

Twumasi Baffour, Priscilla

**Working Paper**

## Selection into employment sectors in urban Ghana and Tanzania

CREDIT Research Paper, No. 13/02

**Provided in Cooperation with:**

The University of Nottingham, Centre for Research in Economic Development and International Trade (CREDIT)

*Suggested Citation:* Twumasi Baffour, Priscilla (2013) : Selection into employment sectors in urban Ghana and Tanzania, CREDIT Research Paper, No. 13/02, The University of Nottingham, Centre for Research in Economic Development and International Trade (CREDIT), Nottingham

This Version is available at:

<https://hdl.handle.net/10419/96353>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



## **Selection into Employment Sectors in Urban Ghana and Tanzania**

by  
**Priscilla Twumasi Baffour**

### **Abstract**

The paper analyses labour market participation and selection into specific employment sectors in urban areas of Ghana and Tanzania using urban worker survey data from 2004 to 2006 in pooled sample models. A baseline participation model and multinomial logit model of occupational choice is applied. Results from the baseline models suggest determinants of participation differ across the two countries pointing out the heterogeneity in the two labour markets. Specifically in Tanzania, education is found to increase the probability of employment but the reverse is the case in Ghana as education increases the probability of not-working (unemployment). In terms of selection into specific sectors, the results suggest both labour markets are characteristic of a preference for formal sector employment by the highly educated as all levels of education are found to monotonically increase the probability of public and private (formal) sector employment and decrease the probability of self-employment (informal sector). Among age cohorts, young people face more challenges in accessing formal sector jobs in both countries compared to older people.

### **JEL Classification:**

**Keywords:** Selection, Sector of Employment, Unemployment, Labour Market, Ghana, Tanzania

---

**Centre for Research in Economic Development and International Trade,  
University of Nottingham**



**Selection into Employment Sectors in Urban Ghana and  
Tanzania**

by

**Priscilla Twumasi Baffour**

**Outline**

1. Introduction
2. Empirical Literature
3. Methodology
4. Conceptual framework / Data
5. Model specification / Results
6. Conclusion

**The Authors**

The author is a Research Student in Development Economics in the School of Economics, University of Nottingham.

Contact e-mail: [lexpt@nottingham.ac.uk](mailto:lexpt@nottingham.ac.uk)

**Acknowledgements**

I am grateful to Professor Oliver Morrissey and Doctor Trudy Owens for valuable guidance and comments.

## **1. Introduction**

The distinction of formal/informal has become the standard way to describe urban labour markets in developing countries, typical of the urban labour market structure in Africa. The formal sector is subject to regulations and the existence of explicit contracts between employers and employees. The informal sector on the other hand is not wholly subject to government regulations and is dominated by individuals and small enterprises that employ few apprentices or hired labourers. An ILO mission to Kenya in 1972 defined economic informality in a broad context by the use of such characteristics as ease of entry, small scale of operation, family ownership, skills acquired outside the formal school system and unregulated and more competitive markets. The concurrent existence of a formal and informal sector is viewed as a result of labour market segmentation with formal sector wages above the market clearing rate that leaves a large pool of people unemployed. As a result, the informal sector is oversupplied with labour at strikingly lower earnings. Human capital theory offers possible explanations for earnings differentials by arguing that worker productivity explains earnings and sector choices (assumed to be utility maximising), despite the possible existence of sector earnings differentials.

The position and nature of the informal sector compared to the formal sector is crucial for the functioning of the labour market and the overall economic structure in countries with a large informal sector. It affects income distribution (inequality) and poverty and has implications for efficiency in terms of allocation of labour and the distortions in the formal sector due to taxes, social security and labour market regulations. This explains why the role of the informal sector has recently been analysed extensively.

The literature on labour markets presents two competing views, segmentation theory (staging hypothesis) and the symmetric market assumption. The traditional staging hypothesis (Fields 1975) is an expansion of the Harris-Todaro model which formalised Lewis (1954) concept of labour market dualism. The theory postulates that formal sector employment is rationed, those unable to obtain formal sector jobs and cannot afford to be unemployed while they search, work in the informal sector. Based on the dual labour market model, the informal sector is considered as a residual component of the market with informal sector workers who suffer from poor labour

market conditions queuing for better jobs in the formal sector. Boudarbat (2004) supports this with the observation of a preference for public sector employment in Africa and a willingness of the educated to engage in 'wait' unemployment to secure well paid and stable public sector jobs. The second view recognizes the informal sector as competitive to the formal sector (symmetric market assumption) with the two sectors characterised by different production functions. As a result of this heterogeneity, some workers are more productive in the formal sector and others in the informal sector. Recent empirical evidence suggests that the real situation is a mixture of these two hypotheses (Maloney, 1999, 2004; Perry *et al.*, 2007).

Empirical evidence on labour market segmentation and sector choice in Africa points to a form of sorting process in order for workers to be situated in a particular employment status. When an individual takes the decision to participate in the labour market, we seek to identify which personal and socioeconomic characteristics are important. Specifically, what labour supply factors are associated with a greater or lesser likelihood of employment in the formal public or private sectors, the informal sector (self-employment) and unemployment. With the informal sector as the fastest growing segment of the labour force in both rural and urban areas in Africa, it is imperative to analyze available data to highlight the characteristics for selection into the different employment sectors to shed light on heterogeneity or any existing discrimination in the labour markets. Consequently, the study applies a baseline probit model of participation to shed light on determinants of participation and further adopt the multinomial logit model to investigate determinants of participation in the formal public and private sectors in addition to the informal sector in both Ghana and Tanzania.

Youth unemployment is an important issue in both developed and developing countries. Youth unemployment is substantially higher than global adult unemployment and has been growing in the last decade (ILO 2006). Existing empirical evidence shows the nature of youth unemployment problem is quite different in developed than developing countries. In developed countries, the difficulties in getting a job as a youth are due to lack of minimum professional skills required in a sophisticated environment with competitive skilled labour supply; in developing countries context, youth unemployment is generally found to rise with educational levels. In the absence of unemployment insurance, only those with family

(economic, social and demographic) resources can afford to wait in order to find a good match between their level of qualification and occupations in the labour market. Consequently, most qualified young people cannot afford to be unemployed and end up in the informal sector where productivity and revenues are low. We therefore seek to further investigate variations by age group (young and old) in employment sector choice in both Ghana and Tanzania.

The study uses the Urban Household Worker Surveys for Tanzania and Ghana collected in three rounds: 2004 (1818 in Ghana and 653 in Tanzania), 2005 (895 in Ghana and 443 in Tanzania) and 2006 (309 in Ghana and 572 in Tanzania). Section 4 provides details on the data.

The following section presents empirical literature, this is followed by the methodology and conceptual framework / data in sections 3 and 4 respectively. Model specification and empirical results are presented in section 5 and finally we conclude in section 6.

## **2. Empirical Literature**

There exist an extensive literature on labour market participation and multi sector labour market modelling. This literature is generally modelled in two strands based on the dualistic labour market assumption by Fields (1975) and the symmetric labour market assumption by Heckman and Sedlacek (1985).

Labour markets in most developing countries are segmented into broadly defined formal and informal sectors (Bourguignon et al., 2003) in a dualistic labour market. Dualism in labour markets arise when earnings differ for workers with similar characteristics depending on the sector of the economy in which they find work; in essence for dualism to exist, different wages must be paid in different sectors to comparable workers. This is grounded in human capital theory developed by Schultz (1961, 1962), Becker (1962, 1964) and Mincer (1962, 1974). This section traces some of the developments in multi sector modelling (occupational choice) and their particular relevance to this study.

Labour market dualism dates back to Lewis (1954) who expressed the view that the rural sector constitutes a stock of potential workers for the urban, formal sector where jobs pay higher wages. This view was subsequently formalized in the Harris and Todaro (1970) model where urban wages are assumed higher than rural wages. Rural workers who choose to search for urban jobs run the risk of becoming unemployed. In equilibrium, expected wages are equated across sectors due to the mass of workers who choose to search in unemployment. Fields (1975) expanded on the Harris-Todaro model by assuming urban workers can choose to become informally employed rather than search for higher paying formal jobs.

According to Fields (1975), formal sector employment is rationed: those who cannot obtain a formal sector job and cannot afford to search from unemployment work in the informal sector. Informal sector jobs are secondary and workers would be better off with a primary job in the formal sector. The informal sector is seen generally as an inferior alternative to formal sector employment in terms of earnings and other contractual arrangements, but informal sector workers are prepared to remain there hoping to enter the formal sector. The informal sector therefore is seen as providing the poorest and most marginalised people who do not have access to the formal sector with opportunities to earn income [and gain some work experience]. Informal sector employment is therefore worse than formal sector employment but superior to unemployment. The free-entry employment opportunities in the informal sector explain the reason why open unemployment rates in developing countries are comparable to those of developed countries and often considerably lower (Turnham, 1971, 1993; World Bank, 1995; ILO, 2003). According to the proponents of this view, efforts should be made to expand employment in the formal sector for the informal sector to eventually disappear (Moser 1984). Conceptually, this allows a distinction between the informal sector (developing countries) and the shadow economy (developed countries), where the latter is an alternative that is typically illegal rather than a stage.

In the context of a developing country, Fields (1975) extension of the Harris and Todaro (1970) model of migration has become the basis for much of the empirical literature that explores the existence of segmentation in the labour market. The model consists of an urban, informal (free-entry sector), in addition to the urban formal sector and rural agriculture. The existence of minimum wage regulations or union

activity in the formal sector creates wage differentials between the formal and informal sectors that make those in the informal sector worse off.

The other strand of the literature classifies the informal and formal sectors as symmetric and competitive with the two sectors characterised by different production functions. Based on this heterogeneity, some workers are more productive in the formal sector and others in the informal sector, this leads to a preference for informal sector employment compared to formal sector employment. Recent empirical evidence confirms this by pointing to a preference contrary to formal sector employment, many informal sector workers favour their current status to formal sector employment (Thomas, 1992).

Pradhan and van Soest (1995) tested the staging hypothesis and the symmetric sector assumption in urban Bolivia to explain the choice between formal sector, informal sector, and not working. The study modelled the choice of employment state in two different choice models, the multinomial logit and the ordered probit model based on the symmetric sector assumption and the staging hypothesis respectively. Subsequently, after estimating the models for both men and women, they find that an ordered model performs better for men while an unordered model fits the data better for women. Pradhan and van Soest (1997) used the same data in a structural labour supply model and find that wage differentials between the formal and informal sector tend to be negative rather than positive, suggesting that non-monetary job characteristics such as job stability, social security, health care access among others are needed to explain why many people prefer formal sector jobs.<sup>1</sup>

Within the formal sector, numerous studies point to the existence of formal labour market segmentation, but lack consensus on the existence of public wage premiums, particularly in developing countries. Results generally depend on how selection is modelled. Tansel (1999) examined how individuals are selected into employment type in a multinomial logit model and further carried out decomposition on a selectivity corrected wage equation by gender in Turkey. The study found all levels of educational attainment increase the probability of joining public administration, state owned enterprises (SOEs) and a covered private sector but reduces the probability of

---

<sup>1</sup> Strassmann (1987) in a similar manner found out that 71% of home workers in Lima would require a considerable financial incentive to move to the formal sector.

participation in other employment categories and subsequently concluded in terms of segmentation of high wage premiums for men in private sector compared to public administration while women in public administration have higher premiums than their counterparts in the private sector.<sup>2</sup>

In Africa, Lindauer and Sobat (1983), Andersson (1993), Van der Gaag and Vijverberg (1988) have discussed issues relating to labour market segmentation and mostly sought answers to why the preference for formal public sector jobs<sup>3</sup>. Lindauer and Sobat (1983) used data for 1971 household survey of the Tanzanian urban wage labour force to determine the pattern of wage differentials across employment sectors. After standardising by worker characteristics, they found a substantial wage premium for public employees (14%) over their private sector counterparts. Van der Gaag and Vijverberg (1988) in a similar study on Côte d'Ivoire, used a switching regression model in sector choice analysis and argued that taking account of sector allocation could reverse the direction of the wage differentials that exist between the public and private wage offers. They further emphasize that in addition to possible wage differentials, there are numerous other factors that make a public sector job preferred to a private job. Andersson (1993) identified differences between public and private wage structures and noted that a distinction between formal and informal sectors is important for men, but not for women in Zambia. Skyt-Neilsen and Rosholm (2001) detected a positive average *ceteris paribus* pay gap in favour of public sector workers in Zambia but noted that at the upper end of the conditional wage distribution it became negative for the highly educated. In part of a study by Teal (2001), earnings determination in the public and private sectors in Ghana was considered with selectivity correction for non-random assignment into the wage sector. The study modelled selection into wage employment in a binary probit model of wage employment for men and women with controls for parental background and education. Findings that emerged from the probit model point to the

---

<sup>2</sup> Examples of other studies on segmentation include Magnac (1991) in Columbia and Grinling (1991) in the analysis of urban labour markets in Costa Rica.

