Selection Into Formal and Informal Employment in Uganda: The Role of Education

Susan Namirembe Kavuma* & Bruno L. Yawe**

,

Abstract

This paper examines the role of educational attainment in being employed in Uganda. Using household panel data for four waves—2005/06, 2009/10, 2010/11 and 2011/12—the study estimates correlated random effects models and analyses the education-employment relationship. We find a positive association between educational attainments and 'not-working' in the static model, but when state dependence is accounted for, the association disappears for all education levels except for degree education which is not significant at conventional levels. Therefore, it can be concluded that the observed positive association between education and 'not-working' is a short-term effect for individuals with less than degree education, but may persist for individuals with degree education. With reference to informal sector employment, the results are mixed. They depict a heterogeneous sector that is hierarchical, whereby more educated individuals are observed in informal salaried work and less educated individuals are observed working in household enterprises. Further, there is a positive association between education and employment status, which is strongest in the formal sector. All results vary by age group, gender and residence.

Keywords: Uganda, employment, formal, informal and education.

JEL Classification: J00, J01, J24

1. Introduction

Labour markets in most developing countries are dualistic: consisting of formal and informal sectors (Hart, 1973; Fields, 1990). The formal sector offers higher wages and better employment terms such as written contracts, social security, leave days and fringe benefits; compared to the informal sector which offers lower wages and flexible employment terms (Fields, 1990; Maloney, 2004). The distinction between the formal and informal sectors has implications for the functioning of the labour market and overall economic development of a country. In particular, the characteristics of these sectors influence wage and employment determination, and therefore have subsequent impacts on poverty reduction and income distribution. According to the 2009/10 household survey data, Uganda's informal sector absorbs four-fifths (80 percent) of the labour force, who are mainly women and the young (below 30 years). The formal sector is small, accounting for 19 percent of the labour force, and continues to shrink over time (Senone, 2013).

In the literature, the commonly adopted definition of informality is the one developed by the International Labour Office (ILO, 1972), which defines the informal sector as

---

* Lecturer, School of Economics, Makerere University: suzkav@yahoo.com
** Senior Lecturer, School of Economics, Makerere University, Email: byawe2010@gmail.com

the sum of non-professional (elementary) self-employed, domestic, unpaid workers and workers employed in enterprises employing five or less employees. In view of the ILO definition, the Uganda Bureau of Statistics (UBoS) classifies the informal sector as comprising the self-employed who include: employers with small businesses, own-account workers, contributing family workers (unpaid) and household farm workers;[1] while the formal sector is classified as comprising employees who are paid a wage or salary or commission. This study adopts the UBoS classification with a slight modification, and defines the informal sector as consisting of household farm workers/unpaid labour in household enterprises, self-employed/own-account workers and informal salaried workers, which is consistent with the ILO's (1993) definition. Informal salaried workers have been included because they do not contribute to the national social security fund as reported in the data, and therefore are informally employed.[2] The formal sector consists of workers in paid employment and contributing to the National Social Security Fund (NSSF).

To the best of our knowledge, only one study by Appleton (2001) has analysed the determinants of employment in Uganda. Appleton's study analysed net entry effects of education on different income sources such as agriculture, non-agricultural self-employment and wage employment using cross-section household survey data for 1992 and 1999/2000. It established that the positive effect of education on different income sources is offset by a negative effect on other income sources. For example, an extra year of education of a household member reduced the probability of receiving income from agriculture by 0.5 percent, which lowered total household earnings by 0.6 percent. Conversely, university education reduced the probability of receiving earnings from agriculture and non-agricultural self-employment. The study noted that the net entry effect of education is positive for all levels. For instance, an extra year of university education increased household earnings in 1992 by 4 percent, and in 1999/2000 by 2 percent. Similarly, an extra year of primary education raised earnings by 0.8 percent in 1992, and by 1.8 percent in 1999/2000.

In the context of the neoclassical labour supply model, this study investigates the probability of an individual being observed as 'not-working', or in 'formal' and 'informal' employment in Uganda's labour market. Specifically, we analyse the factors that influence reservation and market wage, since a worker will only offer her/his labour if the wage rate exceeds their reservation wage. The objective of the paper is to investigate the effect of educational attainment on the participation in the labour market, selection into formal and informal employment, and find out whether it varies by age, gender and residence.

The study data came from four panel waves: the 2005/06 Uganda National Household Survey (baseline survey), 2009/10, 2010/11 and 2011/12 Uganda

---

[1] Family and household workers are often classified as household enterprises and rural farm labour, respectively.

[2] According to the labour laws, all employers and employees working in firms employing more than four employees in Uganda are required to contribute to the national social security fund

National Panel Survey (UNPS) data collected by the Uganda Bureau of Statistics. Five separate labour market states have been constructed: 'not-working', 'household worker', 'self-employed', 'informal salaried' and 'formal salaried'. Random effects models relating the probability of being in each of these states to individuals' education qualifications (levels of education), or years of education, have been estimated. The study was guided by two research questions: (1) Do segments of Uganda's labour market differ with regard to individuals that select into the formal and informal sector of employment? (2) Are there differences in selection into employment states by age group, gender or residence?

The next section discusses related literature, while section 3 presents the theoretical framework used to analyse employment decision. Section 4 describes the empirical model and strategy used in the estimations. The data sources and descriptive statistics are presented in section 5, while section 6 presents the results. Finally, section 7 concludes the paper.

2. Related Literature
Employment decisions are analysed in the context of the standard neoclassical labour supply theory, which assumes competitive labour markets; whereby market forces determine employment and wage levels. The theory assumes that a worker chooses a sector that offers a market wage higher than their reservation wage.[3] Therefore, labour market outcomes in this model are an interaction between demand factors that influence market wage, and supply factors that influence reservation wage. Central to the standard neoclassical labour supply theory is its emphasis on the role of human capital factors in wage and employment determination (Cain, 1976). According to the human capital theory, education is expected to enhance an individual's access to employment by increasing one's productivity (Becker, 1962).

The conventional neoclassical labour supply model is criticised by theorists in support of segmented labour markets (Harris & Todaro, 1970; Doeringer & Piore, 1971). These theorists argue that equally productive workers might have persistent unexplained wage and unemployment differences in the market that are shielded from competitive forces. They assert that the degree of market imperfections and monopoly tendencies may render the perfect competition assumption in the neoclassical theory untenable for empirical work. In the context of developing countries, they posit that the labour market consists of a reserve army of unemployed individuals that has less bargaining power, thereby undermining competitive forces. Therefore, a worker's preference may play a minimal role in wage and employment decisions; rather it is employers who are more likely to have greater influence in these decisions.

---

[3] The reservation wage is the wage which equates the utility an individual receives between working and not-working.

.

In empirical literature, supply decisions are modelled either using structural or reduced-form equations. The structural equations express the amount of labour supplied as a function of an individual's wage and non-wage income (Heckman, 1993). Nonetheless, not all individuals willing to work are observed in the labour market and such individuals have no amount of labour supplied. Therefore, estimates from the structural labour supply function are potentially biased due to missing wages of non-workers. Scholars have addressed this bias by simultaneously determining employment levels and wage returns, but many use reduced form equations that do not require the consideration of economic decisions (Magnac, 1991; Pradhan & van Soest, 1995; Glick & Sahn, 1997; Appleton, 2001; Comola & Mello, 2011; Wamuthenya, 2010; Baffour, 2013). These equations are also considered reduced-form because estimates are for individuals observed in the labour market, since most surveys lack information on preferred labour market state and job search data. The reduced-form model considers factors that influence reservation and market wage since individuals only work if the market wage exceeds their reservation wage.

Many studies (Magnac, 1991; Pradhan & van Soest, 1995; Glick & Sahn, 1997; Wamuthenya, 2010; Comola & de Mello, 2011) simultaneously determine employment levels and wage returns using the selection corrected wage equations framework. Magnac (1991) simultaneously determined employment decisions and wage returns using data for married women in Colombia. He segregated the labour market into four states: non-participating, unemployed, informal and formal sector workers. Employing both bivariate and univariate probit models, he investigated the employment decision as a function of an individual's human capital variables and family characteristics in the two sectors (formal and informal). He found that a one-year increase in education is associated with 4 percent decrease in participation in the informal sector, and 11 percent increase in participation in the formal sector.

