence (urban and rural). Among other things, the survey collected information on individual background and household characteristics, employment segment, industry, occupation, hours worked as well as labour earnings.

We follow conventional practice and restrict our sample to individuals aged between 15 and 64 years, inclusive. Table 2.1 presents summary statistics of the variables used in the analysis. After deleting observations with missing data, the sample consisted of 15,915 men and 20,150 women, distributed between six employment states: wage employment in the public sector, wage employment in the private sector, self-employment, unpaid family work, unemployment and non-participation.

As is expected in developing economies, the proportion of men and women in self-employment is very high. As one can observe at the bottom of the table, 57% of men and 44% of women are self-employed. Wage employment in the public and private sectors, on the other hand, is very low, notably among women. The proportion of women in the public and private wage segments is below 2% for both sectors while for men it is 6% and 11%, respectively. Differently from men who have a high stake in the wage employment segment, women outside self-employment are mostly unpaid family workers (38%).

Turning to education, 16% of men and 43% of women have no formal schooling. For both sexes, but worse for women, the figures are highest among the self-employed and unpaid family workers. For individuals who are wage employed, mainly those in the public sector, lack of formal schooling is very uncommon. Only 2% of men and 3% of women in the public sector have no formal schooling.

There is a major gap in primary school completion across gender. The proportion of women who have completed full primary education is about half (4%) that of men (9%). This gap is again more pronounced among self-employed and unpaid family workers, and almost non-existent within the wage employment segment of the working-age population.

The significance of schooling in determining employment segment selection becomes more obvious at higher levels of education. While the proportion of individuals with full secondary education completed is very low (2% for men and 1% for women), in the public sector they are as high as 14% for men and 20% for women. But in the self-employment and unpaid family worker segments, less than half a percentage of men and women have completed full secondary school. This pattern is consistent with that at the tertiary education level, except that the proportion of women in the public sector with at least some higher education (8%) is well below that of men (12%).

Accordingly, mean years of schooling vary widely among employment segments, pointing to large differences in skill requirements in different parts of the labour market. In the public sector, both men and women tend to be well educated, averaging 8.9 and 9.1 years of schooling, respectively. Among self-employed and unpaid family workers, in contrast, average educational attainment is very low – just, respectively, three and four years for men and two years for women in both sectors. Interestingly, mean years of education among the unemployed and non-participants is as high as in the private sector.

The average age of the working-age population is 32 years for both men and women. Among employed men and women unpaid family workers are the youngest and public sector workers the oldest[1]. Private sector workers are much older than the unemployed and non-participants. The average age of men and women in the private sector is 32 years for both sexes, but it is only 24 and 25 years among the unemployed, and 22 and 26 years among non-participants, respectively. Thus, though men in these three employment segments have equal average educational attainment, private sector workers seem to be more experienced, on average.

The proportion of married men in the public, private and self-employment segments is significantly higher than is the case in the remainder of the segments. It averages 87%, 66%, and 86% in the first three employment segments, but only 15%, 21%, and 11% among unpaid family workers, unemployed, and non-participants, respectively. This is

[1] For men only. For the case of women self-employed workers seem to be the oldest. But confidence intervals overlap, thus suggesting that women are oldest in both sectors (public sector and self-employment).

consistent with the idea that married men are the household providers and therefore are more likely to be employed in sectors that bring external resources into the household. We do not claim this idea to be absolutely true since the reverse can also occur, i.e., men might get married because they are employed and therefore have the means to.

In contrast, the literature on women labour force participation finds that married women are less likely to work than unmarried women, presumably reflecting a higher reservation wage resulting from access to their spouses' incomes as well as due to childbearing and raising children (Glick & Sahn, 1997, p. 799). Our data give some evidences of the above assertion. The proportion of married women is highest among unpaid family workers (81%) followed by self-employed women (61%). The figure for the public and private wage sectors is 54% and 37%, respectively. Married women seem to be less likely to be wage employed.

The number of children in the household averages two for both men and women and does not vary much across gender and across employment segments. Both male and female unpaid family workers average three children in the household while in the remainder of the employment segments the average number of children in the household is two. For both men and women the average number of adults in the household is three and there is not much variation across employment segments.

The proportion of household heads in the sample is highest among men. There is a clear divide between employment segments. For both men and women, the percentage of household heads in the wage and self-employment segments is significantly higher than in the remainder of the segments. This is consistent with the idea that since household heads take the economic responsibility for the household they are likely to work in employment segments that bring external resources home. Again, the opposite might also be true, i.e., well employed household members might end up becoming household heads.

The proportion of rural individuals in wage employment is relatively low. Only 30% of men and 16% of women in the public sector, and 29% and 16% of men and women, respectively, in the private sector live in rural parts of the country. In contrast, for both men and women most of self-employment and unpaid family work occurs in rural areas.

Table 2.1 – Summary statistics of variables used in the analysis; Note: figures for the categorical variables correspond to proportions and should add up to 1; For all figures weighting and clustering effects were taken into account, unless indicated otherwise; linearised standard errors in parentheses.

Employment Segment	Public Sector		Private Sector		Self-employment		Unpaid family work		Unemployment		Non-participation		All	
Variables	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Education level completed														
No formal schooling	0.0176	0.0260	0.1091	0.0825	0.2075	0.5057	0.1432	0.4800	0.0427	0.1048	0.0666	0.2224	0.1555	0.4311
Some primary [1-6]	0.2233	0.1502	0.5406	0.4270	0.6685	0.4491	0.7195	0.4860	0.4463	0.5038	0.5089	0.4569	0.6059	0.4632
Full primary [7]	0.1250	0.1126	0.1221	0.1257	0.0679	0.0246	0.0751	0.0211	0.1524	0.1346	0.1263	0.1121	0.0887	0.0408
Some secondary [8-11]	0.3791	0.4374	0.1681	0.2148	0.0493	0.0196	0.0593	0.0127	0.2852	0.2204	0.2697	0.1848	0.1172	0.0534
Full secondary [12]	0.1388	0.1976	0.0354	0.0928	0.0047	0.0007	0.0028	0.0002	0.0523	0.0248	0.0163	0.0163	0.0197	0.0079
Tertiary education [+13]	0.1162	0.0762	0.0247	0.0573	0.0021	0.0003	0.0000	0.0000	0.0211	0.0117	0.0120	0.0075	0.0129	0.0036
Mean years of schooling	8.94	9.14	5.47	6.69	3.39	1.75	3.87	1.65	6.63	5.63	5.99	4.70	4.43	2.41
	(0.15)	(0.19)	(0.12)	(0.28)	(0.06)	(0.05)	(0.09)	(0.05)	(0.12)	(0.10)	(0.12)	(0.13)	(0.06)	(0.05)
Mean age	38.68	35.96	32.20	31.95	35.63	36.12	20.93	29.59	23.70	25.33	21.77	25.65	31.81	31.98
	(0.36)	(0.54)	(0.30)	(0.57)	(0.22)	(0.24)	(0.36)	(0.19)	(0.29)	(0.18)	(0.39)	(0.42)	(0.15)	(0.13)
Married	0.8675	0.5375	0.6565	0.3665	0.8645	0.6092	0.1533	0.8120	0.2077	0.5494	0.1141	0.2789	0.6544	0.6557
Mean no. of children in the household	2.29	2.12	2.21	2.15	2.22	2.23	2.68	2.60	2.22	2.46	2.32	2.33	2.28	2.40
	(0.08)	(0.10)	(0.06)	(0.13)	(0.03)	(0.04)	(0.17)	(0.09)	(0.07)	(0.06)	(0.11)	(0.07)	(0.04)	(0.04)
Mean no. of adults in the household	3.04	3.44	3.35	3.49	2.55	2.44	3.76	2.85	4.23	3.73	4.09	3.87	3.05	2.83
	(0.06)	(0.10)	(0.06)	(0.12)	(0.02)	(0.03)	(0.16)	(0.08)	(0.07)	(0.06)	(0.07)	(0.07)	(0.03)	(0.04)
Household head	0.8956	0.4402	0.6505	0.3464	0.8809	0.4499	0.0440	0.0083	0.1771	0.0969	0.1334	0.0928	0.6523	0.2260
Rural	0.2955	0.1648	0.2859	0.1643	0.7978	0.7138	0.8052	0.8646	0.1557	0.0866	0.3355	0.2752	0.6297	0.6757