<sup>3</sup> Thomas and Vallee (1996) examine earnings in the informal, formal, and regulated sectors within the manufacturing sector in Cameroon. Appleton, Collier, and Horsnell (1990) considered the private and public sector distinction in Cote d'Ivoire with non-participation as a base category in addition to a three-sector model of wage sector, with the private sector split into union and non-union segments.

important role education plays in job attainment by increasing the probability of wage employment for both men and women with higher effect for men. Such empirical evidence generally indicates an inclination towards segmented labour markets due apparently to existing wage gaps. Boudarbat (2004) confirms this fact by the observation of a preference for public sector employment in Africa and a willingness to engage in 'wait' unemployment to secure well paid and stable public sector jobs among the educated.

A study by Glick and Sahn (1997) in a multi sector labour market setting modelled employment outcomes for men and women in a multinomial logit model of self-employment, private sector wage employment, public sector wage employment and non-participation in Guinea. Individual and household characteristics used as covariates in the models included levels of education, age, children, education of parents and location of residence. The results indicate for both men and women that more education reduces the likelihood of being self-employed while it strongly increases the likelihood of being in the public sector. In terms of private sector wage employment, more education was found to increase the probability that a woman will be a private wage employee while in the case of men, education reduced the probability of private wage employment although the effect was found to be smaller in absolute value than for self employment. A similar study by Mariara (2003) adopted the random utility model in a three way multinomial logit model of selection in a multi sector labour market model of public, private and self-employment as the base for normalisation. Findings from the study showed education and demographic factors are important determinants for the choice of employment sector and earnings in Kenya. Specifically, some levels of education were found to increase the likelihood of working in the public and private sector compared to self-employment for both men and women. For women, with the exception of public sector employees with university education, higher levels of education were found to strongly increase the probability of being in wage employment. Similar patterns were observed for men except for those with secondary education, whose probability of participation were lower than for those with primary education. Results from a pooled model in the study confirmed the importance of education in determining public sector employment in Kenya than private sector employment relative to self-employment.

Kabubo (2003) similarly modelled allocation to either public or private sectors in Kenya and confirmed the importance of education in increasing the probability of participation in employment sectors relative to being inactive in a multinomial logit model of sector choice. The study further found age and its square to influence participation in addition to number of young children in the household.<sup>4</sup>

Rankin, Sandefur and Teal (2010) used the 2004 and 2005 rounds of the Ghana<sup>5</sup> and Tanzania Household Worker surveys in a pooled estimation in the investigation of the role of formal education and time spent in the labour market in explaining labour market outcomes of urban workers. In a multi sector modelling setting, the labour market was divided into self-employed, employment in small private firms, employment in large private firms, public sector employment and not working. The random utility model was adopted based on the symmetric labour market assumption with covariates such as education, age, marital status, household headship, presence children and parental education; results from the models indicated education increases the probability of employment in large firms and the public sector in Ghana but no such evidence was found for Tanzania.

What emerges from the literature on multi sector modelling in employment determination either as an independent study or as part of a modelling framework to correct non random assignment in earnings models, is the lack of consensus on a segmented or symmetric labour market in Africa. This therefore makes it imperative to analyse current available data to highlight the determinants of occupational choice especially within two African countries to contribute to the empirical literature.

### **3. Methodology**

The main objective of the paper is to investigate determinants of labour market participation and employment sector choice in both Ghana and Tanzania. The methods used consist of a baseline binary logit model of participation in the labour

---

<sup>4</sup> Sackey (2005), adopted the random utility model in a study on female labour force participation and fertility in Ghana and found a high rate of participation by women, with education as the key determinant of participation.

<sup>5</sup> In part of a study by Glewwe (1991), the adoption of a random utility model of sector choice led the conclusion that schooling is positively associated with entry into wage employment and among wage employees, those with better education are more likely to be in the public than the private sector.

markets for both countries and a multinomial logit model of participation in the formal (public and private) and informal sectors relative to not-working (unemployment) in both countries.

In the baseline model, participation decision in the labour market is assumed to be a function of variables that influence a person's expected offer wage and reservation wage. An individual chooses to enter the labour market if the offer wage is greater than the reservation wage. Human capital variables are expected to influence the offer wage while household characteristics may influence the reservation wage by affecting productivity in the home and demand for leisure [and other household income]. The choice between working in the formal and informal sectors (employment) and not-working (unemployment) is modelled in a simple binary logit model of participation. In the second model, the choice between employment and not-working (unemployment) is modelled in a multi sector setting of formal public sector employment, formal private sector employment, informal (self-employment) sector and not-working.

#### *Baseline logit model*

Here  $y$  is a binary choice variable which is equal to 1 if the individual is employed and 0 if not-working and  $x$  represents all individual and household characteristics that explain the choice decision.

$$Pr(y = 1|x) = G(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k) = G(\beta_0 + x\beta) \quad (1)$$

$G$  is a logistic function strictly between zero and one for all real numbers  $z$ :

$0 < G(z) < 1$  of the form;

$$G(z) = \exp(z) / [1 + \exp(z)] = \Lambda(z) \quad (2)$$

This follows a cumulative distribution (CDF) for a standard logistic random variable. The logit model is derived from an underlying latent variable model with  $y^*$  as an unobserved variable determined by

$$Pr(y = 1|x) = G(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k) + \epsilon \quad (3)$$

$y = 1$  if  $y^* > 0$  and  $y = 0$  if  $y^* < 0$

The assumption of cumulative distribution function (cdf) leads to the derivation of the response probabilities as

$$\begin{aligned} Pr(y = 1|x) &= Pr(y^* > 0) = Pr(\epsilon > -\beta_0 - \beta_1x_1 - \beta_2x_2 - \dots - \beta_kx_k|x) = \\ &1 - G(-\beta_0 - \beta_1x_1 - \beta_2x_2 - \dots - \beta_kx_k) = G(\beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_kx_k) \end{aligned} \quad (4)$$

Partial effects of continuous variables are derived as follows

$$\frac{\partial Pr(y = 1|x)}{\partial x_i} = g(\beta_0 + x\beta)\beta_i, \text{ where } g(z) \equiv \frac{dG}{dz}(z) \quad (5)$$

The partial effects of dichotomous variables are the difference in probabilities as the variable moves from 1 to 0.

### *Multinomial logit model*

The multinomial logit model extends the binary logit model to more than two choices. This model assumes each individual may select among four mutually exclusive alternatives in the labour market: working in the public sector (indexed  $p_u$ ), working in the private sector (indexed  $p_r$ ), self-employed (informal sector) (indexed  $s$ ) and not-working/unemployment (indexed  $u$ ). An individual compares the maximum utility attainable given each participation alternative and selects the alternative which yields the maximum utility.<sup>6</sup> Preferences are described by a well-behaved utility function whose arguments include the household time of the individual, a Hicksian composite commodity and a vector of exogenous constraints on current decision making.<sup>7</sup> Preferences are maximised subject to time and income constraints with no uncertainty.

Let  $V_{ji}$  be the maximum utility attainable for individual  $i$  if he/she chooses participation status  $j=p_u, p_r, s, u$  and suppose this indirect utility function can be decomposed into a non-stochastic component (S) and a stochastic component ( $\epsilon$ ):

---

<sup>6</sup> The specification does not allow the possibility of working concurrently in more than one sector. This restriction may be unreasonable if individuals work both in a family business and in the formal sector. However, the data does not have information on multiple job holding. Each person reports one current principal employment status.

<sup>7</sup> The model treats non labour income as exogenous. Although this may be an unrealistic assumption (particularly for the informal sector), it is important due to the need to link theory with the data in the absence of appropriate instruments.

$$V_{ji} = S_{ji} + \epsilon_{ji} \quad (6)$$

where  $S_{ji}$  is a function of observed variables and  $\epsilon_{ji}$  is a function of unobserved variables. The probability that individual  $i$  will select the  $j^{th}$  participation status is given by

$$P_{ji} = Pr[V_{ji} > V_{ki}] \text{ for } k \neq j, k = p_u, p_r, s, u \quad (7)$$

or, substituting in from (6),

$$P_{ji} = Pr[S_{ji} - S_{ki} > \epsilon_{ki} - \epsilon_{ji}] \text{ for } k \neq j, k = p_u, p_r, s, u \quad (8)$$

If the stochastic components have independent and identical Weibull distributions, the difference between the errors ( $\epsilon_{ki} - \epsilon_{ji}$ ) has a logistic distribution and the choice model is multinomial logit (McFadden, 1974).<sup>8</sup>

This is a direct extension of a binary logit model to a dependent variable with several unordered categories since the decision to work in a particular sector is assumed not to be sequential or ordered; rather this depends on the sector in which an individual finds a job.<sup>9</sup>

In order to estimate this model, a functional form of the non-stochastic component of the indirect utility function  $S_{ji}$  must be specified. When approximated in a linear form ( $S_{ji} = \beta_j X_i$ ), this yields an empirical specification of the form

$$P_{ji} = \frac{\exp(\beta_j X_i)}{\exp(\beta'_{pu} X_i) + \exp(\beta'_{pr} X_i) + \exp(\beta'_s X_i) + \exp(\beta'_u X_i)} \quad (9)$$

where  $X_i$  is a vector of independent variables that explain labour force participation and  $\beta_j$  is the parameter vector.

In the logit model, it is important to test whether the four sector model simply collapses to the dichotomous model. The dichotomous model is of the form

$$P_{wi} = \frac{\exp(\beta_w X_i)}{\exp(\beta'_w X_i) + \exp(\beta'_n X_i)} \quad (10)$$

---

<sup>8</sup> Weibull distribution has a unimodal bell shape roughly similar to the normal distribution.

<sup>9</sup> Some individuals decide to join the informal sector while awaiting modern wage employment job, others also leave modern sector jobs to become self-employed in the informal sector and vice versa. The choices made do not follow any particular order and this serves as a justification for the MNL model.

where subscripts  $w$  and  $n$  are for working and not-working (unemployed). This simpler model effectively restricts the parameter for self-employment ( $\beta_s$ ) to equal those of the public sector ( $\beta_{pu}$ ) and the private sector ( $\beta_{pr}$ ). The dichotomous model then misspecifies the underlying choice framework unless these coefficients are equal. This can be seen most clearly in terms of the log-odds ratio (Hill 1983). In the logit model, the log-odds ratio of two probabilities is a linear in parameters function of the explanatory variables. The log-odds ratio of working and not working derived from the four choice model is:

$$\ln\left(\frac{P_{pu} + P_{pr} + P_s}{P_n}\right) = \ln[\exp(\beta'_{pu}X_i) + \exp(\beta'_{pr}X_i) + \exp(\beta'_sX_i)] - \beta'_nX_i \quad (11)$$

Then if  $\beta_{pu} = \beta_{pr} = \beta_s$ , a simple dichotomous appropriately specifies the choice model.

$$\ln\frac{p_w}{p_n} = \ln 2 + \beta'_wX_i - \beta'_nX_i \quad (12)$$

where  $P_{pu} + P_{pr} + P_s = P_w$ . If these vectors are not equal, the right hand side of (11) is nonlinear and will be misspecified with a dichotomous dependent variable and the linear relationship implied by (12). The dichotomous model assumes that individuals are indifferent between working in the public, private and the informal sector ( $V_{pu} = V_{pr} = V_s$ ) and that given working, individuals are equally likely to be public, private and informal sector workers ( $P_{pu} = P_{pr} = P_s$ ).

McFadden (1974) suggests several measures of goodness-of-fit for the multinomial logit model, the likelihood ratio statistic is the most commonly used. Accordingly, the null hypothesis for testing that the four employment type model collapses to the dichotomous model is that  $\beta_{pu} = \beta_{pr} = \beta_s$ , is tested using a likelihood ratio test. The test statistic under the null hypothesis is:

$$\lambda - 2[L(\beta_r) - L(\beta_u)] \quad (13)$$

distributed asymptotically as a chi-square variate with  $k$  degrees of freedom, where  $k$  is the number of restrictions.  $L(\beta_r)$  is the log-likelihood function of the four employment status model evaluated under the restriction and  $L(\beta_u)$  is the unrestricted log-likelihood function of the model.