Similarly, Pradhan and Soest (1995) determined wage returns and employment for the formal and informal sector in the Bolivia urban labour market, using household survey data. They analysed the employment decision using two models: the multinomial logit based on the assumption of competitive labour markets, and the ordered probit that regards the informal sector as a buffer zone in support of the staging hypothesis.[4] They found a positive relation between the probability of formal employment and education, and an inverse relation between informal sector employment and educational attainment. Glick and Sahn (1997) investigated the impact of gender and education on employment choices in Guinea using a multinomial logit model with four alternatives: non-participation, self-employment, private and public wage employees. They found that more education is associated with a reduction in the probability of being self-employed, but would strongly increase the probability of working in the public sector. However, the

---

[4]The staging hypothesis assumes that workers are rationed out of the formal sector and choose to either remain unemployed or those who cannot afford to remain unemployed join the informal sector as they wait for an opportunity in the formal sector. Therefore, the informal sector acts as a buffer for the surplus labour (Fields, 1990).

results for private wage employment differed by gender, where more education positively influenced the probability of a woman getting into private wage employment, while the probability was reduced for a man.

Wamuthenya (2010) analysed employment decisions by modelling factors that affect reservation wage and market wage in the formal and informal sectors of the urban areas of Kenya. He estimated a multinomial logit model with four mutually exclusive choices: private, public, informal and unemployed workers. The formal sector was represented by the public and private sector. He found education to have the strongest impact on formal sector employment, particularly the public sector. In particular, the study found that more education is associated with an increase in the probability of participating in the public sector, but the results for the private sector were mixed. In the private sector, more education from primary to secondary school positively influenced participation; but when a worker attained university level education, the impact would be the lowest. Conversely, education was associated with a reduction in the probability of being employed informally.

Comola and de Mello (2011) estimated the determinants of earnings and selection into the labour market in Indonesia. They controlled for selection bias in the wage equations using both a binomial (individuals are either employed or not) and multinomial selection term: individuals select into three labour market states (inactivity, wage-earner and non-salaried work). They found the coefficient on education consistent regardless of the selection term, with wages rising with educational attainment. In estimating selection into the different employment states, they instrumented education, since it was likely to be endogenous. The program intensity computed as the number of schools built in a district between 1973-74 and 1978-79 divided by the number of children aged 5-14 living in that district in 1971 was used as an instrument[5] for years of schooling. They found that an additional year of education was associated with a reduction in the likelihood of non-salaried work and non-participation in reference to salaried work, and the effect greater when educational attainment is instrumented.

In sum, the evidence that more education is associated with an increase in the probability of formal employment and a decrease in informal employment is consistent in the literature, regardless of the estimation strategy and whether education is assumed to be exogenous or endogenous. However, the results from Comola and de Mello (2010) show that an increase in the positive association between education and selection into employment states when education is instrumented. In most cited studies, the strength of the relationship varies by gender and age group, which this study explores. Regarding the relationship between education and unemployment, the available literature provides mixed results, which vary by gender and age group.

---

[5] They used the 2004 wave data collected annually since 1976. The age bracket was selected based on the assumption that most children below the age of 12 still live in their district of birth.

.

3. Theoretical Framework

Selection into employment is analysed as a labour supply decision under the neoclassical model of labour supply (Pencavel, 1986; Killingsworth & Heckman, 1986). An individual is faced with two labour supply decisions. Firstly, the participation or employment decision is referred to as choice at the extensive margin (Heckman, 1993), or the corner solution (Pencavel, 1986). Secondly, the individual decides the number of hours of work (the intensive margin or interior solution), although often these two decisions are made simultaneously.

We adopt the static neoclassical model of labour supply that assumes that an individual aims at maximising utility (U) by consuming optimal combinations of goods (C) and leisure (L), subject to time and income constraints:

$$Max(C,L)U(C,L) \qquad (1)$$

$$\underset{C>0, L>0}{Max}\ U(C,L) \qquad (1)$$

An individual is also endowed with a stock of Time (T), which is allocated between work (h) and Leisure (L):

$$T = h + L \qquad (2)$$

The utility function is also subject to an income constraint:

$$C \leq wh + R \qquad (3)$$

Where w is the real wage rate, and R is non-labour income.

If L lies between 0 and T, and C is positive, implying the individual has entered the labour market but has to decide on the number of hours of work; and assuming U' (first-order derivative of the utility function) is positive and U'' (second-order derivative of the utility function) is negative, which suggests positive and diminishing marginal utilities, an interior solution to the supply optimization problem will be:

$$\frac{\partial U/\partial L}{\partial U/\partial C} = w \qquad (4)$$

Equation (4) implies the marginal rate of substitution between consumption and leisure is equal to the economic rate of substitution.

However, the analysis in this paper investigates the corner solution to the supply optimization problem, i.e., the decision of an individual to enter a labour market (employment decision) where the following condition must be satisfied:

$$\left(\frac{\partial U/\partial L}{\partial U/\partial C}\right)_A < w \qquad (5)$$

The marginal rate of substitution at point A (the extensive margin where L=T) is referred to as the reservation (shadow) wage, denoted as:

$$w_A = \frac{\partial U / \partial L(R, L_0)}{\partial U / \partial C(R, L_0)} \qquad (6)$$

As expressed in equation (6), the reservation wage depends on an individual's preference between consumption of goods and leisure, and the value of non-labour income, assuming the hours devoted to leisure and work are fixed. Thus, the decision to participate in the labour market is influenced by the reservation wage, which should be lower than the market wage (equation (5)). Therefore, the lower the reservation wage, the more likely an individual will be observed in the labour market. Consequently, we model factors likely to influence either reservation wage such as non-labour income, gender, marital status, household size, residence or market wage determined by an individual's educational attainment.

Household characteristics such as marital status and household size are included to account for influence of family relations and time allocated to household production, since labour supply decisions are more complex involving the substitution of wage income for home production, and the influence of interfamilial decisions.

4. Methodology
With some modifications, we adopt Arulampalam et al.'s (2000) estimation strategy and investigate the effect of educational attainment on selection into employment states by estimating random effects probit models specified as:

$$Pr(y_{itj} = 1 | x_{itj}, c_{ij}) = \Phi(\beta x_{itj} + c_{ij}) \qquad (7)$$

where i represents individual 1, 2 …n, t is the time period 1,2…4 and j is employment state j…J; where j_1 is not-working (unemployed), j_2 is household worker, j_3 is self-employed, j_4 is informal salaried and j_5 is formal salaried. The outcome variable y_{itj} is modelled from a latent unobserved variable y_{it}^* that measures an individual's utility derived from being observed in employment state j, and therefore an individual i is observed in employment state j if $y_{itj}^* > 0$, or otherwise the individual is not observed and thus $y_{itj} = 0$, x_{itj} is a vector of observed explanatory variables which includes, age, gender, household size, educational attainment, marital status, non-labour income, household assets, residence, regional and year dummies, c_{ij} is the random intercept accounting for individual unobserved heterogeneity which is assumed to be independent of x_{itj} explanatory variables, Φ is the cumulative distribution function of a standard normal variate, β is a vector of parameters to be estimated.

Although the dependent variable can be constructed as a categorical variable, we could not estimate a multinomial logit model because the independence of irrelevant alternatives (IIA) assumption was violated. This implies that the odds ratio of two alternatives is not independent of other choices, which suggests that

the presence of each choice may alter an individual's selection. An attempt was made to estimate a multivariate probit model that assumes a correlation between the unobservables in the specified models, but the model could not converge with the full specification; thus, a simple model was specified.[6]

Assuming c_{ij} to be independent of the x_{itj} explanatory variables is a strict assumption, given that an individual's propensity to be observed in a particular employment state j may be correlated with any of the explanatory variables such as asset ownership (Arulampalam et al, 2000). Following Wooldridge (2010) the assumption of strict exogeneity of unobserved heterogeneity with the explanatory variables was relaxed by including averages of time varying variables that are correlated with the error term, and the error term is specified as:

$$\varepsilon_{ij} = \alpha_0 + \alpha_1\overline{x_{ij}} + v_{ij} \qquad (8)$$

Where α_0 is the intercept, $\overline{x_{ij}}$ is a vector of averages of time varying explanatory variables for each individual over time, v_i is assumed to be normally distributed and independent of the x_{itj} explanatory variables.