Employment Segment	Public Sector		Private Sector		Self-employment		Unpaid family work		Unemployment		Non-participation		All	
Variables	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Provinces														
Niassa	0.0428	0.0352	0.0155	0.0083	0.0549	0.0282	0.0645	0.0762	0.0138	0.0214	0.0424	0.0312	0.0470	0.0461
Cabo Delgado	0.0771	0.0494	0.0423	0.0139	0.1177	0.0949	0.0613	0.0954	0.0280	0.0364	0.1095	0.0797	0.0941	0.0875
Nampula	0.1577	0.0880	0.1239	0.0711	0.2355	0.1198	0.1879	0.2592	0.1316	0.1456	0.1530	0.1065	0.1994	0.1734
Zambezia	0.0843	0.0791	0.0741	0.0651	0.2415	0.2199	0.1885	0.2012	0.0628	0.0408	0.1353	0.1196	0.1870	0.1872
Tete	0.0668	0.0728	0.0559	0.0206	0.0790	0.0568	0.1357	0.1014	0.0157	0.0209	0.0451	0.0493	0.0746	0.0703
Manica	0.0686	0.0604	0.0484	0.0171	0.0643	0.0724	0.1041	0.0628	0.0275	0.0328	0.0716	0.0802	0.0652	0.0651
Sofala	0.1089	0.0466	0.1446	0.0879	0.0722	0.1211	0.0665	0.0434	0.0676	0.0735	0.1086	0.1164	0.0844	0.0854
Inhambane	0.0560	0.0879	0.0724	0.0642	0.0505	0.0984	0.1071	0.0855	0.0642	0.0394	0.0738	0.0604	0.0623	0.0855
Gaza	0.0387	0.1077	0.0891	0.1011	0.0379	0.1025	0.0546	0.0519	0.0896	0.0525	0.0706	0.0798	0.0518	0.0776
Maputo Province	0.1373	0.1444	0.1614	0.2197	0.0268	0.0518	0.0228	0.0176	0.2324	0.2443	0.0590	0.0808	0.0644	0.0599
Maputo City	0.1619	0.2285	0.1725	0.3311	0.0197	0.0344	0.0072	0.0053	0.2668	0.2924	0.1310	0.1959	0.0698	0.0620
No. of unweighted observations	1,219	505	2,481	483	7,226	8,546	1,529	6,052	1,660	2,602	1,800	1,962	15,915	20,150
Weighted proportion in each segment	0.0557	0.0157	0.1105	0.0149	0.5699	0.4380	0.1084	0.3849	0.0692	0.0798	0.0864	0.0667	1.0000	1.0000

Lastly, wage employment seems to be concentrated in Maputo City. This province alone employs 16% and 23% of the public sector men and women, respectively. Unemployment is also highest in Maputo City, as 26% and 29% of the unemployed men and women, respectively, are from Maputo City.

For the wage sectors, the survey contains information on the monetary value earned from the main occupation in the last 30 days. It also contains information on the number of hours worked in the last seven days. We obtained the measure of wage earnings used in this paper by calculating the ratio of daily earnings to daily hours worked. The results were individual figures of hourly earnings from main occupation in Meticais (MTn). For self-employed workers the survey reports weekly and monthly net income. This variable is calculated as the difference between the monetary amount obtained from the self-employment activities in question and the amount invested or spent to generate it. We converted these net figures to daily values and divided them by daily hours worked, thus obtaining figures for net hourly Meticais.

Table 2.2 reports mean hourly earnings for men and women across three labour market sectors and by education level. To better describe our statistics we also include two figures (Figure 2.1 and Figure 2.2) illustrating mean hourly earnings as well as their respective 95% confidence intervals across employment sectors, for both men and women. Within the wage sector men´s hourly earnings are significantly higher within the public sector. While they, on average, earn 23 MTn per hour within the public sector, they earn only 9 MTn per hour within the private sector, and the confidence intervals do not overlap, thus suggesting high attractiveness of the civil service relative to private wage sector employment. Also for men, self-employment seems to produce the highest pay. Self-employed men earn, on average, 29 MTn per hour. However, because the confidence interval within this sector overlaps with the public sector´s confidence interval, self-employment hourly earnings are not significantly greater than public sector wages. Nevertheless, self-employment hourly earnings are significantly higher than private sector pay. For women, though mean hourly earnings seem to be highest within the public sector, they are not significantly different from hourly earnings within the private and self-employment sectors, since the confidence intervals overlap.

The premium for men´s self-employment pay relative to men within the private sector, however, should be viewed with caution. For the sake of ensuring a large enough sample size we lumped together two types of self-employed workers: self-employed workers with employees and self-employed workers without employees. The first group is rather small, constituting only 14% and 6% of self-employed men and women, respectively[2]. However, men and women are relatively better educated in this group. They average seven and six years of education completed, respectively, while the figures for self-employed men and women without employees are, correspondingly, five and four only.

Therefore, mean hourly earnings within the first group are significantly higher, notably for men. Self-employment hourly earnings average 88 MTn and 50 MTn for men and women with employees, and 20 MTn and 13 MTn for men and women without employees, in that order. Nevertheless, men's average self-employment hourly earnings for those workers without employees are still double that of men´s average private sector hourly earnings (20 MTn versus 9 MTn). Results such as these are not uncommon in the literature. Glick and Sahn found similar results for Guinea. In their view this does not mean that self-employment pays best but instead that self-employed workers underestimate the amount they spend or invest in their activities (Glick & Sahn, 1997, p. 821; Glick & Sahn, 1993).

Across all sectors average hourly earnings grow with educational attainment. For those in self-employment the increase in average hourly earnings is not monotonic, as is the case for wage employment. In general, across all levels of education, men are better paid than women. But women with tertiary education in the private sector are, on average, much better paid than men with the same education level.

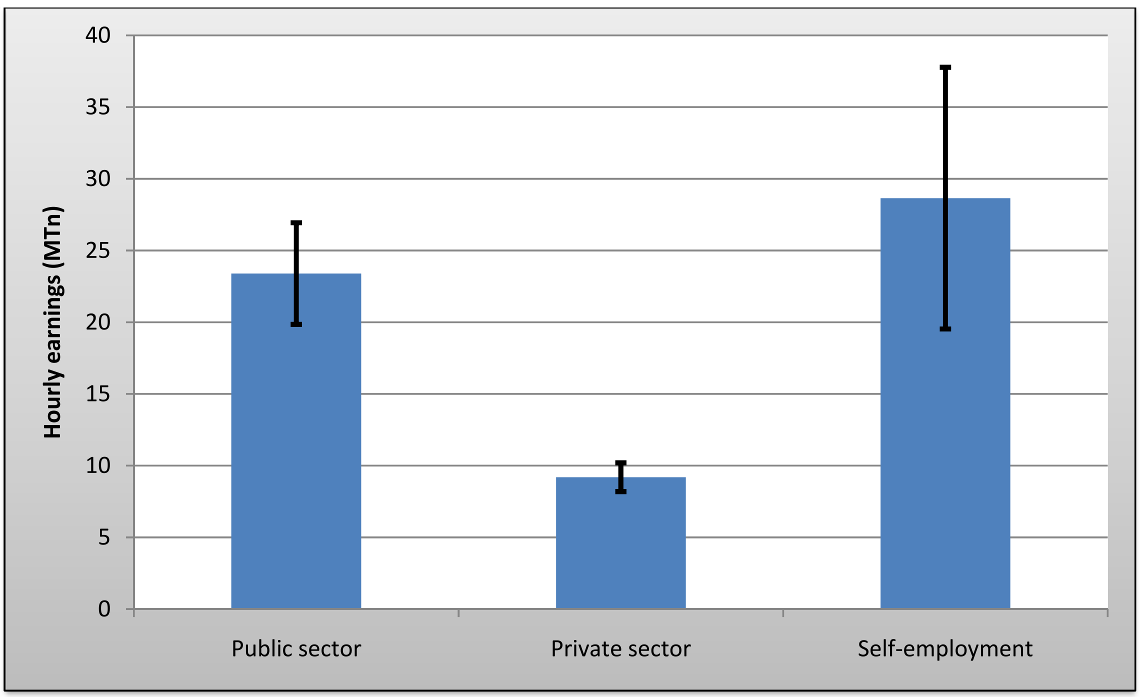
The differentials across employment segments in the men's and women's characteristics reported in Table 2.1 and Table 2.2, notably in schooling, may explain the differentials in access to employment and also earnings. We next explore these issues further using multivariate methodologies.

[2] These percentages are in relation to those self-employed workers who reported earnings. Considering the entire sample of self-employed individuals reported in Table 2.1, self-employed men and women with employees correspond, respectively, to 7% and 2% only.

Table 2.2 – Mean hourly earnings across sectors and by education level completed; Weighting and clustering effects were taken into account; Note: at the time of the survey USD1.00=25.00MT; R1.00=3.50MT.