Coefficients obtained in the logistic estimation serve to provide a sense of the direction of the effects of the covariates on participation and sector choice in the

labour market and cannot be used to indicate the magnitude of impact. To examine the magnitude of impact we calculate the marginal impact of the covariates on the probability of participation and sector choice.

The multinomial logit which allows for more than two categories assumes the errors are independent for each category (employment sector), as a result suffers from the “independence of irrelevant alternatives” (IIA) assumption (Greene, 2003). Under the IIA, no systematic change in coefficients is expected with the exclusion of one of the outcomes from the model. Violation of the IIA assumption implies multinomial logit model is not an efficient and consistent estimator. Consequently, to ascertain the validity of the multinomial logit model, a test for IIA is conducted.

#### **4. Conceptual framework / Data**

The theory underlying participation decision has typically been modelled within the standard neoclassical microeconomic framework. This has been the foundation of most empirical labour supply analysis (Gray et al., 2002; Heckman, 1979; Killingsworth, 1983; Becker, 1964). Labour market outcome is an interaction between demand and supply of labour. Labour demand is a function of marginal productivity which can be improved through skills acquisition by education and experience according to the theory of human capital, in addition to unobserved skills that contribute to productivity improvement. Within the neoclassical framework, individuals are rational actors who maximise utility (with income, market commodities and leisure as arguments in the utility function) and are willing to enter into employment only when the market wage exceeds the reservation wage.<sup>10</sup> The chances of finding employment are determined by whether an individual can find suitable employment. This is determined by the productivity of the individual, the minimum wage and conditions under which they are prepared to accept employment (reservation wage) and the attitude of employers towards employing a given individual, such as women with children (Gray et al., 2002).

---

<sup>10</sup> The amount of money an individual would have to be given to induce him/her to work for the first hour or the minimum wage at which an individual will be willing to enter into employment. This is consequently dependent on the level of education, accumulated work experience and length of career breaks.

Within this framework, the labour force status of an individual is determined in a two-stage process. First of all, the decision is made pertaining whether or not to supply labour in the labour market. The second stage is whether the individual is employed or not and in which sector, depends on a combination of factors which include labour demand (employer preferences for certain characteristics like education, skills, experience, and sex), incentives to actively search for a job and to go for any job offers available. The decision is complex and involves many factors, for example the composition and the dynamics within the household are important as the labour supply decision needs to be considered in terms of household or family needs. A limitation of the model however, is the fact that it ignores the interdependence between members of households and consequently their decisions, in addition to its failure to distinguish between productive market activities and recreation activities.<sup>11</sup>

Equilibrium is determined in the market through interaction of demand and supply of labour in the labour market. This implies participation in the labour force involves a decision by an individual to allocate their time between work and leisure and requires a simultaneous decision on the part of an employer to offer that individual a job. This paper concentrates on the supply-side of the labour market by investigating the determinants of employment in particular sectors (formal public, formal private and self-employed) relative to unemployment.

Urban labour markets in Africa mainly consist of four broad categories; formal modern wage employment in the public or private sector, informal (self-employed) sector and not-working category of which unemployment is a part. We consequently model the labour market choices in Ghana and Tanzania to suit this structure. Public sector workers are individuals who work for the state (government employees), private sector employees are individuals under employment contracts which include manufacturing, non-manufacturing and other formal salaried/wage employees not in the public sector. The exact definition of informal sector (self-employed) according to Hart (1985) is arbitrary to some extent and depends on the specific research objectives, similar to the broad definition of the informal sector by the ILO mission to

---

<sup>11</sup> A typical example is that, married women's labour supply decisions are made by factoring in the decisions by other members of the household. Various extensions have therefore been made to the standard individual labour supply model which includes game theory models, bargaining models, individual utility models and new household economics models, but not widely applied due to huge data demands of these models.

Kenya in 1972. Consequently in this study, the informal sector (self-employed) consists of individuals with earnings from informal activities, this includes self-employed with and without employees (own account workers). Unemployment is categorised as individuals who have taken action in the past four weeks prior to the survey period to seek work or do other business. Such individual are within the statutory working age and excludes students, the incapacitated and unpaid family workers. The not-working category includes the searching unemployed, discouraged workers and unpaid family workers.

### *Data*

The study uses the Urban Household Worker Surveys for Tanzania and Ghana collected in three rounds<sup>12</sup>from 2004 to 2006 in pooled sample estimations. There exist a small panel and recall dimension of the data but are not used in this study due to the short interval between the survey periods and the large attrition that cannot be explained. In addition, we do not expect significant changes in labour markets particularly in the African context in such a short period of time. Consequently for Ghana number of observations by waves are 1818 in wave 1, 895 in wave 2 and 309 in wave 3. For Tanzania number of observations are, 653 in wave 1 443 in wave 2 and 572 in wave 3.

The surveys collect information on incomes, education, labour market experience, household and individual characteristics. For Ghana, the survey covered the four largest urban areas in the country, which are Accra and neighbouring Tema, Kumasi, Takoradi and Cape Coast. In Tanzania, the sample includes six of the largest urban areas, which are Dar es Salaam, Arusha, Iringa, Morogoro, Mwanza and Tanga. The samples are based on a stratified random sample of urban households from the 2000 census in Ghana and the 2000 Household Budget survey in Tanzania with the individual as the unit of analysis.

Tables 1 and 2 illustrate the composition of the urban labour force of Tanzania and Ghana for the three years pooled together for each country. Table 1 shows the distribution of the sample by labour force status according to age, years of education,

---

<sup>12</sup> Surveys were conducted by the Centre for the Study of African Economies, Oxford University

mean, standard deviation and median monthly earnings (converted from local currency to US dollars for comparison) for urban workers in Tanzania. As shown in Table 1 and Figure A1 in appendix, the informal sector is the largest sector, employing 42% of the sample. This is followed by the not-working category (33.3%) of which 16% are searching unemployed, and the formal sector (24%), with private formal sector constituting 17% of the labour force and the public sector (7.5%). Average earnings are notably highest in the public sector, followed by private formal employment and then self-employed.

**Table 1: Employment Status by Age, Education and monthly Earnings for Tanzania**

<b>Employment Status</b>	<b>N</b>	<b>% Share</b>	<b>Age(yrs)</b>	<b>Education(yrs)</b>	<b>Mean(\$)</b>	<b>Std. Dev.</b>	<b>Median</b>
<b>Public</b>	105	7.5	43	12.1	135.58	138.93	107.84
<b>Private</b>	236	16.8	34	8.4	58.16	61.24	42.03
<b>Self employed</b>	594	42.4	36	7.5	36.69	65.91	22.95
<b>Not-working</b>	467	33.3	27	6.6	-	-	-
<b>Total</b>	<b>1402</b>	<b>100</b>	<b>34</b>	<b>7.6</b>	<b>53.25</b>	<b>82.45</b>	<b>31.00</b>

*Source:* Calculations from UHWS 2004, 2005 and 2006 earnings are converted into U.S. dollars by respective official average yearly exchange rate.

Average years of schooling are higher for workers compared to those not-working (average years of education of the searching unemployed is 6.3 years) and highest for public sector workers (who are also older). The average private sector worker is more educated and earns more than the self-employed although relatively younger. The young and least educated are most likely to be in the not-working (unemployed) category in urban Tanzania as in other places in Africa.

Figure 1 shows the distribution of log of earnings in the three sectors of employment in Tanzania. As expected, there are large sector differences in earnings with the public sector dominating the private sector and self-employment. The self-employed have the lowest earnings with wide variation (standard deviation), reflecting the likelihood that those running a business with employees or self-employed professionals earn more than the standard informal worker.

**Figure 1**

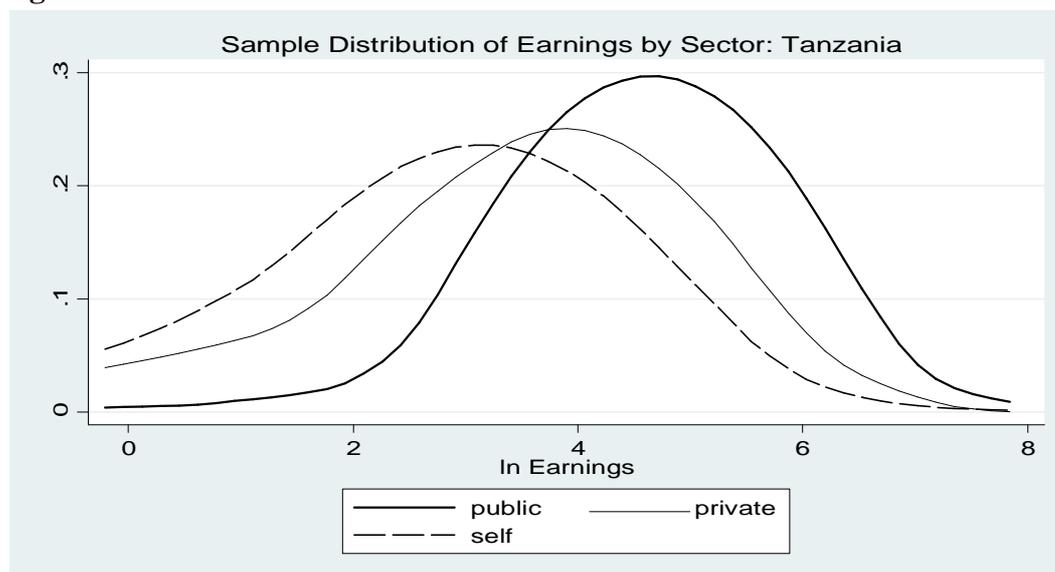


Table 2 provides similar information on employment status by age, years of education and monthly earnings in dollars for Ghana. The distribution is quite different from that for Tanzania: the private sector is the largest in terms of employment, followed by the not-working category (with unemployment constituting about 20%), self-employed and the public sector. The public sector workers, as in Tanzania, are older and more educated with higher earnings than the private sector worker and the self-employed. However, in Ghana, although the average not-working individual is younger, they are not the least educated. This suggests that in the Ghanaian labour market people are willing (or able) to wait in unemployment to get employment in the formal sector rather than move into the informal sector (self-employed). Particularly as the average years of education of the searching unemployed is 8.6 years. This is apparently due to existing earnings gaps, prestige and other non monetary benefits associated with formal sector employment.

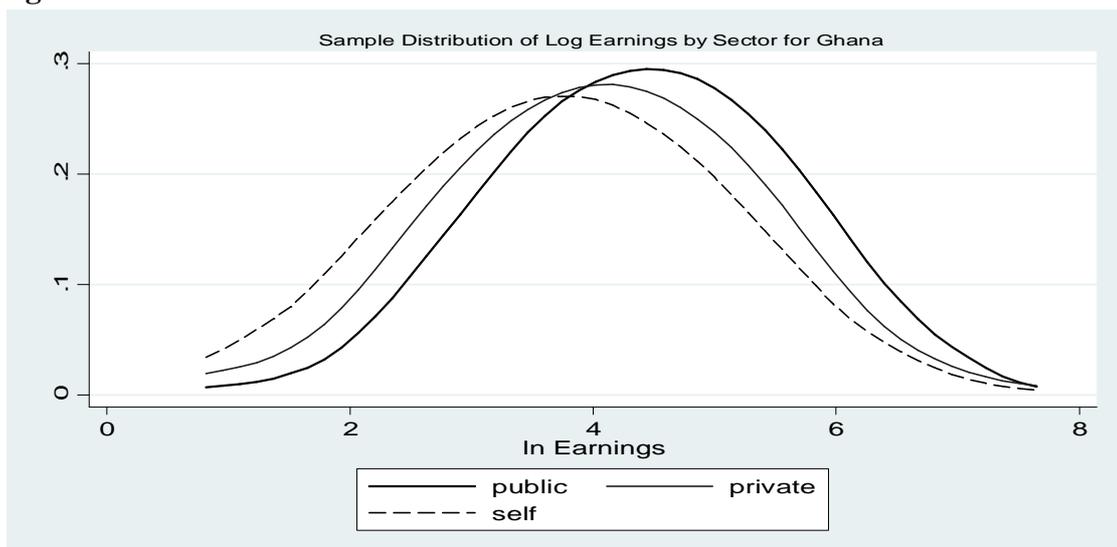
**Table 2: Employment Status by Age, Education and monthly Earnings for Ghana**

<b>Employment Status</b>	<b>N</b>	<b>% Share</b>	<b>Age(yrs)</b>	<b>Education(yrs)</b>	<b>Mean(\$)</b>	<b>Std. Dev.</b>	<b>Median</b>
<b>Public</b>	152	6.3	39	10.6	111.18	100.18	84.60
<b>Private</b>	989	41.3	34	9.1	87.63	124.44	57.03
<b>Self employed</b>	542	22.6	36	6.6	74.19	180.20	44.14
<b>Not-working</b>	713	29.8	26	8.2	-	-	-
<b>Total</b>	<b>2,396</b>	<b>100</b>	<b>34</b>	<b>8.2</b>	<b>84.60</b>	<b>145.10</b>	<b>55.17</b>

*Source:* Calculations from UHWS 2004, 2005 and 2006 mean and median earnings converted into U.S. dollars by respective official average yearly exchange rate.