Therefore equation (7) can be re-written as:

$$Pr\big(y_{itj} = 1|x_{itj},\overline{x_{ij}}\big) = \Phi\big(\beta x_{itj} + \alpha_1\overline{x_{ij}} + v_{ij}\big) \qquad (9)$$

In literature, it is suggested that an individual's current and future employment state could be influenced by their history of labour market experience, which is referred to as state dependence (Heckman, 1981). Consequently, a lagged dependent variable was included in each model to test the correlation between current and past employment state, and the model specified as:

$$Pr\big(y_{itj} = 1|x_{itj},\overline{x_{ij}},\overline{y_{ijt-1}}\big) = \Phi\big(\beta x_{itj} + \gamma y_{ijt-1} + \alpha_1\overline{x_{ij}} + v_{ij}\big) \qquad (10)$$

The marginal effects in equations (9) and (10) for each individual at the mean of the explanatory variables are estimated, and averaged for the entire sub-sample to get the average marginal effects. This is done for the entire sample, and later the sample is disaggregated by age group, gender and residence (urban vs rural).

5. Data Sources and Description

The data came from four waves of the Uganda National Panel Surveys (UNPS): 2005/06, 2009/10, 2010/11 and 2011/12. These are on-going annual panel surveys conducted by the Uganda Bureau of Statistics, a government statutory institution with the mandate to collect and analyse national statistics. The multi-topic panel household surveys started in 2009, with the 2005/06 wave as the baseline survey

---

[6] A multivariate probit model was specified for four employment states excluding the not-working group and therefore analyses only selection into sectors of employment. The model was specified with four covariates: age, age squared, years of education and female.

.

that covered 3,123 households, which the 2009/10 survey set out to track. We adopted the two-stage sampling design used in the 2005/06 UNHS, where in the first stage we selected the UNPS Enumeration Areas (EAs) from the 2005/06 EAs with equal probability, and with the implicit stratification of rural/urban and district categorisation. The surveys have a national representative sample and collected data on household and individual characteristics under five modules: household, agriculture, woman, community and market. In this paper, data is specifically drawn from the household module, which has information on labour market statistics, demographic characteristics, income and expenditure data. The estimation sample is unbalanced and has a total of 17,417 observations for individuals aged 14-64 years who have reported employment state and educational attainment as Table 1 illustrates.

Table 1: Distribution of the Estimation Sample Across Panel Waves

	Employment State	Pooled	2005/06	2009/10	2010/2011	2011/12
	Not-working	604	163	155	144	142
		(3.47)	(3.41)	(3.47)	(3.81)	(3.23)
Informal sector	Household worker	8,572	1,242	2,427	2,169	2,734
		(49.22)	(26.01)	(54.38)	(57.37)	(62.16)
	Self-employed	4,722	2,326	970	715	711
		(27.11)	(48.71)	(21.73)	(18.91)	(16.17)
	Informal salaried	2,802	720	764	628	690
		(16.09)	(15.08)	(17.12)	(16.61)	(15.69)
Formal sector	Formal salaried	717	324	147	125	121
		(4.12)	(6.79)	(3.29)	(3.31)	(2.75)
	Total	**17,417**	**4,775**	**4,463**	**3,781**	**4,398**

Source: Author's construction based on World Bank (2014a-d). Percentages in parentheses.

As Table 1 shows, individuals in the sample are categorised into five mutually exclusive employment states. These include not-working, household worker, self-employed, informal salaried and formal salaried. The not-working group constitutes of individuals not-working but were searching for a job in the previous week of the survey, and also those not searching either because they are waiting for a reply or are discouraged. Individuals not working and not searching are included because those who were searching are few in the sample. Workers categorised as household workers are those either employed on agricultural farms, or are employed in a household enterprise but are not paid. The self-employed workers are own-account workers who have no employees (those with employees were few and are excluded from the analysis). Informal salaried workers are those in paid employment but are not contributing to the national social security fund (NSSF),[7] or were temporary

---

[7] In Uganda it is mandatory for all employers and employees working in a firm with at least five employees to contribute to the national social security fund. Therefore, we regard those not contributing to this fund as informally employed, and those contributing as formal employees.

.

workers in the 2005/06 survey for which data on NSSF is missing. Most of these workers (90 percent) said they worked for a private firm. The formal salaried are workers in paid employment and contributing to NSSF, or are permanent workers in the case of the 2005/06 survey, where most of these workers (90 percent) reported to be working for a public institution. The employment variable was constructed from the following binary response questions:

(1) *"In the last week, did you work for a wage, salary, commission or any payment in kind including doing domestic work even if it was for only one hour?*

(2) *"In the last week did you run a business of any size, for self or another household member, even if it was for only one hour?"*

(3) *"In the last week did you help without being paid in any kind of business run by this household, even if it was only for one hour?"*

(4) *"In the last week were you an apprentice?"*

(5) *"In the last week did you work on the household farm?"*

(6) *"During the last seven days have you worked on a land owned or occupied by a member of your household either in cultivating crops or in farming tasks, or have you cared for livestock belonging to a member of your household or fishing at least for one hour in any day?"*

Question (6) was only asked in the baseline wave (2005/06). Individuals who answered in the affirmative to question (1) were categorised as workers in paid employment, while household workers are those who answered in the affirmative to questions (5) and (6), and the self-employed who answered 'yes' to question (2). The responses to question (d) are not used in this paper, because of the limited number of such individuals.

The statistics in Table 1 show that most workers are observed in household enterprises except in the baseline wave where there are more individuals self-employed than those working in household enterprises. This was a result of differences in questions asked in the baseline and subsequent panel surveys as highlighted above. Also, the estimation sample has a small proportion of the unemployed and individuals working in the formal sector. The small proportion of unemployed is attributed to the large size of the informal sector; which provides job opportunities for workers not able to find a job in the formal sector, and may not be able to wait for a job vacancy in the absence of unemployment benefits to support them. Therefore, in Uganda's case the greater concern may be underemployment rather than unemployment; in reference to the large proportion of the labour force employed in the low productivity sector—the informal sector. The large informal sector in Uganda is mainly a consequence of a small formal sector created by the implementation of the World Bank structural adjustment programmes (SAPs) in the 1980s, which aimed at reducing public expenditure, promoting private sector development, and implementing market oriented economic policies. These programmes led to the downsizing of the public sector.

Uganda runs a 7-4-2-3(4/5) education system, where primary education has a duration of seven years, lower secondary runs for four years, upper secondary for two years and university education can take three to five years depending on the course pursued. Table 2 presents the means of key explanatory variables, and the means of all covariates in the models are provided in Table A1 in the Appendix. The statistics in Table 2 demonstrate that, on average, workers in the formal sector are the most schooled with 13 years of education (equivalent to senior Form VI, which is an advanced level certificate of education).

Table 2: Means of Key Explanatory Variables by Employment State

	Employment state	Education (Years)	Age	Female	Urban	Household assets (USh)
	Not-working	8.97	22.99	0.51	0.51	319,473
		(4.30)	(5.28)	(0.50)	(0.50)	
Informal sector	Household worker	6.03	24.28	0.61	0.09	109,643
		(2.99)	(6.00)	(0.49)	(0.29)	
	Self-employment	7.07	27.58	0.41	0.31	126,782
		(3.41)	(4.99)	(0.49)	(0.46)	
	Informal salaried	8.29	24.64	0.29	0.45	211,921
		(4.04)	(5.67)	(0.46)	(0.50)	
Formal sector	Formal salaried	12.70	28.75	0.32	0.54	249,887
		(3.30)	(4.30)	(0.47)	(0.50)	

Source: Author's construction based on World Bank (2014a-d). Standard deviation in parentheses.