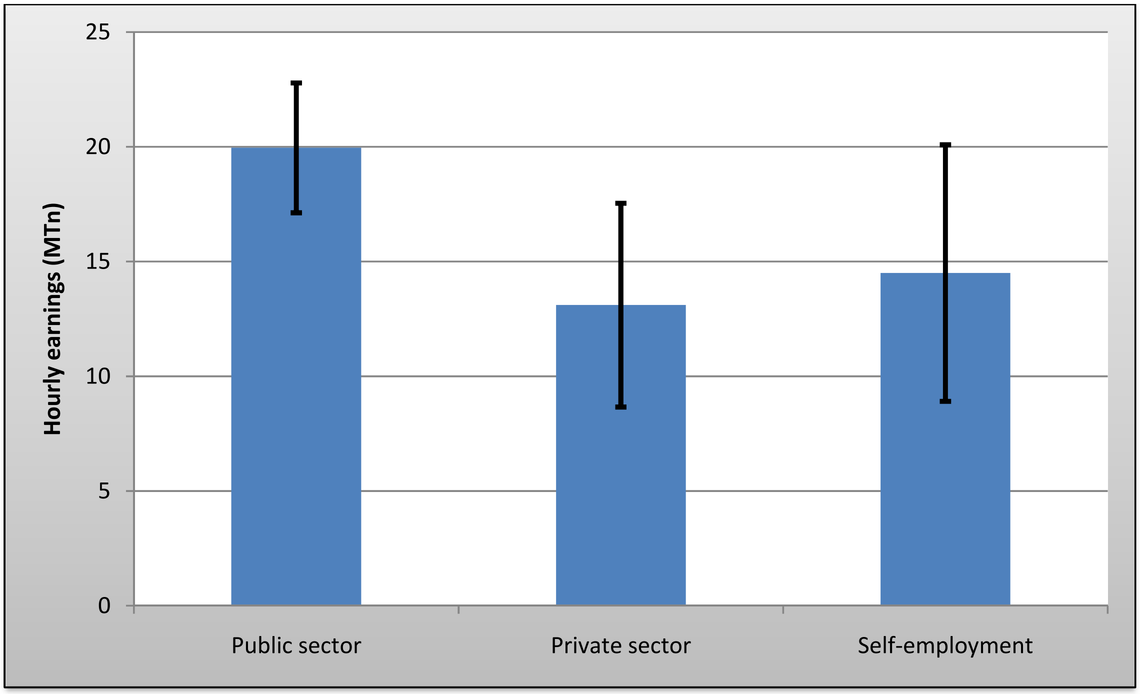
Employment Segment	Public Sector		Private Sector		Self-employment	
Variables	Men	Women	Men	Women	Men	Women
No formal schooling	9.32	6.59	5.18	3.81	12.95	16.81
Some primary [1-6]	9.65	8.09	5.92	3.92	22.92	9.71
Full primary [7]	11.27	10.91	7.17	5.32	21.20	25.14
Some secondary [8-11]	16.55	16.33	11.37	12.92	72.20	20.95
Full secondary [12]	29.75	26.64	35.26	31.88	60.71	174.13[3]
Tertiary education [+13]	74.93	64.35	52.80	82.15	232.44	31.61[4]
Hourly earnings	23.40	19.96	9.20	13.12	28.66	14.50
No. of sample observations	1,053	438	2,177	440	2,455	1,655

Figure 2.1 – Men´s mean hourly earnings and 95% confidence intervals across sectors; Weighting and clustering effects were taken into account; Note: at the time of the survey USD1.00=25.00MT; R1.00=3.50MT.



[3] Only 9 observations in the sample.
[4] Only 2 observations in the sample.

Figure 2.2 – Women´s mean hourly earnings and 95% confidence intervals across sectors; Weighting and clustering effects were taken into account; Note: at the time of the survey USD1.00=25.00MT; R1.00=3.50MT.



3. EMPLOYMENT AND EDUCATION

In this part of the study we estimate a model of labour market participation and determination of segment of employment. As discussed above, a plausible view of the labour market in Mozambique would allow for four distinct employment segment alternatives, apart from non-participation and unemployment. These include wage employment in the private sector, wage employment in the public sector, self-employment and unpaid family work.

Given the number of discrete choices involved, the multinomial logit model provides the most suitable econometric approach to estimating the probabilities that an individual will be found in each employment segment. Let y denote a random variable taking on the values {1, 2, ..., J}, where J is a positive integer representing the assignment to a particular labour market segment. Let x denote a set of conditioning variables such as the individual, household and regional characteristics that might have a bearing on the employment segment determination. We are interested in how changes in the elements of x affect the response probabilities, $P(y=j|x)$, j=1, 2, ..., J. Given that probabilities must sum to one, $P(y=1|x)$ is determined once we know the probabilities for j=2, 3, ..., J

(Wooldridge, 2002, p. 497). If x is a 1*K vector with first-element unity, the multinomial logit model has response probabilities

$$P(y = j|x) = \frac{e^{x\beta_j}}{1 + \sum_{h=1}^{J} e^{x\beta_h}}, j = 2, 3, \dots, \qquad J \tag{1}$$

where β_j is K*1, j=2, ..., J. Because the response probabilities must sum to unity,

$$P(y = 1|x) = \frac{1}{1 + \sum_{h=1}^{J} e^{x\beta_h}}, j = 2, 3, \dots, \qquad J \tag{2}$$

In our particular case, modelling employment sector determination results in a six-way multinomial logit model. Schooling and other characteristics are expected to determine differently the fall into each labour market sector. We run models for men and women separately, and control for place and province of residence, among other things. In order to ensure model identification the parameter vector associated with one of the employment sectors must be set to zero (Cameron & Trivedi, 2009, p. 484). For both men and women we normalise by setting the parameter vector associated with non-participation equal to zero. The coefficient estimates are then interpreted with respect to this category.

Parameter estimates from multinomial logit models of sector of employment determination for men and women are presented in Table 3.1. Wald tests of the null hypothesis that all regression coefficients associated with each of the twelve pairs of employment segments are equal to zero were computed and rejected at 0.0001 level. This suggests that the employment sectors need not be combined, i.e., the six-way split provides a useful view of the labour market structure in Mozambique.

The coefficient estimates in Table 3.1 do not indicate the impact of a change in an explanatory variable on the probability of entering a labour market segment. Instead, they are interpreted as the impact of the associated independent variable on the log odds of the particular labour market segment relative to non-participation (the base category). Because interpretability of the coefficients in this way is rather difficult, we compute marginal effects of changing the values of the covariates on the probability of observing a labour market segment determination outcome while the other covariates are kept fixed. For categorical covariates such as level of education completed, a marginal effect shows how the probability of entry into a particular employment sector is predicted to change as the variable shifts from no formal schooling (the reference group) to, say, full primary school. For continuous variables (such as age) it measures the instantaneous rate of change in the probability of entry in an employment segment as age changes.

Very often studies in which the dependent variable is categorical and the samples are large report marginal effects at means (of other covariates). However, whenever samples are small or moderate they can largely deviate from the marginal effects at values other than the mean. Therefore, according to Greene, current practice supports computing instead the sample average of individual marginal effects to obtain the overall marginal effect (Greene, 2003, p. 668).

Table 3.2 and Table 3.3 report the average marginal effects from changes or shifts in the covariates for the men and women sample, respectively. The estimated marginal effects of schooling on employment sector allocation presented in the tables are broadly intuitive and in line with the descriptive statistics. Controlling for other determining factors of employment sector choice, for both men and women more education compared to no formal schooling increases the chances of employment in the public sector. For example, if a man has completed full primary education the average probability of public sector employment rises by 0.07 and by 0.30 for a man with full secondary education. The corresponding figures for women are 0.06 and 0.53, respectively.

As the number of years of education completed rise women are also more likely to be employed in the private sector. However, the marginal effects for men within the private sector are negative and statistically significant, except for those with some primary education only. This suggests that men who have completed full primary school or further levels are less likely to work in the private sector than men without any formal schooling.

This contrasting impact of education across gender might result from the occupational choices open to each sex. The data set indicates that more than half (54%) of men in the private wage sector are (blue collar) labourers, craftsmen or unskilled workers, while a very low proportion are, say, clerks (4%). In contrast, the corresponding proportions for women are 23% in the low skilled occupations and 14% in clerical occupations. Clearly, contrary to the case of men, entry into the type of occupations open to, or chosen by, women requires relatively higher levels of education.

Table 3.1 – Multinomial logit estimates of employment segment determination (Non-participation is the base category); Note: model estimation accounts for survey design; * p<0.1, ** p<0.05, *** p<0.01; Model includes dummies controlling for province of residence.

	Men					Women				
Variables	**Public Sector**	**Private Sector**	**Self-Employment**	**Unpaid Family Work**	**Unemployment**	**Public Sector**	**Private Sector**	**Self-Employment**	**Unpaid Family Work**	**Unemployment**
Education level completed										
Some primary	1.0968***	−0.2822	−0.2209	−0.1454	0.1292	2.0482***	1.2402***	0.0788	−0.0747	0.5256***
Full primary	2.2175***	−0.5859**	−0.5792**	−0.5918**	0.0172	4.0180***	1.5815***	−0.5585***	−0.6178***	0.4860***
Some secondary	2.4141***	−1.3190***	−1.7347***	−1.6763***	−0.3053	4.9593***	1.4536***	−1.2263***	−1.4948***	0.3875**
Full secondary	3.2252***	−0.9474***	−1.9582***	−1.9703***	0.2210	6.3054***	2.5332***	−2.3384***	−3.1204***	0.2618
Tertiary education	2.8652***	−1.5882***	−2.5095***	−26.9029***	−0.9089**	5.5380***	2.4016***	−2.9574***	−30.6927***	0.2199
Age	0.7801***	0.6040***	0.5641***	0.2985***	0.4047***	0.6424***	0.5533***	0.3549***	0.2188***	0.2767***
Age squared	−0.0095***	−0.0080***	−0.0073***	−0.0040***	−0.0055***	−0.0076***	−0.0073***	−0.0044***	−0.0028***	−0.0044***
Rural	0.2804	−0.4330**	1.2264***	1.7165***	−0.6330***	0.6674***	−0.0122	1.1373***	2.2334***	−1.0812***
Household headship	1.4310***	0.9646***	1.5246***	−2.6147***	−0.1718	2.1505***	1.5135***	1.8247***	−2.9052***	0.8935***
No. of children in the household	0.0038	−0.0054	−0.0028	0.0311	0.0112	−0.0024	−0.0317	0.0536**	0.1257***	0.0283
No. of adults in the household	−0.1121***	−0.0568*	−0.1053***	−0.0570*	−0.0511**	0.0321	−0.0251	−0.1519***	−0.1405***	−0.0400*
Married	1.4788***	1.1632***	1.4715***	0.9614***	0.4305**	0.9134***	0.0089	1.3444***	1.2635***	1.0719***
Constant	−16.7794***	−8.2403***	−8.5490***	−5.9399***	−5.1095***	−18.2387***	−11.3890***	−6.0021***	−5.0401***	−3.8405***
No. of observations	15,915					20,150				
Pseudo R-squared	0.3841					0.3637				
Log-likelihood	−3,390,430					−4,109,178				
Wald χ^2 (d.f)	20,284.26 (110)					19,719.47 (110)				
Prob > χ^2	0.0000					0.0000				

