Revisiting the determinants of informal sector in Burkina Faso

Traoré, Jean Abel

University Ouaga II, Laboratory of Economic Policy Analysis (LEPA)

November 2012

Revisiting the Determinants of Informal Sector in Burkina Faso[*]

Jean Abel Traoré[†], Université Ouaga II, Burkina Faso

Abstract

The main objective of this paper is to highlight the determinants of informal sector activity with a special focus on informal self-employment opposed to informal wage employment. For this purpose we build a theoretical framework to show the important role of education in occupational choice. To empirically test our hypothesis, we estimate a two-stage degenerated nested Logit model after rejecting the IIA hypothesis. Our results show that the probability of formal sector employment increases with the level of education. In the informal sector education attainment tends to reduce the probability of self-employment as opposed to wage employment. Other relevant determinants are related to gender, age and place of residence.

JEL Classification: J24, J31, O12

Key words: informal sector, education, nested logit

INTRODUCTION

The general objective of this paper is to revisit the determinants of informal sector activity in Burkina Faso. The paper specifically seeks to investigate the drivers of informal self-employment relative to formal and informal wage employment. Since one would expect education and the basic understanding of the way the economy functions to affect people's decision to enter one sector or another, the paper first develops a simple theoretical model to shed light on that potential link. Then, using household survey data from Burkina Faso for 2007 it empirically assess the impact of various factors on the probabilities for three employment options: formal sector wage employment, informal sector self-employment and informal sector wage employment. The probabilities are estimated using a two-stage degenerated nested logit model after rejecting the IIA[1] hypothesis.

The results show that education positively impacts the likelihood of wage employment whether in the formal or in the informal sector of the economy in Burkina Faso. Likewise, education increases the probability of formal sector employment over informal sector employment. Findings also show that some other factors do not play a clear role in driving people into the formal sector of the economy. Those include place of residence, access to credit, access to drinking water and to phone lines. The results also show that even in the informal sector men are still more likely to land job offers than women.

The concept of informal sector finds its origins back in the 1970s with the emergence of heterogeneous good and service production activities out of the usual framework of the economy (BIT, 1970; Keith Hart, 1972).Since then the concept of informal economy is sparking off interesting debates on its definition, determinants and links with formal economy, job creation and economic growth. If there is a consensus on the difference between informal activities and "prohibited" activities, some authors still argue that the development of informality is due to the will to escape from heavy taxes and regulations (Loayza, 1996; Loayza and al., 2006). According to Botero and al (2003), the regulation of labor market through a minimum wage for example prevents the economy from using all the available labor. Thus, the unemployed are obliged to take refuge in the informal sector.

The evolution of the informal sector in developing countries is intimately related to the socioeconomic crisis of 1970s and the following Structural Adjustment Programs (Roubaud and Seruzier, 1991). In fact, most governments in developing countries had to reduce the supply of public jobs as part of their strategy to reduce budget deficits. Moreover, privatizations shifted the burden to occupy a growing workforce to an embryonic private sector. The failing public and private sectors, and the seasonality of rural activities led to the emergence of a new type of activities called "survival activities" generating incomes even in urban areas. That explains the reference to the informal sector as a low productive one deemed to disappear as the economy develops (Schneider and Enste, 2000).

---

[1]Independence of Irrelevant Alternatives

However, over the last decade, the informal sector has shown persistence in Sub-Saharan economies and positioned itself as an alternative source of income and professional insertion. The informal sector stands as the second source of job creation behind the rural sector by occupying more than 70% of the non-farm workforce and by contributing about 42% to the GDP of these economies[2]. In Burkina Faso, the unstructured economy[3] represents 25% of the country's GDP and occupies about 70% of the non-farm workforce[4]. Moreover the 1-2-3 survey carried out in Ouagadougou (city capital of Burkina Faso) has highlighted that only 9.3% of informal sector entrepreneurs have chosen this activity because they did not find a job in the formal sector. However, most of the informal sector's actors consider informal self-employment as a best way for professional insertion. In fact as pointed out by Günther and Launov (2006), depending on their characteristics, some individuals may have a comparative advantage to work in the informal sector. According to the same survey informal self-employment represented 72.2% of informal units in 2001 and more than 81.5% of new jobs created in the informal sector from 1990 to 2000. However over the last decade self-employment represented only 52% of new jobs in this sector. Thus, the unstructured economy progressively reveals his capacity to hire new labor under salaried status.

These facts point to the need to revisit the factors that drive people into the informal sector. While informal self-employment used to be dominant, the informal sector is increasingly offering wage employment opportunities. Thus, shedding light into the dynamics of the informal sector requires that one first understands the determinants of informal self-employment. The picture can then be complemented with an investigation of the intrinsic, economic and social characteristics leading or constraining people to informal self-employment. The answer to those questions can contribute to rationalize the persistence of informal activities and enlighten policies aimed at addressing the issue of informality. In fact, the basic statistics computed with the data base show that the average monthly earning for the formal salaried employees is 81,071.01 FCFA against 44,424.53 FCFA for the informal wage earners and 43,000 FCFA for the self-employed. From these differences it is puzzling that the majority of the informal sector actors would declare their strong preference for the informal sector as revealed by the 1-2-3 survey in Ouagadougou.

Several studies on the labor market of Developed countries have pointed to individual socioeconomic characteristics, business environment, and government policies as the main determinants of self-employment (Rees and Shah 1986; Bernhardt 1994; Millàn and al, 2010). Some authors as Yamada (1996), Combarnous (1999), find similar results for sub-Saharan Africa after distinguishing informal and formal sectors.

---

[2] International Labor Organisation (ILO) : 12th African Regional Meeting ; Johannesburg, South Africa, October 11-14 2011
[3] According to the ISND (the National Agency of Statistics in Burkina Faso), the informal sector refers to all the units of production without a fiscal identification number **and/or** which do not keep written books
[4] INSD (2003)

Studies of the determinants of informality in Burkina Faso are typically limited to the few major cities (Ouagadougou, Bobo-Dioulasso and Koudougou). This paper uses the annual survey of households' standards of living 2007 which is a recent database covering the whole country of Burkina Faso. Thus this paper goes deeper in featuring the particularities of informal activities (non-agricultural) in rural areas and adds to the understanding of informality trends in Burkina Faso.

The paper is structured as follows. The first section provides a literature review. The second section develops a theoretical framework around the role of education in choosing one sector of activity or another. Section 3 presents the database used in our empirical analysis as well as the variables used in our estimation model. Section 4 presents estimation results along with the related discussion. The last section concludes.

Section 1: Literature Review

Standard models of occupational choice based on individual rationality argue that the choice between self-employment (formal or informal) and formal wage employment depends on the difference of utility between these two statuses. Then an individual would choose informal or formal self-employment if the expected utility, most of time measured by the expected income, exceeds the utility from formal wage employment (Yamada, 1996; Rees and Shah, 1986; Bernhardt, 1994; Johansson, 2000; Constant and Zimmermann, 2006). However, some authors as have found that the difference in expected incomes may not be a significant factor in some instances (see, e.g. El Aynaoui, 1997 in Morocco). El Aynaoui (1997) emphasizes intrinsic characteristics as gender, age, education and extrinsic characteristics such as familial and economic environments. Do and Duchene (2008) also show that women in Vietnam have to rely on self-employment mainly in the informal sector despite the attractiveness of formal salary jobs.

In Sub-Saharan Africa (SSA), relatively low levels of women's education and social norms often limit women to unpaid domestic cleaning work. For example according to the 1-2-3 survey in WAEMU[5] capitals, 64% of women did not complete primary school compared to45% for men (Kuepie, Nordman and Roubaud, 2009). In those conditions, engaging into small income-generating activities at home or on streets remains entry point into the labor market. In Ouagadougou, women represent about 47% of informal self-employed in catering and retail trade from which they derive the most precarious incomes (INSD, 2003). This contrasts with the case of well-educated women who are likely to be wage-earner in the formal sector (Kabubo-Mariara, 2003). However, Hammouda and Lassassi (2011) find that women who attended vocational school are still likely to be self-employed in the informal sector. These results suggest an ambiguous role for human capital in the occupational choice of individuals, including women, men and the young alike.