Figure A2 in appendix depicts the proportion of different occupation categories in Ghana. Contrary to the Tanzanian urban labour market, private sector is the dominant sector in the Ghana sample in terms of employment; this is followed by not-working, self-employed and the public sector respectively. Distribution of log earnings in Ghana for the three main employment sectors is presented in figure 2. The public sector is the dominant sector in terms of earnings followed by private and self-employed. However, comparing Ghana and Tanzania shows sector earnings differentials are not as large in Ghana.

**Figure 2**



Comparison of pooled earnings across sectors between the two countries (figure A3 in appendix) indicates earnings in Ghana are more than earnings in Tanzania with mean overall earnings of \$53 in Tanzania and \$85 in Ghana although there is overlap in the earnings distribution with the distribution for Ghana skewed to the right of that of Tanzania.

Figure A4 shows the pooled distribution of earnings in both countries for the different sectors of employment. A pooled distribution of earnings across sectors of employment (appendix A3) point to the hierarchy that exist in earnings by sector with the public sector at the top followed by the private and self-employed respectively but important to note the wide dispersion of earning in self-employment relative to the formal sector shows the existence of heterogeneity in the sector consistent with Fields (1990).

A breakdown of labor market states into two age cohorts of young (15-35) and older (36-64) by country are presented in appendix figures A1 to A4. The pattern of distribution of older cohort for Tanzania shows more older people are in self-employment, followed by private sector employment, public and unemployment respectively. This distribution of young cohort on the other hand shows majority of them are in self-employment private and public sector respectively. For Ghana, the pattern is fairly similar to the total distribution as older people in Ghana are in private sector employment, self-employment, public sector employment and unemployment in that order. Most young people in Ghana are in private sector employment, this is followed by unemployment, self-employment and public sector in order. Overall, the proportion of unemployment within the young cohort is more than the proportion of unemployment within the older cohort in both countries and points to the inherent differences between age cohorts in both labour markets.

## **5. Model Specification and Empirical Results**

With the standard neoclassical microeconomic framework underlying participation and sector choice decisions, we estimate equations (2) and (8). Variables used in the model influence employment choice by influencing expected earnings and the reservation earnings. The structure of the family has significant effects on participation, particularly for women in the labour market. Most empirical studies

find a negative relationship between number of children in the family and the probability of participation by the wife (Kaufman, 1994). This is within the context of division of labour in the household where the husband is the breadwinner and the woman engaged in (reproductive) childrearing and household work. The effect on participation depends on the level of education of the woman; better educated women in particular are less constrained by the need to care for children because they expect higher earnings and can afford to hire domestic help. This is especially relevant in a developing country context (Ghana and Tanzania) where domestic service is affordable among educated women in full-time employment in the formal sector and some self-employed. This notwithstanding, with the rigidities in wage employment, there is the tendency for women to select themselves into the informal sector or unemployment, especially if they have lower levels of education.

Studies that include children in participation models typically disaggregate number of children based on level of dependence; we are unable to do this because the data do not have information on the ages of children (and there is a low response rate on number of children), subsequently, we use a dummy variable for individuals who reported having children and living with children who depend on them. Other household characteristics included are dummy variables for access to non-labour income, parent's highest education to capture the effect of networks in access to jobs, and town (region) of residence to control for location opportunities in accessing jobs. Parent's education is measured as the highest level of education attained by both mother and father separately. It is expected that more educated parents will have more social capital in terms of networks that facilitate job search. Personal characteristic variables included are age, sex, level of education and marital status. Age controls for life cycle and any potential labour market experience. A home ownership variable as a proxy for assets was initially included but dropped given low response rate (less than 50% of sample), and was insignificant when included, instead resource based measure of access to non-labour income is used (from the question do you receive any income apart from labour income). All estimations include year dummies to control for macroeconomic trend. A summary of all variables used is reported in Appendix Tables A1 and A2.

## *Results*

The baseline model gives an indication of the main determinants of participation in the labour market. We estimate the model separately for both countries due to the diversity in the samples. Results from the baseline binary logit model of labour force participation for Ghana and Tanzania with not-working as the reference category are presented in Table 3 (coefficients and average marginal effects), results of the models estimated with unemployment as the reference category and years of education are presented in Appendix Tables A3 and A4). A test of multiple exclusion restriction is conducted using the likelihood ratio test, the null hypothesis tested is that all slope coefficients in each model are simultaneously equal zero. Across models, we reject the null hypothesis at 1% level of significance (results in Appendix A5).

All variables in the Tanzania model have the expected signs. Primary and secondary levels of education compared to no education increases the probability of employment. Age raises the probability of employment, but at a decreasing rate with age as the threshold at which the probability of employment start to decrease is at 40 years. Men in the Tanzanian labour market have a higher probability of employment. In terms of parents' education, father's education is the most important as it has a positive though minimal impact on the probability of employment (marginal effect is 1.1%). This gives an indication that although social network (parent's status) is important the effect is limited; employment through contacts may be a combination of factors in addition to education of parents in urban Tanzania. Individuals with access to non-labour income have lower probability of employment compared to those with no form of non-labour income. This means such people have access to other resources to sustain them and wait longer hence are more likely to remain in the not-working category (unemployment). These results are generally mirrored when unemployment (searching) is used as the reference category as shown in Table A3.

The participation model for Ghana reveals interesting asymmetries. Secondary education increases the probability of not-working. Results with unemployment as the reference category (Table A3) confirm this as secondary and tertiary levels of education reduce the probability of employment and increase the probability of unemployment in Ghana. Most highly educated people prefer to be in unemployment to secure jobs in the formal sectors. This explains why the unemployed in Ghana are

not the least educated as is the case in Tanzania. Age and age squared variables indicate probability of employment increases with age but at a decreasing rate, after 41 years, the probability of employment decreases with age. Marriage increases the probability of employment in Ghana, possibly because people wait to get employed before they marry. Marriage also increases one's social capital particularly if the other partner is in the (formal) labour market. Also, children decrease the probability of employment due mainly to the burden of having to provide care. Father's education decreases the probability of employment relative to not-working; this effect however disappears with unemployment as the reference category (Table A3). In Ghana, living in Accra compared to living in other urban areas increases the probability of employment, consistent with economic activity being concentrated in Accra as the capital city.

#### *Multinomial logit model*

Although the multinomial logit model intuitively seems to be an appropriate methodology for the analysis of occupational choice, the Independence of Irrelevant Alternatives - IIA property (i.e. the odds ratios of choosing existing alternatives are assumed to be independent of the other alternatives.) is tested by using McFadden et al. (1977) likelihood ratio statistic. The test statistic is calculated as  $2[\text{the maximized log likelihood value of unrestricted model} - \text{the maximized log likelihood value of restricted model}]$ . It is an asymptotic chi-square distribution with degrees of freedom equal to the number of parameters in the restricted model. Across the two countries, the tests indicate that the multinomial logit model does not violate the IIA, hence appropriate for our sector choice.

Results from multinomial logit model of occupational (sector) choice estimated for Tanzania and Ghana are presented in Tables 4 and 5 (average partial effects), coefficients are presented in appendix tables 6 and 7. A likelihood ratio test of the null hypothesis of equality of coefficients between any pair of employment sectors ( $\beta_{pu} = \beta_{pr} = \beta_s$ ) in both models is rejected at 1% level of significance. This is an indication that the labour market is heterogeneous and the decomposition into public, private, self-employment and not-working is suitable. In addition, a likelihood ratio test is also used to test for multiple exclusion restriction; the null hypothesis tested is

that all slope coefficients in each model are simultaneously equal to zero. Across models, we reject the null hypothesis at 1% level of significance (Table A5).

Table 3: Baseline model of labour market participation; (Not-working as reference category)

	Tanzania		Ghana	
	Coefficients	Average Partial Effects	Coefficients	Average Partial Effects
Age	0.479*** (0.045)	0.013*** (0.001)	0.324*** (0.035)	0.013*** (0.001)
Age <sup>2</sup>	-0.006*** (0.001)	-	-0.004*** (0.000)	-
Primary	0.676** (0.329)	0.068** (0.030)	0.082 (0.166)	0.012 (0.022)
Secondary	0.755** (0.358)	0.076** (0.033)	-0.484** (0.195)	-0.069** (0.028)
Tertiary	0.745 (0.852)	0.075 (0.111)	-0.137 (0.244)	-0.019 (0.035)
Sex	1.464*** (0.190)	0.147*** (0.017)	0.235* (0.121)	0.033* (0.017)
Marriage	0.481 (0.392)	0.048 (0.039)	0.719*** (0.167)	0.102*** (0.023)
Children	0.226 (0.357)	0.023 (0.038)	-0.583*** (0.194)	-0.083*** (0.026)
Household head	0.025 (0.547)	0.002 (0.061)	0.455*** (0.165)	0.064*** (0.023)
Non-Labour income	-1.805*** (0.640)	-0.181*** (0.062)	-0.077 (0.146)	-0.011 (0.021)
Father's education	0.101*** (0.032)	0.010*** (0.003)	-0.052** (0.022)	-0.007** (0.003)
Mother's education	-0.002 (0.035)	-0.000 (0.003)	0.015 (0.023)	0.002 (0.003)
Dar es Salaam / Accra	0.419 (0.305)	0.042 (0.029)	2.538*** (0.138)	0.360*** (0.023)
Constant	-11.053*** (0.918)		-6.579*** (0.597)	
$\chi^2$ (D.F)	271.70 (15)		629.75 (15)	
Log-likelihood	-420.67		-993.47	
Pseudo-R <sup>2</sup>	0.52		0.29	
Observations	1,355		2,297	

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries, in addition to two year dummies not reported for brevity.

Results from both models indicate that age increases the probability of participating in all employment sectors in both Tanzania and Ghana. However, the probability of participation in these employment sectors increases at a decreasing rate. Average

partial effects for the model for Tanzania in Table 4 indicate, older people are less likely to be found in the not-working category, as age increases the probability of employment in all three sectors. The different education level dummies relative to no education decrease the probability not-working and self-employment in Tanzania and increase employment probabilities in the two formal sectors particularly in the public sector, the same pattern is found in the specification with years of education (Appendix Table A8). In Tanzania, being a man increases the likelihood of self-employment and private sector employment by 5% and 12% respectively but at the same time reduces the probability of not-working and public sector employment. Father's highest education level decreases the probability not-working but increases self-employment although at 10% in Tanzania. Access to non-labour income is associated with an increased probability of no-working and a reduced probability of self-employment in Tanzania. These results are generally consistent when we estimate the model with the narrower searching unemployed category as the reference group (Appendix A8 and A9 for Tanzania and Ghana respectively). In the model with searching unemployed as the reference group, we do not include tertiary education level as there is no observation with tertiary education in unemployment in the Tanzania sample.

Table 5 presents average partial effects for the Ghana model. The results reinforce findings in the baseline model that informal (self-employed) sector is the least preferred sector of employment by the educated and significantly so by the highly educated as all levels of education reduces the probability of self-employment. Notably, education is a vital determinant of employment in the public sector in Ghana similar to Tanzania as all levels of education are found to monotonically increase the probability of employment in the sector. This same pattern is found in the specification with years of education (Appendix Table A9) as years of education is found to increase the probability of not-working and public sector employment and decrease the probability of self-employment. Age decreases the probability of not-working and increases probability of employment in all sectors.