For both sexes schooling is significantly and negatively associated with self-employment choice, except for less than full primary completion. This means that for both men and women having completed full primary education or further schooling lowers the chances of being self-employed relative to men and women without any formal schooling, respectively, and the effects are stronger amongst women. Similar patterns are observed with respect to unpaid family work, and again education seems to have a much stronger effect on women than men. For instance, in the case of men, the coefficients on some primary and full primary school are not statistically different from zero. Further, having completed secondary education lowers the average probability of unpaid family work by 0.09 only. For women, however, even some primary education is associated with (statistically significant) lower chances of unpaid family work, and the corresponding fall in average probability for full secondary education relative to no schooling is as high as 0.24.

Though significant (except for men at the tertiary education level) and positive the average marginal effects of schooling associated with unemployment are very small for both men and women. For instance, for both sexes having completed full secondary school increases the average probability of unemployment only by 0.07 for men and 0.04 for women. The positive relationship between schooling and the likelihood of unemployment might result from a mismatch between skills demanded by the private wage sector and self-employment activities, on the one hand, and the type of jobs and remunerations educated workers were hoping to get, on the other hand. They may perhaps rather stay unemployed than enter any of these employment segments while waiting for the possibility of entry into the wage public sector (Krishnan, Sellasie, & Dercon, 1998, p. 17).

Previous studies in developing countries applying similar methodologies as in this study also find that education significantly determines employment segment selection and are generally consistent with our results. In a Guinean study, Glick & Sahn (1997) split the labour market into three employment segments, apart from non-participation: Self-employment, private wage employment and public wage employment. They concluded that for both men and women more education lowers the chances of self-employment while it strongly increases the likelihood of public sector employment. Further, only for women did education impact positively on the likelihood of private sector employment. In an Ivorian study, Vijverberg (1993) found that education raised the probability of wage employment while it lowered the probability of non-agricultural self-employment, for both sexes. In Ghana, Glewwe (1991) found that education was positively associated with wage employment, and among the wage employed, the better educated were more likely to be in the public sector than in the private sector. Using Kenyan data, Wambugu (2002) found that additional schooling lowered the probability of agricultural and informal sector employment, but increased the chances of private (for females only) and public sector employment (for both males and females).

Krishnan, Sellassie & Dercon (1998) used a combination of cross-section and panel data based on three household surveys conducted in 1990, 1994 and 1997, to study employment allocation in the Ethiopian urban labour market. In close similarity to this study, they split the labour market into five employment segments: the public sector, the private sector, self-employment, unemployment and non-participation. In sum, for both men and women they found that schooling increased the likelihood of employment in the public sector. The same was found to be true for men in the private sector, but not entirely true for women. Women with completed primary education were less likely to work in the private sector relative to uneducated women. But women with at least secondary education completed (relative to no formal schooling) were more likely to work in the private sector. Lastly, for both men and women, schooling in comparison to no formal schooling, was found to lower the chances of self-employment.

All in all, as this study, developing country studies employing similar methodologies end up concluding that education, relative to no formal schooling, increases the chances of wage employment and lowers the chances of employment outside the wage labour segment.

The impact of age on employment segment allocation seems to be concave for both men and women and across most sectors. The older a man or woman is the more likely he or she is to be in public, private or self-employment, but the rate of change tends to flatten as he or she gets older. With respect to unpaid family work the reverse tends to occur, however. Individuals become less likely to fall into this employment segment as they grow older, but at a decreasing absolute rate.

Rural men are less likely to fall into non-participation, to be unemployed and to work in the public and private wage employment segments, but more likely to be self-employed and unpaid family workers relative to urban men. The same occurs with rural women, except that they do not significantly differ from urban women with respect to chances of public sector employment and self-employment.

Headship of household is positively associated with a mode of employment that brings resources into the household. In other words, for both men and women, being a household head relative to being an ordinary household member significantly increases the chances of working in the public, private (for women only) and self-employment segments, but significantly lowers the chances of falling into non-participation, unemployment (for men only) and unpaid family work. A conspicuous result is that household headship seems to drive both men and women into self-employment, as the average marginal effects are much larger for this segment than for the public and private sectors. Given the lower chances of entry into the public and private sectors, self-employment seems to be the measure of last resort for those with the responsibility of providing external resources to the household.

The number of children in the household does not have any significant impact on employment segment determination for men. But for women, though the average marginal effects are rather small, a greater number of children in the household makes them less likely to fall into non-participation and private wage segment, while it increases their chances of falling into unpaid family work. This makes evident the gender biases with respect to child rearing in such countries. Results such as these are not uncommon in the literature. For instance, Bhorat & Leibbrandt (2001) also found that the number of children in the household was insignificant in determining African men's labour force participation decisions in South Africa, but significantly deterred women from participation.

The marginal effects of a greater number of adults in the household are also very small for both sexes. Nevertheless, it significantly lowers the chances of self-employment for both men and women, and increases the likelihood of unemployment and public sector work for women. Lastly, married men are significantly more likely to fall either into public sector or self-employment segments of the labour market, and less likely to be unemployed or to fall into non-participation than unmarried men. Married women, in contrast, are significantly less likely to work in the private sector segment, but more likely to be self-employed, unpaid family workers or unemployed. These results are consistent with African societies where women have to balance domestic responsibilities with the need to augment household income.

Now we describe the model prediction by education level completed for men and women separately. We use the method of predictive margins in which we vary education across the whole sample of men and women and then average the predictions within each sex sample. In other words, we first pretend that all people in our men (women) sample have no formal schooling but hold their other characteristics constant. Then we calculate their probability of falling into each employment segment. Then we pretend that all people in our men (women) sample have only some primary schooling, still holding their other characteristics constant, and calculate the probabilities of each outcome. We repeat this process up to tertiary education. The difference between those six sets of calculated probabilities for men and women within each labour market segment, then, is the difference associated with schooling, holding other characteristics constant, i.e., the difference correspond to the average marginal effects of education presented in Table 3.2 and Table 3.3.

Table 3.4 reports the predicted probabilities of entry into each labour market segment. Naturally, they mirror the average marginal effects of education discussed before. Clearly, for both sexes education is the key to public service employment. For those with less than primary education completed, chances of public sector employment are extremely low. But those with completed full secondary or further education are very likely to fall into this employment segment.

Table 3.2 – Multinomial logit average marginal effects of employment segment allocation for men; Note: * p<0.1, ** p<0.05, *** p<0.01; Model includes dummies controlling for province of residence.

	Men					
Variables	**Public Sector**	**Private Sector**	**Self-Employment**	**Unpaid Family Work**	**Unemployment**	**Non-Participation**
Education level completed						
Some primary	0.0140***	−0.0165	−0.0198	0.0012	0.0144*	0.0067
Full primary	0.0686***	−0.0270**	−0.0646***	−0.0212	0.0208**	0.0233**
Some secondary	0.1791***	−0.0365***	−0.2032***	−0.0605***	0.0385***	0.0826***
Full secondary	0.3045***	−0.0276	−0.3182***	−0.0934***	0.0667***	0.0680***
Tertiary education	0.3570***	−0.0343*	−0.3165***	−0.1785***	0.0177	0.1546***
Age	0.0071***	0.0118***	0.0235***	−0.0172***	−0.0005	−0.0247***
Age squared	−0.0001***	−0.0002***	−0.0003***	0.0002***	0.0000	0.0003***
Rural	−0.0116*	−0.1457***	0.1584***	0.1179***	−0.0709***	−0.0482***
Household headship	0.0149***	0.0134	0.3095***	−0.2646***	−0.0469***	−0.0263***
No. of children in the household	0.0001	−0.0009	−0.0024	0.0033	0.0004	−0.0006
No. of adults in the household	−0.0009	0.0023	−0.0079**	0.0022	0.0005	0.0039***
Married	0.0087**	0.0067	0.0940***	−0.0184	−0.0335***	−0.0576***

Table 3.3 – Multinomial logit average marginal effects of employment segment allocation for women; Note: * p<0.1, ** p<0.05, *** p<0.01; Model includes dummies controlling for province of residence.

