Gender Earnings Differentials and Power: Some African Evidence[*]

Tayo Fabusuyi[†]

Abstract

We examine a micro-data set from the Ghana Household Worker Survey (GHWS) to address two questions: what informs the observed occupational segregation in the informal labor market and what explains the earnings differential even within narrowly defined activities? This paper advances the notion that the differential outcomes observed across gender is explained in part by the prevailing socio-cultural norms that impose extra costs on women. These costs manifest themselves in two ways – a non-pecuniary dimension that reflects the penalty a woman puts on jobs that are not flexible enough for her to carry out her household responsibilities and a pecuniary dimension that is related to access and the differential cost of productive resources. We showed that these costs influence occupational choice and that they disappear with tenure in the self employed sector but persist within the cohort of wage earners.

---

[*]I would like to express my gratitude to Francis Teal for comments and suggestions. I am also indebted to Justin Sandefur at the Center for the Study of African Economies (CSAE) for providing the data.

[†]Alumni, Oxford University Department of Economics, tayo.fabusuyi@oba.co.uk

1 Introduction

Of late, there has been strong interest in examining earnings differentials across gender in developing countries. Increasingly, much of this attention is being focused on the informal sector. One of the key drivers of this interest is the disproportionate presence of women in the informal sector. In Sub Saharan Africa (SSA), 84% of women in are informally employed compared to 63% of men, with most of the women earning significantly less compared to their male peers and often concentrated in the low income sectors (Chen et al, 2002).

Historical analysis as to the sources of gender earnings differentials have focused on both the demand side and supply side explanations. Demand side explanations, which include both taste and statistical discrimination, are premised on the notion that males and females have different opportunities for employment and that this results in unequal pay for the same set of productivity related characteristics. Supply side explanations, on the other hand, emphasize the characteristics and decisions of individual workers. They attribute differences in earnings to motivation, qualification and labor preferences of individual workers.

The approach of this study builds on the aforementioned but introduces yet another dimension to the discourse. In looking at what obtains in a typical Sub-Saharan African economy, the study examines the prevailing social and cultural norms within these societies and investigates the extent of the bias by the society against women. Specifically, our research focuses on two key questions: what informs the observed occupational segregation in the informal labor market and what explains the earnings differential even within narrowly defined activities? In addressing the former, we focus on accessibility issues and gender specific preferences. With regards to the latter, we focus on the theory of statistical discrimination and we establish a conceptual link with labor market attachment to ascertain if the differential treatment between the sexes tapers off when a woman demonstrates market attachment.

To investigate these issues, the study makes use of a micro dataset from the Ghana Household Worker Survey (GHWS) to analyze the sources of gender earnings differentials with a focus on the informal sector. The micro-data set is a panel collected over a three year period across four distinct geographical areas. The dataset is divided into two categories; wage employees and self-employed. Within the self-employed, we distinguish between business owners and traders. These important distinctions within the self employed cohort of individuals allow us to investigate a recent strand of thought on the dualistic structure of the informal sector in developing countries.

The rest of the study is organized thus. Section 2 conducts an analysis of the existing literature on this debate and gives some background information on the Ghanaian society. It touches on the importance of socio cultural norms in explaining the forces that shape some of the outcomes that we observe. Section 3 focuses on the theoretical foundation and constructs the theoretical model that underlines the empirical analysis. Section 4 details the empirical strategy and gives the summary statistics of the key variables. Section 5 presents the key results while the last section concludes and gives some insight into the policy implications of the findings.

2 Background Context

A plethora of literature on gender earnings differentials exists. However, we focus on the theories that the present research effort seeks to improve upon in explaining the gender earnings differential within the informal sector of a Sub Saharan African country – Becker's (1971) analysis of employers "taste for discrimination," and Phelps' (1972) model of "statistical discrimination".

Becker posits that employers behave as if there is an added cost from hiring workers from a less favored group. He sees discrimination as emanating from the utility maximizing behavior of employers who bear some costs by recruiting certain type of workers. A simple analogy to illustrate this is to assume that an employer faces a decision of hiring either a male or a female. However, he has some aversion to hiring women such that given a wage rate of g for women, the effective wage becomes $g + d$, where d represents the discrimination coefficient. Consequently, the employer hires fewer women. In Phelps's model, uncertainty in the labor market leads employers to make hiring decision under conditions of uncertainty and incomplete information. If they perceive women to be less productive, they can use gender as a means of predicting the worker's productivity and in making future promotion decisions.

These issues may have more relevance in developing countries. Becker, for example, argued that in a perfectly competitive market, discrimination will be eroded in the long run since the cost advantage gained by non discriminating employers allows them to drive out the prejudiced ones. However, in markets that are not competitive, discrimination can exist indefinitely since the opportunity for the non-discriminatory employer to make profit becomes eroded if market agents are willing to boycott the employer. In most SSA communities, the multilingual and multiethnic identities of the population means that such societies are balkanized or physically segmented based on ethnicity, a situation that finds much similarity with non – competitive markets.

Ghana, for example, has a complex ethnic structure and this typifies what obtains in other SSA countries. There are six main ethnic groups: the Akan (comprising primarily of the Ashanti and Fanti), the Ewe, the Ga-Adangbe, the Mole-Dagbani, the Guan, and the Gurma. These groups have distinct languages and norms. However, there are commonalities too. The Ghanaian society is a patriarchal society and the men have been able to use this lopsidedness in power distribution to their own advantage. This is aptly illustrated by Goldstein and Udry (2005) in a study of the Akwapim farming people of Southern Ghana. We also see a manifestation of this within Ghana's urban labor market. Chen et al (2005) showed that women earn 70% of what men earn and they are disproportionately concentrated in activities that give them the flexibility to attend to household responsibilities.

The aforementioned brings to the front burner how non economic factors come into play in determining economic outcomes. Certain institutional, social and cultural norms are instrumental in creating the perceptions that are biased against women and this is also applies outside of the Ghanaian society. Ethnographic evidence, for example, exists that shows the prevalence of patriarchal institutions in most SSA ethnic groups and a high degree of heterogeneity across these groups. Quisumbing and Maluccio (2000), in a study in Ethiopia, stated that anthropological evidence shows that as one moves towards the southern part of the country, the power structure within the household becomes more skewed against the female gender.

In Burkina Faso, Kevane and Gray (1996) documented that the divorce rate is as high as 50% for the Bwaba and as low as 10% for the Mossi – a reflection that the Bwaba woman has a much lower exit cost compared to a Mossi woman. A similar picture exists in Tanzania between the Chagga and Makonde ethnic groups. Among the Chagga, women cannot inherit land and they have a high exit cost in case of divorce with the converse being the case for the Makonde (Gopal and Salim, 1998). This observation is not confined to SSA. For example, Bobonis (2004), in a study of how income transfers affect marital dissolution rates and intra- household resource allocation in rural Mexico,

observed that dissolution rate varies across ethnic groups and it is influenced by the degree of heterogeneity in social norms across these groups.

Such unwritten rules shape societies' perception and expectation from each of the genders and constrain the options a woman has. Browne (1997) opined that "to sustain shared values about gender roles, working women make choices consistent with the expectations of them as wives and mothers". Pagan and Sanchez (2000) in a study of the micro-enterprise sector in Mexico also showed that there are differences in expectations when a female engages in a business activity as compared to when a male does. It is thus obvious that there is the need to look beyond conventional theories of earnings differentials since existing theories do not take into consideration the peculiarities of what obtains within these environments. The present research effort addresses this concern.

3 Theoretical Foundation and Modeling

3.1 Theoretical Foundations

We address two key questions in this study: what informs the observed occupational segregation by gender in the Ghanaian informal sector and what explains the earnings differential across gender even within narrowly defined activities?

In answering these questions, we examine existing theories of the gender wage gap, albeit from a slightly different perspective. We focus on two theories: Gary Becker's (1971) analysis of employers' taste for discrimination and Edmund Phelps' (1972) model of statistical discrimination. We examine them not from an employer's perspective but from the lens of society. This serves two distinct purposes. First, it allows us for a more tractable framework regarding the job choices an individual makes and how his or her gender influences this choice. Second, it provides for a sufficiently broad definition of discrimination that finds relevance even for cohorts of individuals who are self employed.

We perceive a job as a bundle of attributes and assume that there are gender specific differences to these attributes. Since women generally shoulder more of the housework responsibilities compared to men, it is safe to assume that they will be more sensitive to these attributes. This imposes an extra cost on women that men do not bear. *We argue that this cost could manifest itself in a non-pecuniary manner, an example being the stigma associated with a woman undertaking a specific activity that is not culturally acceptable, or in a pecuniary form, with differential access and cost of capital across gender being a prime example.*

The non-pecuniary cost is associated with Becker's taste of discrimination. There are parallels we can paint between this line of thinking and Becker's theory. Just as the employer has an aversion to hiring women, so also does the society has aversion against women being in some specific job types. And as the employer bears some cost from an inefficient utilization of productive resources, so too does society bear a cost in its differential treatment of women. The stronger the bias against women, caused by the prevailing socio-cultural norms, the more costly inefficiencies in productive activities will be within that society.

Let us assume a society where there are only two job types – an individual can either work at home (self employed) or work away from home (wage earner). The disutility of labor is given as $g(h, a)$ where h represents number of hours worked at home and, a the number of hours worked away from home and $g'(.), g''(.) > 0$. For women, we reflect an extra cost to working away from home by including a cost of the form $f(a).I(\bar{a} > a_{min})$ where I represents an indicator function. This cost is non pecuniary in nature[3] but it has a lot of relevance in setting her reservation wages.

We relate the pecuniary cost to Phelps model of statistical discrimination. We assume there are agents who control the conditioning variables of employment opportunities. Instead of offering

[3] An example could be the stigma associated with a woman working away from home or the reduced flexibility that comes with such a job.

different job contracts to men and women, we can perceive of the situation as these agents offering, for example, different loan contracts to men and women[4]. We assume that this added cost manifests itself in form of a decreased per hour earnings. This cost will not be relevant if she works at home. The cost would only kick in if she is a wage employee or if she has her own business, where we have assumed that owning a business belongs to the upper tier of the informal sector. This particular construct develops Field's (1990) notion that the informal sector exhibits dualistic properties with one of the sectors being a free entry sector and a second having barriers to entry.