---

[5] West African Economic and Monetary Union

With regard to the role of education in occupational choices two main strands emerge from studies of the labor market in developed countries where the distinction between informal and formal sector have been overlooked. The first one holds that a high level of education favors self-employment (Le 1999, Rees and Shah, 1986; Evans and Leighton, 1989; Borjas and Bronars, 1989; Blanchflower, 2000; Congregado and al, 2009). Education may be a good channel to acquire managerial skills to choose entrepreneurship. However, Casson (2003) questions this idea and claims that formal education is not a panacea for sound managerial skills. Thus the second theoretical strand which suggests that higher level education increases the chance of wage employment in the less risky formal sector (Le, 1999; Evans and Jovanovic, 1989; Kidd, 1993; Taylor, 2001; Destré and Henrard, 2004).

Results from many studies of the informal sector in SSA lend support to the second strand above (House et al, 1993; Taylor, 1996; Rankin, Sandefur and Teal, 2010). In fact, in most West African cities the formal sector (public and private) values education (Kuepie, Nordman and Roubaud, 2009) by providing social protection and a better access to financial services. In contrary less educated people are more likely to be self-employed or wage-earners in the informal sector since the alternative to create a formal firm involve some financial and administrative cost many of them cannot bear. Cogneau (2001) and Zerbo (2006) point to job rationing in the formal sector as an important factor that limits the absorption capacity of the structured economy for qualified labor. This explains the increasing number of young graduates working in the informal sector.

Moreover, education increases the chances of getting better paid positions in the formal sector (Kuepie, Nordman and Roubaud, 2009; Nguetse Tegoum, 2009). This contrasts with the case of uneducated young individuals who are more likely to be unpaid workers as apprentices or family-aids. For those young individuals there are therefore many barriers to the labor market. Limited by their low level of education or their lack of experience most of them are relegated to the informal sector although their hope is for a better job in the structured economy (Zerbo, 2006; House and al, 1993).

Similarly, family background as including marital status, household size and parental occupation impacts occupational choice and the risk of informality. Rees and Shah (1986) analyze that married individuals mentally supported by their family are more likely to take risk by electing for self-employment. However, Dolton and Make (1990) consider that the same family background can bring about risk aversion because of family responsibilities. Zerbo (2006) argues that when the household relies only on the head's income to meet end needs any negative shock to his work conditions (e.g. decrease of income) can lead family members to consider income generating activities generally in the informal sector, regardless of how precarious those activities may be. In such instances, household size and the expansion of the informal sector are positively correlated. However, El Aynaoui (1997) and Njikam and al (2005) respectively find that in the case of Morocco and Cameroon individuals whose parents work in the formal sector are more

likely to get a job in the formal sector as well whereas those whose parents are in the informal sector are more likely to seek opportunities in the informal sector.

External factors to individuals such as place of residence (rural or urban), the economic and institutional environments also play a role in the occupational choice. With regard to the place of residence, we must recognize that informal activities are originally an urban phenomenon. The availability of infrastructure and the demand for services and consumption goods in urban areas increases the probability of self-employment or informal sector employment for those living in those areas. However informal activities relating to trade, transportation, manufacturing have lately been expanding in the rural areas as well. These activities represent an opportunity to diversify the sources of household income and are consistent with the dynamics of countryside development.

With regards to individuals' economic environment this refers to physical and financial endowment. It is well known that entry into the formal sector is conditional on some level of education whereas informal sector is generally characterized by its open access. However starting an informal business may require a minimum level of investment capital (Jütting and al, 2008). The level of available capital in turns determines the type and the branch of activities one can enter in. Most employers or self-employed are also better endowed with physical and financial capital compared to wage workers. Therefore, liquidity constraints to self-employment in either the formal or the informal sector do exist (Congregado and al, 2009). Available evidence from Burkina Faso shows that most informal actors derive their capital from personal savings, inheritance and family donations (INSD, 2003).

The business cycle may also play a role regarding informal sector activity. Zerbo (2006) shows that during recessions the supply of formal jobs decreases and more people have to earn a living in the informal sector. Loayza and Rigolini (2006) also find evidence of a counter-cyclical informal sector. However, Fiess, Fugazza and Maloney (2010) document episodes of pro-cyclical informal employment in Latin America countries, especially when driven by relative demand or productivity shocks to the non-tradable sector.

Also important in understanding occupational choices is the institutional context, i.e. labor market regulation, intellectual property rights and the business environment. Many authors, including Maloney (2004), Oviedo (2008), Ingram and al. (2007), DCED (2009) and Andrews and al. (2011) analyze that entry into the informal sector is based on a cost/benefit analysis. An individual would choose informal sector if the expected benefit exceeds the net benefit of activity in the formal sector. The benefits of informality include not only the expected income but also tax avoidance, autonomy, and greater flexibility in terms of working hours which sometimes proves useful for women often constrained by social norms (World Bank, 2004; Jütting and al, 2008). The costs involve non access to public infrastructure (electricity, telecommunication), limited access to formal financial market, lack of social protection (social security,

insurance), and lack of formal contracts which leads to occupational hazards, long working hours and uncertainty of future earnings (Jütting and al, 2008).

However other authors argue that the cost/benefit analysis requires perfect information and skills that the majority of informal workers do not have (DCED, 2009; Jütting and al, 2008). Available evidence shows that 60% to 80% of own-account workers in Ouagadougou don't know how to register their activity whereas only 1% to 3% consider the process is heavy and expensive (INSD, 2003). Thus, entry into the informal sector is not necessarily voluntary as implied by the cost/benefit approach.

Whereas the cost/benefit perspective assumes that individuals make informed decisions to enter the informal sector, the empowerment approach posits that governance failure is the main cause of exclusion from the structured economy, especially for the poor (Perry and al, 2007).Unequal access to basic social services like education and health translates into differentiated access to formal labor markets. Initially low capabilities combine with limited access to social and judicial services and lack of economic opportunities to restrict the poor to the informal sector. Thus the empowerment approach points to the fact engagement in the informal sector is beyond people's control (i.e. not voluntary) and seems to be more in line with the realities of Sub-Saharan Africa where significant disparities in access to public services and opportunities exist between agents from the informal and formal sector of the economy.

Section 2: Theoretical Framework

Consider a small closed economy with perfectly competitive markets and two sectors of activity: formal (F) and informal (I). Individuals are endowed with one unit of labor which they inelastically supply in the labor market, whether formal or informal. We assume a more productive technology relative to the informal sector. This may be because of a better access to infrastructure and financial markets. In other words the total factor productivity is higher in the formal sector compared to the informal one.

A- Production

Production of goods and services in the formal sector is given by a neo-classical production function[6]:

$$Q^F = AF^F(H,K) \quad \text{with} \quad A > 1 \tag{1}$$

Q : Total production of goods and services in the formal sector

H : human capital (quantity and quality of labor) used in the formal sector

K : physical capital used in the formal sector

A : Total factor productivity (global productivity of factors) in the formal sector

---

[6] F(.) at least twicely differentiable; ii) strictly decreasing partial derivatives; iii) concavity ; iv) constant returns to scale ; v) INADA conditions

In the short term, physicaly capital is fixed. Letting $K = K_0 = 1$ then equation (1) becomes:

$$Q^F = AF^F(H) \tag{2}$$

At the competitive equilibrium of the economy the marginal productivity of human capital is equal to the real wage rate[7]:

$$AF_H^F(H) = w^F \tag{3}$$

Production in the informal sector is also given by a similar neo-classical function:

$$Q^I = BF^I(h, k) \tag{4}$$

Q^I:Total production of goods and services in the informal sector

h :human capital (quantity and quality of labor) used in the informal sector

k:physical capital used in the informal sector

B: Total factor productivity (global productivity of factors) in the informal sector

As the formal sector is more productive technology relative to the informal sector then $A > B$