	Women					
Variables	Public Sector	Private Sector	Self-Employment	Unpaid Family Work	Unemployment	Non-Participation
Education level completed						
Some primary	0.0060***	0.0092***	−0.0076	−0.0225***	0.0230***	−0.0082
Full primary	0.0706***	0.0209***	−0.1042***	−0.0404***	0.0405***	0.0126
Some secondary	0.2149***	0.0226***	−0.2296***	−0.1049***	0.0537***	0.0434***
Full secondary	0.5309***	0.0587***	−0.4557***	−0.2386***	0.0381**	0.0666***
Tertiary education	0.4803***	0.0839***	−0.4551***	−0.3023***	0.0713***	0.1219***
Age	0.0061***	0.0037***	0.0145***	−0.0085***	0.0006	−0.0163***
Age squared	−0.0001***	−0.0000***	−0.0002***	0.0001***	−0.0001***	0.0002***
Rural	−0.0002	−0.0107***	0.0095	0.1902***	−0.1233***	−0.0656***
Household headship	0.0223***	0.0064*	0.4767***	−0.4551***	−0.0057	−0.0447***
No. of children in the household	−0.0008	−0.0011*	−0.0031	0.0091***	−0.0008	−0.0034***
No. of adults in the household	0.0026***	0.0008	−0.0088***	−0.0028	0.0026**	0.0056***
Married	−0.0014	−0.0187***	0.0525***	0.0180*	0.0119***	−0.0624***

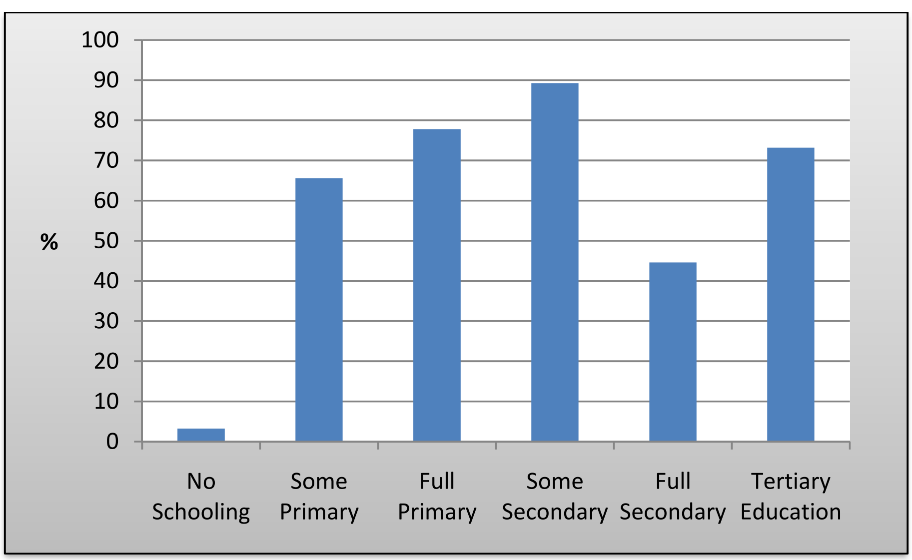
Private sector average entry probabilities are very low, notably for women. For men they tend to fall with educational attainment while the opposite happens for women. But at all educational levels women are less likely to be employed in private sector than men. Men and women with no formal schooling are very likely to be self-employed. The average probability of falling into this sector is 0.56 for both sexes. In contrast, for men and women with university education the probabilities of self-employment are, correspondingly, 0.24 and 0.10 only. The chances of doing unpaid family work also drop with educational attainment. For instance, both men and women with tertiary education have a zero average probability of falling into this segment.

As discussed before, the chances of falling into unemployment tend to increase with schooling, notably for women, representing a possible mismatch between demand and supply of skills, and/or the preference for queuing for public sector jobs as educational attainment increases. The chances of non-participation for both men and women also rise with education. As shown in Figure 3.1 this pattern might be due to the fact that the proportion of non-participant men and women that are still in school also tends to increase with educational attainment.

Table 3.4 – Predicted labour market sector entry probabilities of men and women, by level of education completed; Note: all probabilities significant at the conventional levels of significance, 1%, 5%, or 10%.

	Education Level Completed / Labour market sector	No Schooling	Some Primary	Full Primary	Some Secondary	Full Secondary	Tertiary Education
Men	Public Sector	0.0054	0.0194	0.0741	0.1845	0.3099	0.3624
	Private Sector	0.1410	0.1245	0.1140	0.1044	0.1134	0.1067
	Self-Employment	0.5605	0.5407	0.4959	0.3573	0.2423	0.2441
	Unpaid Family Work	0.1785	0.1798	0.1573	0.1181	0.0851	0.0000
	Unemployment	0.0535	0.0679	0.0743	0.0919	0.1202	0.0712
	Non-Participation	0.0611	0.0678	0.0844	0.1437	0.1291	0.2156
Women	Public Sector	0.0010	0.0070	0.0716	0.2159	0.5320	0.4813
	Private Sector	0.0050	0.0142	0.0259	0.0275	0.0637	0.0889
	Self-Employment	0.5581	0.5505	0.4539	0.3285	0.1024	0.1030
	Unpaid Family Work	0.3023	0.2798	0.2619	0.1973	0.0636	0.0000
	Unemployment	0.0594	0.0824	0.0999	0.1131	0.0975	0.1307
	Non-Participation	0.0742	0.0660	0.0868	0.1176	0.1408	0.1961

Figure 3.1 – Percentage of non-participants currently in school.



4. EARNINGS AND EDUCATION

Education is conventionally understood to bring many benefits. In this study we are interested in its economic benefits, specifically labour market benefits. As we saw in the foregoing section, one of the labour market benefits of education is to increase the chances of wage employment in the public and private sectors. Education is also believed to lead to increases in labour productivity and earnings (Cohn & Geske, 1990). Acknowledging these benefits, several methods have been developed to quantify their significance. The returns to education approach gained relevance and is based on the premise that education results in direct measurable returns to the individual and society.

Two methods stand out as the most important for calculating rates of returns to education (Psacharopoulos, 1981; Johnes, 1993, p. 28). The first method – the algebraic method – consists in finding the discount rate that equates the stream of education benefits to the stream of education costs, where these costs include only the opportunity costs of staying in school (foregoing earnings), at a given point in time. In other words, if Y stands for labour earnings, c for cost years, n for benefit years, and h and s are subscripts for higher and secondary education, in that order, the rate of return (r) is found by solving the following equation for r:

$$\sum_{t=1}^{n}(Y_h - Y_s)_t(1+r)^{-t} = \sum_{t=1}^{c}(Y_s)_t\,(1+r)^{t} \qquad (3)$$

The problem with this approach is that it requires detailed data on age-earnings profiles by educational level, which may not be readily available, notably in the case of developing economies(Johnes, 1993, p. 29).

The second method of calculating returns to education is the earnings function approach. It started in the 1970s with the publication of Mincer's (1974) book and since then it became widely used(Psacharopoulos, 1981, p. 321). This approach relies on the human capital theory that links investment in education to higher labour productivity and thus higher earnings. In his model Mincer explains inequality in earnings by focusing on schooling and post-schooling investment as the central explanatory variables. Schooling investment is measured by years of education completed and post-schooling investment is given by labour experience, in the absence of more specific measures. Formally, his basic model is described by the following equation:

$$\ln Y = a + bS + cE + dE^2, \quad (4)$$

where Y represents earnings, S stands for years of education, E stands for years of work experience, and a, b, c, and d are parameters to be estimated. Equation (4) can be estimated using standard OLS techniques and cross-section data. Using U.S. census data of 1960 and different versions of Equation (4), Mincer found that schooling and post-schooling investment accounted for close to two-thirds of the inequality of earnings of adult, white, urban men(Mincer, 1974, p. 96).

The model in its simplest form is a powerful tool for explaining earnings inequality, but it is often expanded to control for other factors that might have effects on worker pay, such as worker occupation, gender, sector of employment, and location, among other things. It is convenient to look at the raw coefficients of education in the extended function to report returns to education, but some scholars argue that they are in fact wage effects (Psacharopoulos & Patrinos, 2004, p. 116).

4.1 Standard Earnings Functions

We nevertheless follow the Mincerian approach to analyse labour earnings outcomes associated with schooling across the different employment segments for men and women separately. In other words, we wish to know, separately for men and women, whether schooling and post-schooling investments yield equivalent returns in three employment segments for which reliable earnings data are available: the public wage sector, private wage sector and self-employment.

Aiming for more flexibility we amend Equation (4) in order to estimate separately the Mincerian returns to some primary, full primary, some secondary, full secondary and some tertiary or further education. In other words, rather than using continuous education we use dummy variables for the level of education completed. Experience is approximated by age.

The dependent variable is log of hourly earnings. Apart from education and age, we also include dummies to control for place and province of residence. The omitted categories are, respectively, urban and Maputo City. Control dummies for occupational choice and industry were also initially included. However, due to the relatively small sample size resulting from splitting the labour market across sectors and gender, the number of individuals in some occupations and industries was too small to produce reliable estimation results. Therefore, these controls were dropped from the regressions.