3.2 Theoretical Modeling

Given the function specified below:

$$min: G = g(h,a) + f(a).I(\bar{a} > a_{min})$$

$I(\bar{a} > a_{min})$ represents an indicator function and the objective function is:

subject to

$$h + a + l = 1 \ (time\ constraint)$$

(h is the time the individual spends working at home, i.e. self employed and a, the time spent working away from home, i.e. wage earner and l represents leisure. T is normalized to 1[5]. We further assume that these job options are mutually exclusive.)

$$w_h h + w_a a \geq \bar{=} Q \ (budget\ constraint)$$

(Q defines the dollar worth of the minimum requirements less income from other sources. We also normalize Q to 1.)

$$h \geq 0$$
$$a \geq 0$$

The following Kuhn Tucker conditions hold for some optimal (h^*, a^*) and non negative multipliers

$$\lambda_1(w_h h + w_a a - Q) = 0$$
$$\lambda_2(1 - h - a - l) = 0$$
$$\lambda_3 h = 0$$
$$\lambda_4 a = 0$$
$$(G_h, G_a) + \lambda_1(w_h, w_a) + \lambda_2(-1,-1) + \lambda_3(1,0) + \lambda_4(0,1) = 0$$

We ignore the trivial case of $h = a = 0$ and examine two possible scenarios: one where the individual only works at home and the other, where she works away from home. If we assume that the worker works solely at home, we have $h > 0 \ and \ a = 0$. Consequently, $\lambda_3 = 0, \lambda_4 \geq 0$ and:

$$G_h + \lambda_1 w_h - \lambda_2 = 0$$
$$\Rightarrow \quad w_h = \frac{\lambda_2 - G_h}{\lambda_1} = \frac{\lambda_2 - g_h}{\lambda_1} \ or \ \lambda_1 w_h = \lambda_2 - g_h$$

On the other hand, if we assume that the individual works away from home, we have $a > 0 \ and \ h = 0$. Thus, $\lambda_4 = 0, \lambda_3 \geq 0$ and:

$$G_a + \lambda_1 w_a - \lambda_2 = 0$$
$$\Rightarrow \quad w_a = \frac{\lambda_2 - G_a}{\lambda_1} = \frac{\lambda_2 - g_a - f_a|_{\bar{a} > a_{min}}}{\lambda_1} \ or \ \lambda_1 w_a = \lambda_2 - g_a - f_a|_{\bar{a} > a_{min}}$$

Assuming a simple functional form[6] as specified below:

$$G = h + a + a^2/2$$

then;

$$\lambda_1 w_h = \lambda_2 - 1 \ and, \ \lambda_1 w_a = \lambda_2 - 1 - a$$

[4] These costs cover all form of differential access to productive resources. For wage earners, the costs may be incurred when hiring decisions are made or when promotions decisions are taken.

[5] We have assumed that $a_{min} = 0$ which may not be the case.

[6] This functional form is only being employed to reveal some stylized facts or buttress existing anecdotal evidence in line with the hypotheses that follow. The empirical analysis is not premised on this functional specification.

$$\Rightarrow \qquad \frac{w_a}{w_h} = \frac{\lambda_2 - 1 - a}{\lambda_2 - 1}$$

If we interpret the lagrangian multipliers as the shadow prices and assume no utility to leisure, then the following hold:

$$\frac{w_a}{w_h} = 1 + a \Rightarrow \; w_a = (1 + a)w_h$$

If wages are equalized ($w_a = w_h$) then the optimality condition above necessarily implies that $a = 0$, i.e. the woman would only work at home. On the other hand, if working outside is positive, i.e. a > 0, then the optimality condition implies that we must have $w_a > w_h$. In a graphical form, w_a increases in "a", where w_h is both the intercept and the slope of the w_a line. This suggests that home wages responds to wages from working away from home and vice versa.

Using simple derivatives, the optimality conditions imply "a" increases (positive) in w_a and decreases (negative) in w_h i.e;

$$a = \frac{w_a}{w_h} - 1 \;\Rightarrow\; \frac{\delta a}{\delta w_a} > 0, \quad \frac{\delta a}{\delta w_h} < 0$$

We use the optimality conditions to show that w_a is bounded above: it can be as high as twice the home wage w_h. It is also bounded below: it can be as low as home wage w_h. To see this, set $a = 1$ in i.e. assuming zero leisure, and fully working away from home. Then the optimality conditions imply that $w_a = 2w_h$. That means that working away from home must sufficiently rise, by twice the home wage, in order to induce the woman to fully abandon working at home. Also if we set $a = 0$, then $w_a = w_h$, i.e. the woman works fully at home.

If the woman works away from home, her reservation wage will increase with the increase in the non pecuniary cost associated with the activity since a higher reservation wage is needed to compensate her for the disutility incurred on the job. She will also demand a higher reservation wage if there are better alternative opportunities to earning income.

We can easily extend this to the 3 sector model we have in the present study. We make a distinction within the self employed sector by making a provision for an upper tier sector (business owner) that is not a free entry sector while still keeping to the assumption that no non-pecuniary cost is incurred by women in this sector. It does necessarily imply that a woman will only be indifferent between trading and owning her business if $w_h = w_b - c$, where the number of hours spent working within the upper tier sector is b and $0 \leq c < w_b$. Using the expression $w_a = (1 + a)w_h$ derived earlier, we have:

$$w_a = (1 + a)(w_b - c)$$

In this case, we see that a limited outside option lowers the reservation wage for a woman working away from home. This does not however, exhaust the gamut of options that exist. It is not impossible that a wage job for a woman incurs both the pecuniary and non-pecuniary cost as specified below:

$$w_a - c = (1 + a)(w_h)$$

The pecuniary cost increases the woman's reservation wage of working away from home by shifting the w_a line vertically by c. In all the scenarios specified above, we see that the inclusion of these unobservable costs influence the labor supply decisions a woman makes.

4 Empirical Strategy and Data Source

4.1 Hypotheses

We have already presented the notion that not only is there differences across gender in terms of preferences to specific job types, but there are also differences in the costs of productive resources. We investigate this by framing two hypotheses:

- Differences in job preferences reflect unevenly borne non-pecuniary costs across gender. Thus, one would expect a more egalitarian society to have similar profiles regarding occupation choices.
- The magnitude of c, the pecuniary cost incurred by women falls and at the limit approach that of a man for a woman who can demonstrate labor market attachment. Consequently, the gender earnings gap should not exist if a woman can establish labor market attachment.

The empirical strategy outlined below investigates these hypotheses.

4.2 Empirical Strategy

This section documents the various approaches used to address the questions and the hypotheses raised above. In the first stage, we deliberate on the determinants of gender earnings differentials without making provisions for selectivity and endogeneity concerns. The second stage takes these issues into consideration. It addresses these concerns through the use of two approaches.

4.2.1 Preliminary analysis of Gender Earnings differentials

We begin the analysis by examining the wage regressions using the pooled cross sectional dataset. A log wage is fitted to each observation using the Mincerian earnings function of the form:

$$lnw_i = \alpha_0 + \alpha_1 gender_i + \alpha_2 X_i + \alpha_3 Y_i + \varepsilon_{1i} \quad (3.1)$$

w_i represents the monthly earnings of the individual in Ghanaian cedis. The explanatory variables are divided into two broad sub groups: a vector of personal characteristics (X_i) and a vector of job characteristics (Y_i). The vector of job related characteristics include the sector to which the individual belongs; tenure, which was specified in a non-linear manner; log of hours, log of employees number for the self-employed cohort of observations and log of firm size for the wage employees. Gender is classified separately to show that it is the variable of interest. It takes the value 1 if the individual is male and 0 otherwise.

While keeping to this specification, we progressively add to the number of control variables to see if it in any way affects the magnitude of the gender dummy, assuming it is significant. Separate regressions were analyzed based on this specification both on the gender of the individual and the sector to which he or she belongs in. We precede these disaggregations with Wald tests to determine if the disaggregations are justified. We also use the specification to determine if labor market attachment plays any role in reducing the gender differential by dividing our dataset into different cohorts of observations based on tenure.

Our approach to determining if labor market attachment plays any role in reducing the gender differential is to run a splined tenure regression. We ask the question, if a woman could over time demonstrate labor attachment, does that lower the cost at which she could procure essential productive resources and at the limit, does this cost approach that of a man on the average? We test this hypothesis by running separate regressions based on different tenured cohorts to establish if indeed the gender dummy could be rendered insignificant with the passage of time.

However, these previous steps only address one side of the coin. A determinant of the gender earnings gap is the difference in the sectoral distribution of male and female. If women are concentrated in low return sectors, then some of the difference in the average returns that accrues to each of the gender could be attributable to this skewed distribution. Using the Duncan index of

occupational dissimilarity, we establish if there is occupational segregation within the labor market. The Duncan index is specified below:

$$D = .5\sum_{j} |f_i/F - m_j/M| \tag{3.2}$$

j indexes the occupation, the lowercase characters, f, m represent number of women and number of men in each occupation type, and the upper case letters represent total number of each gender in all occupations. If $D = 0$, the same occupational distribution obtains across the sexes. If $D = 1$, men and women are totally segregated into different occupations. Thus, we interpret the fraction as the fraction of men who will have to change jobs for each occupation to have the same percentage of women. For our data, the Duncan index between the occupational distribution between men and women is 0.43[7].

If preferences differ across gender, then it is of relevance asking what shapes these preferences. We argue that there would not be a significant difference in the distribution of preferences across gender if no socio-cultural norms exist that are biased sufficiently enough to crea te an uneven playing field. To address this point of contention, we extended on earlier models by explicitly making provision for job attributes as a determinant in the occupation choice process. We argue that job attributes are of paramount importance in occupational choice, especially for a woman. We model occupational choice by means of a multinomial logit as specified in the spirit of Glick and Sahn (1997) though we have made explicit provision for job related attributes in this study which does not reflect in their original construct.

We propose:

$$V_{ij} = \beta X_i + u_{ij} \tag{3.3}$$

V_{ij} represents the indirect utility function, and X_i represents a vector of the individual's characteristics.

However, we can decompose the error term above thus:

$$u_{ij} = \gamma Z_j + \epsilon_{ij} \tag{3.4}$$

Where Z_j represents the vector of job attributes. This format allows us to incorporate benefits or costs that are associated with the job[8].

Substituting equation (3.3) in (3.2) gives:

$$V_{ij} = \beta X_i + \gamma Z_j + \epsilon_{ij} \tag{3.5}$$

[7] This figure needs to be interpreted within the context of our data set and the focus of this study. We are only looking at the informal sector, thus one would expect that if the assumptions of free entry and mobility of labor hold, then the Duncan index should be much lower when compared to a broader data set that includes the public and the formal private sector where these barriers may be more visible. Secondly, most studies use three-digit occupation codes to calculate the Duncan index but we have done the calculation at a much higher level of aggregation by using the sector only as the point of reference. Though our approach underestimates the Duncan index, we have a value high enough that it begs the question of why there relatively high segregation within the informal sector where traditionally there has been very little entry cost and relatively unhindered mobility.