In the short term, physical capital is considered fixed so $k = k_0 = 1$ and relation (4) becomes:

$$Q^I = BF^I(h) \tag{5}$$

At the competitive equilibrium of the economy the marginal productivity of human capital is equal to real wage rate:

$$BF_h^I(h) = w^I \tag{6}$$

Thus human capital receives a higher wage in the formal sector. In fact,

$$H = h \Rightarrow AF_H^F(H) > BF_h^I(h) \quad \text{because } A > B \text{ so } \quad w^F > w^I \tag{7}$$

In other words, due to technology differences qualified individuals receive higher wages in the formal sector compared to the informal one. For simplicity we assume that wages in the formal sector capture all other benefits from formal employment including social security contributions (family allowance, pension, insurance…)

B- Consumption

Consider an individual i whose labor income is mainly allocated to consumption of goods and services. If i is employed in the formal sector his utility is as follows:

$$U_i^F = U^F(E_i, M_i, w_i^F) \tag{8}$$

E_i : level of education (proxy of human capital)

M_i : intrinsic and familial characteristics (gender, age, marital status, household size)

w_i^F : gross income of individual i in the formal sector

Similarly if i is employed in the informal sector his utility is written as:

$$U_i^I = U^I(E_i, M_i, w_i^I) \tag{9}$$

So the utility of agent i in the whole economy can be expressed like:

---

[7]$F(.)$is a function at least twicely differentiable. In this paper $\frac{\partial F}{\partial x}$ is noted F_x and $\frac{\partial^2 F}{\partial x^2} = F_{xx}$

$$U_i = U(x, E_i, M_i) \tag{10}$$

With

$$x = [0, 1, 2, 3, \ldots, N]^8 \tag{11}$$

$$\longleftarrow \qquad \longrightarrow$$

Informal   Formal

It is possible to group together the workers in different occupational classes ranked according to the importance of incomes and benefits related to each class. For example in the formal sector on can distinguish employers, wage earners, interns and sector, employers, own-account workers, wage earners, apprentices and family-aids in the informal.

x is a variable that represents the different occupational groups(professional categories) in the labor market. As x increases individual i moves toward the formal sector with higher wages and benefits. On the other hand, when x decreases individual i moves toward informality with growing vulnerability and precariousness. Hence:

$$U_x > 0 \tag{12}$$

Intuitively, for a given level of education and other socioeconomic characteristics, if the consumer is already in a formal occupational class providing consistent incomes and social advantages, the marginal utility for reaching a better class should be declining. Thus the following assumption:

The marginal utility of x is declining $\Rightarrow U_{xx} < 0$ \hfill (13)

However, as the level of education increases, the chances to access a better occupational class in the formal sector increases as well. Therefore, the marginal utility associated with the progression toward these better classes may tend to increase. Thus the following assumption:

The marginal utility of x is increasing in the level of education $\Rightarrow U_{xE} > 0$ \hfill (14)

Under the assumption of rational individuals, as the level of education increases and given intrinsic and familial characteristics, a consumer should prefer the occupational class that maximizes his utility subject to the constraint of entry costs into the formal sector. That is expressed by the following maximization program:

$$\text{Max}_x \, U(x, E, M) \tag{15}$$

$$ucC(x, E) \leq R$$

---

[8]We follow Mbaye and Benjamin (2012) who depart from the traditional formal vs. informal distinction and develop an approach to informality in continuous terms.

C is the for entry in the formal sector cost function[9].It captures all the costs pertaining to the search of information and administrative process necessary to land a formal sector job for example. Those costs are increasing in x. Thus:

$$C_x(x,E) > 0 \qquad (16)$$

Moreover individuals with a higher education level may have a comparative advantage relative to unqualified individuals in their job search. For instance they may have better ability to use ICT[10], to draft a CV and hold a job interview. So we assume that the entry cost to formal sector decreases as the level of education increases:

$$C_E(x,E) < 0 \qquad (17)$$

Also, we assume that:

- The marginal cost to enter in the formal sector increases with the occupational class

$$\Rightarrow C_{xx} > 0 \qquad (18)$$

- The marginal cost to enter in the formal sector decreases with the level of education

$$\Rightarrow C_{xE} < 0 \qquad (19)$$

R : represents the maximum wealth or assets that a consumer can allocate to the job search.

The Lagrangian associated to program (15) can be written as:

$$L(.) = U(x,E,M) + \lambda\big(R - C(x,E)\big) \text{ with } \lambda > 0 \qquad (20)$$

The first order conditions imply that.

$$L_x(.) = 0 \Rightarrow U_x(x,E,M) - \lambda C_x(x,E) = 0 \qquad (21)$$

Next, consider a function $\Gamma(.)$ defined as: $\Gamma(x,E,M,\lambda) = U_x(x,E,M) - \lambda C_x(x,E)$ $\qquad$ (22)

According to the Implicit Functions Theorem, if $\Gamma_E(x,E,M,\lambda) \neq 0$ then there exists a function, $\Phi(.)$ such that $\Phi \equiv x(E,M,\lambda)$ and $\Phi_E = \frac{\partial x}{\partial E} = -\frac{\Gamma_x}{\Gamma_E}$ $\qquad$ (23)

$\Gamma_E = U_{xE}(x,E,M) - \lambda C_{xE}(x,E)$ and since $U_{xE} > 0$ and $C_{xE} < 0$ according to (14) and (19)

It follows that $\Gamma_E > 0$ $\qquad$ (24)

In application of the Implicit Functions Theorem: $\Phi_E = \frac{\partial x}{\partial E} = -\frac{\Gamma_x}{\Gamma_E}$ and since

$$\Gamma_x = U_{xx}(x,E,M) - \lambda C_{xx}(x,E) \text{ with } U_{xx} < 0 \text{ and } C_{xx} > 0 \text{ so } \Gamma_x < 0 \qquad (25)$$

It follows that: $\frac{\partial x}{\partial E} > 0$ $\qquad$ (26)

Proposition1: (Interpretation) The more educated an individual, the higher the occupational category which he aims to in the formal sector. This is because higher-ranked occupational categories imply higher labor income, and thus, utility, on average.

The above theoretical result underlies the hypothesis to be tested empirically in the next section:

---

[9] The function $C(x,E) = \frac{ax^2}{bE}$ with $a > 0, b > 0$ is an example of function meeting the assumptions (16) (17) (18) and (19)

[10]Information and Communication Technology

H_0: The probability to work in the formal sector increases with the level of education.

In other words a low level of education increases the probability to work in the informal sector.

Section 3: Empirical Framework

The annual households' standards of livingsurvey2007 group together the non-agricultural workforce in five (5) social and occupational classes[11]: Wage earners of the public sector, wage earners of the private (structured) sector, employers or own-account workers of the informal sector, the salaried of the informal sector and family-aids or apprentices. Since the goal of this paper focuses on informality, we restructure those classes into 3:

    i)   The salaried of formal sector: wage earners of public and private sector

    ii)   The employers and own-account workers of the informal sector: informal self-employment

    iii)   The salaried of informal sector

Family-aids and apprentices who receive no money are combined with non-workers in a group called "the non-paid workers". According to relation (11) x can be defined as:

$$x = \begin{cases} 0 & \text{Non-paid workers} \\ 1 & \text{Salaried of informal sector} \\ 2 & \text{Informal self-employment} \\ 3 & \text{Salaried of formal sector} \end{cases}$$

Thus the factors that take a working age individual into one of these 4 situations can be estimated using a multinomial logit model. The utility associated to class x is assumed to be linear in a vector of observable characteristics of the individual i as follows:

$$U_{ix} = z_{ix}\beta_x + \varepsilon_{ix} \quad i = 1, \dots, N \qquad x = 0, 1, 2, 3 \tag{27}$$

z_{ix} : vector of observable characteristics of the individual i

β_x : vector of parameters to be estimated

ε_{ix} : error term

The characteristics of individual i pertain mainly to his level of education, age, gender, familial background and the earnings derived from his occupational category.