Ordinary Least Squares (OLS) estimates from earnings equations for men and women in public wage employment, private wage employment, and self-employment are presented in Table 4.1. The education point estimates are in line with the descriptive statistics reported above in Table 2.2. Overall they are positive and increase with attainment across all sectors.

To better scrutinise the impact of education on earnings we calculate percentage changes in earnings associated with educational attainment and report them in Table 4.2 (using the statistically significant education dummies only). In semilog specifications such as these whereby education is represented by dummies for each level of completion, the percentage change in earnings due to schooling is calculated as 100*(e^c – 1), where c stands for the education dummy coefficient at each level (Glick & Sahn, 1997, p. 819).

Across all labour market sectors and controlling for other factors, hourly earnings of men and women with only some primary education do not differ significantly from uneducated men's and women's hourly earnings, except for women in the public sector, whose hourly earnings increase by 24%. The impact of schooling on hourly earnings matters only from full primary completion onwards, except for women in the private sector, for whom primary school does not significantly increase hourly earnings, and self-employed women for whom secondary education and above does not put them in a significantly better off position compared to uneducated women.

Returns to full secondary schooling are much higher in the private sector, particularly for women. Controlling for other determining factors of earnings, men and women in this sector earn, on average, five and nine times better than their uneducated counterparts, respectively. In contrast, both men and women with full secondary school in the public sector, and men in self-employment, earn only four times better than men and women without any formal schooling, in the respective sector.

In the private sector, women with some or completed tertiary education[5] have higher returns to education than men at the same education level. While women's hourly earnings at this level are fifteenfold that of women with no formal schooling, men's hourly earnings are sevenfold only that of uneducated men. In the public sector, things

[5] To keep an acceptable sample size, the dummy on tertiary education was set equal to 1 for those who completed at least 13 years of education. However, this has the disadvantage of not differentiating returns to, say, a completed degree from returns to just one year of university education.

are in favour of men, but more balanced. Tertiary educated men and women get, respectively, eightfold and sevenfold the hourly earnings uneducated men and women get.

Going back to Table 4.1, the concavity of experience proxied by age is seen across all sectors, but it is statistically significant at the conventional levels only for men. This means that men's hourly earnings increase with experience but at a decreasing marginal rate.

Table 4.1 – OLS earnings functions; Note: dependent variable is log of hourly earnings; model estimation accounts for survey design; * p<0.1, ** p<0.05, *** p<0.01.

Employment Sectors	Public Sector		Private Sector		Self-Employment	
Variables	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Education Level Completed						
Some primary	0.2774	0.2171**	0.0874	0.1199	0.1428	−0.0072
Full primary	0.5258**	0.5998***	0.2882***	0.3144	0.3082**	0.4719***
Some secondary	0.9131***	0.9395***	0.6968***	0.9556***	0.5748***	0.7382***
Full secondary	1.3780***	1.4358***	1.6370***	2.2018***	1.3757***	0.9469
Tertiary education	2.0959***	2.0157***	1.8905***	2.7356***	1.9004***	0.9803
Age	0.0368**	−0.0029	0.1134***	0.0496*	0.0918***	0.0307
Age squared	−0.0003*	0.0002	−0.0012***	−0.0003	−0.0011***	−0.0004
Rural	−0.0582	0.0169	−0.1835**	−0.1957	−0.1317	0.1210
Province						
Niassa	−0.2352**	−0.0215	−0.4912***	−0.2241	−0.6949***	−0.2937
Cabo Delgado	−0.0046	−0.1854	−0.0772	0.2243	−0.4719***	−0.1254
Nampula	−0.0969	−0.0829	−0.7262***	−0.3237	−1.1367***	−0.8512***
Zambezia	−0.1855**	−0.1241	−0.4158***	−0.4657	−0.8890***	−0.1183
Tete	−0.1178	−0.0495	−0.3222***	0.4213*	0.0617	0.7098***
Manica	−0.2048**	0.0541	−0.2107**	0.3304	−0.5008***	−0.1960
Sofala	−0.1927**	0.0314	−0.3412***	−0.2016	−0.5610***	−0.1372
Inhambane	−0.0808	−0.0860	−0.0740	−0.5407***	−0.2599*	0.3124***
Gaza	−0.1497*	−0.0022	−0.1669*	−0.2097	−0.0355	0.1894
Maputo Province	−0.0726	−0.1742	−0.0357	0.0138	0.0347	0.8195***
Constant	0.9486***	1.6180***	−0.6707***	−0.0274	0.5824*	0.7829*
No. of observations	1,053	438	2,177	440	2,455	1,655
R-squared	0.4607	0.4249	0.4216	0.6351	0.1477	0.1476

Concerning the control variables, the rural dummy is overall negative but statistically significant for men in the private sector only. Thus only in this sector, rural individuals (men) earn less than their urban counterpart, on average. The provincial dummies indicate that in general hourly earnings are significantly higher in Maputo City (the omitted category), except for women in the private wage sector and self-employment in Tete province, and self-employed women in Inhambane and Maputo Province who tend to earn significantly better than Maputo City's women.

The R-squared, a measure of goodness of fit, is reasonably high in the private and public wage sectors, indicating good model fit, but it is relatively low in the self-employment segment.

Table 4.2 – Percentage change in earnings associated with education; Note: calculated from the table above as $100*(e^c – 1)$, where c stands for the education dummy coefficient at each level.

Employment Sectors	Public Sector (%)		Private Sector (%)		Self-Employment (%)	
Categories	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Some primary	-	24	-	-	-	-
Full primary	69	82	33	-	36	60
Some secondary	149	156	101	160	78	109
Full secondary	297	320	414	804	296	-
Tertiary education	713	651	562	1442	569	-

4.2 Multinomial Logit Selection Bias-Corrected Earnings Functions

In the previous section, we assumed that workers were randomly assigned to each employment sector. However, if this is not the case and we simply employ OLS regression techniques, there is potential for sample selection bias in the coefficients estimates of the earnings functions. This occurs because there might be unobserved worker characteristics that affect both the determination of employment sector and earnings, i.e. there is a correlation between the error terms of the earnings functions and the process that determines employment sector choice. If this correlation is statistically significant then the OLS returns to education reported in Table 4.1 would be biased.

To get around sample selection bias studies often employ Heckman's (1979) two-step procedure. The first step involves modelling the probability of observing the outcome variable, say, the probability of being wage employed (i.e., the selection equation). The wage employment estimates are used to derive estimates of the inverse Mills ratio (Lambda). The Lambda is then included in the right-hand side of the wage equation (earnings function), thus making it conditional on selection into wage employment. Statistical significance of Lambda indicates that the wage equation needed correction for sample selection.

However, Heckman's two-step procedure allows for binomial outcomes only. In other words, the individual is either wage employed or not. Given the multinomial nature of our particular employment selection model, a different procedure that takes this multitude of options into account must be applied instead.

Bourguignon, Fourier & Gurgand (2004), formally published by Bourguignon, Fourier & Gurgand (2007) in the *Journal of Economic Surveys*, reviewed the set of methods available in the literature on selectivity bias correction using multinomial specifications and used Monte-Carlo experiments to test their performance.

Lee (1983) proposed an approach whereby Heckman's (1979) two-step procedure is generalised and selectivity is modelled as a multinomial logit. His approach is simple and the selectivity correction term requires the estimation of one parameter only. But his correlation assumption is highly restrictive for most empirical cases. It implies that unobservable determinants of, say, self-employment against any other employment alternative should be correlated in the same direction with unobservable determinants of self-employment hourly earnings (Bourguignon, Fournier, & Gurgand, 2004, p. 6; Schmertmann, 1994).

Dubin & McFadden (1984) also proposed sample selection correction based upon multinomial logit models, in which M – 1 correction parameters are added to each OLS earnings function, where M stands for the number of employment sector choices. In contrast to Lee (1983), they make no assumption regarding the correlation of the unobservable determinants of the wage and selection equations, but their assumption regarding the linearity of the wage equation's error term restricts the class of allowed distributions for this error term (Bourguignon et al., 2004, p. 7). This might be problematic when the independence of irrelevant alternatives (IIA)[6] property of the multinomial logit model is not satisfied (Koch & Ntege, 2008).

According to Dahl (2002) whenever the multinomial logit model is warranted for sample selection correction, a large number of correction parameters has to be estimated, thus making practical implementation intractable. Therefore, he proposed a non-parametric approach wherein the set of probabilities used to correct the wage equation for selectivity bias are restricted to a chosen subset of particular interest. He

[6] The IIA assumption states that the ratio of the probabilities of any two alternatives being chosen is independent of the number of available alternatives.

went further and proposed a special case of his approach in which the probability of falling into, say, wage employment in the public sector, is the only information required to correct the public sector OLS earnings functions for sample selection bias. In contrast to Lee's (1983) approach, Dahl's allows for any sign structure between the correlation of the error terms of the wage and selection equations (Bourguignon et al., 2004, p. 9). But in this procedure the correction terms have no structural interpretation (Koch & Ntege, 2008, p. 18).