The next most schooled are those not-working with 9 years of education (equivalent to senior two – lower secondary). This is surprising but suggests that there are a number of more educated individuals choosing to wait for an opportunity in the formal sector rather than join the informal sector. The least schooled workers are those working in a household enterprise that is consistent with the fact that most of these workers reside in the rural area (91 percent) as shown in the column for urban.

In reference to age, the average age for a worker in the formal sector is the highest, which signals that entry in the formal sector increases with age; and the youth may have explicit and implicit barriers to enter this sector, for example, the lack of experience which is normally a requirement for public sector jobs. Similarly, workers in self-employment are older compared to workers in household enterprises, informal salaried and those not-working. Perhaps this signals the lack of accumulated financial and human capital that may be required to start up a business. Generally, there is almost a gender balance for individuals not-working; but one can note that there are more females in household work, which perhaps suggests that females have a lower reservation wage and therefore are more willing to offer their labour in a household enterprise even when they are not paid. Alternatively, work in a household enterprise may give females an opportunity to be more involved in home production. The statistics further show that females are less represented in self-employment, informal and formal salaried.

In the estimation sample, there is almost a gender balance for individuals not working (unemployed) in urban and rural areas. Household labour is predominantly rural, and self-employment as well as informal salaried work are more prevalent in rural than urban areas. It is interesting to note that those not working come from wealthier families, since this group has the highest mean for household assets, which partly explains why these individuals may choose to wait for an opportunity in the formal sector rather than join the informal sector. The statistics reveal that individuals working in the formal sector come from wealthier families, given that the mean for household assets is the second highest in the sample. It is not surprising to note that, on average, household workers have the lowest levels of household assets. This suggests that workers in this segment are less resource-endowed, and are more likely to constitute the working poor.

6. Results

Initially, a static model is estimated to establish the relationship between educational attainment and being observed in an employment state by estimating equation (9), and the results for all covariates are presented in Table 3. As a robustness check, equation (9) is re-estimated using years of education and also a quadratic term for education to test for non-linearity in the association between education and employment is included. For brevity, only the results for educational attainment are presented in Table 4. As earlier noted, the paper investigates whether the association between educational attainment and employment state varies by age group, gender and residence by re-estimating equation (9) for the respective sub-samples. In the interest of brevity, only the results for educational attainment are presented in Tables 5, 6 and 7 for the respective sub-samples. Later state dependence is allowed and a dynamic model specified in equation (10) is estimated and the results are provided in Table 8.

The static model results presented in Table 3 show that most of the variables have the expected signs except the education dummies in the not-working (unemployed) model. According to the human capital theory, investment in education is expected to increase an individual's level of productivity, and therefore increase the probability of being employed. Therefore, it is surprising that this study finds education raises an individual's probability of not-working (unemployed) which increases with level of education. The results contradict findings of similar studies (Glick & Sahn, 1997; Wamuthenya, 2010; Baffour, 2013 for the Tanzanian labour market) but are consistent with findings by Baffour (2013) for the Ghanaian labour market.

These results suggest that more educated individuals would rather wait for a job opportunity in the formal sector rather than join the informal sector. For informal employment, the paper finds mixed results. A negative relationship between education and household employment (agricultural workers) is observed, which is consistent with available literature. The relation between education and self-employment is mixed. While similar studies find a negative association between education and self-employment (Glick & Zahn, 1997), in this paper this association is only observed with post-secondary education, and a positive association is found at low levels of education. We also observe a strong positive relationship between education and informal salaried work which increases with education.

Table 3: Estimated Probabilities of Selection into Informal and Formal Employment: Average Marginal Effect (AME)

Variable	Not-working		Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	AME	Std. Err	AME	Std. Err	AME	Std. Err	AME	Std. Err	AME	Std. Err
Age	-0.080***	0.028	-0.110***	0.021	0.107***	0.025	0.063**	0.025	0.193***	0.067
Age squared	0.001***	0.000	0.001***	0.000	-0.001***	0.000	-0.001***	0.000	-0.003***	0.001
Household size	0.011	0.028	0.009	0.017	0.007	0.019	0.005	0.021	-0.078*	0.041
Primary education	0.154*	0.084	-0.115*	0.060	0.141**	0.071	0.051	0.079	0.766***	0.258
Lower secondary	0.216***	0.083	-0.511***	0.062	0.399***	0.073	0.319***	0.077	1.430***	0.242
Upper secondary	0.302**	0.117	-1.033***	0.097	0.126	0.110	0.849***	0.108	2.708***	0.291
Diploma	0.262*	0.139	-1.758***	0.116	-0.371***	0.128	1.487***	0.117	3.399***	0.307
Degree	1.090***	0.173	-1.541***	0.218	-0.850***	0.238	0.842***	0.201	4.036***	0.394
Female	-0.069	0.064	0.749***	0.056	-0.047	0.066	-1.028***	0.071	-0.293**	0.149
Married	-0.312***	0.072	0.288***	0.056	0.257***	0.067	-0.679***	0.071	0.011	0.150
Head	-0.377***	0.089	-0.548***	0.064	0.828***	0.076	-0.096	0.078	0.040	0.160
Youth	-0.393***	0.126	0.139	0.092	-0.037	0.109	-0.318***	0.110	-0.143	0.248
Non-labour income	0.109*	0.061	-0.112***	0.042	0.017	0.049	0.125**	0.052	0.198*	0.107
Household assets	-0.063*	0.036	0.045**	0.021	-0.017	0.023	-0.055**	0.025	0.008	0.054
Urban	0.359***	0.071	-1.419***	0.065	0.846***	0.067	0.836***	0.070	0.514***	0.131
Eastern region	0.292***	0.080	0.530***	0.068	-0.184**	0.079	-0.574***	0.085	-0.006	0.166
Northern region	-0.006	0.084	0.455***	0.066	0.117	0.076	-0.651***	0.082	-0.379**	0.180
Western region	-0.151*	0.088	0.533***	0.067	-0.463***	0.081	-0.237***	0.077	-0.017	0.166
Year dummy - 2010	0.086	0.072	0.108**	0.044	-0.181***	0.050	0.031	0.055	0.067	0.119
Year dummy - 2011	0.067	0.079	0.360***	0.055	-0.414***	0.063	-0.073	0.067	-0.111	0.149
Log likelihood	-1423.4		-5637.67		-4588.01		-3952.96		-895.50	
Wald chi2 (23)	153.33		1149.68		553.11		589.95		175.80	
Observations	11,769		11,408		11,408		11,408		11,408	

Note: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Uncompleted primary and the central region are the reference group.

Table 4: Estimated Probabilities of Selection into Employment Given Years of Education

Variable	Not-working		Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	CRE (1)	CRE (2)	CRE (1)	CRE (2)	CRE (1)	CRE (2)	CRE (1)	CRE (2)	CRE (1)	CRE (2)
Years	0.041***	-0.016	-0.118***	0.089***	0.003	0.185***	0.100***	-0.099***	0.343***	0.040
of education	(0.009)	(0.032)	(0.007)	(0.025)	(0.008)	(0.031)	(0.009)	(0.031)	(0.028)	(0.090)
Education		0.003*		-0.013***		-0.011***		0.011***		0.014***
squared		(0.002)		(0.001)		(0.002)		(0.002)		(0.004)
Log likelihood	-1441.6	-1439.85	-5718.73	-5682.7	-4653.11	-4633.2	-4001.24	-3980.11	-907.61	-902.66
Wald chi2 (19/20)	148.39	149.63	1116.56	1146.15	517.58	536.28	556.48	580.38	168.56	181.87
Observations	11,850	11,850	11,488	11,488	11,488	11,488	11,488	11,488	11,488	11,488

Note: Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. Other controls include: age, gender, marital status, household head dummy, youth dummy, non-labour income, household assets, residence and year dummies.