The first issue we are faced with is that the earnings are observed only if the occupational class is chosen. So there is a potential risk of endogeneity bias relative to observed earnings. To cope with that issue we estimate the earnings equation of each category in order to predict the associated expected incomes. Then these expected earnings will be included in a second step as explanatory variables.

The selection bias associated with the estimation of the earnings equation stems from the fact that income-earners are only observed when they work buy not everybody works or receives compensation

---

[11] According to the database used the formal self-employment represents only 0.2% of the workforce; So this negligible group has been excluded of our analysis

when they do. Also wage-earners may have a comparative advantage or some unobserved attributes that pre-select them into formal wage employment or informal jobs (Roy, 1951). To address the selection bias issue we use the Heckman's two-step procedure generalized by Lee (1983) for the multinomial logit case (see in Appendix). At the first stage the multinomial logit model is used to compute the correction terms λ_{ix} from the predicted probabilities for individual i to be in category x. The λ_{ix} (generalized forms of the inverse Mill's ratios) are then introduced into the earnings equation for each professional category, x, as follows:

$$y_x = z_x \beta_x + \theta_x \lambda_{ix} + w_x \tag{28}$$

y_x : logarithm of the monthly income provided by professional category x
z_x: vector of observable elements of human capital (education[12], age) and family environment
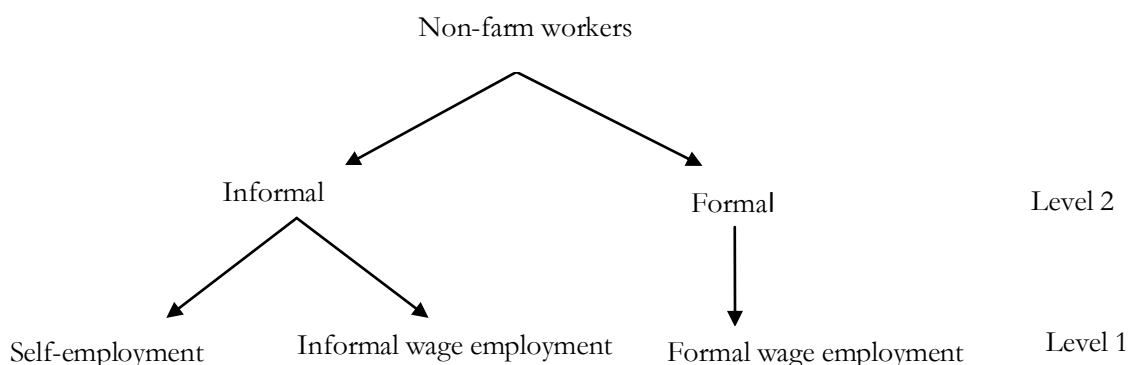w_x : mean-independent of the regressors

θ_x : correlation parameter between the errors terms of earnings and the one of each class

Lee's method has been criticized for relying on a strong assumption for the joint distribution of error terms (Schmertmann, 1994; Dahl, 2002; Bourguignon, Fournier and Gurgand, 2007). However, given the small size of our sub-samples we prefer Lee's method than Dubin and McFadden's or Dahl's method. In fact, based on Monte-Carlo simulations, Bourguignon and al. (2007) conclude that "Lee's method is adapted to very small samples ".

Another potential problem is that the multinomial logit may suffer from the violation of the Independence of Irrelevant Alternatives assumption (IIA). According to Bourguignon, Fournier and Gurgand (2007) "when the focus is on estimating an outcome over selected populations rather than on estimating the selection process itself the multinomial logit is consistent even in case of violation of IIA". But our ultimate objective is to shed light on the factors that select individuals into each occupational class and predicted earnings maybe one of the relevant factors. Based on our interest for the selection process itself we have performed the Hausman test and found evidence that the IIA assumption was violated (see Table 11 in Appendix).

Next, we use a *degenerated nested logit* which relaxes the IIA assumption. Its computation was facilitated by the fact that the different occupational choices in the labor market can be represented by the following tree structure:

---

[12]Several authors have shown that education can be endogenous when estimating an earnings equation and use instrumental variables relative to parental education and occupation, distance to the closest school…are used. Since our analysis already focuses on heads of households it is difficult to find relevant instrumental variables for education.

```
                              Non-farm workers



              Informal                    Formal              Level 2



   Self-employment      Informal wage employment   Formal wage employment    Level 1
```

The above representation allows us to highlight at the first level the relevant determinants of self-employment compared to the informal wage employment. At the second level, the factors affecting the probability to work in the formal sector compared to the informal one will be pointed out.

Level 1: $x = \begin{cases} 1 & \text{Self-employment} \\ 0 & \text{Informal wage employment} \end{cases}$

Level 2: $x = \begin{cases} 1 & \text{Formal sector} \\ 0 & \text{Informal sector} \end{cases}$

A- Definition of variables

We define some binary dependent variables as follows:

Level 1 - the "self-employment" variable (S_empl):

$$S\_empl_i = \begin{cases} 1 & if\,i\,is\,own-account\,worker \\ 0 & if\,i\,is\,salaried\,of\,informal\,sector \end{cases}$$

Level 2 - the "Job" variable:

$$Job_i = \begin{cases} 1 & if\,i\,works\,in\,the\,formal\,sector \\ 0 & if\,i\,works\,in\,the\,informal\,sector \end{cases}$$

The explanatory variables used at the 2nd level of our model are:

i)   Intrinsic characteristics, family background and place of residence

The Gender variable for the household head is equal to 1 when the head is a man and 0 otherwise. The age of the head is stratified in 4 dummies variables: 15-24, 25-34, 35-44 and 45-64. The reference variable is the age bracket 45-64. This stratification allows us to measure the impact of the different generations. The marital status of the head is equal to 1 if married and 0 otherwise. There are also household size and level of education[13]. The level of education is stratified in 4 dummies modalities: No education, 1-6 years

---

[13]Vocational training is not included in the analysis because of the lack of consistent data on this aspect of education

of schooling (Primary school), 7-14 (Secondary school) and more than 14 years (+14, high education). The dummy variable No education is the variable of reference. Finally, the variable "place of residence" takes the value 1 if the household head lives in urban area and 0 otherwise.

ii) Expected Earnings and branch of activities

The expected earnings refer, to the predicted earnings based on the earnings equation of each occupational option. There are noted E_W_{fwe} for formal wage employment; E_W_{iwe} for informal wage employment and E_W_{se} for self-employment. To minimize the risk of colinearity, we use the interactions between the expected earnings and the level of education (see Tables 6 and 8 in Appendix). The branches of activities are represented by the following dummies: Industries, Tradable Services, Non Tradable Services, and Financial Services[14]. The variable of reference is Financial Services.

iii) Access to public services, financing services and social protection

The variable "Possession of fixed phone line" delivered by the National Telecommunication Agency is equal to 1 if the head of the household has a phone line and 0 if not. The other variable is "Access to electricity" which is equal to 1 if the head of household has access to electricity delivered by the National Electricity Agency and 0 otherwise. The variable "Access to drinking water"[15] delivered by the National Water Agency is equal to 1 if the head of the household has access to drinking water and 0 otherwise.

The dummy variable "Access to credit" is equal to 1 if the head of the household has applied and obtained a credit and 0 otherwise. Variables relative to social protection are: the "Right to old-age pension" equals to 1 if the head of the household benefits from this right and 0 otherwise; and the "Right to paid leave" equal to 1 if the head of the household benefits from this right and 0 otherwise.

B- Descriptive statistics

We use the household standards of living annual survey for 2007 in Burkina Faso. The database provides demographic and socioeconomic information on a representative sample of 8,500 households across the country. Our focus on the non-farm workforce reduces the database to 2,513household heads, including 86.3% of men and 13.7% of women.About29.6% of the household heads are wage workers in the formal sector; 10.5% of them are wage workers in the informal sector whereas self-employed represent 59.9%.