Bourguignon et al. (2004) suggested a modification of Dubin & McFadden's (1984) model which allows the error term of the wage equation to be normal. Their Monte-Carlo experiments comparing all of the surveyed approaches to correct for sample selection bias suggest that Dubin & McFadden's (1984) model performs relatively better. But, when the IIA assumption of the multinomial logit model is not satisfied, their version of the model is to be preferred (Bourguignon et al., 2004, p. 17).

4.2.1 The model

Our preferred approach to correct the OLS earnings functions for selectivity bias is the modified version of Dubin & McFadden's (1984) model. Below we formally describe the estimation problem and correction procedure. Consider the following two models:

$$y_1 = x\beta_1 + \mu_1 \tag{5}$$

$$y_j^* = z\gamma_j + \eta_j, \quad \text{j = 1 ... M} \tag{6}$$

where the vector x contains all determinants of the variable of interest y_1 (which in our case stands for hourly earnings within a specific sector), and the vector z denotes the observable exogenous factors influencing the choice of the employment sector. The terms μ_1 and η_j are idiosyncratic shocks. The former satisfies the following conditions: $E(\mu_1|x,z)=0$ and $V(\mu_1|x,z)=\sigma^2$. The subscript j is a categorical variable describing the assignment of the workers among M labour market sector alternatives.

A worker falls into option j among M alternatives to maximise y_j^*, which is a latent variable capturing the discrete observation of whether he or she works in that particular sector or not. In other words, hourly earnings in, say, the public sector are only observed if the individual is assigned to work there, and this only happens if the individual utility derived from selecting to work in that sector is higher than the utility derived from working in alternative sectors. Formally, y_1 is only observed if:

$$y_1^* > \max_{j \neq 1} y_j^* \tag{7}$$

If we define:

$$\varepsilon_1 = \max_{j \neq 1}(y_j^* - y_1^*) \tag{8}$$

$$\varepsilon_1 = \max_{j \neq 1}(z\gamma_j + \eta_j - z\gamma_1 - \eta_1),$$

it implies that ε_1 should be negative. Assuming that the η_j's follow the IIA property, McFadden (1973) has shown that the discrete employment sector determination component can be consistently estimated with a multinomial logit model. Simply put:

$$P(\varepsilon_1 < 0|z) = \frac{e^{z\gamma_1}}{\sum_j e^{z\gamma_j}}$$

While estimating consistent parameters (γ_j) of the latent variable model (6) is straightforward, the problem lies in estimating the parameter β_1 of the wage equation (5). The disturbance terms of the wage equation (5) and selection equation (6) may not be independent. As a result the error term and explanatory variables of the wage equation (5) become correlated, leading to inconsistent OLS estimates of β_1 (Bourguignon et al., 2004, p. 4).

Given the above estimation problem, Bourguignon et al. (2004) propose the following assumption regarding the correlation (r_j^*) between the error terms of the wage and selection equations:

$$E(\mu_1|\eta_1 \ldots \eta_M) = \sigma \sum_{J=1 \ldots M} r_j^* \eta_j^* \tag{9}$$

With the help of this assumption, they show that the parameter β_1 of the wage equation (5) conditional on choosing employment sectors j=1 can be consistently estimated by the following equation:

$$y_1 = x_1\beta_1 + \sigma\left[r_1^* m(P_1) + \textstyle\sum_{J=2 \ldots M} r_j^* m(P_j)\frac{P_j}{(P_j - 1)}\right] + w_1, \tag{10}$$

where w_1 is an independent error term, P_1 the probability that alternative 1 will be preferred, and $m(P_j) = \int J(v - \log P_j)g(v)dv$, $\forall j$. This method adds selection correction terms to each wage equation. Unlike Heckman's (1979) procedure whereby there is one choice involved only, in this case there are many alternatives, which result in as many selection correction terms as there are alternatives. These terms are consistent estimators of the conditional expected values of the residuals derived from the first stage multinomial selection equation. In the Heckman's two-step procedure notation, the coefficients on these variables constitute the lambdas.

Parametric identification of the selection equation requires the addition of exclusion restrictions, i.e., variables that affect the chances of employment in a particular sector but not hourly earnings. In this study we use four such variables: the number of children in the household, the number of adults in the household, marital status and household headship status of the worker.

4.2.2 Empirical results

As in the case of the standard OLS earnings functions reported above we estimate separate models for men and women in the public wage employment, private wage employment, and self-employment. In the first stage (selection) equations we consider those working in the above mentioned sectors with or without earnings data[7] as well as unemployed individuals[8]. Once the selection correction terms are extracted we include them in the second stage (wage) equations, thus correcting them for sample selection bias.

The models are estimated using a Stata (Statacorp, 2009) user-written ado-file named SELMLOG developed by Bourguignon et al. (2004), and results are reported in Table 4.3[9]. Within the public sector the selection correlations reported at the bottom of the table show that the unobserved determinants of hourly earnings are correlated in a statistically significant manner with the unobservables of both the process that determines selection into public sector and into unemployment, for men, and with the process that determines selection into self-employment, for women. Within the private sector, the earnings function's error term is significantly correlated with the self-employment equation's error term, for men. For women the unobserved determinants of hourly earnings are significantly correlated also with the unobserved determinants of public sector employment. The selection correlation terms within the self-employment sector are not statistically significant at the conventional levels of significance, except for

[7] This introduces another type of sample selection bias into our regression estimates, since unobserved characteristics of those who reported earnings might differ from those who did not. Nevertheless, for the sake of avoiding further complications to our analysis, in this paper we ignore this issue.

[8] Non-participants were excluded since they do not participate actively in the labour market. Unpaid family workers were excluded because none of them had tertiary education. Including them in the regressions would result in (unreliable) zeros for the coefficient estimates of the "tertiary education" dummy.

[9] We report the OLS selectivity corrected earnings functions only, since results and interpretations of the selection equations are qualitatively similar to those seen above in Table 3.1.

men, for whom the unobserved determinants of earnings are significantly correlated with the unobserved determinants of the process that determines selection into unemployment.

Thus, the model results indicate that across the three employment sectors and for both sexes´ earnings OLS functions, except for self-employed women, sample selection bias was indeed a problem and was corrected for by the inclusion of selectivity correction terms[10]. Sample selection bias was not found to be a major problem in the earnings function corresponding to self-employed women. Hence, the hourly earnings of a woman with average characteristics in the self-employment sector would not differ significantly from the hourly earnings of a woman randomly selected into the various sectors. Therefore, in this case the OLS estimates should be trusted.

The schooling estimates show a similar pattern as the OLS ones, whereby educated individuals earn significantly better than uneducated ones across all labour market sectors. Like we did above, we calculate percentage changes in earnings associated with educational attainment and report them in Table 4.4. Compared to a similar table in the OLS earnings functions case, this table shows that the correction for selectivity bias resulted in greater returns to education for men and women in the public sector and for men only in the private sector, but in slightly smaller returns to education for women in the private sector and for self-employed men. Further, for men even some primary education provides greater returns than no schooling. Furthermore, men's returns to education are greater than women's within the public sector, but the inverse occurs within the private sector, at least from secondary education onwards. Lastly, the convexity of the returns to education is very clear across all sectors, i.e., returns to education are greatest at the tertiary education level, followed by secondary schooling and then primary schooling.

[10] The fact that the correction terms for men´s earnings functions within the private sector and within self-employment are not statistically significant might suggest that there selectivity bias was not too much of a problem. In fact, this is also given by the selection correlations within the respective earnings functions which are only statistically significant at 10% level of significance.

Table 4.3 – Selectivity-corrected earnings functions; Note: dependent variable is log of hourly earnings; Statistical significance based on bootstrapped standard errors (1,000 repetitions); * p<0.1, ** p<0.05, *** p<0.01.