Table 5: Estimated Probabilities of Selection into Employment Given Educational Attainment by Birth Cohort

Variable	Not-working		Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	Adult	Youth	Adult	Youth	Adult	Youth	Adult	Youth	Adult	Youth
	0.209*	0.043	-0.148*	-0.062	0.100	0.161	0.096	0.000	0.906***	0.442
Primary education	(0.112)	(0.120)	(0.079)	(0.090)	(0.090)	(0.110)	(0.111)	(0.108)	(0.314)	(0.405)
	0.161	0.237**	-0.696***	-0.262***	0.403***	0.313***	0.531***	0.109	1.595***	0.855**
Lower secondary	(0.128)	(0.107)	(0.087)	(0.087)	(0.097)	(0.105)	(0.114)	(0.101)	(0.301)	(0.356)
	-0.242	0.531***	-1.387***	-0.605***	-0.002	0.247	1.282***	0.405***	3.014***	1.647***
Upper secondary	(0.252)	(0.147)	(0.137)	(0.132)	(0.145)	(0.160)	(0.155)	(0.146)	(0.358)	(0.440)
Diploma	-0.003	0.452**	-2.146***	-0.979***	-0.464***	-0.246	1.909***	0.805***	3.678***	2.296***
	(0.202)	(0.200)	(0.149)	(0.188)	(0.152)	(0.245)	(0.152)	(0.199)	(0.370)	(0.497)
Degree	0.803***	1.344***	-1.800***	-1.067***	-1.126***	-0.390	1.496***	-0.021	4.185***	3.234***
	(0.238)	(0.254)	(0.269)	(0.367)	(0.286)	(0.429)	(0.255)	(0.338)	(0.469)	(0.681)
Log likelihood	-545.05	-842.05	-3388.68	-2202.56	-3110.74	-1454.53	-2100.66	-1793.57	-680.46	-200.58
Wald chi2 (22)	65.93	87.19	741.87	415.73	332.89	181.46	341.76	260.29	141.94	34.79
Observations	7,069	4,700	6,950	4,458	6,950	4,458	6,950	4,458	6,950	4,458

Note: Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. Other controls include: age, gender, marital status, household head dummy, youth dummy, non-labour income, household assets, residence and year dummies. Uncompleted primary is the reference group.

The age variable has the expected sign except in the household worker model. In the not-working model, individuals employed are the reference group, and therefore the results suggest that the probability of being observed not working decreases with age at a decreasing rate, with a threshold at 40 years. Conversely, employment increases with age as individuals accumulate human and financial capital required at the workplace. These results are consistent with exiting literature (Glick & Sahn, 1997; Baffour, 2013), which finds a concave age-employment relationship when those not working (unemployed) are the reference group. The age relation is consistent with the coefficient for the youth variable, which confirms the inverse relationship between age and not-working. It is surprising that the relationship between age and household work is convex, but perhaps the results signal the transition of older workers from the formal sector after accumulating human and financial capital with the threshold at 55 years.

The results in Table 3 suggest that household size is not a predictor of an individual's employment state, except for the formal salaried where a weak negative association is noted. This could be due to the failure to account for the ages of family members, especially those under five years, who require more care as several studies show (Magnac, 1991; Glick & Sahn, 1997; Wamuthenya, 2010). It was not possible to include a variable that captures the age of a child because such information is missing in the data. The paper finds that being female increases the probability of being observed in household work, but decreases the likelihood of selecting into self-employment, informal and formal salaried work. Perhaps due to the low levels of human capital, a large proportion of females are observed in household work, and are not able to access salaried jobs; although it is possible there could be gender discrimination in paid employment. Consistent with findings by Glick and Sahn (1997), this paper finds a positive correlation between being married and being observed in household work or self-employment. In contrast, the paper finds that being married decreases the likelihood of being observed in informal salaried work. Although these results contradict the predictions of the pooling theory,[8] they are not surprising in the African context. They suggest that family businesses rely on family resources (both human and financial), while employment terms such as long hours of work in informal salaried jobs could be unattractive to married individuals. It is also possible that being married may be an outcome rather than a predictor of employment.

It is plausible to observe that individuals heading households are more likely to work and to work as self-employed as compared to their counterparts, which is consistent with findings by Wamuthenya (2010) for the Kenyan labour market. However, it is surprising that being a household head reduces the chance of working as household labour. Perhaps, these results signal that most workers in household enterprise are labourers, and household heads could be holding second jobs from which they earn.[9] Interestingly, the paper finds that the relation of non-

---

[8] Assumes sharing of income between spouses and one of the spouses is likely not to work if s/he benefits from the working spouse' income.

[9] About a third of individuals in the sample have a second job.

labour income with employment state is weak except for household and informal salaried work. The coefficient on non-labour income has the expected sign for household work, since an increase in an individual's non-labour income is expected to raise one's reservation wage, and consequently lower participation. In contrast, the association for informal salaried work is positive. Perhaps, in the case of informal salaried work receiving non-labour income may be an outcome rather than a determinant of employment.

Surprisingly, household assets or wealth is not a strong predictor of employment state. Interestingly, individuals from wealthier families are more likely to work. Perhaps, the results signal the underlying factor, the role of family or social networks in accessing jobs in Uganda's labour market. Similarly, individuals from wealthier families are more likely to work in household enterprises, which signal a higher probability of wealthy households to run household enterprises. Further, it is interesting to note that individuals from wealthy families find informal salaried work less attractive, which suggests that jobs in this segment could be offering low remuneration that are unattractive to individuals from wealthier families. Parental education and occupation are common predictors of labour participation in the literature; however, these variables could not be included because they had many missing values in the data.

Residing in an urban area increases an individual's chance of being observed as not-working, in self-employment, informal and formal salaried work; but reduces the likelihood of working as a household worker. These results are plausible, suggesting that workers in urban areas have a higher reservation wage compared to their counterparts in rural areas; and are therefore less likely to enter the labour market or accept job offers with low remuneration. On the other hand, workers in urban areas may be more productive due to higher levels of human and financial capital than their counterparts in rural areas, which increase their chance of selecting into self-employment, informal and formal salaried work. The negative relation between residing in an urban area and working in a household enterprise is not surprising because most of the enterprises are agricultural related.

The results for the regional categorical variable suggest regional disparities, especially in the northern region which is distinct from the other regions. For example, an individual residing in the northern region is less likely to work in the formal sector compared to a counterpart in the central region (the results for the eastern and western region are insignificant). These results are not surprising since the northern region experienced economic and social decline during the twenty-year Lord's Resistance Army (LRA) civil war, which led to institutional break-down. Interestingly, the probability of not-working increases when an individual resides in the eastern region compared to the central region. These results are surprising and therefore, it would be instructive to know the underlying factors increasing the reservation wage for workers in the eastern region. Lastly, the time dummies suggest that there are strong time effects influencing the decision to work in a household enterprise or being self-employed. The results imply that over time the probability of working as a household worker has

increased, and the probability of being self-employed has decreased conditioned on all variables in the model. These results signal the expansion of household work and contraction of sole proprietorship segments of the Uganda's labour market, which requires further investigation to establish the predictors for this trend.

As a robustness check equation (9) is re-estimated using years of education and also a quadratic term for education is included to further test for non-linearity in the relationship between education and being observed in a particular employment state. The results are provided in Table 4 and they are consistent with earlier results in Table 3. For example, we still find a positive relation between education and not-working, or working as informal salaried or formal salaried. Similarly, we find a negative relation between education and working as a household worker. With reference to non-linearity, there is a strong evidence for concavity for the relation between education, and working as a household worker and being self-employed. These results suggest that individuals with lower levels of education have a higher probability of being observed working in household enterprises or self-employment, compared to their counterparts conditioned on the variables in the models. In contrast, strong convexity for the relation between education and working as informal salaried is found. Therefore, less educated individuals are less likely to be observed in informal salaried work compared to their counterparts, *ceteris paribus.* We argue that the results for non-linearity for not-working and formal salaried are inconclusive because of the small sample size for this group of workers as shown in Table 1. Given that the IIA assumption was violated, implying that it may not be appropriate to analyse the choices separately, we estimate a multivariate probit model and the results are provided in Table A2 in the Appendix. A comparison of the results in the two tables shows they are consistent, which validates the results in Table 4.