---

[14]Industries refer mainly to manufacturing, mining, and building. Tradable Services refer to trade, repairing, and transportation. Non Tradable Services refer to public administration, education and health.
[15]This variable includes the individual line, the shared lines and the public drinking fountains

As shown in Table 1 in the Appendix, men are strongly predominant in all 3 occupational classes whereas women are mostly Self-employed. In fact women represent 18.6% of the self-employed but only 8.5% of formal wage earners and 0.4% of informal wage earners. There are also strong disparities in education. In fact, about 66.3 % of self-employed have no education compared to 35.1 % of informal sector workers and only 7% of formal sector's wage earners.

Workers with a higher level of education (more than 14 years of schooling) represent 23.6% of wage workers in the formal sector, 7.2 % of wage workers in the informal sector and only 1.8 % of self-employed. Those figures suggest that high level of education increases the probability of work in the formal sector. In the informal sector educated workers seem to prefer a salaried status at the expense of self-employment. With regard to the place of residence, wage earners from both the formal and informal sectors are more likely to live in urban areas. About 88.7 % of wage earners of the informal sector live in urban areas. Most of the employers in the informal sector also live in urban areas. The distribution of workers and employers across regions thus seems to be driven by the urban bias of national development strategies as urban areas provide the best business opportunities due to available public infrastructure and consumption goods and services.

The average monthly wage income for formal sector's workers is 81,071.01 FCFA compared to 44,424.53 FCFA in the informal sector and 43,000 FCFA for the self-employed. However, there are strong inequalities within each occupational group. In fact, half of the salaried employees in formal sector earn less than 60,000 FCFA per month compared to 30,000 FCFA for the informal wage earners and 30,500 FCFA for the self-employed.

Regarding access to public services, only 3.7% of self-employed people have a fixed phone line compared to 6.8% of informal wage earners and about 20.1% of the formal wage earners. About 77.4% of the formal wage earners have access to electricity compared to 53.2 % of the informal wage earners and only 31.3% of the informal salaried. The informal salaried are also disadvantaged in terms of access to drinking water. These disparities stem from the fact that almost of the wage earners both formal and informal live in urban areas where there is a better provision of public services.

There are also strong inequalities in access to social protection between the different occupational categories. In fact, about 91% of the formal wage earners benefit from an old-age pension opposed to7.5% of informal salaried and 2.3% of the self-employed. Likewise, only 2.3% of the latter can claim paid leaves against 19.2% in the informal sector and 93.9% of the salaried in the formal sector.

Regarding access to financial services, in 200741% of the formal wage earners have applied for a credit compared to 7.5% of salaried from the informal sector and 13% of the self-employed. The data show that almost all applicants received were granted a loan regardless of their occupational class.

Section 4: Results and discussion

A- The earnings equation

Abstracting away from Lee's selection term in the estimation equation, we find that the level of education, the size of the household and urban residency positively and significantly affect (at 1% level of significance) the earning profile of the self-employed. Age and marital status have no significant effect whereas the gender variable is significant (at 10%).

With the selection term we find the same relevant variables but with higher returns. In fact, the return to primary education (1-6 years of schooling) for self-employed is 13.15% compared to 38.48% for the secondary education and 99.18% for higher education. This suggests a convex return to education for self-employed people. The results are in line with the conclusions of Rankin, Sandefur and Teal (2010) for Ghana and Tanzania. Kuepie and al (2009) and NguetseTegoum (2009) find similar results for 6 capitals of WAEMU and Cameroon respectively.

Moreover, there is evidence to suggest that gender has an impact on self-employment income. Being a man increases earnings by 11.5% as compared to being a woman. Clearly, men earn more out of self-employment relative to women. In fact, according to the results above, the level of education increases the earnings in the self-employment category. And yet the education attainment of women is low as opposed to men. So this difference of qualifications may explain the earnings gap between men and women. Moreover the women are confined in less productive activities (retail, catering) whereas men are more likely to work in the high productive one (manufacturing). Likewise, living in urban area enhances the earning profile by a factor of 44.9% compared to rural area. The coefficient on the correction-selection terms in the equation is negative (but not significant). This result suggests that the unobservable characteristics that select people into self-employment do not impact their earning profile.

With regard to the earning equation for wage employment in the informal sector, the Lee selection term is positive but only significant at 10%.That suggests those selected into wage employment in the informal sector have some unobservable characteristics that make them more suitable for this employment status compared to the other workers. Those characteristics can be related to their low reservation wages or attributes like resourcefulness. In fact, it has been shown in the descriptive statistics that informal wage earners are relatively more educated than the self-employed. According to the same statistics the formal sector provides the highest earnings. So one can assume that the informal wage earners after failing to get a job in the formal sector, accept lower wages in the informal sector.

Age, household size, marital status and primary school education are not significant determinants of earning profiles. In contrast, the return to 7-14 years of schooling is about 39.52% and 76.01% for higher education. Thus there is also some evidence that returns to education are convex for informal wage

employment. Rankin, Sandefur and Teal (2010) find similar results for Ghana although their results for Tanzania differ. Furthermore, there is some evidence to suggest that gender matters in this sub-sector as well, since being a man increases the earnings to 190% compared to being a woman[16]. This result suggests that even in the informal sector men are still more likely to land an offer than women. The main reason of this finding may be the weakness of women's education relatively to men. Also, living in urban area enhances the incomes to 71.41% compared to urban area. In fact the urban area contrary to rural area provides better infrastructure and a higher consumption demand that stimulates the production. So more productive and more competitive the urban informal sector may provide higher ages than the rural one.

Without Lee selection term the returns to education are convex for wage income in the formal sector. Except for the household size which is significant at 10% all other variables are insignificant.
However, including the correction-selection term in the estimation equation suggests negative and significant (at 1%) returns to education, regardless of the level of education. The unobservable characteristics that select people into the formal sector seem to also affect their incomes negatively in that sector. These results align with Rankin,Sandefur and Teal (2010) for Tanzania, but contrast with Kuepie and al (2009) and NguetseTegoum (2009).

Reason for the above results may be that in Burkina Faso the public sector is the main entry to formal sector employment. The public sector provides jobs far more than the structured private sector. Entry to the public sector is subject to the successful completion of national exams depending on their level of education. With scarce job openings it is common for postgraduate to sit for exams requiring an undergraduate or secondary school diploma. Thus in the public sector it is usual to see high educated individuals working at positions requiring less education.

Besides education, gender has a significant and negative impact on earnings at 5%. Results suggest that women are discriminated against in the formal sector. Household size is a positive and significant (at 10%), and thus appears as a determinant of wage employment in the formal sector. Formal sector wage earners living in urban areas are likely to earn more than the ones living in rural areas.

We use the 3 estimated earnings equations above to predict the expected incomes of each paid workers in each occupational category. Then these expected incomes are used as explanatory variables in the nested logit.

---

[16] The findings may suffer from some bias due to the negligible number of women in the sub-ample of informal wage employment (only 0.4%)

B- Level 1 : the determinants of informal self-employment

As it summarized in Table 5 in Appendix the main determinant of self-employment is education. In fact the higher education and primary school education variables are significant at 5% whereas the secondary school education is significant at 1%. All the coefficients for education levels are negative. Thus, the probability of self-employment (relative to wage employment) in the informal sector decreases with the level of education. For instance being high educated reduces by 5.73% the probability to be self-employed whereas the secondary level decreases it by 7.18% (see Table 9 in Appendix). This result differs slightly from Rankin and al (2010) where a multinomial logit is used to show that the probability of either self-employment or employment in a small firm (in the informal sector) decreases with education. However, these authors do not compare between the two types of employment in the informal sector, thus the value-added of a nested logit. In fact, the first level of our model allows us to show that in the informal sector the level of education increases the probability of wage employment.

Concerning the gender, the men are less likely to be self-employed than women. In fact, being a man reduces by 11.68% the probability to work in self-employment category. So we confirm results obtained by Do and Duchene (2008) in the case of Vietnam and Kabubo-Mariara (2003) in the case of Kenya. Despite the weakness of the expected earnings the women are constrained to work in the informal sector as self-employed.