[8] Due to data limitation, the present study will only use dummies for the sectors to proxy these job attributes. While this limits the quality of the research findings, it represents a movement in a direction that will ultimately provide for an explicit modeling of these job specific attributes.

The individual will choose the sector for which V_{ij} is the highest. Therefore, the probability that an individual i choose sector j will be:

$$P_{ij} = Pr(V_j > V_k) \forall j \neq k \tag{3.6}$$

$$= Pr(\beta_j X_i + \gamma_j Z_j + \epsilon_{ij} > \beta_k X_i + \gamma_k Z_j + \epsilon_{ik}) \tag{3.7}$$

$$= Pr(\beta_j X_i - \beta_k X_i) + (\gamma_j Z_j - \gamma_k Z_j) > \epsilon_{ik} - \epsilon_{ij} \tag{3.8}$$

We assume that the errors are independently and identically distributed with a Weibull distribution. Given this assumption, $(\epsilon_{ik} - \epsilon_{ij})$ will have a logistic distribution.

If;

$$V_{ij} = \begin{cases} 1 & \text{for } V_{ik} \leq V_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

and given the assumption on the distribution of the error term, the parameters of the utility model could be estimated through a multinomial logistic approach as specified thus:

$$\Pr(V_{ij}|_j) = \frac{\exp(\beta_j X_i + \gamma_j Z_j)}{\sum_{j=1}^{j} \exp(\beta_j X_i + \gamma_j Z_j)} \tag{3.10}$$

The equation above differs from conventional multiple logistic models in that it makes a provision for job characteristics and reflects the weight an individual places on these attributes. To the extent that the sensitivity to these attributes varies across individuals, they will in principle trade one job attribute against another.

In our specific situation, we considered four categories (3 sectors and individuals who are not participating in the labor market). Thus, the equation above will be written thus:

$$\Pr(V_{ij}|_j) = \frac{\exp(\beta_j X_i + \gamma_j Z_j)}{\sum_{j=0}^{3} \exp(\beta_j X_i + \gamma_j Z_j)} \tag{3.11}$$

To address the indeterminacy problem, we need to select a base group and set the parameters equal to zero. Consequently, the coefficients of the other group will be measured relative to the base group. Assuming the $j = 0$ is selected as the base group, we have the probabilities specified thus:

$$\Pr(V_{ij}|_j) = \frac{\exp(\beta_j X_i + \gamma_j Z_j)}{1 + \sum_{j=1}^{j} \exp(\beta_j X_i + \gamma_j Z_j)} \tag{3.12}$$

and

$$\Pr(V_{ij}|_{j=0}) = \frac{1}{1 + \sum_{j=1}^{j} \exp(\beta_j X_i + \gamma_j Z_j)} \tag{3.13}$$

However, the price we pay for this simplification is that it becomes more difficult making a comparison of estimates across sectors. What we have with the estimates are relative rather than absolute measurements. To correct for this, we calculate the marginal effects of the explanatory variables on the probabilities as shown below:

$$\delta P_i / \delta X_j = P_j(1 - P_j)\beta \tag{3.13b}$$

$$\delta P_i / \delta X_k = -P_j P_k \beta \tag{3.13c}$$

4.2.2 Gender Earnings differentials correcting for potential problems

Given that an appreciable number of the respondents are out of the labor force, selectivity becomes an issue if unobservables in the wage and in the participation equations are correlated and therefore this needs to be addressed. Of our present sample, 37.9% self report to be out of the labor force. Thus, there is a compelling case to take selectivity into consideration.

We use Heckman (1976, 1979) correction method to address this concern using the equations below:

$$lnw_i = x_{1i}\beta_1 + \epsilon_{1i}, \tag{3.14}$$

$$e_i{}^* = x_{2i}\beta_2 + \epsilon_{2i} \tag{3.15}$$

Equation 3.14 determines the individual's market wage and equation 3.15 estimates the individual's propensity to work.

As specified in earlier equations, x_{1i} and x_{2i} are vectors of the explanatory variables, w_i is the market wage if the individual works and $e_i{}^*$ captures the propensity to work. The error terms are assumed mean zero and the βs are the parameters of interest.

Heckman's approach was to define a dummy variable, e_i that assumes the values specified below:

$$e_i = \begin{cases} 1 & \text{for } e_i{}^* \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{3.16}$$

This construct allows us to observe market wage for values of $e_i = 1$. However, the cohort of workers may not be a true representation of the population if the error terms, ϵ_{1i} and ϵ_{2i} are positively correlated. Failure to correct for this problem will lead to inconsistent estimates of the parameters in the wage equation.

We rerun the Mincerian earnings function using the Heckman procedure. Essentially, what the Heckman process does is to estimate the parameters in equation (3.15) using the entire sample. These estimates are then subsequently used to compute the selectivity term, λ_i for each observation. The λs are now included in the expectation form of equation (3.14). An insignificant λ means the OLS estimate of the wage equation will not produce significantly biased results while the converse holds true if λ is significant.

We extend this approach to the case of the multinomial logit process using the generalized form of Heckman approach proposed by Lee (1983). The multinomial logistic process is attractive as a tool for estimating a selection process over a range of exclusive choices since it guarantees simplicity and makes little demand on computing requirement. However, this comes at the cost of fairly restrictive assumptions. Lee addressed this by arguing that what obtains in the classical Heckman case could be extended to polychotomous choice selectivity models.

If it is indeed the case that participation in a given sector is non-random, then unmeasured characteristics of the individual may influence both the wage and sector selection process. Biased estimates will be obtained if these characteristics are correlated with the right hand side variables in the earnings function. To address this problem, Lee proposed a two stage method based on the model outlined below.

Given:

$$\begin{cases} lnw_s = y_s = x\beta_s + \epsilon_s \\ lnw_s{}^* = y_s{}^* = z\gamma_s + u_s \end{cases} \tag{3.17}$$

Where $(s = 1.....m)$ represent the number of sectors, including the choice of not participating in the labor market. The disturbance terms are assumed to have mean values of 0 and constant variance.
Sector s will be chosen if and only if

$$y_s^* > max_{j \neq s}(y_j^*) \tag{3.18}$$

Define

$$e_s = max_{j \neq s}(y_j^* - y_s^*) \tag{3.19}$$

From equations 3.15 and 3.16, it must be the case that $e_s < 0$. For each pair (ϵ_s, e_s), assume the marginal distribution of ϵ_s is $G_s(\epsilon)$ and the marginal distribution of e_s is $F_s(e)$ where $F_s(e) \equiv Pr[e_s < e]$. Using this relation we have:

$$F_s(e) \equiv Pr[e_s < e] = Pr[max_{j \neq s}(y_j^* - y_s^*) < e] \tag{3.20}$$

$$= \frac{\exp(e)}{\exp(e) + \sum_{j=1}^{m} \exp(z_j \gamma)} \tag{3.21}$$

Assuming $\epsilon_s \sim N(0,1)$, we can estimate the earnings function through a two stage method specified below:

$$lnw_s = y_s = x\beta_s + \tau\lambda_s + \epsilon_s \tag{3.22}$$

The equation above is not any different compared to our earlier Mincerian earnings function except for the introduction of $\tau\lambda_s$. Inclusion of λ_s, the selectivity term, corrects for the selectivity bias and gives a consistent parameter estimation. τ is the coefficient of the selectivity term and a significant estimate indicates the presence of selectivity.

4.3 Data Source and Descriptive Statistics

A micro dataset from the Ghana Household Worker Survey (GHWS) was used to analyze the sources of the gender income differentials. The survey, conducted by the Center for the Study of African Economies at the University of Oxford, is a panel data set collected initially over the period from October 2003 to July 2004 and resurveyed during the summers of 2005 and 2006. The geographical areas covered by the survey include Accra, Kumasi, Tokoradi and Cape Coast.

For the purpose of this study, only two waves of the panel (2003/2004 and the summer of 2005) were employed. The relevant variables were stacked and classified into these two broad categories:

- Household information: This category includes information on household size, demographic composition, religion, ethnicity, gender, migrant status, and educational status.
- Employment information: This section contains information on labor force participation, unemployment, self and wage employment, occupation, firm size, number of employees, tenure and age of business.

The pooled dataset has 1199 observations[9]; 523 men and 676 women. The number of women may be higher for two reasons. Firstly, men may be represented more proportionately within the formal sector. However, we have no information on this. Secondly, it may be the case that girls drop out of school earlier and compared to boys, are less likely to be in school beyond the age of 15.

Individuals are classified into three main categories: wage employee, self-employed, or non-participants in the paid labor force. All respondents who reported to work on a wage for time basis

[9] This figure does not include the number of respondents who reported being out of the labor force.

are classified as wage employees. A distinction was made within the cohort of the self employed individuals by further disaggregating it into two sectors – traders, which correspond to the free entry sector and business owners, which represent the upper tier of the informal sector.

The dataset has nine broad categories for ethnicity. We subsequently perform F-tests on the estimates of these ethnicities to see if they are conceptually related enough as to be treated as homogenous. Based on the outcome of the F-test, we reclassified the ethnicity into five groups with the Hausas and the Mandes as the excluded group.

Some of the control variables were specified in a non-linear manner. Thus, the vector of personal characteristics includes squared terms for education, tenure and age. We also include dummies for religion, married, gender and ethnicity. Since a woman's labor supply decisions are more sensitive if she has to cater to young children, we made use of the number of children under the age of 12 in the regression analysis. We use mother's education as an instrument variable since we have more observations for mother's education compared to father's education and since the mother's education has a superior explanatory power.

Workers characteristics include tenure and the square of tenure to capture the non linear relationship specified, number of employees for the self employed cohort, and log of firm size for the wage employees, log of hours and occupational dummies. Unfortunately, our dataset is not fine enough to disaggregate the occupation to a finer analysis thus heterogeneity of activities that may occur within these sectors are not captured. Tables 1b – c gives the descriptive statistics of the key variables both in a pooled form and based on the sectors to which the observations belong.

5 Analysis, Discussion And Findings

Earnings differentials for the Informal Sector

To find answers to what informs the observed occupational segregation by gender in the Ghanaian informal sector and what explains the earnings differential across gender even for narrowly defined activities, we start by estimating a Mincerian specification as specified in section 3 using the pooled cross sectional dataset. An F – statistic was computed to test pooling the data across the two time periods. Pooling the data was accepted at the 95% confidence level. Running the pooled dataset on the specification above revealed an important insight – without controlling for other relevant characteristics, the raw earnings differential is more than 50% on the log point scale. Thus, it will be expected that a woman's earnings will increase by more than half were she a man. Models 2 to 5 of Table 2a give the OLS estimates of the control variables as they are progressively added to the regression equation.