6.1 Estimated Probabilities by Birth Cohort
The paper also investigates whether the relationship between education and selection into employment in Uganda's labour market varies by birth cohort. The sample is divided into the youth who were born from 1981 onwards, and adults born before 1981. Therefore, the youth were aged 14-24 years in the baseline wave. We can observe a wide variation in the results for the not-working group between the two age groups. For instance, in the case of an adult, it is only an individual with degree education who is associated with an increased probability of not-working, while a youth with post-primary education is more likely not to work. These results signal a higher reservation wage for the young educated individuals compared to the older counterparts, which could be due to the availability of family support and/or the lack of job opportunities for more educated individuals in Uganda's labour market over time. The results for the household worker segment show that an educated adult is less likely to be observed working in a household enterprise compared to the youth counterpart conditioned on the covariates in the models. Similarly, the results for self-employment are driven by the adult cohort, with most of the education dummies associated with selection into self-employment.

As expected, an adult with lower secondary education is more likely to be self-employed than a youth counterpart, *ceteris paribus.* Perhaps, this is due to

differences in accumulated human and financial capital, although cohort effects such as differences in the economic and social environment may be influencing the results. Interestingly, we note that an educated adult is more likely to be observed as an informal salaried worker than a youth counterpart, conditioned on the variables in the model. These results are surprising, but perhaps reinforce the results for not-working, which suggest that an educated youth has a higher reservation wage than their adult counterparts, and are less likely to select into jobs with low remuneration. In reference to formal salaried work, an educated adult has a higher probability of selecting into formal employment. These results could be signalling the oversupply of educated young individuals that does not match their demand.

6.2 Estimated Probabilities by Gender
The hypothesis that the role of education for selection into employment varies by gender is tested and the results are presented in Table 6. We observe that the results for a female and male are all in the same direction, but the magnitude of the effect varies. For instance, when gender is accounted for, the results show a strong positive association between degree education and not-working for both genders, although the effect is greater for females compared to males. These results suggest that an educated female has a higher reservation wage than the male counterpart, which could be because society considers males as bread winners for their families, and therefore males are more motivated to work than females. In case of household workers, an educated female is less likely to be observed working in a household enterprise than a male counterpart.

Interestingly, a female with lower secondary education is more likely to be self-employed compared to a male counterpart. Perhaps, males with the same education have a higher probability of being observed in formal employment, which is supported by the results for formal salaried workers in the last two columns in Table 6. The results for the formal salaried are interesting, showing a different pattern along the education profile. For instance, at lower levels of education females have a higher probability of selecting into formal salaried work compared to males, while this is reversed for degree education. These results mirror both supply and demand-side factors. On the supply-side, females with higher education may not be a good match with their male counterparts. On the demand-side, employers may have a preference for male workers who have less work disruptions: for example, paternity leave is for only four days, while maternity leave is for sixty days.

6.3 Estimated Probabilities by Residence
Given that there is a wide disparity in the socio-economic characteristics between urban and rural areas in Uganda, this paper investigates whether the association between education and selection into employment varies by residence. Table 7 presents the average marginal effects of educational attainment. The association between education and not-working is stronger for the urban sub-sample, except for individuals with diploma education. With reference to workers in household enterprises, overall the association between education and employment is stronger for rural areas compared to urban areas.

Table 6: Estimated Probabilities of Selection into Employment Given Educational Attainment by Gender

Variable	Not-working		Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male
Primary education	0.083	0.104	-0.260**	-0.028	0.261**	0.067	0.079	-0.016	0.856**	0.662**
	(0.105)	(0.128)	(0.102)	(0.075)	(0.118)	(0.089)	(0.143)	(0.096)	(0.435)	(0.324)
Lower secondary	0.149	0.154	-0.715***	-0.392***	0.547***	0.283***	0.602***	0.130	1.074**	1.480***
	(0.101)	(0.127)	(0.105)	(0.079)	(0.122)	(0.092)	(0.133)	(0.096)	(0.430)	(0.303)
Upper secondary	0.283*	0.231	-1.518***	-0.763***	0.302	0.020	1.204***	0.653***	3.356***	2.187***
	(0.151)	(0.170)	(0.171)	(0.116)	(0.198)	(0.131)	(0.194)	(0.130)	(0.536)	(0.355)
Diploma	0.209	0.102	-2.191***	-1.485***	-0.291	-0.459***	1.946***	1.201***	3.632***	3.212***
	(0.171)	(0.218)	(0.195)	(0.145)	(0.221)	(0.156)	(0.206)	(0.146)	(0.547)	(0.386)
Degree	1.082***	0.939***	-2.418***	-1.116***	-1.111**	-0.701***	2.138***	0.144	3.904***	4.098***
	(0.199)	(0.256)	(0.427)	(0.249)	(0.507)	(0.266)	(0.370)	(0.251)	(0.689)	(0.500)
Log likelihood	-691.32	-706.76	-2317.04	-3257.27	-1897.73	-2644.34	-1327.1	-2563.17	-286.84	-584.05
Wald chi2 (22)	212.44	75.02	479.54	525.58	258.22	281.71	203.99	304.75	62.11	103.43
Observations	5,850	5,919	5,667	5,741	5,667	5,741	5,667	5,741	5,667	5,741

Note: Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. Other controls include: age, gender, marital status, household head dummy, youth dummy, non-labour income, household assets, residence and year dummies. Uncompleted primary is the reference group.

Table 7: Estimated Probabilities of Selection into Employment Given Educational Attainment by Residence

Variable	Not-working		Household Labour		Self-employed		Informal Salaried		Formal Salaried	
	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban	Rural
Primary education	0.251	0.098	-0.211	-0.090	0.224	0.115	0.068	0.022	0.838*	0.722**
	(0.174)	(0.089)	(0.179)	(0.065)	(0.200)	(0.075)	(0.188)	(0.086)	(0.506)	(0.303)
Lower secondary	0.278*	0.166*	-0.307*	-0.548***	0.139	0.461***	0.135	0.329***	1.356***	1.403***
	(0.149)	(0.093)	(0.165)	(0.069)	(0.187)	(0.080)	(0.167)	(0.087)	(0.475)	(0.288)
Upper secondary	0.364**	0.237*	-0.582**	-1.171***	-0.210	0.250**	0.440**	0.983***	2.174***	2.911***
	(0.184)	(0.142)	(0.233)	(0.109)	(0.250)	(0.126)	(0.217)	(0.125)	(0.554)	(0.351)
Diploma	0.245	0.332**	-1.480***	-1.879***	-0.591**	-0.352**	0.850***	1.790***	3.107***	3.533***
	(0.220)	(0.165)	(0.272)	(0.134)	(0.259)	(0.161)	(0.219)	(0.143)	(0.591)	(0.372)
Degree	1.205***	0.877***	-1.632***	-1.527***	-0.998***	-1.104**	0.220	1.574***	4.095***	3.935***
	(0.253)	(0.277)	(0.403)	(0.298)	(0.378)	(0.463)	(0.304)	(0.312)	(0.745)	(0.509)
Log likelihood	-403.31	-1004.96	-1015.51	-4573.17	-1110.56	-3422.1	-1042.16	-2861.42	-377.36	-498.3
Wald chi2 (22)	49	82.08	127.99	747.72	95.12	392.57	127.2	401.32	46.69	119.01
Observations	2,254	9,515	2,123	9,285	2,123	9,285	2,123	9,285	2,123	9,285

Note: Standard errors in parentheses, * p < 0.10, ** p < 0.05, *** p < 0.01. Other controls include: age, gender, marital status, household head dummy, youth dummy, non-labour income, household assets, residence and year dummies. Uncompleted primary is the reference group.

For the self-employed, generally the relation between education and employment is greater for rural areas compared to urban areas. For example, an individual residing in a rural area with secondary education is more likely to be self-employed than the counterpart in an urban area, *ceteris paribus*.