Table 4: Average Partial effects from multinomial logit estimates: Tanzania

	Not-working	Self	Private	Public
Age	-0.013*** (0.001)	0.006*** (0.001)	0.003*** (0.001)	0.004*** (0.001)
Primary	-0.068** (0.030)	-0.001 (0.046)	0.003 (0.039)	0.067 (0.041)
Secondary	-0.073** (0.033)	-0.109** (0.047)	0.019 (0.040)	0.163*** (0.041)
Tertiary	-0.043 (0.109)	-0.304*** (0.086)	0.132* (0.068)	0.215*** (0.044)
Sex	-0.143*** (0.017)	0.051** (0.023)	0.124*** (0.020)	-0.032** (0.013)
Marriage	-0.040 (0.040)	0.021 (0.036)	0.003 (0.037)	0.016 (0.019)
Children	-0.006 (0.038)	0.062 (0.043)	-0.022 (0.041)	-0.034 (0.029)
Non-Labour income	0.167*** (0.061)	-0.144*** (0.043)	-0.026 (0.040)	0.003 (0.018)
Father's education	-0.010*** (0.003)	0.006* (0.003)	0.002 (0.003)	0.002 (0.002)
Mother's education	0.000 (0.003)	-0.001 (0.004)	0.002 (0.003)	-0.002 (0.002)
Dar es Salaam	-0.043 (0.029)	0.047* (0.027)	0.026 (0.027)	-0.030** (0.015)
Test of IIA $\chi^2(24)$	1.75	78.59	26.73	7.77
N	1,355			

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries, in addition to two year dummies not reported for brevity.

Men are more likely than women to be in formal private and public sectors in the Ghanaian labour market. Although this is subject to further research, it may connote some form of discrimination in the labour market whereby formal sector employers may have a preference for men conditioning on education and all other demographic and household characteristic. Also, with little flexibility and fixed hours of work in wage employment particularly in the private sector in the African context, women may select themselves into the informal sector in order to cope with the need to care for children and domestic work to the extent that this sector enables them to combine productive work and care work. Similarly, being married is associated with increased probability of self-employment and decreased probability of not-working in addition to a 7% increased probability of not-working and a 9% decreased probability of private sector employment by individuals with children. This highlights the need to provide care at home and the demands of formal sector employment particularly in the private sector that deters individuals with children. Residence in Accra is

associated with increased probability of employment in all sectors and given that Accra is the capital city, this is an indication of differences in employment opportunities that exist in the country. In contrast to Tanzania, parent's highest education level particularly father's education is found to increase the probability of formal sector employment mainly in the private sector.

Table 5: Average Partial effects from multinomial logit estimates for Ghana

	Not-working	Self	Private	Public
Age	-0.013*** (0.001)	0.006*** (0.001)	0.005*** (0.001)	0.002*** (0.001)
Primary	-0.030 (0.023)	-0.070*** (0.019)	0.054* (0.028)	0.047** (0.022)
Secondary	0.057** (0.028)	-0.188*** (0.028)	0.048 (0.034)	0.083*** (0.023)
Tertiary	0.014 (0.035)	-0.189*** (0.037)	0.074* (0.041)	0.101*** (0.024)
Sex	-0.038** (0.016)	-0.171*** (0.015)	0.184*** (0.017)	0.025** (0.011)
Marriage	-0.092*** (0.023)	0.064*** (0.020)	0.025 (0.027)	0.003 (0.015)
Children	0.069*** (0.026)	0.008 (0.024)	-0.088*** (0.030)	0.011 (0.018)
Non-Labour income	0.010 (0.021)	0.038** (0.019)	-0.039* (0.022)	-0.009 (0.011)
Father's education	-0.007** (0.003)	-0.005 (0.003)	0.008** (0.004)	0.004* (0.002)
Mother's education	-0.002 (0.003)	-0.005 (0.004)	0.009** (0.004)	-0.002 (0.002)
Accra	-0.358*** (0.023)	0.052*** (0.017)	0.279*** (0.019)	0.027*** (0.010)
Test of IIA $\chi^2(24)$	17.75	667.39	32.99	364.01
N	2,297			

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey, in addition to two year dummies not reported for brevity.

Analysis of variation in sector choice according to age cohort is carried out to investigate heterogeneity in job attainment by age group. Due to the very small samples in some of the categories when we split the data by age group, we are unable to estimate the models separately for young (age 15-35) and older (age 36-64) cohorts for each country. Instead, a dummy variable to capture age group effect is introduced into the initial models for the respective countries.

Average partial effects of results with young dummy for both Tanzania and Ghana are presented in Appendix Tables A15 and A16 respectively. In both countries, being

young is associated with a higher probability of not-working. Given that the sample does not include students, since it can be argued that this group will most likely consist of individuals in school who are not working or looking for jobs, the results shed light on the challenges young people face in obtaining jobs in both countries due possible to lack of experience. In Tanzania, young people have decreased probability of employment in the public sector (10%), however in Ghana, the probability of employment in all three employment sectors are low (negative) among young compared to older people (although that of self-employment is at 10% level of significance). While marginal effects of self-employment and private sector employment for young are insignificant in the case of Tanzania, the results suggests young people in the Ghanaian labour market may have more difficulties in terms of job attainment than their counterparts in Tanzania.

Finally, conditioning on all the covariates (education and family background), the probability of selection into the four labour market statuses by age are presented in Tables 6 and 7 for Tanzania and Ghana respectively. In Tanzania, the probability of not-working decreases with age, while that of self-employment increases with age. In the two formal sectors, the results confirm earlier findings of the important role of age in job attainment particularly in the public sector as the probabilities are very high for older compared to a younger individuals.

**Table 6: Predicted individual probabilities of occupation type by Age: Tanzania**

Age in years	Not-working	Self-employment	Private	Public
20	0.834*** (0.021)	0.094*** (0.017)	0.071*** (0.015)	0.002 (0.001)
30	0.381*** (0.014)	0.380*** (0.025)	0.188*** (0.022)	0.051*** (0.019)
40	0.189*** (0.009)	0.421*** (0.123)	0.126** (0.051)	0.265 (0.170)
50	0.080*** (0.005)	0.252 (0.231)	0.056 (0.037)	0.613** (0.261)
60	0.022*** (0.002)	0.111 (0.124)	0.023 (0.014)	0.843*** (0.135)

Source : estimated from multinomial logit model for Tanzania

In the case of Ghana (Table 7), the pattern found for the probability of not-working at specific age points is similar to that of Tanzania although the probabilities for older individuals are much lower. Age is very important in determining employment in the informal sector in Ghana as the probabilities significantly increase with age, this

connotes the possibility of access to funds to start up own business which older people would have more channels than younger people. Also, this could also be due to the fact that when people get older in the labour market, they require some level of flexibility on the job which may be available in the informal sector. Within the two formal sectors, the results suggests the importance of age in determining employment in these sectors cannot be over emphasised as although we do not find a monotonic increase in probabilities with age the probabilities are higher for older than younger individuals.

Table 7: Predicted individual probabilities of occupation type by Age: Ghana

Age in years	Not-working	Self-employment	Private	Public
20	0.739*** (0.032)	0.033*** (0.003)	0.213*** (0.031)	0.015** (0.007)
30	0.357*** (0.012)	0.301*** (0.019)	0.304*** (0.018)	0.039*** (0.004)
40	0.142*** (0.010)	0.620*** (0.028)	0.200*** (0.022)	0.038*** (0.009)
50	0.042*** (0.004)	0.824*** (0.019)	0.106*** (0.014)	0.028*** (0.008)
60	0.006*** (0.001)	0.929*** (0.012)	0.047*** (0.008)	0.017** (0.008)

Source : Estimated from multinomial logit model for Ghana

## 6. Conclusion

The study investigated the determinants of selection into employment sectors in the urban labour markets of both Ghana and Tanzania. To this end, a baseline participation model was estimated to give an indication of the main determinants of participation in the two labour markets. In addition, the multinomial logit model of employment determination in specific sectors was used to ascertain a more insight into individual sector choice. Consistency checks on the result were further conducted by a re-specification of the models with searching unemployed in the labour markets as the reference category instead of the broader not-working category.

Results from the baseline participation model suggest determinants of participation differ for the two countries pointing out the heterogeneity in the two labour markets. In Tanzania, education is found to increase the probability of employment but this is not the same for Ghana. In Ghana, the results indicate higher levels of education increases the likelihood of not-working. Suggesting Boudarbat's (2004) observation of a preference for public sector employment in Africa and a willingness to engage in

'wait' unemployment to secure well paid and stable public sector jobs among the educated may be evident in Ghana given the. These results are consistent with searching unemployed as the reference category.

Results from the multinomial logit models suggest on the other hand that both labour markets are characteristic of a preference for formal sector employment as all levels of education are found to increase the probability of public and private sector (formal sector) employment and decrease the probability of self-employment (informal sector) and not-working in both labour markets although in Ghana, education (years and secondary level) increases the probability of not-working. These advantages with education are also reflected along gender lines in both labour markets with men having high probability of selection into formal sector employment with the reverse for self-employment and not-working (unemployment). Results based on age differentials in labour market attainment point to the existence challenges young people face in accessing jobs in both Ghana and Tanzania as we find high probabilities of not-working to be associated with young relative to older people.

Previous studies in different parts of Africa confirm the fact that education is important in determining which sector of the labour market an individual works. Consistent for most part of our results, Glick et al. (1997) found that for men and women in Guinea, more education reduces the likelihood of being in self-employment while it strongly increases the likelihood of being in the public sector. Similarly, Vijverberg (1993) found that for men and women in Cote d'Ivoire, the probability of being in wage employment rises with education level while the probability of non-agricultural self-employment falls with additional schooling. On Ghana, Glewwe (1991) confirms our finding of a strict preference for formal sector jobs and hierarchical nature of the preference by the finding that schooling is positively associated with entry into wage employment and among wage employees, those with better education are more likely to be in the public sector than in the private sector.

Finally in both countries, we find that, the probability of self-employment increases with age. This is an indication that older people require some level of flexibility on the job which is not available in the formal sector. Consequently in both labour markets, conditioning on education and all explanatory variables, older people are more likelihood to be in self-employment.

## References

- Andersson, P. (1993). Labour market structure in a controlled economy: The case of Zambia. *Ekonomiska Studier* no. 37, Department of Economics, University of Gothenburg.
- Appleton, S., Collier, P., and Horsnell, P. (1990). Gender, education and employment in Cote D'voire. *Social Dimensions working paper* no. 8, The World Bank.
- Becker, G.S. (1962). Investment in human capital. *Journal of Political Economy*, LXX Supplement: 9-49.
- Becker, G. S. (1964). Human capital: A theoretical and empirical analysis. (*New York, National Bureau of Economic Research*).
- Boudarbat, B. (2004). Employment Sector Choice in a Developing Labour Market. Mimeo, *Department of Economics*, University of British Columbia.
- Bourguignon, F., Robilliard, A. and Robinson, S. (2003). Representative versus real households in the macro-economic modelling of inequality. *DIAL Working Paper* no. DT/2003/10. Development, Institutions and Analysis de Long Term, Paris.
- Fields, G.S., (1975). Rural urban migration, urban unemployment and underemployment, and job-search activity in LDCs. *Journal of Development Economics* 2, no 2, 165-187.
- Fields, G.S., (2005). A guide to multi sector labour market models. *Social Protection Discussion paper series* no. 0505, The World Bank.
- Glewwe, P. (1991). Schooling skills and returns to government investment in education. *Living Standards Measurement Study Working paper* no. 76
- Glick, P. And Sahn, D.E. (1997). Gender and education impacts on employment and earnings in West Africa: Evidence from Guinea. *Economic Development and Cultural Change*, vol. 45, no.4, pp.793-823.
- Greene, W. H., & Hensher, D. A. (2010). Modelling ordered choices: A primer. *Cambridge University Press*.
- Hart, K., (1985). The informal economy. *Cambridge Anthropology* 10, no. 2, 54-58
- Heckman, J. and Sedlacek, G. (1985) Heterogeneity, aggregation and market wage functions: an empirical model of self-selection in the labour market. *Journal of Political Economy* 93, pp. 1077-1125.
- Hill, M. A. (1983). Female Labour Force Participation in Developing and Developed Countries--Consideration of the Informal Sector. *The Review of Economics and Statistics*, 65(3), 459-468.
- Hausman, J. and D. McFadden (1984). Specification Tests for Multinomial Logit Model. *Econometrica*, 52(5), 1219-1240.
- International Labour Office (2003), *Global Employment Trends 2003*, (Geneva: ILO).
- International Labour Office (2006), *Global Employment Trends for Youth*, (Geneva: ILO).
- Kabubo, J. (2003). Wage determination and the gender wage gap in Kenya: Any evidence of gender discrimination. *AERC Research paper* no.132. African Economic Research Consortium, Nairobi.