From the pooled cross sectional regressions, one observes some results that are common irrespective of the particular specification. The quadratic specification of the education earning profile shows a convex relationship for all specifications – at low levels of education, a negative return to education obtains and it becomes positive at higher levels. The converse is the case for the age – earnings profile.

It is also of relevance to ascertain if there are differences in returns to factors of production across sectors. For example, do investments in schooling yield the same returns in different sectors of the informal labor market? Tables 2b to 2d give the results of this analysis. The sector-specific earning function regressions exhibit similar results to the pooled results, although some subtle differences were observed across the sectors. The control variables have the highest explanatory power for the cohort of individuals who are wage employees (with R – square valuess ranging from 0.391 to 0.488).

The low explanatory powers of the control variables observed in both the trading and the own business sector indicates that other explanatory variables may have been omitted from the regression equation. Ideally, a variable such as entrepreneurial abilities should have been added to the list of control variables, however, it is not easily observed. Consequently, observable characteristics may not

have as high an explanatory power within the self employed sector as what obtains in the wage sector. We surmise that the screening variables often used in the formal sector finds much relevance in the wage sector and much less within the self employed labor market.

With regards to the study's parameter of interest, the male dummy was significant for all specifications and across all the sectors. Its magnitude is greatest among traders and smallest among the wage employees. Excluding the ethnic dummies and the interaction terms between the gender and the ethnic dummies, the differential in earnings across the gender ranges between a low of 0.213 to 0.295 for wage employees and a high of 0.349 to 0.403 for traders. Among the cohort of business owners, the magnitude of the male dummy is within 0.243 to 0.282.

What is obvious is that the self employment sector which is comprised of both traders and business owners has a higher gender earnings differential compared to the wage employees. The reason for this differential cannot be attributed to discrimination in the conventional sense. However, there could still be some form of discriminatory practices that the society imposes on women that may be responsible for this. *The relatively high raw earnings differential among traders particularly begs the question: in a sector with narrowly defined activity and one characterized by free entry, how is it that such an appreciable difference in earnings persists between the genders?*

We empirically test this by running a splined tenure regression with the dataset being divided into separate cohorts. Table 4b shows the results for individuals who are traders. The parameter of interest, the male dummy, is 0.65 and significant at the 99% confidence level for cohorts of traders with less than 6 years of tenure. The male dummy is not significant for individuals with more than 6 years of tenure. This will lead us to believe that a labor attachment of that duration within the trading sector is sufficient to guarantee that men and women are extended the same treatment.

The converse is the case for wage employees. What we observe from the regression results of Tables 4c is that there are no earnings differentials between the sexes for the cohort of wage employees with less than 3 years of tenure. However, with individuals with more than 9 years of tenure, the male dummy is 0.305 and is significant at the 99% confidence level. By adding more controls through the introduction of ethnic dummies, the male dummy is 0.26 and significant at the 90% confidence level for wage employees with 3 to 9 years of tenure.

This finding runs contrary to Pinkston (2003) who showed that wages of the disadvantaged group rise faster with tenure on the job. What we observe here is that new workers are initially offered similar wage contracts and over time, the wages of women falls relatively compared to men. Unobserved heterogeneity of job types within this sector may explain this outcome. Men may be represented more in job types where the earnings rises with tenure but women may be more visible in activities where earnings are capped after a specific time.

5.2 Occupational Choice

We have been able to offer some insights as to what explains the wage differential across gender even for narrowly defined activities. We subsequently address the question of what informs the observed occupational segregation by gender. We empirically test what determines participation and occupational choice by running the multinomial logistic regression and we present the estimates in Table 5a. The comparison or base group is the wage employees. *Out* represent individuals who are not part of the workforce[10] and the other labels are self descriptive. The religion dummy shows that compared to Christians, Muslims have a higher probability of being in the self employed sector. If we look at the gender dummy, we observe that it is significant across all the occupation types and negative. A statistically significant but negative coefficient across the occupation categories shows that, ceteris paribus, a man is less likely to select these occupations compared to being a wage earner.

[10] The age of respondents falls between 15 and 65. Thus, at the lower end of the age spectrum, an appreciable number of individuals are not in the workforce.

From the estimates in Table 5a, one observes that the single most determining factor that explains the choice of an occupation is the gender dummy.

We investigate this further by computing the marginal effects of age, years of schooling and the number of children less than the age of 12 on occupational outcomes while imposing a linear specification on earnings determinants. Table 5d shows that one year increase in age will increase men's participation in both the wage and own business sectors compared to women. Excluding the non participating sector, increase in schooling impacts positively only on the wage employee sector with men more likely to have a higher increase in probability compared to women. The number of children less than the age of 12 is strongly significant with a negative coefficient for both men and women, although women are more likely to reduce their participation in the wage employee sector compared to men.

We further show the differences between the sexes by examining the marginal fixed effects not only at the mean as was done above, but by graphing the probabilities of an individual selecting specific sectors based on the control variable of interest. This presents a more readable option and shows the stark differences between the two genders by stacking on the same graph the responsiveness of each of the sexes to changes in the explanatory variables. We present the effect on occupational outcomes to the control variables of interest in Figures 1a-c.

We explore two dimensions of the variation. One dimension compares women to men across all of our samples to see if there are differences in sensitivities to determinants of occupation outcome. The other dimension compares women across distinct ethnic groups to ascertain if there are differences to these explanatory variables once an individual's ethnicity changes.

The rationale for proceeding along this line is two-fold. Looking at the gradient of the marginal effect functions, we can determine how responsive these individuals are to taking an activity in a specific sector as a result of a marginal increase in the explanatory variable. We can subsequently use this information to proxy their preferences. *We contend that if there are differences in preferences, a candidate explanation will be the presence of socio- cultural norms that are biased against one of the sexes*[11]. Secondly, comparing women across ethnic groups allows us to determine if the cost that the society imposes on women for participating in a specific sector changes depending on an individual's ethnicity.

Our focus will be on age and education as determinants of the occupational choice an individual makes. With age as the determinant of occupation outcome, the chart on the probability of owning a business reveals that men are more likely to own a business compared to women and that the momentum of this trend picks up approximately 10 years earlier compared to women; approximately 20 years for men and 30 for women. At all levels of the workforce, a man has a higher probability of being a business owner compared to a woman. The trading sector reveals a pattern similar to what obtains with the business owners. The only difference is that a woman actually overtakes a man with the passage of time. An individual older than 35 years has a higher probability of being a trader if a woman compared to a man.

One result with the most variation is how the number of years of education influences the choice an individual makes regarding an occupation. At all education levels, an increase in years of schooling leads to a decrease in the probability that an individual will choose the trading sector. This decrease manifests itself much later compared to men. For both sexes, the probability of joining the wage employment sector increases with education though women tends to be more sensitive to this change at higher levels of education. And while men have a higher rate of owning their own business at lower levels of education, women catch up with them as they increase their year of schooling eventually surpassing men once they have a equivalent of a secondary school leaving certificate.

[11] We say a candidate explanation since differences in preferences may arise for reasons other than socio cultural norms. For example, women may be more involved in raising children because they have a comparative advantage in this activity compared to men – a specific example will be breastfeeding.

We could see the validation of Field's thinking regarding the existence of a dualistic structure for the informal sector. But more importantly, the findings reveal that there are some gender specific patterns in this dual structure. A man has a head start in being able to access the upper tier of the informal sector, an indication that he may be more privilege accessing seed funds compared to a woman. This pattern too obtains at lower levels of education for both genders.

Taking a look at Chart 1b, we observe that the accelerated increase in women taking up a job with the wage employment only kicks in shortly after 12 years of education. It is at this threshold too that their probability of owning a business surpasses that of men. These observations portend two things. For one, getting 12 years of education or more is enough to attract wages in excess of a woman's reservation wages of working as a wage employee. Secondly, it allows her to scale the barrier associated with getting access to credit to start a business. Thus, that certificate, or her being able to demonstrate that she has been in school sufficiently long, serves as a screening mechanism that differentiates her from other women.

In our comparison across ethnic groups, we used a sampling of just two ethnic groups: the Akans and the Ewes/Ga. The reason for this was that these two groups were large enough in the dataset to allow for detailed analysis. For an Akan woman, the probability of being a wage employee increases proportionately more at each level of education compared to an Ewe or a Ga woman. The difference in gradient is particularly noticeable after 12 years of schooling. At approximately 18 years of education, an Ewe woman has a probability of 0.28 of being a wage employee compared to 0.58 for an Akan woman. A plausible explanation is that the fixed cost f which we alluded to earlier in the loss aversion function specified in Section 3 is lower for an Akan woman compared to a Female who is Ewe or Ga. Consequently, her (Akan) reservation wages for working away from home will be less compared to an Ewe woman. One can surmise that less eyebrows are raised if an Akan woman works away from home compared to an Ewe female[12].

5.3 Controlling for Selectivity and Endogeneity

From our dataset, approximately 38% of the respondents self-declare to be out of the labor force. Consequently, there are legitimate concerns over possible sample selection biases. We address the issue of selectivity and endogeneity using the standard Heckman approach and we extend on this by using Lee's approach in estimating wage equations for each sector of interest.

Recollect equations 3.14 and 3.15 which we repeat below:

$$lnw_i = x_{1i}\beta_1 + \epsilon_{1i} \tag{5.3}$$

$$e_i^* = x_{2i}\beta_2 + \epsilon_{2i} \tag{5.4}$$

The first step involves using a probit method to estimate the parameters in equation 5.4.

We now use this to compute the selection term for each individual in the sample. We subsequently use the computed selection term as an added regressor in estimating the earnings function over the sample of working individuals only.

To achieve this, we need to address the identification issue – i.e. find the variable(s) that will be used as exclusion restrictions. Ideally, these are variables that affect whether an individual participate in the labor market but does not affect her earnings. We use the number of children less than the age of 12 and mother's education to identify the λs. Under12 variable was only significant for the cohorts of individuals who are out of the labor force, however, mother's education has no explanatory power in predicting an individual's job choice.

[12] A few words of caution are necessary here. One would have been able to make a more objective comparison if the profile for men were the same for both ethnic groups. However, this is not the case and unless this variation is controlled for, whatever conclusion is drawn from the graph needs to be interpreted with caveat.

Table 6a presents the outcome from the regression of the earnings equation on the pooled data using Heckman's two stage method. The selection term, λ is insignificant, showing that a simple ordinary least square estimate will not produce a significantly biased result. From Table 5a, we observe that the value of the coefficients are not significantly different compared to the ones obtained by OLS method though there are appreciable differences in the standard error.