6.4 Dynamic Effects
A dynamic model specified in equation (10) is estimated, and the paper finds evidence of state dependence in all employment states with the effect greatest for formal employment and smallest for not-working as Table 8 shows. These results suggest that in a period of one year, an individual is more likely to remain in the same employment state. In view of the results, we can infer that formal employment is the most desired employment state, followed by informal employment, while not-working is the least desired state. Overall, the results are consistent with the results for the static model, but with slightly less significant variables and a reduction in the magnitude of the effect. For example, the relation between age and employment state becomes weaker, and only significant for the not-working and formal salaried category. These findings reinforce the observation that the probability of not-working reduces with age, and the likelihood of working in the formal sector increases with age.

With reference to education, the key explanatory variable of interest, the results are consistent, except for the signs switching for the education dummies for the not-working group. Interestingly, the positive effect of education on not-working generally disappears for all education levels except for degree education.

However, the relation is not significant at conventional levels, except with diploma education, which would weakly reduce the probability of not-working. In view of these results, we conclude that the observed positive relation between education and not-working is a short-term effect for individuals with less than degree education, which disappears when state dependence is considered.

7. Conclusion
Using four household panel waves—2005/05, 2009/10, 2010/11 and 2011/12 – this paper has investigated the relationship between educational attainment and employment states for individuals in Uganda's labour market. This relationship was estimated using the correlated random effects model, which accounts for individual heterogeneity and also relaxes the assumption of strict exogeniety of the fixed effects with the covariates in the model by including averages of time varying covariates (Wooldridge, 2010). The paper could not use the multi-sector models, such as the multinomial logit, because the IIA assumption was violated. However, an attempt was made to investigate the role of educational attainment in selection of employment using a multivariate probit model with few covariates because the full model could not converge. Therefore, the use of the multivariate probit model had limited scope, and consequently the results were used as a robustness check for the model with years of education.

Table 8: Estimated Probabilities of Selection into Employment from the Dynamic Model

Variable	Not-working		Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	AME	Std. Err	AME	Std. Err	AME	Std. Err	AME	Std. Err	AME	Std. Err
Lagged dependent variable	0.789***	0.172	1.238***	0.047	1.387***	0.050	1.329***	0.060	1.414***	0.122
Age	-0.061**	0.030	-0.005	0.025	-0.015	0.027	0.013	0.029	0.171**	0.068
Age squared	0.001**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.002***	0.001
Household size	0.015	0.015	-0.017	0.021	0.017	0.022	0.013	0.025	-0.056	0.046
Primary education	0.073	0.106	-0.058	0.057	0.033	0.062	0.085	0.075	0.288	0.217
Lower secondary	-0.158	0.132	-.288***	0.061	0.213***	0.064	0.222***	0.075	0.692***	0.188
Upper secondary	-0.239	0.216	-.625***	0.094	0.010	0.100	0.714***	0.097	1.096***	0.201
Diploma	-0.535*	0.303	-.950***	0.105	-0.233**	0.114	1.039***	0.098	1.445***	0.192
Degree	0.168	0.276	-.937***	0.235	-0.494**	0.222	0.514***	0.192	1.870***	0.255
Female	-0.011	0.098	0.438***	0.053	-0.079	0.059	-.578***	0.064	-0.183	0.125
Married	-0.165	0.101	0.159***	0.055	0.101	0.062	-.353***	0.064	-0.036	0.129
Head	-0.069	0.118	-0.076	0.061	0.264***	0.068	-0.158**	0.073	-0.027	0.137
Youth	-0.162	0.183	-0.148	0.092	0.156	0.101	-0.025	0.109	-0.146	0.232
Non-labour income	0.059	0.090	-0.111**	0.047	-0.002	0.051	0.173***	0.056	0.172*	0.103
Household assets	-0.008	0.028	0.035	0.024	-0.009	0.026	-0.050*	0.028	-0.017	0.055
Urban	0.231**	0.110	-.768***	0.061	0.496***	0.059	0.306***	0.064	0.327***	0.104
Eastern region	0.294***	0.111	0.176***	0.062	0.018	0.065	-.203***	0.074	-0.279**	0.142
Northern region	-0.170	0.133	0.428***	0.060	-.218***	0.065	-.246***	0.072	-0.278**	0.142
Western region	-0.027	0.129	0.172***	0.063	-0.086	0.069	-0.071	0.072	-0.073	0.134
Year dummy - 2011	0.045	0.085	0.155***	0.055	-0.102*	0.059	-0.098	0.065	-0.175	0.128
Log likelihood	-479.77		-2195.96		-1848.58		-1451.33		-379.99	
Wald chi2 (20)	91.87		1716.49		1118		1046.62		420.66	
Observations	5,300		5,098		5,098		5,098		5,098	

Note: * p < 0.10, ** p < 0.05, *** p < 0.01. Uncompleted primary and the central region are the reference group.

Initially, dummies for educational attainment are included in the models and the paper finds that the probability of not working (unemployed) increases with education. These results are surprising and contrast evidence from similar studies (Glick & Sahn, 1997; Wamuthenya, 2010; Baffour, 2013 for the Tanzanian labour market), except Baffour (2013) who found similar results for the Ghanaian labour market. Later a dynamic model is estimated and the relationship between education and not working becomes negative for all education levels except primary and degree education (though they are not significant at conventional level). Therefore, we conclude that the observed positive relationship is a short-term effect that may not be observed after a period of one year in Uganda's labour market. In reference to sector employment, the paper finds that the probability of being observed in household work decreases with education, while the probability of being observed in paid employment (informal and formal salaried) increases with education. The relationship between education and self-employment is mixed, suggesting that individuals with low education increase the probability of being observed in self-employment, while those with post-secondary would decrease the likelihood of being in self-employment.

As robustness check equation (9) is re-estimated using years of education and also a quadratic term for education to test for non-linearity in the relationship between education and employment is included. The paper finds the results are consistent showing that individuals with more education are more likely to be unemployed conditioned on covariates in the model. Perhaps these results are signalling the lack of jobs for more educated individuals in Uganda's labour market. Further, in this paper I find that less educated individuals are more likely to be observed in household work and self-employment than their counterparts *ceteris paribus*. These results are plausible suggesting that household work and self-employment are low productivity sectors attracting low productivity individuals. In addition, more educated individuals are more likely to be observed in paid employment; informal and formal salaried work compared to their counterparts *ceteris paribus*. The results from the multivariate model reinforce findings from the random effects model indicating that individuals with less education are more likely to be observed in household work and self-employment, while individuals with more education are more likely to be observed in paid employment conditioned on the variables in the models. These results are consistent with available literature, signalling the selection of high productivity individual into high productivity sectors (segments) of the labour market in Uganda.

In this paper the informal sector is defined to consist of household work, self-employment and informal salaried. In view of the results for education, the informal sector is heterogeneous and could be hierarchical with household work occupying the lower position, self-employment in the middle and informal salaried the upper position. I suggest that future work on employment in the informal sector could adopt this ordering. This paper was not able to disentangle the effect of entry barriers and individual preferences in selecting into employment and therefore, it is not possible to know whether individuals select into the informal sector by choice

or chance. However, the results signal that more productive workers (more educated) prefer paid employment especially formal employment.

With reference to the parameter estimates by birth cohort there are distinct differences in the results for the youth (born in 1981 onwards) and the adult (born before 1981) for the not-working (unemployed) employment state. For the adults it is only degree education which is associated with an increase in the probability of not working, while for the youth this positive association is strong post primary education. I conclude that these are cohort not age effects that mirror changes in the social and economic environment in Uganda over time. For example, the SAPs which were implemented in the 1980s and the universal primary and secondary education programmes implemented in 1997 and 2007 respectively. Therefore, the observed trend could be a result of the supply of more educated individuals outstripping their demand. The parameter estimates by gender, show that the association between education and selection into employment is stronger for females than for males *ceteris paribus* except for individuals with a degree selecting into formal salaried work. This is an interesting finding which could be signalling gender discrimination in formal firms.

With regard to parameter estimates by residence; urban vs rural areas, the paper finds that the positive relation between educational attainment and being observed not working is stronger in urban compared to rural areas conditioned on the covariates in the models. These results signal the differences in economic development for urban and rural areas with the former more developed and therefore individuals in these areas are expected to have a higher reservation wage than their counterparts in rural areas. In reference to employment in the different segments of Uganda's labour market, overall the parameter estimates for rural areas are stronger than urban areas except for individuals with degree education selecting into formal employment. This difference in results could be attributed to the concentration of large firms in urban areas which offer jobs attractive to more educated individuals.