Employment Sectors	Public Sector		Private Sector		Self-Employment	
Variables	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Education Level Completed						
Some primary	0.7722***	0.3462	0.1458*	0.0944	0.2276**	0.0202
Full primary	1.4272***	0.9898***	0.3634***	0.3657*	0.5080***	0.1671
Some secondary	2.3630***	1.6164***	0.9070***	0.9414***	0.6176***	0.4087
Full secondary	3.0688***	2.3879***	1.9226***	2.1775***	1.2185***	0.4208
Tertiary education	3.7752***	3.0692***	2.2387***	2.7297***	1.5557***	0.1877
Age	0.0795**	– 0.0048	0.0975***	0.0468	0.0625***	0.0641***
Age squared	– 0.0006	0.0002	– 0.0010***	– 0.0003	– 0.0007**	– 0.0007**
Rural	– 0.2440	– 0.3590**	– 0.3530***	– 0.0873	– 0.2098	0.2802
Province						
Niassa	– 0.0577	– 0.2404	– 0.8014***	– 0.3018	– 0.4968	0.3190
Cabo Delgado	0.1131	– 0.2266	– 0.2540	– 0.0994	– 0.4126	0.3217
Nampula	– 0.1009	– 0.1921	– 0.9328***	– 0.4711**	– 1.0766***	– 0.6708***
Zambezia	– 0.1677	– 0.3426	– 0.7364***	– 0.3339	– 0.8012***	0.2540
Tete	– 0.1588	– 0.0842	– 0.4199***	– 0.0144	0.0206	0.7408***
Manica	– 0.1423	– 0.1659	– 0.4594***	0.0714	– 0.4176*	0.2963
Sofala	– 0.3089***	– 0.2129	– 0.4203***	– 0.3473**	– 0.6021***	0.0236
Inhambane	– 0.0235	– 0.1475	– 0.2344**	– 0.8638***	– 0.2113	0.6015***
Gaza	– 0.1097	– 0.0583	– 0.2120**	– 0.5200***	0.1257	0.5314***
Maputo Province	0.0238	– 0.1243	– 0.0229	– 0.1023	0.0097	0.8282***
Public sect. correction term	0.9964****	0.2110	0.2742	– 0.8867*	0.1985	– 0.3612
Private sect. correction term	– 0.6765	0.1392	0.1405	0.1604	0.4798	– 0.5464
Self-employ. correction term	0.5290	– 0.8380**	– 0.7063	– 0.7825**	0.5153	0.3475
Unemploy. correction term	2.0820***	0.6031	0.2352	– 0.2156	1.2449	– 0.4262
Constant	– 2.2262	0.8200	– 0.6877	– 0.5869	1.0905**	– 0.5702
Selection Correlations						
Public sector	0.7655***	0.3018	0.3213	– 1.2041**	0.1357	– 0.2764
Private sector	– 0.5198	0.1991	0.1647	0.2178	0.3281	– 0.4182
Self-employment	0.4065	– 1.1984**	– 0.8278*	– 1.0626**	0.3524	0.2659
Unemployment	1.5997***	0.8626	0.2757	– 0.2928	0.8513*	– 0.3262

Glick & Sahn (1997) employed instead the Lee (1983) method mentioned above to correct their earnings functions for sample selection. They found absence of sample selectivity bias. For Kenya, Wambugu (2002) also followed the Lee (1983) approach. Selectivity bias was also absent, except for women in the public sector. Koch & Ntege (2008) based the sample selection correction of their South African earnings functions on the original Dubin & McFadden (1984) method. In this case, however, they found the presence of selectivity bias in all of the analysed sectors (private sector, public sector, and self-employment).

Table 4.4 – Percentage change in earnings associated with education; Note: Figures derived from selectivity-corrected earnings functions are underscored and are calculated from the table above as 100*(e^c – 1), where c stands for the education dummy coefficient at each level; the remainder of the figures are the same as in Table 4.2, calculated from the standard OLS earnings functions.

Employment Sectors	Public Sector (%)		Private Sector (%)		Self-Employment (%)	
Categories	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Some primary	_116_	-	_16_	-	_26_	-
Full primary	_317_	_169_	_44_	_44_	_66_	60
Some secondary	_962_	_403_	_148_	_156_	_85_	109
Full secondary	_2,052_	_989_	_584_	_782_	_238_	-
Tertiary education	_4,261_	_2,052_	_838_	_1,433_	_374_	-

5. CONCLUSION

The objectives of this study were twofold. First we wanted to identify the impact schooling and other characteristics have on the chances of employment in Mozambique. Then, we wanted to link education to earnings in a Mincerian type approach.

Mozambique is no different from other developing countries such as Kenya, Ethiopia, Guinea, etc., whereby a big slice of the working-age population works outside the wage sector. In such countries self-employment, unpaid family work, informal sector employment, subsistence agriculture employment and the like are more common than employment in the public and private wage sectors, where the latter two in normal circumstances tend to be the preferred ones. Therefore, analysing the impact of schooling on the chances of employment and earnings in the wage sector only is nonsensical. Instead we have done the analysis in a number of labour market sectors for which reliable data are available.

Using the only available Mozambican Labour Force Survey conducted over the years 2004 and 2005, and employing a multinomial logit model, firstly, we found that for both men and women schooling relative to no formal education increases the chances of employment in the public wage sector. Within the private wage sector schooling was found to be important for women only, though men's chances of employment were found to be greater at all education levels. For both sexes schooling was found to lower the chances of self-employment and unpaid family work.

Secondly, education has a bearing on earnings across all sectors for which reliable earnings data are available, and its impact on earnings are convex, i.e., returns to education are greatest at the tertiary education level, followed by secondary schooling and then primary schooling. Within the public wage sector, private wage sector, and self-employment sector, more educated men and women are paid significantly better than uneducated men and women, respectively. Returns to education are greater for women than men in the private wage sector, but greater for men than women in the public sector. For individuals in self-employment secondary education or further schooling matters significantly for men only, suggesting that self-employed women engage in activities for which advanced education attainment does not bring much more income.

Thus, education is important to increasing the chances of wage employment in the public sector, which in developing countries such as Mozambique can be considered a luxury. Educated men and women would rather non-participate or stay unemployed while waiting for the possibility of getting into the public wage sector. However, only a very small sample of the working-age population is educated enough to enter this sector. Only those few who can get enough schooling end up benefiting from it. The remainder of the working-age population resort to activities outside the wage labour market. But even within these less attractive sectors, education does pay.

While in this paper we get around sample selection bias in face of multiple employment statuses, our earnings functions estimates might be prone to other type of bias which is originated by the omission of some important variables – the omitted variable bias. Out of these variables the related literature devotes much attention to ability, since it is believed to affect both education and earnings positively, therefore creating an upward bias in the schooling estimates. Correcting our earnings functions for ability bias is a possible avenue for further research.

6. REFERENCES

Bhorat, H., & Leibbrandt, M. (2001). Modelling Vunerability and Low Earnings in the South African Labour Market. In H. Bhorat, M. Leibbrandt, M. Maziya, S. Van der Berg, & I. Woolard (Eds.), *Fighting Poverty: Labour Markets and Inequality in South Africa* (pp. 107-129). Lansdowne: UCT Press.

Bhorat, H., & McCord, A. (2003). Employment and labour market trends in South Africa. In HSRC (Ed.), *Human Resources Development Review.* Cape Town: HSRC Press.

Bourguignon, F., Fournier, M., & Gurgand, M. (2007). Selection bias corrections based on multinomial logit model: Monte-Carlo comparisons. *Journal of Economic Surveys , 21* (1), 174-205.

Bourguignon, F., Fournier, M., & Gurgand, M. (2004). Selection Bias Corrections based on the Multinomial Logit Model: Monte-Carlo Comparisons. *Départment et Laboratoire d'Économie Théorique et Appliquée , Working Paper No. 20*.

Cameron, A., & Trivedi, P. (2009). *Microeconometrics Using Stata.* Texas: Stata Press.

Cohn, E., & Geske, T. (1990). *Economics of Education.* Oxford: Pergamon Press.

Dahl, G. (2008). Mobility and the returns to education: Testing a Roy Model with multiple markets. *Econometrica , 70* (6), 2367-2420.

Demery, L., & Grootaert, C. (1993). Priority Analysis. In L. Demery, M. Ferroni, C. Grootaert, & J. Wong-Valle (Eds.), *Understanding the Social Effects of Policy Reform.* Washington D.C.: The World Bank.

Dubin, J., & McFadden, D. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica , 52* (2), 345-362.

Glick, P., & Sahn, D. (1997). Gender and Education Impacts on Employment and Earnings in West Africa: Evidence from Guinea. *Economic Development and Cultural Change , 45* (4), 793-823.

Glick, P., & Sahn, D. (1993). Labor Force Participation, Sectoral Choice, and Earnings in Conakry, Guinea. *Cornell Food and Nutrition Policy Program Working Paper Series , Working Paper no. 43*.

Greene, W. (2003). *Econometric Analysis.* New Jersey: Prentice Hall.

Heckman, J. (1979). Sample selection bias as a specification error. *Econometrica , 47* (1), 153-161.

Johnes, G. (1993). *The Economics of Education.* London: Macmillan Press.

Koch, S., & Ntege, S. (2008). Returns to Schooling: Skill Accumulation or Information Revelation? *University of Pretoria, Department of Economics Working Paper Series , Working Paper No. 12*.

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

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

Mansur, E., Mendelsohn, R., & Morrison, W. (2008). Climate Change Adaptation: A study of fuel choice and consumption in the US energy sector. *Journal of Environmental Economics and Management , 55*, 175-193.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka (Ed.), *Frontiers in Econometrics.* New York: Academic Press.

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

Psacharopoulos, G. (1981). Returns to Education: An Updated International Comparison. *Comparative Education , 17* (3), 321-341.

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

Schmertmann, C. (1994). Selectivity bias correction methods in polychotomous sample selection models. *Journal of Econometrics , 60* (1), 101-132.

Statacorp. (2009). Stata Statistical Software: Release 11.0. Texas: Stata Corporation.

UNESCO. (2009). *Overcoming inequality: why governance matters.* Paris and Oxford: United Nations Educational Scientific and Cultural Organization (UNESCO) and Oxford University Press.

Vijverberg, W. (1993). Educational Investments and Returns for Women and Men in Côte d'Ivoire. *Journal of Human Resources , 28* (4), 933-974.

Wambugu, A. (2002). Education, Employment and Earnings in Kenya. *Göteborg University, Department of Economics* .

Wooldridge, J. (2002). *Econometric Analysis of Cross Section and Panel Data.* Cambridge: The MIT Press.