However, we also need to take into consideration the endogeneity of the sector of employment since individuals who are in, for example, trading, may differ from individuals who are wage employees in some unobservable manner. We address this problem by using Lee's extension of Heckman's approach to polychotomous choice models. In the first stage, we model probability of employment using multinomial logit. Identification problems are addressed using the approach outlined above. The predicted probabilities of an individual being in a specific sector obtained from the multinomial logit are then used in constructing the selectivity correction terms.

The selectivity correction terms, λ_{ij} of individual i in sector j are subsequently used as an added regressor in the form specified below:

$$lnw_{ij} = x_i\beta_j + \tau_j\lambda_{ij} + \epsilon_{ij} \quad \textbf{(5.5)}$$

We run this regression for each of the sectors and observed that the selectivity term is not significant for any of the sectors as shown in Table 6b. Consequently, a simple OLS estimate will not provide inconsistent estimates of the earnings function and findings from earlier regressions are still valid. The insignificance of the selectivity terms revealed that individuals are not self selecting into these sectors based on some unobservable characteristic. Thus, a plausible way to explain the occupational segregation observed are that there are other forces at play that conditions the options individuals have, especially women.

6 Conclusion and Policy Implications

This study examines the question of gender earnings differential from a broader perspective that incorporates not only the elements of discrimination as defined in extant literature but also aspects of socio-cultural norms, that though non-economic in nature, shape the environment in which most economic decisions are made. The study argues that these unwritten rules have significant consequences in what determines women's labor force participation, the sector of employment she partakes in and the earnings she makes from such activity.

The research findings show that there are differences in preferences across gender. We attribute these differences in preferences to differences in costs which we modeled as having both a pecuniary and a non pecuniary dimension. With regards to the non-pecuniary costs, the findings show that individuals consider these costs and will in principle trade one job attribute against the other. For example, we showed that women are much less likely to be wage employees relative to men and more likely to be traders. Our explanation is that this reflects the extra premium women put on having a job that is flexible enough to allow them attend to household chores and raise children.

We use a spline regression on tenure to refute or validate if labor market attachment helps a woman to avoid the pecuniary costs. We run regressions on different cohorts of individuals categorized based on tenure or length of time on the activity, we see that in trading there is a gender difference in earnings of 65% on the log point scale for individuals with less than 6 years of tenure and this dummy is rendered insignificant for cohorts of individuals with tenure in excess of 6 years. We surmise that that is the length of time needed for a woman to demonstrate labor attachment sufficiently enough to allow her to be treated at par with a man. The hypothesis is however refuted for the wage employees where an outcome diametrically opposite to the one observed for trading

holds. A plausible explanation for this is that employers predominantly hire women into occupations in which pay and prospects for advancement are poor.

In addressing the policy implications, we will like to address two salient issues. One, we observed that the second hypothesis is validated in the self employment sector but refuted for the cohort of wage employees. This finding reveals that employers still have preferences regarding job types that they employ women for. These job types are the traditionally defined jobs for women that give little room for pay and growth prospect. A policy that addresses this issue will go a long way in removing the gender wage differential for the cohort of wage employees.

We also observed that there are gender specific differences with the age when an individual owns a business with men being approximately 10years ahead of women. This shows the comparative ease with which these sexes can access capital. Even at lower levels of education, men do better compared to women in being able to float their own business. It is thus obvious that a level playing does not exist in the market for loan-able funds. Our finding is extremely crucial in that it gives an insight into the demographics that underline some of the issues of interest and allows policies to be targeted specifically to these cohorts of individuals.

We have also shown that there are differences in preferences on occupational choice between male and female. However, addressing this is more problematic since it speaks to people's values, conviction and beliefs. Consequently, there may not be a quick way to address this problem. Having said this, providing information that reveals the degree of differences in preferences that are conditioned by societal norms represents a movement in the right direction. We can subsequently build on this to identify possible approaches through which the distortionary effects an individual's gender and ethnicity creates on her labor opportunities may be minimized.

References

Anderson S. & Baland Jean-Marie (2002), 'The Economics Of Roscas And Intrahousehold Resource Allocation,' *The Quarterly Journal of Economics*, MIT Press, vol. 117(3), 963-995.

Becker, Gary (1971), *The Economics of Discrimination*, Chicago, University of Chicago Press.

Bobonis, G. (2004), 'Income Transfers, Intra Household Resource Allocation and Marriage Markets,' mimeo: University of California, Berkeley.

Browne, Irene (1997), 'Explaining the Black-White Gap in Labor Force Participation among Women Heading Households,' *American Sociological Review*, vol. 62 (2), 236-252.

Chen, Martha Alter, Renana Jhabvala and Frances Lund (2002), 'Supporting Workers in the Informal Economy: A Policy Framework,' *Working Paper on the Informal Economy*, Number 2. Geneva: Employment Sector, International Labour Office.

Chen, Martha Alter, Joann Vanek, Francie Lund, James Heintz with Christine Bonner and Renana Jhabvala. (2005), 'The Progress of the World's Women 2005: Women, Work and Poverty,' New York: UNIFEM.

Fields, Gary S. (1990), 'Labour Market Modeling and the Urban Informal Sector: Theory and Evidence,' in David Turnham, Bernard Salomé, and Antoine Schwarz, eds., The Informal Sector Revisited, Paris: Development Centre of the Organisation for Economic Co-Operation and Development.

Glick, P. and D. Sahn (1997), 'Gender and education impacts on the employment and earnings in West Africa: Evidence from Guinea,' *Economic Development and Cultural Change*, 45: 793-824.

Goldstein, M. and Udry, C. (2005). 'The Profits of Power: Land Rights and Agricultural Investment in Ghana', Mimeo: LSE, London. http://www.econ.yale.edu/~cru2/pdf/goldsteinudry.pdf

Gopal, Gita, and Salim, Maryam, eds. (1998), Gender and Law: Eastern Africa Speaks. Proceedings of a World Bank-Economic Commission for Africa Conference. Washington DC. World Bank.

Heckman J. J. (1976), 'The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models,' *Annals of Economic and Social Measurement* 5, 475–492.

Heckman J. J. (1979), 'Sample Selection Bias as a Specification Error,' *Econometrica* 47, 153–161.

Kevane M. and Gray L (1999), 'A Woman's Field Is Made At Night: Gendered Land Rights And Norms In Burkina Faso,' *Feminist Economics*, Taylor and Francis Journals, vol. 5(3) 1-26.

Lee, L.-F., (1983), .Generalized Econometric Models with Selectivity,. *Econometrica*, 51, 507-512.

Pagan, Jose A & Sanchez, Susana M, 2000, 'Gender Differences in Labor Market Decisions: Evidence from Rural Mexico,' Economic Development and Cultural Change, University of Chicago Press, vol. 48(3), 619-37

Phelps, Edmund S. 'The Statistical Theory of Racism and Sexism,' *American Economic Review* 62 (1972), 659-661.

Pinkston, Joshua C. (2003), 'Screening Discrimination and the Determinants of Wages,' *Labour Economics,* vol. 10, pp 643-58.

Quisumbing A. and Maluccio J. (2000), 'Intrahousehold Allocation and Gender Relations: New Empirical Evidence,' FCND Discussion Paper No. 84.

Sánchez, Susana M. (1998), 'Gender Earnings Differentials in the Microenterprise Sector: Evidence from Rural and Urban Mexico,' Unpublished Manuscript, Report No. 36149, The World Bank.

Table 1a: Definition of Variables

Variable	Definition
learn	Log of Earnings
age	Age of Respondent
agesq	Age of Respondent squared
educ	Years of Schooling
educsq	Years of Schooling squared
tenure	Duration of respondent in specific activity
tenuresq	duration squared
lhours	Log of Hours
momeduc	Mother's education
under12	Number of Children < 12 years
lemploys_num	Log of employees' number
lfirmsize	Log of firm size
dum_muslim	Religion dummy with Christians as the reference group
dum_emp	Dummy for individual in wage employment
dum_ownbus	Dummy for individual who has her own business
dum_trad	Dummy for traders
dum_ethakan	Ethnic dummies - dummy for Akan
dum_ethewega	Dummy for Ewes and Gas
dum_ethgurmaguan	Dummy for Gurma and the Grusis.
dum_ethmdagbanigrusi	Dummy for Moledagbani and the Grusis.

Table 1b - Pooled breakdown by both sector and gender

	Pooled			Sectors									Gender					
	All sectors			Trade			Wage			Business Owners			Male			Female		
Variables	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev
learn	1199	12.97	0.94	462	12.79	0.97	419	13.23	0.84	318	12.9	0.94	523	13.26	0.95	676	12.75	0.86
age	1199	34.74	10	462	36.61	9.78	419	32.79	10.18	318	34.58	9.6	523	35.17	9.96	676	34.4	10.02
agesq	1199	13.07	7.47	462	14.36	7.45	419	11.78	7.51	318	12.87	7.14	523	13.36	7.56	676	12.84	7.39
educ	1199	8.6	4.42	462	6.95	4.44	419	10.63	3.78	318	8.33	4.1	523	9.95	3.66	676	7.55	4.67
educsq	1199	93.46	70.75	462	67.92	58.89	419	127.18	77.68	318	86.06	58.43	523	112.39	69.07	676	78.81	68.57
tenure	1199	10.25	9.51	462	12.13	9.93	419	7.74	8.3	318	10.84	9.66	523	10.12	9.04	676	10.36	9.86
tenuresq	1199	1.95	3.36	462	2.46	3.31	419	1.29	2.47	318	2.11	4.22	523	1.84	2.87	676	2.04	3.7
lhours	1199	3.86	0.27	462	3.89	0.26	419	3.84	0.24	318	3.85	0.31	523	3.88	0.24	676	3.84	0.29
momeduc	1199	4.65	5.24	462	3.74	4.82	419	5.91	5.56	318	4.32	5.08	523	4.8	5.15	676	4.54	5.31
under12	1199	1.02	1.25	462	1.16	1.4	419	0.77	1.03	318	1.13	1.25	523	0.97	1.22	676	1.06	1.28
lemploys_num	1199	0.09	0.34	462	0.04	0.21	N/A	N/A	N/A	318	0.29	0.55	523	0.14	0.42	676	0.06	0.24
lfirmsize	1199	0.37	0.66	N/A	N/A	N/A	419	1.07	0.7	N/A	N/A	N/A	523	0.6	0.76	676	0.19	0.5