- Kaufman, B. E. (1994). The economics of labour markets. *Fort Worth: Dryson Press*
- Krishnan, P., Sellasie, T., and Dercon, S. (1998). The urban labour market during structural adjustment: Ethiopia 1990-1997. *CSAE WPS/98-99*. Centre for the Study of African Economies, Oxford University.
- Lewis, W.A. (1954). Economic development with unlimited supplies of labour. *The Manchester School*, 22,141-145.
- Lindauer, D. and Sabot, R. (1983). The public-private wage differential in a poor urban economy. *Journal of Development Economics* 12, no. 3, pp137-152.
- Lubell, H., (1990). The informal sector in the 1980s and 1990s. *Development centre, O.E.C.D.*, Paris.
- Maddala, G.S. (1990). Limited dependent and qualitative variables in econometrics. (Cambridge University Press, Cambridge).
- Maloney, W.F, (1999). Does informality imply segmentation in urban labour markets? Evidence from sectoral transitions in Mexico. *World Bank Economic Review* 13(2), pp. 275-302.
- Maloney, W.F, (2004). Informality revisited. *World Development* 32 (7), pp. 1159-1178.
- McFadden, D., K. Train and W. Tye (1977). A application of diagnostic tests for the Independence of Irrelevant Alternatives property of the Multinomial Logit model. *Transportation Research Record*, 637, 39-46
- Mincer, J. (1962). On-the-job training: Costs, returns, and some implications. *The Journal of Political Economy*, 70(5), 50-79.
- Mincer, J. A. (1974). Schooling and earnings. In *Schooling, experience, and earnings* (pp. 41-63). Columbia University Press.
- Okpukpara, B.C. and Odurukwe, N. (2006). Incidence and determinants of child labour in Nigeria: Implications for poverty reduction. *AERC Research paper no.156*. African Economic Research Consortium, Nairobi.
- Pradhan, M., A. And van Soest, A. (1995). Formal and Informal Sector Employment in Urban areas of Bolivia. *Labour Economics*, 2:275-97
- Pradhan, M. and van Soest, A. (1997). Household labour supply in urban areas of Bolivia. *Review of Economics and Statistics* 79, 300 –310.
- Perry, G.E., Maloney W.F, Arias, O.S., Fajnzylber, P., Mason, A.D. and Saavedra-Chanduvi, J. (2007). Informality: Exit and Exclusion. *Washington DC: The World Bank*, World Bank Latin America and Caribbean Studies.
- Rankin N., Sandefur J. and Teal F. (2010). Learning and earning in Africa: Where are the Returns to Education High. *CSAE WPS/2010-02*.
- Strassmann, W., (1987). Home-based enterprises in cities in developing countries. *Economic Development and Cultural Change* 36, 121 – 144.
- Tansel, A., (1999). Public-private employment choice, wage differentials and gender in Turkey. *Economic Growth Centre*, Yale University, Discussion paper no. 797.
- Teal, F. (2001). Education, incomes, poverty and inequality in Ghana in the 1990s. *CSAE WPS/2001-21*, Oxford: Centre for the Study of African Economies.

- Teal, F. (2002). Education, incomes and inequality in Ghana in the 1990s. *CSAE WPS/2001-21*, Centre for the Study of African Economies
- Thomas, M. and Vallee, L. (1996). Labour market segmentation in Cameroonian manufacturing. *Journal of Development Studies*, vol. 32, No. 6, pp 876-898.
- Todaro, M.P. (1969). A model of labour migration and urban unemployment in Less Developed Countries. *American Economic Review* 39: 138-148
- Turnham, D. (1971). The employment problem in Less Developed Countries. *Paris Development Centre of the Organisation for Economic Co-operation and Development*.
- Van der Gaag, J. and W. Vijverberg, (1988). A switching regression model for wage determinants in the public and private sectors of a developing country. *The Review of Economics and Statistics*, volume 70, no. 2, 244-252.
- Vijverberg, W. (1993): 'Educational investment and returns for women and men in Cote d'Ivoire', *The Journal of Human Resources, Issue 4, Special issue: Symposium on investments in women's Human Capital and Development* (Autumn, 1993), 933-974.
- World Bank (1995), *World Development Report 1990* (Washington: World Bank).
- Xiaodong, G. and van Soest, A. (2002): 'Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico', *Labour Economics* 9, (2002) 513-529

## Appendix

**A1:** Sample means by labour market state for Tanzania; (sample standard deviation of variables other than dummies in parenthesis)

	Unemployment	Self Employment	Private Sector	Public Sector
Education (yrs)	6.6 (3.24)	7.5 (4.22)	8.4 (4.15)	12.1 (4.13)
Education level:				
None	0.19	0.12	0.09	0.03
Primary	0.73	0.77	0.63	0.36
Secondary	0.06	0.02	0.12	0.08
Tertiary	0.02	0.09	0.16	0.53
Married	0.16	0.60	0.35	0.70
Children	0.24	0.77	0.42	0.81
Age (in years)	28 (13.01)	37.56 (11.12)	33.95 (11.15)	42.90 (9.98)
Parent education	5.56 (4.40)	6.07 (4.18)	6.53 (4.61)	6.54 (5.03)
Non-labour income	0.03	0.17	0.12	0.17
Dares Salaam	0.28	0.40	0.36	0.31
Number of Observations	467	594	236	105

**A2:** Sample means by labour market state for Ghana; (sample standard deviation of other variables other than dummies in parenthesis)

	Unemployment	Self Employment	Private Sector	Public Sector
Education (yrs)	8.56 (4.39)	6.60 (4.78)	9.09 (4.10)	10.62 (3.68)
Education level:				
None	0.17	0.32	0.13	0.06
Primary	0.50	0.57	0.54	0.43
Secondary	0.24	0.08	0.22	0.30
Tertiary	0.09	0.04	0.11	0.22
Married	0.24	0.70	0.54	0.66
Children	0.34	0.75	0.56	0.72
Age (in years)	26.96 (9.91)	36.17 (9.52)	34.43 (11.55)	38.58 (11.40)
Parent education	5.85 (4.34)	3.91 (3.93)	5.31 (4.40)	5.79 (4.74)
Non-labour income	0.26	0.30	0.37	0.36
Accra	0.02	0.26	0.49	0.44
Number of Observations	424	542	989	152