Concerning the household's age all age brackets are not significant. Regarding family background, while the size of the household is not significant, married individuals are more likely to be self-employed compared to single people. The place of residence is also a relevant for informal self-employment. In fact living in urban area reduces the probability of self-employment relatively to rural area where small activities of trade, manufacturing, repairing are increasingly used to cope with the uncertainty of farming and to diversify income sources.

With regard to the branch of activities, being in non-tradable services and industries increases the likelihood of self-employment. Likewise, access to credit enhances the probability to be self-employed as opposed to informal wage employment. In fact undertaking a business requires a minimum of capital. Access to credit may surely stimulate entrepreneurship. Also, while access to drinking water reduces the probability to be self-employed, access to electricity is not significant. Finally, individuals who are entitled to paid leave are more likely to be wage workers than self-employers[17].

In order to highlight the effect of expected earnings relative to self-employment (by minimizing the risk of colinearity) we compute again the first level of the nested logit after having replaced the education

---

[17]This result may not be useful since self-employers basically pay themselves, and so cannot claim paid leaves.

dummies by the cross education and these expected earnings. Table 9 (in Appendix) shows that the expected earnings associated to each level of education reduce significantly the probability to be self-employed. For instance, earnings expected by the high educated reduces this probability by 0.03% whereas the earnings associated to secondary level reduce it by 0.05%. So regardless to the level of education self-employment does not provide attractive expected incomes.

C- Level 2: The determinants of formal employment as opposed to informal employment

As shown in Table6 (Appendix) the level of education increases the probability of employment in the formal sector. The higher education and secondary school variables are significant at 1% whereas the primary school variable is significant at 5%. Indeed, high education attainment reduces by 47.95% the probability to work in the informal sector where as the secondary level decreases it by 34.51% (Table 9 in Appendix). The probability to work in the informal sector is also reduced by only 8% when moving from no education to the primary level. These results lend support to our theoretical finding from Section 2. Kuepie et al, (2009), Rankin and al (2010), and NguetseTegoum (2009) find similar results.

Contrary to place of residence, age bracket 25-34, marital status, household's size and, gender are significant. In fact, individuals between 25 and 34 years old are less likely to work in the formal sector than those in the range of 45-64 years old. This can be explained by the low education attainment of the young. Likewise men are more likely to work in the formal sector than women. This result may also find his explanation to the weakness education attainment of women as opposed to men. Even if the national competition to enter the public sector is indifferently opened to both men and women, there are fewer women, who meet the minimum qualification requirements.

Regarding the branch of activities only ``tradable services" (trade, repairing) significantly reduce the probability to work in the formal sector. While firms involved in international trade are required to be registered in order to do their business, the one devoted to the domestic trade can escape to this requirement. And yet as the 1-2-3 survey has shown the informal units are domestic market-oriented. They generally subcontract with the international traders to insure the retail inside the country. But the retailers are subject to commercial card different from the fiscal identification number. Then retailers with commercial card are considered are informal according to the definition of informal sector used in this paper. The variables measuring the access to public services and credit have no impact on the propensity to work in the formal sector whereas the variables relative to social protection are significant at 1%. In fact, individuals who benefit from old-age pension and/or are entitled to paid leave are less likely to work in the informal sector.

Finally the cross expected earnings-education effect on the probability to work in the formal sector is positive and significant regardless the level of education (see Table 8 in Appendix). Indeed the formal expected earnings associated to the high education attainment increase by 15.69% the probability to work

in this sector whereas the secondary level of education enhances it by 16.21% (see Table 9 in Appendix). But the earnings expected by individuals with primary school level in the formal sector increase only by 0.05% the probability to work in this sector. These finding demonstrate that formal sector gives more value to education.

CONCLUSION

The main goal of this paper was to revisit the factors that select individuals into the informal sector of the economy in Burkina while investigating the drivers of occupational choice between self-employment and wage employment. The paper first developed a theoretical framework to shed light on the potential role of education in agents' decision-making process. Ranking occupational options from low-paid and precarious to high-paid with benefit, we find that the more educated an individual, the higher the occupational category which he aims to. We then take this prediction to the data, using annual households' standards of living survey 2007 for Burkina Faso. Empirical results provide evidence for educated preference for wage employment (whether in the formal or in the informal sector). They also confirm significant gender imbalances as women are less likely than men to get a job (whether in the formal or in the informal sector). Empirical findings also show the important role of individual characteristics like family background and urban / rural residential status as well as institutional factors, including access to credit, to public infrastructures, safe water and electricity.

Table 1: Descriptive statistics of independant variables

Independant Variables	Formal wage employment	Informal wage employment	Self-employment
Intrinsic and familial			
Man	91.5%	99.6%	81.4%
Woman	8.5%	0.4%	18.6%
Mean Age	38.81%	36.16	38.53
Mean household size	4.37%	3.81	4.77
Single	17.8%	25.3 %	12.5%
Married	82.3%	74.7 %	87.5%
Education			
No education	7%	35.1 %	66.3 %
1-6 years of schooling	8.2%	23.4 %	20.5 %
7-14 years of schooling	61.2%	34.4 %	11.4 %
+14 years of schooling	23.6%	7.2 %	1.8 %
Place of residence			
Urban	79.3 %	88.7 %	54.2 %
Rural	20.7 %	11.3 %	45.8 %
Branchs			
Industries	7.5 %	29.4 %	30.4 %
Tradable Services	8.2 %	40.4 %	57.1 %
Non-tradable Services	76.4 %	21.1 %	8.6 %
Financial services	7.8 %	9.1 %	3.9 %
Access to public services			
Fixed phone line (yes)	20.1 %	6.8 %	3.7 %
Access to electricity	77.4 %	53.2 %	31.3 %
Access to drinking water	81.9 %	84.2 %	53.6 %
Access to financial services			
Applied for a credit (yes)			
Positive answer			
	41 %	7.5 %	13 %
	39.7 %	7.2 %	10.6 %
Social Protection			
Old-age pension (yes**)**	91%	7.5 %	2.3 %
Paid leave	93.9 %	19.2%	2.3 %

Source: computed by the author from the Annual Survey on the household standard of living conditions/ 2007 database

Table 2: Earnings equation with Lee's correction term in informal self-employment

Log (earnings)	Coef	z
Individual and familial characteristics		
Age	0.0201573	1.46
Age²	-0.0002263	-1.36
Gender	0.1149578	**2.11****
Marital status	-0.0997227	-1.27
Household size	0.0726237	**8.19*****
Education		
1-6 years of schooling	0.1315545	**2.79*****
7-14 years of schooling	0.3848081	**2.69*****
+14 years of schooling	0.9918264	**3.34*****
Place of residence	0.4490648	-1.27***
λ_{isempl}	-0.0813299	-0.48
constant	9.230063	30.83***

Number of obs = 1505
Wald chi2(10) = 291.03
Prob> chi2 =0.0000
R-squared = 0.1570
Adj R-squared = 0.1514

Significant at: *** 1%;** 5%;* 10%. Standard errors are bootstrapped with 100 replications

Table 3: Earnings equation with Lee's correction term in informal wage employment

Log (earnings)	Coef	z
Individual and familial characteristics		
Age	-.0126741	-0.38
Age²	0.0002804	0.68
Gender	1.905854	**1.94***
Marital status	0.2023358	1.63
Household size	0.014026	0.53
Education		
1-6 years of schooling	0.085138	0.84
7-14 years of schooling	0.3951843	**4.38***
+14 years of schooling	0.7601822	**3.78***
Place of residence	0.7141545	**2.63***
λ_{iwempl}	2 .5474412	**1.81***
Constant	6.653549	4.37***

Number of obs = 265
Wald chi2(10) = 291.03
Prob> chi2 =0.0000
R-squared = 0.1570
Adj R-squared = 0.1514

Significant at: *** 1%;** 5%;* 10%. Standard errors are bootstrapped with 100 replications