Table 1c - Sectoral breakdown by gender

	Traders						Business Owners						Wage Employee					
	Male			Female			Male			Female			Male			Female		
Variables	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev	Obs	Mean	Std.Dev
learn	94	13.17	1.18	368	12.7	0.88	158	13.11	1.022	160	12.69	0.81	271	13.39	0.8	148	12.94	0.82
age	94	35.26	10.37	368	36.96	9.61	158	36.29	9.79	160	32.89	9.11	271	34.49	9.9	148	29.68	9.99
agesq	94	13.5	7.85	368	14.58	7.33	158	14.12	7.47	160	1.64	6.59	271	12.87	7.5	148	9.8	7.14
educ	94	9	3.97	368	6.43	4.4	158	9.21	3.23	160	7.45	4.64	271	10.72	3.62	148	10.46	4.08
educsq	94	96.58	61.56	368	60.66	55.97	158	95.19	48.83	160	77.03	65.48	271	127.9	77.66	148	125.85	77.97
tenure	94	11.15	9.05	368	12.38	10.14	158	12.33	9.45	160	9.36	9.66	271	8.46	8.47	148	6.42	7.86
tenuresq	94	2.06	2.88	368	2.57	3.41	158	2.41	3.14	160	1.8	5.06	271	1.43	2.64	148	1.03	2.12
lhours	94	3.97	0.19	368	3.87	0.28	158	3.88	0.25	160	3.81	0.35	271	3.85	0.24	148	3.81	0.24
momeduc	94	4.98	5.44	368	3.43	4.61	158	3.99	4.76	160	4.64	5.38	271	5.19	5.23	148	7.22	5.93
under12	94	1.2	1.66	368	1.14	1.32	158	1.03	1.1	160	1.21	1.37	271	0.84	1.07	148	0.62	0.94
lemploys_num	94	0.09	0.34	368	0.02	0.15	158	0.4	0.65	160	0.18	0.42	N/A	N/A	N/A	N/A	N/A	N/A
lfirmsize	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	271	1.16	0.67	148	0.89	0.73

Table 2a - Pooled Regression (All Sector)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Male	0.508	0.386	0.246	0.256	0.261	0.318
	(0.048)***	(0.051)***	(0.052)***	(0.054)***	(0.054)***	(0.085)***
Age		0.039	0.053	0.052	0.051	0.049
		(0.017)**	(0.017)***	(0.017)***	(0.017)***	(0.017)***
Agesq		-0.041	-0.062	-0.059	-0.060	0.059
		(0.023)*	(0.022)***	(0.023)***	(0.023)***	(0.023)***
Educ		-0.039	-0.033	-0.032	-0.035	-0.035
		(0.016)**	(0.015)***	(0.016)**	(0.016)**	(0.016)**
Educsq		0.006	0.004	0.004	0.004	0.004
		(0.001)***	(0.001)***	(0.001)***	(0.001)***	(0.001)***
Lhours		0.388	0.485	0.487	0.415	0.009
		(0.090)***	(0.089)***	(0.089)***	(0.089)***	(0.089)***
Tenure		0.031	0.027	0.026	0.027	0.029
		(0.007)***	(0.007)***	(0.007)***	(0.007)***	(0.007)***
Tenuresq		-0.060	-0.048	-0.045	-0.052	-0.055
		(0.019)***	(0.019)***	(0.019)**	(0.019)***	(0.019)***
Lemploys_num			0.347	0.344	0.336	0.347
			(0.073)***	(0.076)***	(0.075)***	(0.075)***
Lfirmsize			0.373	0.426	0.411	0.42
			(0.043)***	(0.061)***	(0.060)***	(0.061)***
Trade				0.114	0.128	0.142
				(0.087)	(0.087)	(0.087)
Own business				0.086	0.111	0.103
				-0.090	-0.090	(0.090)
Dum_ethakan					0.266	0.256
					(0.056)***	(0.073)***
Dum_ethgur~n					0.562	0.46
					(0.154)***	(0.240)*
Dum_ethmda~i					0.165	0.224
					(0.114)	(0.149)
Dum_ethewega					0.353	0.492
					(0.069)***	(0.091)***
maleetho						0.022
						(0.109)
maleeth1						0.161
						(0.311)
maleeth2						-0.313
						(0.136)**
maleeth3						-0.153
						-0.231
_cons	12.737	10.078	9.487	9.438	9.570	10.739
	(0.032)***	(0.451)***	(0.443)***	(0.448)***	(0.443)***	(0.308)***
R-Square	0.073	0.200	0.253	0.255	0.280	0.282
# of Obs.	1199	1199	1199	1199	1199	1199

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 2b - Regression for Traders

	Model 1	Model 2	Model 3	Model 4
Male	0.460	0.403	0.349	0.322
	(0.010)***	(0.111)***	(0.110)***	(0.109)***
Age		0.023	0.026	(0.019)
		(0.033)	(0.033)	(0.032)
Agesq		(0.017)	(0.025)	(0.016)
		(0.045)	(0.044)	(0.044)
Educ		(0.013)	(0.015)	(0.004)
		(0.031)	(0.030)	(0.030)
Educsq		0.002	0.002	0.001
		(0.002)	(0.002)	(0.002)
Lhours		0.420	0.410	0.334
		(0.165)**	(0.162)**	(0.161)**
Tenure		0.065	0.062	0.066
		(0.016)***	(0.015)***	(0.015)***
Tenuresq		(0.193)	(0.183)	(0.198)
		(0.049)***	(0.048)***	(0.047)***
Lemploys_num			0.879	0.902
			(0.207)***	(0.204)***
Dum_ethakan				0.179
				(0.097)*
Dum_ethgur~n				0.624
				(0.251)**
Dum_ethmda~i				0.203
				(0.182)
Dum_ethewega				0.511
				(0.136)***
_cons	12.707	11.147	10.921	10.418
	(0.045)***	(0.173)***	(0.362)***	(0.845)***
R-Square	0.038	0.143	0.174	0.186
# of Obs.	462	462	462	462

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 2c - Regression for Wage Employees

	Model 1	Model 2	Model 3	Model 4
Male	0.507	0.295	0.213	0.242
	(0.078)***	(0.071)***	(0.068)***	(.068)***
Age		0.045	0.052	0.045
		(0.024)*	(0.023)**	(.023)**
Agesq		-0.040	-0.055	-0.048
		(0.033)	(0.031)*	(0.031)
Educ		0.037	0.023	0.011
		0.027	(0.026)	(0.026)
Educsq		0.003	0.003	0.003
		(0.001)**	(0.001)**	(.001)**
Lhours		0.187	0.386	0.286
		(0.136)	(0.132)***	(0.132)**
Tenure		0.022	0.009	0.009
		(0.013)*	(0.013)	0.012
Tenuresq		-0.014	0.016	0.011
		(0.045)	(0.043)	(0.042)
Lfirmsize			0.361	0.343
			(0.052)***	(0.052)***
Dum_ethakan				0.305
				(0.075)***
Dum_ethgur~n				0.422
				(0.261)
Dum_ethmda~i				0.121
				(0.182)
Dum_ethewega				0.311
				(0.087)***
_cons	12.892	11.978	10.427	9.880
	(0.063)***	(0.183)***	(0.320)***	(0.641)***
R-Square	0.084	0.392	0.451	0.480
# of Obs.	419	419	419	419

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 2d - Regression for Business Owners

Variables	Model 1	Model 2	Model 3	Model 4
Male	0.396	0.282	0.243	0.246
	(0.091)***	(0.105)***	(0.105)**	(0.105)**
Age		0.065	0.059	0.068
		(0.039)*	(0.039)	(0.038)*
Agesq		(0.072)	(0.068)	(0.083)
		(0.052)	(0.051)	(0.051)
Educ		(0.020)	(0.022)	(0.028)
		(0.038)	(0.038)	(0.038)
Educsq		0.003	0.002	0.002
		(0.003)	(0.003)	(0.003)
Lhours		0.612	0.583	0.513
		(0.164)***	(0.163)***	(0.162)***
Tenure		0.018	0.018	0.021
		(0.014)	(0.014)	(0.014)
Tenuresq		(0.017)	(0.014)	(0.021)
		(0.029)	(0.028)	(0.028)
Lemploys_num			0.253	0.235
			(0.091)***	(0.090)***
Dum_ethakan				0.374
				(0.116)**
Dum_ethgur~n				0.655
				(0.288)**
Dum_ethmda~i				0.181
				(0.230)
Dum_ethewega				0.348
				(0.135)**
_cons	12.678	8.896	9.114	9.091
	(0.065)***	(0.857)***	(0.852)***	(0.841)***
R-Square	0.043	0.162	0.183	0.222
# of Obs.	318	318	318	318

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 3a - Regression for Male (Pooled)

Variables	Model 1	Model 2	Model 3	Model 4
Age	0.071	0.075	0.078	0.076
	(0.031)**	(0.030)**	(0.030)***	(0.030)***
Agesq	-0.079	-0.091	-0.095	-0.094
	(0.042)*	(0.040)**	(0.040)**	(0.040)**
Educ	-0.065	-0.075	-0.073	-0.086
	(0.032)**	(0.031)**	(0.031)**	(0.031)***
Educsq	0.006	0.005	0.005	0.006
	(0.002)***	(0.002)***	(0.002)***	(0.002)***
Lhours	0.446	0.701	0.703	0.637
	(0.166)***	(0.163)***	(0.165)***	(0.166)***
Tenure	0.008	0.008	0.004	0.008
	(0.016)	(0.015)	(0.016)	(0.016)
Tenuresq	0.009	0.019	0.028	0.013
	-0.053	-0.051	-0.051	-0.051
Lemploys_num		0.346	0.344	0.344
		(0.095)***	(0.099)***	(0.099)***
Lfirmsize		0.395	0.484	0.484
		(0.057)***	(0.085)***	(0.084)***
Trade			0.243	0.234
			(0.141)*	(0.139)
Own business			0.129	0.154
			(0.136)	(0.135)
Dum_ethakan				0.291
				(0.092)***
Dum_ethgur~n				0.528
				(0.215)**
Dum_ethmda~i				0.015
				(0.190)
Dum_ethewega				0.160
				-0.111
_cons	9.931	8.853	8.682	8.887
	(0.833)***	(0.812)***	(0.825)***	(0.823)
R-Square	0.121	0.201	0.206	0.227
# of Obs.	523	523	523	523

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 3b - Regression for Female (Pooled)