Several factors such as an individual's motivation to work, intellectual ability, family background, local unemployment rate and employment contracts which are likely to affect selection into employment are unobserved in the estimated models. Therefore, the results should be interpreted with this caveat in mind. Also, the panel data is unbalanced and therefore the results could be biased by attrition if it is non-random. This paper was constrained by the choice of estimators which do not allow use of weights and therefore could not address the attrition bias in the estimates. Lastly, education is widely believed to be endogenous and Comola and de Mello (2010) find that when education is instrumented its positive association with employment increases. However, in this paper education was assumed to be exogenous because of lack of valid instruments in the data.

References

Appleton, S. 2001. Education, Incomes and Poverty in Uganda in the 1990s. University of Nottingham, School of Economics CREDIT Research Paper, No. 01/22.

Arulampalam, W., A. L. Booth & M. P. Taylor. 2000. *Unemployment Persistence.* Oxford Economic Papers 52. 2000), 24–50.

Baffour, P., T. 2013. Selection into Employment Sectors in Urban Ghana and Tanzania. CREDIT Research Paper. No. 13/02.

Becker, G. S. 1962. Investment in Human Capital: A Theoretical Analysis. *The Journal of Political Economy*, Vol.70, Issue 5 Part 2, Investment in Human Beings, 9–49.

Cain, G. 1976. The Challenge of Segmented Labour Market Theories to Orthodox Theory: A Survey. *Journal of Economic Literature.* 14 (4, 1215–1257.

Comola, M. & L.R. de Mello. 2011. Salaried Employment and Earnings in Indonesia: New Evidence on the Selection Bias. Paris School of Economics, OECD Economics Department

Doeringer, P. & M. Piore. 1971. *Internal Labour Markets and Manpower Analysis.* London: Sharpe Publishers.

Fields, S. G. 1990. Labour Market Modelling and the Urban Informal Sector: Theory and Evidence. *http:// digitalcommons.ilr.cornell.edu/articles.*

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

Harris, R. J. & M. P. Todaro. 1970. Migration, Unemployment and Development: A Two-sector Analysis. *American Economic Review*, 1970, 60(1): 126–42

Heckman, J. 1993. What Has Been Learned About Labour Supply in the Past Twenty Years?' *The American Economic Review*, 83(2). Papers and Proceedings of the Hundred and Fifth Annual Meeting of the American Association, 116–121.

International Labour Office (ILO). 1972. *Employment, Income and Equality: A Strategy for Increasing Productivity in Kenya.* Geneva.

—. 1993. Resolution Concerning Statistics of Employment in the Informal Sector, Adopted by the Fifteenth International Conference of Labour Statisticians.

Killingsworth, M. & J. Heckman. 1986. Female Labour Supply: A Survey. *Handbook of Labour Economics.* Elsevier Science Publishers.

Magnac, 1991. Segmented or Competitive Labor Markets. *Econometrica*, 59(1): 165–187.

Maloney, W. 2004. Informality Revisited. *World Development.* 32(7): 1159–1178. Elsevier Limited.

Pencavel, J. 1986. Labour Supply of Men: A Survey. *Handbook of Labour Economics.* Vol.1, Chapter 1, Elsevier Science Publishers.

Pradhan, M. & A. van Soest. 1995. Formal and Informal Sector Employment in Urban Areas of Bolivia. *Labour Economics,* 2: 275–297 Elsevier.

Ssennono, F. V. 2013. The Uganda National Household Survey 2012/13.

Uganda Bureau of Statistics. 2010. Uganda National Household Survey Report: Socio-Economic Module.

Wamuthenya, W. R. 2010. Determinants of Employment in the Formal and Informal Sectors of the Urban Areas of Kenya. African Economic Research Consortium (AERC) Paper 194.

Wooldridge, J. 2010. Correlated Random Effects Models with Unbalanced Panels. Paper presented at the 15th Conference on Panle Data, Bonn, Germany, July 3–5, 2009.

World Bank. 2014a. *2005/06 Uganda National Household Survey Data. Databank*. Washington, DC: World Bank.

—. 2014b. *2009/10 Uganda National Household Survey Data. Databank*. Washington, DC: World Bank.

—. 2014c. *2010/11 Uganda National Household Survey Data. Databank*. Washington, DC: World Bank.

—. 2014d. *2011/12 Uganda National Household Survey Data. Databank*. Washington, DC: World Bank.

Appendix

Table A1: Sample Means of Explanatory Variables Across Waves

Variable	Pooled	2005/06	2009/10	2010/11	2011/12
Age	22.30	21.25	22.05	22.03	23.61
	(4.78)	(4.38)	(4.99)	(4.57)	(4.76)
Uncompleted primary	0.51	0.54	0.51	0.51	0.48
	(0.49)	(0.50)	(0.50)	(0.50)	(0.50)
Primary	0.17	0.19	0.18	0.16	0.17
	(0.38)	(0.39)	(0.38)	(0.37)	(0.37)
Lower secondary	0.21	0.20	0.21	0.20	0.22
	(0.41)	(0.40)	(0.40)	(0.40)	(0.41)
Upper secondary	0.06	0.04	0.06	0.07	0.07
	(0.21)	(0.21)	(0.23)	(0.26)	(0.26)
Diploma	0.03	0.02	0.03	0.03	0.04
	(0.17)	(0.14)	(0.17)	(0.18)	(0.20)
Degree	0.01	0.005	0.02	0.02	0.02
	(0.09)	(0.07)	(0.13)	(0.14)	(0.12)
Household size	6.49	6.47	6.36	6.55	6.58
	(3.37)	(3.64)	(3.27)	(3.38)	(3.29)
Female	0.53	0.54	0.52	0.53	0.53
	(0.50)	(0.50)	(0.50)	(0.50)	(0.50)
Head	0.20	0.20	0.17	0.20	0.21
	(0.40)	(0.40)	(0.38)	(0.40)	(0.41)
Non-labour income	0.36	0.45	0.30	0.37	0.34
	(0.48)	(0.50)	(0.46)	(0.48)	(0.47)
Household assets (log)	9.95	8.68	10.36	10.35	10.27
	(1.94)	(1.80)	(1.82)	(1.83)	(1.80)
Urban	0.23	0.27	0.27	0.21	0.19
	(0.42)	(0.44)	(0.44)	(0.41)	(0.39)
Household consumption	331.35	341.481	302.86	301.50	371.12
	(540.44)	(550.88)	(492.31)	(427.15)	(638.82)
Education policy	0.23	0.07	0.24	0.27	0.33
	(0.42)	(0.26)	(0.43)	(0.44)	(0.47)

Source: Author's construction based on World Bank (2014a-d). Standard deviation in parenthesis, household assets and household consumption were deflated by the CPI for the respective years.

Table A2: Estimated Probabilities of Selection into Employment from Multivariate Probit Model (MVP).

Variable	Household Worker		Self-employed		Informal Salaried		Formal Salaried	
	MVP (1)	MVP (2)	MVP (1)	MVP (2)	MVP (1)	MVP (2)	MVP (1)	MVP (2)
Years of education	-0.077***	0.018	-0.003	0.112***	0.057***	-0.018	0.183***	0.017
	(0.003)	(0.012)	(0.003)	(0.012)	(0.003)	(0.013)	(0.006)	(0.029)
Education squared		-0.006***		-0.007***		0.004***		0.008***
		(0.001)		(0.001)		(0.001)		(0.001)
Log likelihood	-23113.41	22940.37	-23113.4	22940.37	-23113.41	22940.37	-23113.41	22940.37
Wald chi2 (16/20)	3925.85	4331.34	3925.85	4331.34	3925.85	4331.34	3925.85	4331.34
Observations	16,813	16,813	16,813	16,813	16,813	16,813	16,813	16,813

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Other controls include: age, gender, marital status, household head dummy, youth dummy, non-labour income, household assets, residence and year dummies.