Figure A1

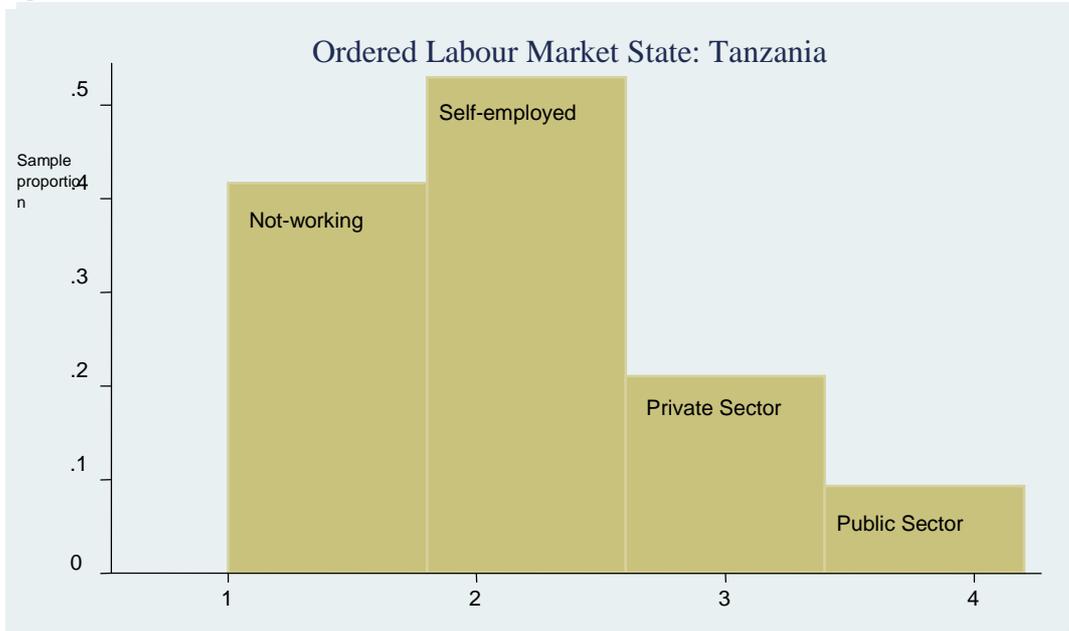


Figure A2



Figure A3

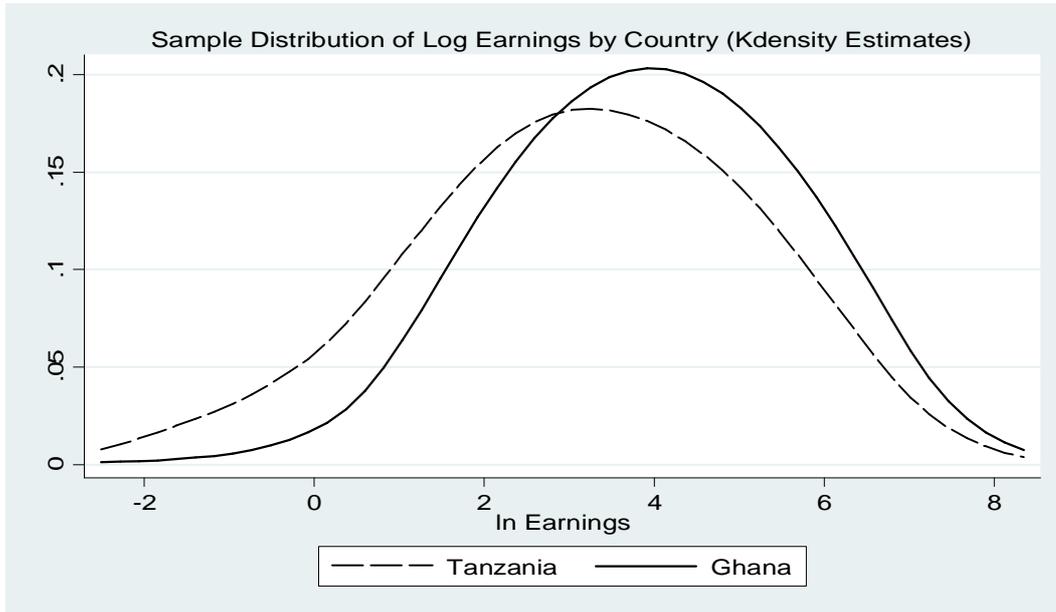


Figure A4

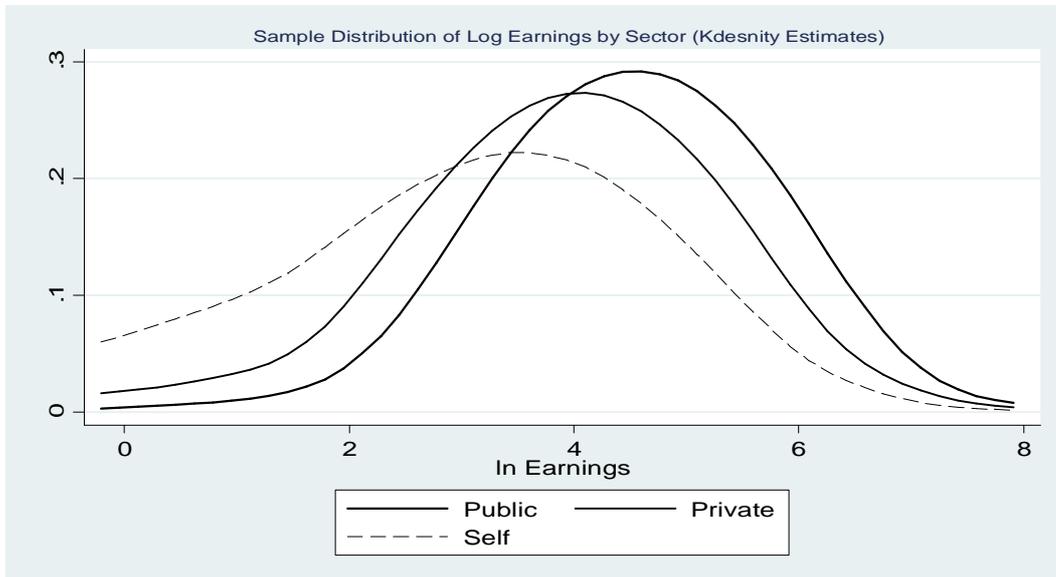


Figure A5

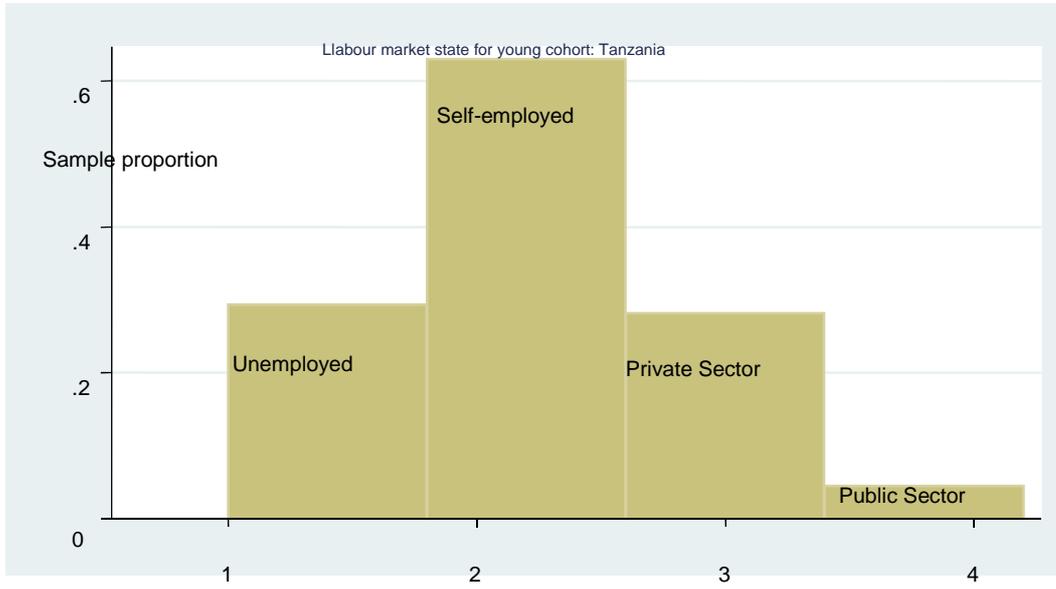


Figure A6

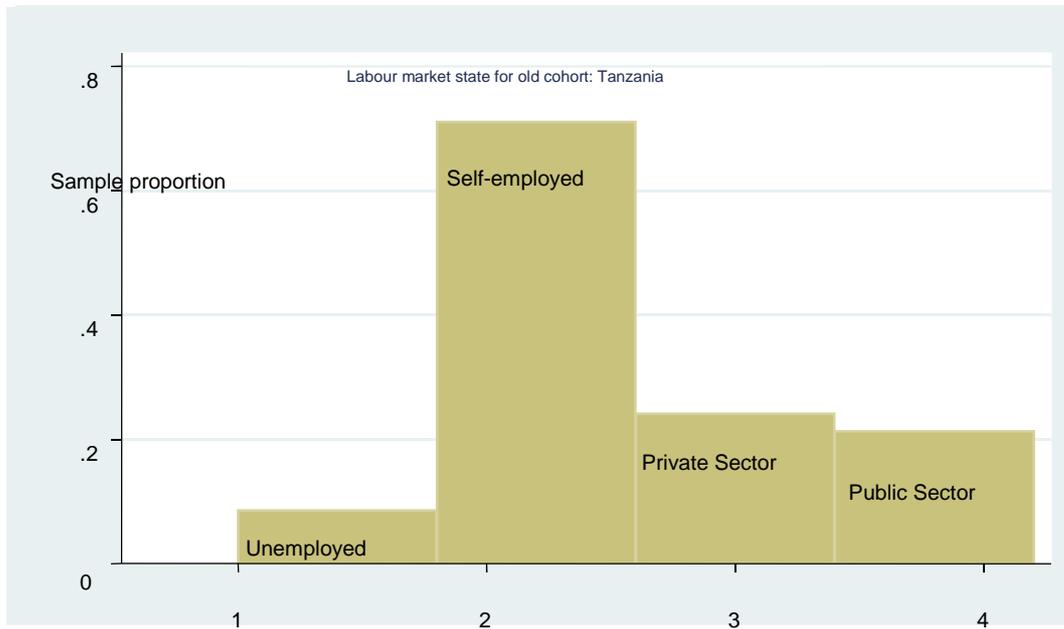
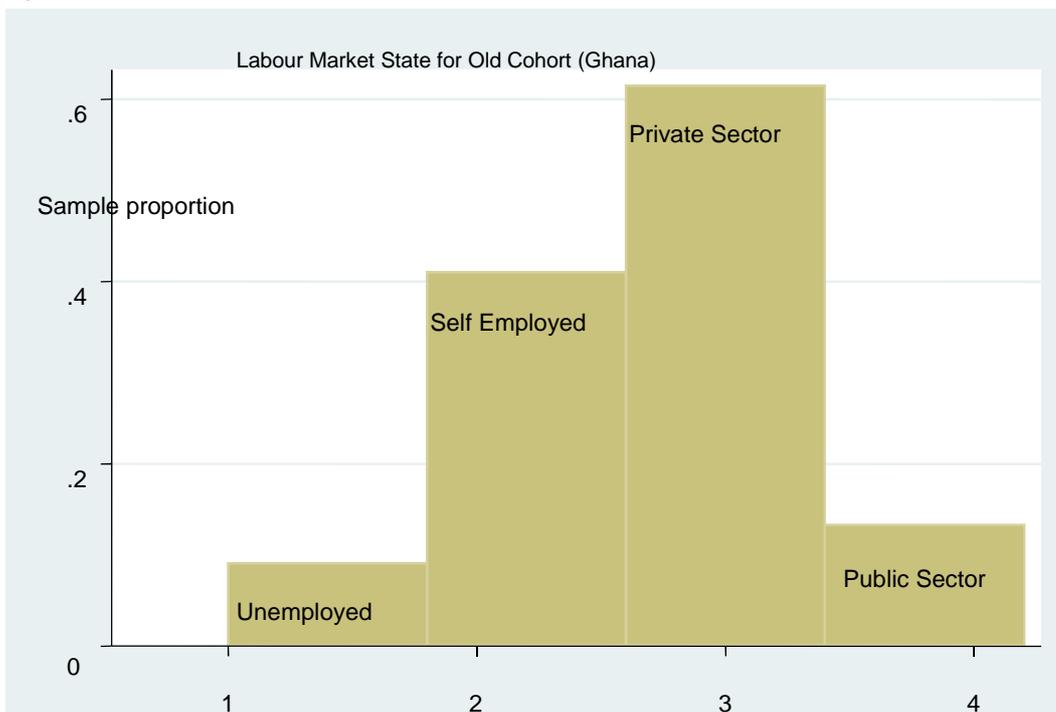


Figure A7



Figure A8





A3: Baseline model of labour market participation (Unemployment as reference category)

	Tanzania		Ghana	
	Coefficients	Average Partial Effects	Coefficients	Average Partial Effects
Age	0.322*** (0.054)	0.011*** (0.002)	0.243*** (0.043)	0.010*** (0.001)
Age <sup>2</sup>	-0.004*** (0.001)	-	-0.003*** (0.001)	-
Primary	0.970*** (0.336)	0.097*** (0.033)	-0.028 (0.192)	-0.003 (0.022)
Secondary	1.265*** (0.384)	0.127*** (0.037)	-0.784*** (0.227)	-0.091*** (0.026)
Tertiary	-	-	-0.533** (0.271)	-0.062** (0.031)
Sex	1.035*** (0.216)	0.104*** (0.021)	-0.133 (0.144)	-0.015 (0.017)
Marriage	0.480 (0.328)	0.048 (0.033)	1.022*** (0.195)	0.118*** (0.022)
Children	0.179 (0.320)	0.018 (0.032)	-0.460** (0.227)	-0.053** (0.026)
Household head	-0.289 (0.440)	-0.029 (0.044)	0.213 (0.194)	0.025 (0.022)
Non-Labour income	-0.426 (0.486)	-0.043 (0.049)	-0.001 (0.172)	-0.000 (0.020)
Father's education	0.018 (0.037)	0.002 (0.004)	-0.034 (0.031)	-0.004 (0.004)
Mother's education	0.036 (0.039)	0.004 (0.004)	0.015 (0.031)	0.002 (0.004)
Dar es Salaam / Accra	0.381 (0.260)	0.038 (0.026)	3.676*** (0.389)	0.424*** (0.042)
Constant	-6.919*** (1.038)		-3.128*** (0.802)	
$\chi^2$ (D.F)	129.16 (14)		317.43 (15)	
Log-likelihood	-323.35		-723.046	
Pseudo-R <sup>2</sup>	0.323		0.298	
Observations	1,023		2,028	

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries, in addition to two year dummies not reported for brevity.

A4: Baseline participation model with years of schooling (Reference category; Not-working)

	Tanzania		Ghana	
	Coefficients	Average Partial Effects	Coefficients	Average Partial Effects
Age	0.477*** (0.047)	0.013*** (0.001)	0.318*** (0.035)	0.013*** (0.001)
Age <sup>2</sup>	-0.006*** (0.001)	-	-0.004*** (0.000)	-
Education (yrs)	0.064** (0.029)	0.006** (0.003)	-0.018 (0.014)	-0.003 (0.002)
Sex	1.471*** (0.194)	0.148*** (0.017)	0.256** (0.121)	0.037** (0.017)
Marriage	0.462 (0.392)	0.046 (0.039)	0.741*** (0.161)	0.106*** (0.023)
Children	0.267 (0.383)	0.027 (0.038)	-0.565*** (0.186)	-0.081*** (0.027)
Household head	-0.024 (0.604)	-0.002 (0.061)	0.462*** (0.161)	0.066*** (0.023)
Non-Labour income	-1.807*** (0.620)	-0.181*** (0.062)	-0.075 (0.146)	-0.011 (0.021)
Father's education	0.100*** (0.031)	0.010*** (0.003)	-0.059*** (0.021)	-0.008*** (0.003)
Mother's education	-0.010 (0.033)	-0.001 (0.003)	0.020 (0.022)	0.003 (0.003)
Dar es Salaam / Accra	0.397 (0.287)	0.040 (0.029)	2.444*** (0.183)	0.350 (0.023)***
Constant	-10.787*** (0.882)		-6.305*** (0.577)	
$\chi^2$ (D.F)	900.54 (13)		790.28 (13)	
Log-likelihood	-421.155		-1008.504	
Pseudo-R <sup>2</sup>	0.517		0.282	
Observations	1,355		2,301	

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries, in addition to two year dummies not reported for brevity.

A5: Likelihood-ratio test of exclusion restriction

	Tanzania	Ghana
$\chi^2$	901.51	805.85
Degrees of freedom	15	15
p-value	0.000	0.000

A6: Multinomial logit model of Sector Choice; Tanzania (Reference category: Not-working)

	Self	Private	Public
Age	0.497*** (0.049)	0.453*** (0.053)	0.855*** (0.118)
Age <sup>2</sup>	-0.006*** (0.001)	-0.005*** (0.001)	-0.009*** (0.001)
Primary	0.658* (0.346)	0.591 (0.385)	1.820** (0.832)
Secondary	0.446 (0.381)	0.770* (0.416)	3.511*** (0.837)
Tertiary	-0.478 (0.849)	1.150 (0.844)	3.973*** (1.133)
Sex	1.256*** (0.202)	1.738*** (0.218)	0.664** (0.319)
Marriage	0.431 (0.400)	0.336 (0.446)	0.668 (0.498)
Children	0.234 (0.361)	-0.084 (0.418)	-0.498 (0.598)
Non-Labour income	-1.877*** (0.657)	-1.433** (0.687)	-1.540** (0.724)
Father's education	0.108*** (0.034)	0.089*** (0.035)	0.135*** (0.044)
Mother's education	-0.009 (0.037)	0.009 (0.038)	-0.042 (0.050)
Dar es Salaam	0.468 (0.307)	0.455 (0.335)	-0.138 (0.394)
Constant	-12.126*** (0.967)	-11.171*** (1.104)	-24.178*** (2.891)
$\chi^2$ (D.F)	477.31(42)		
Log-likelihood	-1086.69		
Pseudo-R <sup>2</sup>	0.35		
Observations	1,355		

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are; no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries. Regression includes two year dummies not reported for brevity.

A7: Multinomial logit model of Sector Choice; Ghana (Reference category: Not-working)

	Self	Private	Public
Age	0.507*** (0.048)	0.228*** (0.039)	0.279*** (0.069)
Age <sup>2</sup>	-0.006*** (0.001)	-0.002*** (0.001)	-0.003*** (0.001)
Primary	-0.253 (0.180)	0.380** (0.180)	1.021** (0.410)
Secondary	-1.378*** (0.256)	-0.050 (0.216)	1.181*** (0.436)
Tertiary	-1.170*** (0.339)	0.274 (0.272)	1.739*** (0.471)
Sex	-0.794*** (0.153)	0.835*** (0.128)	0.805*** (0.221)
Marriage	0.839*** (0.188)	0.554*** (0.183)	0.537* (0.294)
Children	-0.304 (0.221)	-0.633*** (0.203)	-0.240 (0.346)
Non-Labour income	0.171 (0.183)	-0.188 (0.155)	-0.239 (0.237)
Father's education	-0.045 (0.029)	0.063*** (0.024)	0.026 (0.044)
Mother's education	-0.020 (0.030)	0.036 (0.024)	-0.022 (0.045)
Accra	2.123*** (0.214)	2.779*** (0.194)	2.554*** (0.259)
Constant	-9.369*** (0.817)	-6.426*** (0.648)	-11.507*** (1.358)
$\chi^2$ (D.F)	1416.59(42)		
Log-likelihood	-2136.73		
Pseudo-R <sup>2</sup>	0.25		
Observations	2,297		

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are; no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries. Regression includes two year dummies and country dummy for the pooled model not reported for brevity.