	Model 1	Model 2	Model 3	Model 4
Age	0.028	0.048	0.045	0.041
	(0.022)	(0.021)**	(0.022)	(0.022)*
Agesq	-0.030	-0.055	-0.052	-0.049
	(0.029)	(0.029)*	(0.030)	(0.029)*
Educ	-0.044	-0.029	-0.028	-0.030
	-0.019	-0.019	-0.019	-0.019
Educsq	0.007	0.004	0.004	0.004
	(0.001)***	(0.001)***	(0.001)	(0.001)***
Lhours	0.361	0.381	0.382	0.324
	(0.106)***	(0.104)***	(0.105)	(0.103)***
Tenure	0.038	0.032	0.031	0.034
	(0.008)***	(0.008)***	(0.008)	(0.008)***
Tenuresq	-0.078	-0.062	-0.061	-0.070
	(0.020)***	(0.020)***	(0.020)	(0.020)***
Lemploys_num		0.388	0.384	0.403
		(0.123)***	(0.129)	(0.126)***
Lfirmsize		0.362	0.366	0.370
		(0.091)***	(0.092)***	(0.095)***
Trade			0.181	0.174
			(0.114)	(0.116)
Own business			0.100	0.096
			(0.123)	(0.123)
Dum_ethakan				0.248
				(0.070)***
Dum_ethgur~n				0.516
				(0.229)**
Dum_ethmda~i				0.262
				(0.142)*
Dum_ethewega				0.492
				(0.087)**
_cons	10.344	9.914	9.897	10.003
	(0.536)***	(0.534)***	(0.540)***	(0.528)**
R-Square	0.161	0.200	0.201	0.243
# of Obs.	676	676	676	676

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 4a - Splined Tenure Regression (All Sector)

	Bottom 10%	Tenure < 3	3 ≤ tenure < 9	Tenure ≥ 9	Top 10%
Male	0.347	0.334	0.217	0.231	0.213
	(0.143)**	(0.095)***	(0.088)**	(0.107)**	(0.214)
Age	0.005	0.074	0.049	-0.049	0.196
	(0.054)	(0.032)**	(0.035)	(0.059)	(0.078)**
Agesq	0.001	-0.103	-0.053	0.069	-0.214
	(0.083)	(0.049)**	(0.050)	(0.077)	(0.085)**
Educ	0.020	-0.043	-0.012	-0.033	-0.037
	(0.044)	(0.028)	(0.029)	(0.036)	(0.052)
Educsq	0.001	0.005	0.004	0.003	0.001
	(0.003)	(0.002)***	(0.002)**	(0.002)	(0.004)
Lhours	-0.386	0.080	0.451	0.424	0.967
	-0.239	-0.146	(0.153)***	(0.213)**	(0.328)***
Lemploys_num	(dropped)	1.797	0.074	0.494	0.346
		(0.322)***	(0.108)	(0.140)***	(0.241)
Lfirmsize	0.337	0.343	0.414	0.389	0.744
	(0.126)***	(0.096)***	(0.090)***	(0.159)**	(0.409)*
Trade	-0.303	-0.021	0.239	0.271	0.113
	(0.171)*	(0.123)	(0.135)*	(0.237)	(0.668)
Own business	-0.126	-0.168	0.271	0.069	0.326
	(0.194)	(0.135)	(0.141)*	(0.239)	(0.684)
Dum_ethakan	0.353	0.268	0.208	0.354	0.287
	(0.150)**	(0.100)***	(0.092)**	(0.116)***	(0.209)
Dum_ethgur~n	-0.773	0.379	0.925	0.202	0.163
	(0.718)	(0.348)	(0.235)***	(0.372)	(0.422)
Dum_ethmda~i	0.539	0.063	0.362	0.435	-0.778
	(0.434)	(0.197)	(0.178)**	(0.240)*	(0.469)*
Dum_ethewega	0.559	0.359	0.276	0.406	0.455
	(0.182)***	(0.359)***	(0.110)**	(0.136)***	(0.257)*
_cons	13.182	10.641	9.353	11.634	4.633
	(1.320)***	(0.775)***	(0.847)***	(1.411)***	(2.354)*
R-Square	0.433	0.354	0.324	0.246	0.246
# of Obs.	124	347	370	482	167

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 4b - Splined Tenure Regression (Traders)

	Tenure < 6	6 ≤ tenure < 15	Tenure ≥ 15
Male	0.557	0.259	0.244
	(0.178)***	(0.168)	(0.232)
Age	0.048	0.034	0.095
	(0.050)	(0.081)	(0.090)
Agesq	-0.058	-0.040	-0.113
	(0.073)	(0.114)	(0.100)
Educ	-0.003	-0.047	0.061
	(0.048)	(0.051)	(0.070)
Educsq	-0.001	0.006	-0.006
	(0.001)	(.0033946)*	(0.006)
Lhours	0.061	0.129	0.860
	(0.215)	(0.350)	(0.328)***
Lemploys_r	1.778	0.750	0.811
	(0.478)***	(0.299)**	(0.355)**
Dum_ethaka	0.032	0.321	0.215
	(0.158)	(0.151)**	(0.200)
Dum_ethgu	0.582	(dropped)	0.612
	(0.344)*		(0.429)
Dum_ethmc	0.214	0.560	-0.101
	(0.259)	(0.334)*	(0.384)
Dum_ethew	0.367	0.842	0.466
	(0.220)*	(0.238)***	(0.264)*
_cons	11.205	11.401	7.319
	(1.176)***	(1.894)***	(2.408)***
R-Square	0.224	0.254	0.123
# of Obs.	168	172	122

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 4c - Splined Regression (Wage Employees)

	Tenure < 3	3 ≤ tenure < 9	Tenure ≥ 9
Male	0.127	0.263	0.294
	(0.117)	(0.155)*	(0.102)***
Age	0.097	0.093	0.009
	(0.044)**	(0.058)	(0.039)
Agesq	-0.144	-0.108	0.011
	(0.018)**	(0.084)	(0.046)
Educ	0.018	-0.069	0.042
	(0.037)	(0.069)	(0.047)
Educsq	0.003	0.007	0.000
	(0.002)	(0.003)**	(0.002)
Lhours	0.335	0.360	0.353
	(0.236)	(0.314)	(0.186)*
Lfirmsize	0.386	0.390	0.370
	(0.089)***	(0.095)***	(0.096)***
Dum_ethakan	0.362	0.298	0.124
	(0.127)***	(0.145)**	(0.126)
Dum_ethgur~n	0.559	0.860	0.032
	(0.474)	(0.669)	(0.348)
Dum_ethmda~i	-0.135	0.471	0.504
	(0.273)	(0.411)	(0.317)
Dum_ethewega	0.221	0.350	0.195
	(0.149)	(0.177)**	(0.136)
_cons	9.059	9.069	10.329
	(1.203)***	(1.670)***	(1.074)***
R-Square	0.446	0.471	0.448
# of Obs.	158	131	130

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 4d - Splined Regression (Business owners)

	Tenure < 5.5	5.5 ≤ tenure < 14	Tenure ≥ 14
Male	0.276	0.154	0.394
	(0.159)	(0.167)	(0.234)*
Age	0.154	-0.112	0.206
	(0.064)	(0.084)	(0.159)
Agesq	-0.224	0.190	-0.229
	(0.096)	(0.115)	(0.183)
Educ	-0.076	0.036	-0.069
	(0.058)	(0.076)	(0.076)
Educsq	0.006	-0.001	0.003
	(0.004)	(0.005)	(0.005)
Lhours	0.171	0.666	0.726
	(0.235)	(0.272)**	(0.371)*
Lemploys_n	0.312	0.260	0.156
	(0.155)	(0.136)*	(0.191)
Dum_ethaka	0.413	0.020	0.729
	(0.174)	(0.186)	(0.251)***
Dum_ethgu	0.911	0.084	0.869
	(0.410)	(0.515)	(0.668)
Dum_ethmd	0.031	0.111	0.504
	(0.350)	(0.338)	(0.588)
Dum_ethew	0.570	0.014	0.467
	(0.204)	(0.217)	(0.299)
_cons	9.177	11.559	5.479
	(1.242)	(1.759)***	(3.552)
R-Square	0.231	0.221	0.209
# of Obs.	116	105	97

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 5a - Pooled Mlogit N = 1199

Occupation*	Out	Trader	Own Business
Age	-0.329	0.228	0.140
	(0.056)***	(0.067)***	(0.067)**
Educ	0.065	0.018	0.048
	(0.069)	(0.071)	(0.075)
Agesq	0.400	-0.292	-0.195
	(0.076)***	(0.089)***	(0.091)**
Educsq	-0.006	-0.009	-0.010
	(0.004)*	(0.004)**	(0.005)**
under12	0.153	0.089	0.084
	(0.084)*	(0.087)	(0.091)
dum_muslim	-0.124	0.666	0.717
	(0.254)	(0.269)**	(0.268)***
Migrant	-0.409	-0.482	-0.547
	(0.212)*	(0.223)**	(0.238)**
Married	-0.134	0.659	0.601
	(0.223)	(0.227)***	(0.240)**
Mother's educ	-0.006	0.000	-0.015
	(0.017)	(0.018)	(0.020)
Male	-0.753	-1.856	-0.667
	(0.178)***	(0.207)***	(0.206)***
_cons	6.955	-3.047	-2.127
	(.956)***	(1.177)***	(1.185)*

**base outcome = Wage earner*

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 5b - Mlogit for men, N = 523

Occupation*	Out	Trader	Own Business
Age	-0.538	-0.063	-0.049
	(0.082)***	(0.105)	(0.096)
Educ	0.024	0.105	0.222
	(0.114)	(0.116)	(0.127)*
Agesq	0.661	0.070	0.064
	(0.110)***	(0.137)	(0.124)
Educsq	0.000	-0.002	-0.020
	(0.006)	(0.006)	(0.008)***
under12	0.162	0.202	0.075
	(0.133)	(0.135)	(0.131)
dum_muslim	0.068	0.146	0.569
	(0.322)	(0.373)	(0.325)*
Migrant	-0.156	-0.232	-0.690
	(0.324)	(0.360)	(0.346)**
Married	-0.916	0.095	0.429
	(0.369)**	(0.400)	(0.367)
Mother's educ	-0.020	0.017	-0.005
	(0.024)	(0.030)	(0.027)
_cons	9.628	0.672	-0.343
	(1.459)***	(1.882)	(1.762)

**base outcome = Wage earner*

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 5c - Mlogit for women, N = 676

Occupation	Out	Trader	Own Business
Age	-0.141	0.461	0.417
	(0.088)	(0.099)***	(0.109)***
Educ	0.160	0.147	0.013
	(0.097)*	(0.100)	(0.105)
Agesq	0.186	-0.585	-0.578
	(0.124)	(0.137)***	(0.154)***
Educsq	-0.016	-0.021	-0.009
	(0.006)***	(0.006)***	(0.007)
under12	0.234	0.136	0.195
	(0.132)*	(0.137)	(0.145)
dum_muslin	0.192	1.517	1.295
	(0.507)	(0.522)***	(0.543)**
Migrant	-0.607	-0.634	-0.576
	(0.310)**	(0.323)**	(0.353)
Married	0.767	1.379	0.988
	(0.351)**	(0.349)***	(0.376)***
Mother's edu	-0.007	-0.025	-0.039
	(0.027)	(0.029)	(0.031)
_cons	3.543	-7.558	-6.459
	(1.400)***	(1.655)***	(1.797)***