Table 4: Earnings equation with Lee's correction term in formal wage employment

Log (earnings)	Coef	z
Individual and familial characteristics		
Age	-0.0197704	-0.72
Age²	0.0002835	0.85
Gender	-0.1995532	**-2.12****
Marital status	-0.0715734	-0.84
Household size	0.0287195	**1.87***
Education		
1-6 years of schooling	-0.5503897	**-2.37****
7-14 years of schooling	-1.045977	**-3.35*****
+14 years of schooling	-0.1995532	**-2.12****
Place of residence	0.1479463	**1.97****
λ_{fwempl}	-0.8921949	**-5.59*****
constant	12.91301	**15.91*****

Number of obs = 743
Wald chi2(10) = 198.70
Prob> chi2 =0.0000
R-squared = 0.1968
Adj R-squared = 0.1859

Significant at: *** 1%;** 5%;* 10%. Standard errors are bootstrapped with 100 replications

Table 5: The determinants of self-employment (compared to informal wage employment)

Self-employment	Coef	z
Individual and familial characteristics		
15-24 yearsold	0.0075017	0.02
25-34yearsold	-0.1105696	-0.47
35-44 yearsold	-0.0596904	-0.27
45-64 yearsold (ref)		
Gender	-4.522155	**-4.23*****
Marital status	0.77483	**3.48*****
Household size	0.0305217	0.84
Education		
No education (ref)		
1-6 years of schooling	-0.3813448	**-1.86***
7-14 years of schooling	-0.9259177	**-4.36*****
+14 years of schooling	-0.7310485	**-2.10****
Place of residence	-1.314027	**-4.10*****
Branchs of activities		
Industries	-4.10	**2.30****
Tradable Services	1.237836	**4.15*****
Non- tradable Services	0.2941083	0.85
Financial services (ref)		
Access to credit	0.8456315	**2.49****
Access to public services		
Fixed phone line	0.0813634	0.24
Access to electricity	0.2349714	1.30
Access to drinking water	-0.5010378	**-1.71***
Social protection		
Old-age pension	-0.6012056	-1.34
Paid leave	-1.931864	**-6.90*****
Constant	6.112622	**5.34*****
Number of obs = 1770		
Wald chi2(19) =239.80		
Prob> chi2 =0.0000		
Pseudo R² =0.2682		
Log pseudo likelihood = -546.87291		

*** 1%;** 5%;* 10%.

Table 6: The determinants of self-employment (compared to informal wage employment). The cross effect of earnings and education

Self-employment	Coef	z
Individual and familial characteristics		
15-24 years old	0.0066268	0.02
25-34years old	-0.1102132	-0.47
35-44 years old	-0.0580791	-0.27
45-64 years old (ref)		
Gender	-4.52077	**-4.23***
Marital status	0.7744339	**3.48***
Household size	0.0329466	0.90
Education		
No education (ref)		
α_{Prim}	-0.035634	**-1.81***
α_{Sec}	-0.086631	**-4.34***
α_{Sup}	-0.0642332	**-2.07****
Place of residence	-1.305182	**-4.06***
Branchs of activities		
Industries	0.7133952	**2.29****
Tradable Services	1.237418	**4.14***
Non- tradable Services	0.2930967	0.85
Financial services (ref)		
Access to credit	0.8441839	**2.48****
Access to public services		
Fixed phone line	0.079735	0.23
Access to electricity	0.2333202	1.29
Access to drinking water	-0.5017851	**-1.71***
Social protection		
Old-age pension	-0.6013111	-1.34
Paid leave	-1.930875	**-6.90***
Constant	6.091446	**5.33***

Number of obs = 1770
Wald chi2(19) =240.08
Prob> chi2 =0.0000
Pseudo R² =0.2681
Log pseudo likelihood = -546.95086

$\alpha_{Prim} = EW_{se} \times Primary\ school\ dummy$
$\alpha_{Sec} = EW_{se} \times Secondary\ school\ dummy$
$\alpha_{Sup} = EW_{se} \times Superior\ school\ dummy$

Table 7: The determinants of formal sector compared to informal sector

Job	Coef	z
Individual and familial characteristics		
15-24 years old	-0.3950061	-0.56
25-34 years old	-0.8804565	**-2.27****
35-44 years old	-0.5465154	-1.49
45-64 years old (ref)		
Gender	2.820918	**1.76***
Marital status	-0.6633913	**-1.78***
Household size	0.1196081	**2.57****
Education		
No education (ref)		
1-6 years of schooling	1.039589	**2.46****
7-14 years of schooling	3.112254	**7.94*****
+14 years of schooling	3.130433	**6.16*****
Place of residence	0.1210718	0.25
Branchs of activities		
Industries	-0.7271668	-1.38
Tradable Services	-1.135806	**-2.17****
Non- tradable Services	0.7501837	1.57
Financial services (ref)		
Access to credit	-0.2577847	-0.77
Access to public services		
Fixed phone line	-0.2212572	-0.55
Access to electricity	1.132036	**3.40*****
Access to drinking water	-0.1324575	**-0.31**
Social protection		
Old-age pension	3.877862	**8.89*****
Paid leave	2.905785	**7.07*****
IV_{formal_sector}	0.7150867	**1.99****
$IV_{informal\_sector}$	0.2071439	0.72
Constant	-8.410456	**-3.55*****

Number of obs = 2513
Wald chi2(21) =647.89
Prob> chi2 =0.0000
Pseudo R² =0.8274
Log pseudo likelihood = -263.3733

*** 1%;** 5%;* 10%.
IV_{formal_sector}:inclusive value of the degenerated nest formal sector
$IV_{informal\_sector}$: inclusive value of the nest informal sector
The both inclusive values are in the interval [0-1] which reinforces the consistency of the nested logit.

Table 8: The determinants of formal sector compared to informal sector. The cross effect of earnings and education

Job	Coef	z
Individual and familial characteristics		
15-24 years old	-0.2816919	-0.40
25-34 years old	-0.8476615	**-2.19****
35-44 years old	-0.5301698	-1.45
45-64 years old (ref)		
Gender	2.906264	**1.82***
Marital status	-0.6886988	**-1.84***
Household size	0.1141519	**2.41****
Education		
No education (ref)		
ω_{Prim}	0.097601	**2.45****
ω_{Sec}	0.2858879	**8.00*****
ω_{Sup}	0.2767164	**6.20*****
Place of residence	0.1619622	0.34
Branchs of activities		
Industries	-0.7238285	-1.37
Tradable Services	-1.136047	**-2.16****
Non- tradable Services	0.7524047	1.57
Financial services (ref)		
Access to credit	-0.2593175	-0.77
Access to public services		
Fixed phone line	-0.2488794	-0.62
Access to electricity	1.061853	**3.18*****
Access to drinking water	-0.121266	-0.28
Social protection		
Old-age pension	3.888026	**8.91*****
Paid leave	2.914325	**7.07*****
IV$_{formal\_sector}$	0.728164	**2.03**
IV$_{informal\_sector}$	0.2236562	0.78
Constant	-8.498611	**-3.59*****

Number of obs = 2513
Wald chi2(21) =649.14
Prob> chi2 =0.0000
Pseudo R² =0.8275
Log pseudo likelihood = -263.12171

Significant at:*** 1%;** 5%;* 10%.
IV$_{formal\_sector}$:inclusive value of the degenerated nest formal sector
IV$_{informal\_sector}$: inclusive value of the nest informal sector
The both inclusive values are in the interval [0-1] which reinforces the consistency of the nested logit
$\omega_{Prim} = EW_{fwe} \times Primary school\ dummy$
$\omega_{sec} = EW_{fwe} \times Secondary\ schoo\ l dummy$
$\omega_{sup} = EW_{fwe} \times Superior\ school dummy$

Table 9: Marginal effect of relevant variables on the probability to be self-employed

Self-employment	dy/dx
Gender	-0.1168021
Marital status	0.0573744
1-6 years of schooling	-0.0242968
7-14 years of schooling	-0.0718646
+14 years of schooling	-0.0573309
α_{Prim}	-0.0020558
α_{Sec}	-0.0049978
α_{Sup}	-0.0037057
Place of residence	-0.0713025
Industries	0.0368384
Tradable Services	0.0779689
Acces to credit	0.0368202
Access to drinking water	-0.0280244
Paid leave	-0.2351149