A8: Average Partial effects from multinomial logit model with years of schooling: Tanzania

	Not-working	Self	Private	Public
Age	-0.013*** (0.001)	0.005*** (0.001)	0.004*** (0.001)	0.004*** (0.001)
Education (years)	-0.006** (0.003)	-0.013*** (0.003)	0.005* (0.003)	0.014*** (0.002)
Sex	-0.144*** (0.016)	0.054** (0.023)	0.123*** (0.021)	-0.033** (0.013)
Married	-0.038 (0.039)	0.021 (0.036)	0.003 (0.036)	0.014 (0.018)
Children	-0.010 (0.035)	0.072* (0.042)	-0.027 (0.040)	-0.034 (0.028)
Non-labour income	0.169** (0.066)	-0.141*** (0.044)	-0.029 (0.041)	0.000 (0.017)
Father's education	-0.010*** (0.003)	0.006* (0.003)	0.002 (0.003)	0.002 (0.002)
Mother's education	0.001 (0.003)	-0.002 (0.004)	0.003 (0.004)	-0.002 (0.002)
Dar es Salaam	-0.040 (0.030)	0.046* (0.027)	0.025 (0.027)	-0.032** (0.015)
<i>N</i>	1,355			

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey, in addition to two year dummies not reported for brevity.

A9: Average Partial effects from multinomial logit model with years of schooling: Ghana

	Not-working	Self	Private	Public
Age	-0.009*** (0.001)	0.006*** (0.001)	0.001 (0.002)	0.003*** (0.001)
Education (years)	0.004* (0.002)	-0.014*** (0.002)	0.002 (0.003)	0.008*** (0.002)
Sex	0.009 (0.015)	-0.215*** (0.015)	0.184*** (0.019)	0.022* (0.012)
Married	-0.121*** (0.022)	0.087*** (0.023)	0.034 (0.029)	-0.001 (0.015)
Children	0.047* (0.026)	0.006 (0.026)	-0.067** (0.033)	0.014 (0.016)
Non-labour income	-0.001 (0.020)	0.046** (0.021)	-0.034 (0.023)	-0.011 (0.013)
Father's education	-0.007 (0.005)	-0.006 (0.008)	0.008* (0.004)	0.005** (0.003)
Mother's education	-0.003 (0.003)	-0.006 (0.004)	0.011** (0.005)	-0.003 (0.003)
Accra	-0.416*** (0.043)	0.053** (0.022)	0.330*** (0.029)	0.033*** (0.011)
<i>N</i>	2,297			

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey, in addition to two year dummies not reported for brevity.

A10: Multinomial logit model of Sector Choice; Tanzania (Reference category: Unemployed)

	Self-employment	Private Sector	Public Sector
Age	0.345*** (0.055)	0.292*** (0.062)	0.695*** (0.112)
Age <sup>2</sup>	-0.004*** (0.001)	-0.003*** (0.001)	-0.007*** (0.001)
Primary	1.119*** (0.339)	0.551 (0.347)	-0.066 (0.448)
Secondary	1.153*** (0.387)	0.984** (0.398)	1.673*** (0.453)
Sex	0.820*** (0.227)	1.378*** (0.245)	0.516 (0.321)
Married	0.424 (0.335)	0.351 (0.391)	0.607 (0.427)
Children	0.180 (0.326)	-0.104 (0.390)	-0.406 (0.534)
Non-Labour income	-0.512 (0.496)	-0.139 (0.524)	-0.325 (0.587)
Father's education	0.024 (0.037)	0.016 (0.040)	0.080* (0.046)
Mother's education	0.032 (0.040)	0.061 (0.042)	0.029 (0.051)
Dar es Salaam	0.452* (0.263)	0.454 (0.296)	-0.030 (0.357)
Constant	-8.232*** (1.055)	-6.734*** (1.229)	-18.100*** (2.599)
$\chi^2$ (D.F)	290.49 (39)		
Log-likelihood	1014.62		
Pseudo-R <sup>2</sup>	0.21		
Observations	1,071		

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are; no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries. Regression includes two year dummies not reported for brevity.

A11: Average Partial effects from multinomial logit estimates: Tanzania

	Unemployment	Self-employment	Private Sector	Public Sector
Age	-0.011*** (0.001)	0.006*** (0.002)	0.000 (0.001)	0.004*** (0.001)
Primary	-0.082*** (0.030)	0.182*** (0.044)	-0.029 (0.037)	-0.071*** (0.026)
Secondary	-0.108*** (0.035)	0.056 (0.046)	0.005 (0.038)	0.047** (0.022)
Sex	-0.097*** (0.020)	0.004 (0.030)	0.120*** (0.025)	-0.027 (0.017)
Marriage	-0.039 (0.035)	0.022 (0.044)	0.000 (0.042)	0.017 (0.025)
Children	-0.004 (0.035)	0.069 (0.053)	-0.028 (0.047)	-0.037 (0.036)
Non-Labour income	0.035 (0.050)	-0.077 (0.047)	0.037 (0.040)	0.005 (0.023)
Father's education	-0.002 (0.003)	-0.001 (0.004)	-0.001 (0.004)	0.004* (0.002)
Mother's education	-0.004 (0.003)	-0.001 (0.005)	0.006 (0.004)	-0.001 (0.003)
Dar es Salaam	-0.041 (0.025)	0.050 (0.033)	0.024 (0.031)	-0.033* (0.019)
<i>N</i>			1,071	

A12: Multinomial logit model of Sector Choice; Ghana (Reference category: Unemployed)

	Self-employment	Private Sector	Public Sector
Age	0.447*** (0.056)	0.129*** (0.047)	0.200*** (0.075)
Age <sup>2</sup>	-0.005*** (0.001)	-0.001* (0.001)	-0.002* (0.001)
Primary	-0.379* (0.211)	0.275 (0.209)	0.919** (0.423)
Secondary	-1.735*** (0.289)	-0.342 (0.250)	0.868* (0.453)
Tertiary	-1.647*** (0.368)	-0.108 (0.301)	1.351*** (0.488)
Sex	-1.280*** (0.175)	0.444*** (0.150)	0.384 (0.234)
Married	1.244*** (0.222)	0.884*** (0.215)	0.832*** (0.314)
Children	-0.304 (0.259)	-0.516** (0.241)	-0.176 (0.369)
Non-labour income	0.241 (0.206)	-0.103 (0.176)	-0.168 (0.252)
Father's education	-0.032 (0.037)	-0.046 (0.031)	0.048 (0.048)
Mother's education	-0.021 (0.038)	0.039 (0.031)	-0.028 (0.049)
Accra	3.232*** (0.406)	3.874*** (0.392)	3.626*** (0.428)
Constant	-6.383*** (0.965)	-2.453*** (0.820)	-8.056*** (1.485)
$\chi^2$ (D.F)	1248.48 (42)		
Log-likelihood	-1850.683		
Pseudo-R <sup>2</sup>	0.252		
Observations	2,028		

Note: robust standard errors in parenthesis \*\*\* p<0.01, \*\* p<0.05, \* p<0.1. Reference categories for dummy variables are; no education, not-married, no children, female, no other source of income apart from employment, other urban areas covered in the survey for the two countries. Regression includes two year dummies and country dummy for the pooled model not reported for brevity.

A13: Average Partial effects from multinomial logit estimates: Ghana

	Unemployment	Self-employment	Private Sector	Public Sector
Age	-0.009*** (0.001)	0.006*** (0.001)	0.001 (0.002)	0.003*** (0.001)
Primary	-0.015 (0.022)	-0.086*** (0.021)	0.049 (0.031)	0.052** (0.025)
Secondary	0.077*** (0.026)	-0.215*** (0.029)	0.048 (0.037)	0.091*** (0.026)
Tertiary	0.053 (0.032)	-0.229*** (0.039)	0.066 (0.045)	0.111*** (0.027)
Sex	0.012 (0.015)	-0.215*** (0.015)	0.181*** (0.019)	0.023* (0.012)
Married	-0.114*** (0.022)	0.081*** (0.022)	0.031 (0.030)	0.001 (0.017)
Children	0.048* (0.026)	0.006 (0.026)	-0.067** (0.033)	0.013 (0.020)
Non-labour income	0.000 (0.019)	0.043** (0.021)	-0.034 (0.024)	-0.010 (0.013)
Father's education	-0.008 (0.007)	-0.001 (0.004)	0.006* (0.003)	0.003* (0.001)
Mothers education	-0.002 (0.003)	-0.006 (0.004)	0.010** (0.005)	-0.003 (0.003)
Accra	-0.419*** (0.043)	0.056*** (0.022)	0.329*** (0.029)	0.033*** (0.011)
<i>N</i>	2,028			

A14: Likelihood-ratio test of exclusion restriction MNL

	Tanzania	Ghana
$\chi^2$	1153.58	1416.59
Degrees of freedom	42	42
p-value	0.000	0.000

A15: Average Partial effects from multinomial logit model with young dummy:  
Tanzania

	Not-working	Self	Private	Public
Primary	-0.098*** (0.034)	0.014 (0.046)	0.018 (0.039)	0.066 (0.041)
Secondary	-0.115*** (0.038)	-0.082* (0.047)	0.034 (0.040)	0.164*** (0.041)
Tertiary	-0.151 (0.131)	-0.242** (0.095)	0.172** (0.073)	0.221*** (0.044)
Sex	-0.136*** (0.018)	0.048** (0.023)	0.121*** (0.021)	-0.033** (0.013)
Marriage	-0.080* (0.045)	0.046 (0.037)	0.012 (0.038)	0.022 (0.019)
Children	-0.057 (0.042)	0.087** (0.042)	-0.018 (0.041)	-0.012 (0.028)
Non-Labour income	0.169** (0.067)	-0.146*** (0.045)	-0.027 (0.041)	0.004 (0.018)
Father's education	-0.012*** (0.003)	0.007* (0.003)	0.003 (0.003)	0.003 (0.002)
Mother's education	0.006* (0.004)	-0.004 (0.004)	0.000 (0.003)	-0.003 (0.002)
Young (15-35yrs)	0.132*** (0.024)	-0.016 (0.028)	-0.020 (0.024)	-0.097*** (0.018)
Dar es Salaam	-0.062* (0.032)	0.055** (0.028)	0.033 (0.028)	-0.026* (0.015)
<i>N</i>	1,355			

Note: reference category for young age group is Old (36-64 years of age)

A16 Average Partial effects from multinomial logit model with young dummy:  
Ghana

	Not-working	Self	Private	Public
Primary	-0.048** (0.023)	-0.056*** (0.019)	0.057** (0.028)	0.047** (0.022)
Secondary	0.029 (0.028)	-0.168*** (0.029)	0.055* (0.033)	0.084*** (0.023)
Tertiary	-0.023 (0.037)	-0.173*** (0.039)	0.091** (0.041)	0.105*** (0.024)
Sex	-0.042** (0.016)	-0.178*** (0.015)	0.193*** (0.017)	0.028*** (0.011)
Marriage	-0.152*** (0.022)	0.124*** (0.020)	0.024 (0.027)	0.005 (0.015)
Children	0.003 (0.025)	0.063*** (0.024)	-0.083*** (0.029)	0.017 (0.017)
Non-Labour income	0.041* (0.021)	0.014 (0.020)	-0.044** (0.022)	-0.011 (0.011)
Father's education	-0.007** (0.003)	-0.005 (0.004)	0.010** (0.004)	0.002 (0.002)
Mother's education	-0.002 (0.003)	-0.005 (0.004)	0.009** (0.004)	-0.002 (0.002)
Young (15-35)	0.135*** (0.022)	-0.033* (0.018)	-0.063*** (0.024)	-0.039*** (0.013)
Accra	-0.372*** (0.023)	0.063*** (0.018)	0.282*** (0.019)	0.028*** (0.010)
<i>N</i>	2,297			

Note: reference category for young age group is Old (36-64 years of age)