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 5d: Marginal Effects of age and educ on Occupational Outcome

	Males		Females	
	dy/dx	Std. Dev.	dy/dx	Std. Dev.
Prob (Being out of the labor force) Male = 0.351, Female = 0.365				
age	-0.0203089	0.00139***	-0.0185765	0.00184***
educ	0.0060772	0.00332*	0.0089073	0.00411**
under12	-0.0101628	0.01081	0.0023367	0.01324
Prob (Being a trader) Male = 0.248, Female = 0.345				
TRADING				
age	0.0107769	0.00106***	0.0167333	0.00167***
educ	-0.0230941	0.00267***	-0.0186315	0.00387***
under12	0.0237314	0.00872***	0.0183774	0.01303
Prob (Being a wage employee) Male = 0.192, Female = 0.101				
WAGE EMPLOYEE				
age	0.0031196	0.00098***	-0.0014579	0.00098
educ	0.0238111	0.00274***	0.0151123	0.00241***
under12	-0.0275927	0.00941***	-0.035819	0.00975***
Prob (Being a business owner) Male = 0.209, Female = 0.188				
BUSINESS OWNER				
age	0.0064124	0.00099***	0.0033011	0.00132**
educ	-0.0067942	0.00258***	-0.0053881	0.0031*
under12	0.0140242	0.00833*	0.0151049	0.01003

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 6a - Earnings Function with Selectivity*

	Model 1	Model 2
male	0.386	0.398
	(0.051)***	(0.066)
age	0.039	-0.003
	(0.017)**	(0.064)
agesq	-0.041	0.007
	(0.023)*	(0.079)
educ	-0.039	-0.034
	(0.016)**	(0.023)
educsq	0.006	0.005
	(0.001)***	(0.001)***
tenure	0.031	0.035
	(0.007)***	(0.009)***
tenuresq	-0.060	-0.062
	(0.019)***	(0.026)**
lhours	0.388	0.414
	(0.091)***	(0.117)***
λ		-0.192
		-0.379
_cons	10.078	10.826
	(0.451)***	(1.361)***
# of Obs	1199	1199

**Model 1 does not include the selectivity term while model 2 does. R-Squared value for model 1 is 0.2 and model 2 has a Wald Chi squared value of 293.11*

Significant at the .10 level ** at the .05 level and * at the .01 level*

Table 6b - Correcting for Selectivity, a la Lee

	Trade	Ownbus	Wage Employee
age	0.049	0.100	0.082
	(0.157)	(0.056)*	(0.036)**
agesq	-0.063	-0.329	-0.100
	(0.194)	(0.120)***	(0.050)*
educ	-0.056	-0.072	0.009
	(0.143)	(0.109)	(0.051)
educsq	-0.002	0.000	0.004
	(0.008)	(0.009)	(0.002)*
tenure	0.086	0.025	0.030
	(0.019)***	(0.018)	(0.008)***
tenuresq	-0.245		0.012
	(0.060)***		(0.066)
lhours	1.637	0.584	0.569
	(0.804)**	(0.211)***	(0.224)**
lemploys_num	1.243	0.256	
	(0.389)***	(0.149)*	
λ	0.759	0.655	-0.595
	(1.196)	(0.683)	(0.463)
lfirmsize			0.249
			(0.091)***
_cons	8.447	7.885	11.350
	(1.911)	(1.600)	(2.104)***
R-Square	0.105	0.139	0.391
# of Observation	371	239	241

λ *represents the selectivity term for each sector*

Significant at the .10 level ** at the .05 level and * at the .01 level*

Figure 1a: Occupational Choice and Years of Schooling

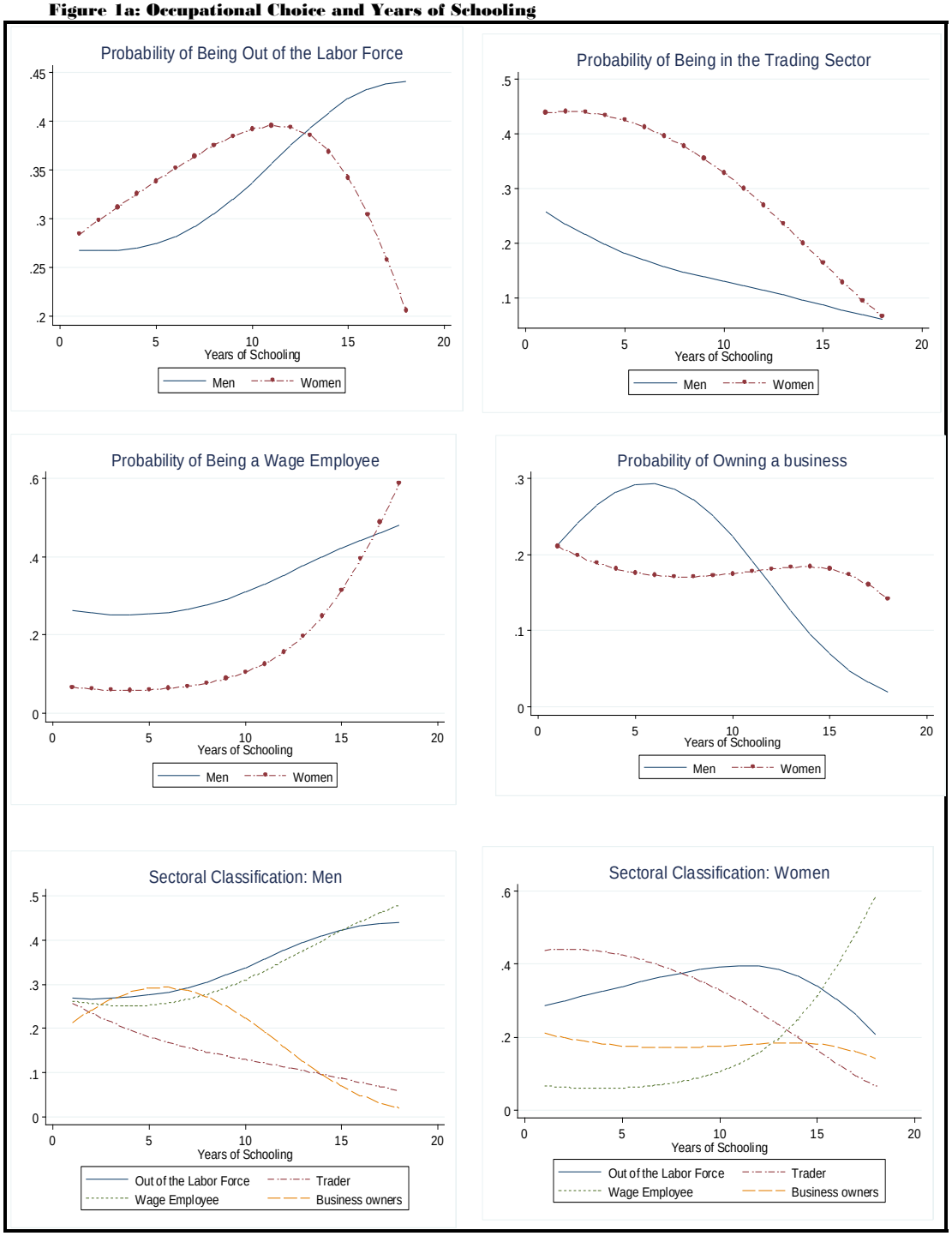

Figure 1b: Occupational Choice and Age of Respondent

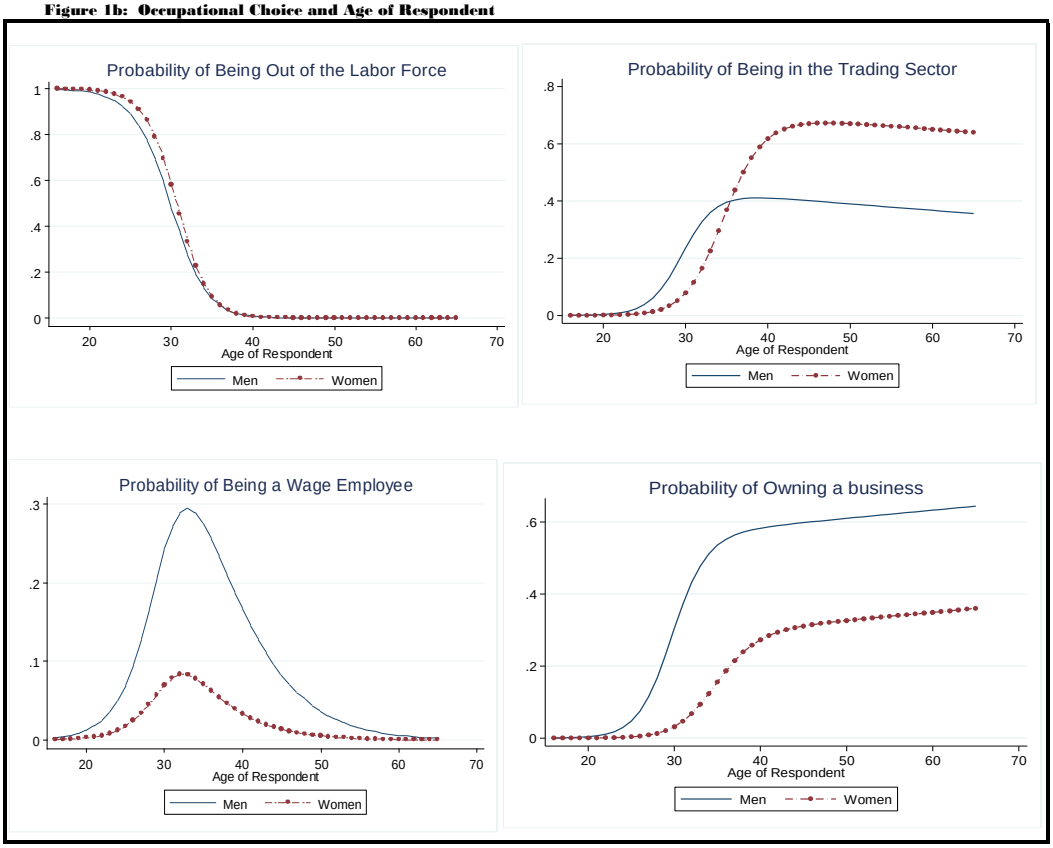

Figure 1c: Comparing Wage Employment Probability and Schooling between Akans and Ewes/Gas

Akan **Ewe/Ga**