Table 10: Marginal effect of relevant variables on the probability to work in formal sector

Self-employment	dy/dx
Gender	0.0802439
Marital status	-0.0460994
Household size	0.0067422
1-6 years of schooling	0.0804775
7-14 years of schooling	0.3451762
+14 years of schooling	0.4795211
ω_{Prim}	0.005536
ω_{Sec}	0.0162157
ω_{Sup}	0.0156954
Industries	-0.0350838
Tradable Services	-0.0605631
Access to electricity	0.0678562
Old-age pension	0.4791918
Paid leave	0.2950685

Table 11: Hausman test on IIA assumption

	(b)		(B)	(b-B)	sqrt(diag(V_b-V_B))
15-24 years old	.4322383		.4716627	-.0394244	.111767
25-34 years old	1.776191		.8152304	.9609607	.8372462
35-44 years old	1.331011	.	.6109859	.7200247	.6831695
Gender	1.262434		2.685301	-1.422867	2.941679
Marital status	-.8276009		-1.322147	.4945459	.9314985
Household size	-.073851		.045981	-.1198319	.1112347
1-6 years of	-.3814543		.5806767	-.962131	.7199598
7-14 years of	-2.373336		.8164566	-3.189793	.5717697
+14 years of	-2.363074		.8467596	-3.209833	.9580352
Residence	1.299506		1.615039	-.3155334	1.976996
Credit	-3.116049		-1.868684	-1.247365	2.876012
Industries	.1338596		-.6783548	.8122144	.4524685
Tradable Services	.2373467		-1.160489	1.397836	.4288987
Non- tradable	-1.007268		-.2807054	-.7265631	.3516044
Fixed phone line	.131557		-.0427554	.1743124	.3792311
Access to	-2.012692		-.4823039	-1.530388	1.329217
Access to drinking	-2.012692		-.4823039	-1.530388	1.329217
Old-age pension	-3.733368		-.3473322	-3.386036	-3.246336
Paid leave	-1.313388	.	2.091549	-3.404937	.2505696
Constant	1.360499		-5.012287	6.372787	5.321032

Test: Ho: difference in coefficients not systematic
chi2 (21) = (b-B)'[(V_b-V_B)^(-1)](b-B)
=   177.10
Prob>chi2 =   0.0000

Lee's model

Consider the following model:

$$y_s = x\beta_s + u_s \tag{1}$$

$$y_j^* = z\gamma_j + \eta_j \text{where } j = 1,2,...M \tag{2}$$

j is a categorical variable that describes the choice of an economic agent among M alternatives based on utilities y_j^*.

The outcome variable y_s is observed if and only if category s is chosen which happens:

$$y_s^* > \max_{s \neq j}(y_j^*) \tag{3}$$

Define:

$$\varepsilon_s = \max_{s \neq j}(y_s^* - y_j^*) = \max_{s \neq j}(z\gamma_j + \eta_j - z\gamma_s - \eta_s) \tag{4}$$

Then under this definition, condition (3) is equivalent to $\varepsilon_s < 0$

The problem is to estimate the parameter vector β_s while taking into account that the disturbance term u_s may not be independent of all η_j. This would introduce some correlation between the explanatory variables and the disturbance term in the outcome equation of model (1). Because of this, least squares estimates of β_s would not be consistent.

Define $\Gamma = (z\gamma_1, z\gamma_2, ... z\gamma_M)$ then

$$E(u_s|\varepsilon_s < 0, \Gamma) = \lambda(\Gamma) = \mu(P_1, P_2, ... P_M) \tag{5}$$

Given that the relation between the M components of Γ and the M corresponding probabilities is invertible, there is a unique function μ that can be substituted for λ.

Therefore, consistent estimation of β_s can be based on the following regression:

$$y_s = x\beta_s + \mu(P_1, P_2, ... P_M) + w_s \tag{6}$$

Where w_s is a residual that is mean-independent of the regressors

In a widely quoted paper, Lee (1983) proposed a generalization of the two-step selection bias correction method introduced by Heckman (1979) that allows for any parameterized error distribution. His method extends to the case where selectivity is modelled as a multinomial logit. This approach is simple and requires the estimation of only one parameter in the correction term. This is, however, achieved at the cost of fairly restrictive assumptions.

Call $F_{\varepsilon_s}(.|\Gamma)$ the cumulative distribution of ε_s. The cumulative $J_{\varepsilon_s}(.|\Gamma)$. Defined by the following transform $J_{\varepsilon_s}(.|\Gamma) = \Phi^{-1}(F_{\varepsilon_s}(.|\Gamma))$ where Φ the standard normal is cumulative, has a standard normal distribution. Lee's model can be written as implying that:

$E(u_s|\varepsilon_s, \Gamma) = \sigma\rho_s J_{\varepsilon_s}(\varepsilon_s|\Gamma)$ and yields an estimating equation of the form:

$$y_s = x\beta_s - \sigma\rho_s \frac{\phi(J_{\varepsilon_s}(0|\Gamma))}{F_{\varepsilon_s}(0|\Gamma)} + w_s \tag{7}$$

BIBLIOGRAPHY

Andrews D, Caldera S. A et Johansson A. (2011) Towards a better understanding of the informal economy. *Economics department* working papers no.873 OECD

ILO. (1972) Employment, Incomes and Equality. A Strategy for Increasing Productive Employment in Kenya, *ILO, Geneva*

Benjamin N, Mbaye A. et al. (2012) The Informal Sector in Francophone Africa: Firm Size, Productivity, and Institutions. *World Bank*

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

Cogneau, D. (2001) Formation du revenu, segmentation et discrimination sur le marché du travail d'une ville en développement : Antananarivo fin de siècle ». Document de Travail, 2001/18. DIAL

Combarnous F (1999) La mise en oeuvre du modèle logistique multinomial emboîté dans l'analyse de la participation au marché du travail, Centre d'économie du développement, Université Montesquieu-Bordeaux IV-France.

Destré G. et Henrard.V. (2004) The determinants of occupational choice in Colombia: an empirical analysis. Cahiers de la Maison des Sciences Economiques, Maison des Sciences Economiques, Université Paris 1 Panthéon-Sorbonne

Donor Committee for Enterprise Development, (2008) Supporting Business Environment Reforms: Practical Guidance for Development Agencies, report from Donor Committee for Enterprise Development, August.

El Aynaoui J (1997). Participation, choix occupationnel et gains sur un marché du travail segmenté : une analyse appliquée au cas du Maroc. Centre d'économie du développement, Université Montesquieu-Bordeaux IV-France.

Hammouda N., Lassassi M. (2009) Déterminants de la participation au marché du travail et du choix occupationnel : une analyse micro économétrique appliquée au cas de l'Algérie », GDRI DREEM Working paper.

Hart K. (1972) Informal Income Opportunities and Urban Employment in Ghana, *Journal of Modern African Studies, vol. II*

INSD. (2003b) Le secteur informel dans l'agglomération de Ouagadougou : performances, insertion, perspectives .Premiers résultats de la phase 2 de l'enquête 1-2-3. Version provisoire. Institut National de la Statistique et de la Démographie. Ouagadougou; Burkina Faso.

Johansson E. (2000) "Self-employment and the predicted earnings differential-evidence from Finland. *Finnish Economic Paper, 13:45_55*.

Jütting, J. J. Parlevliet, et T. Xenogiani (2008). Informal Employment Re-loaded. *OECD Development Centre Working Paper* No. 266, Paris: OECD, January

Kabubo-Mariara (2003) Wage determination and the gender wage gap in Kenya: Any evidence of gender discrimination? AERC Research Paper 132 African Economic Research Consortium, Nairobi May 2003

Kuepie, Nordman et Roubaud (2009) Education and earnings in urban West Africa. *Journal of Comparative Economics* 37 (3) (2009) 491–515. IRD, DIAL, France; CEPS/INSTEAD, Luxembourg

Lachaud, J.-P. (1997) Pauvreté, vulnérabilité et marché du travail au Burkina Faso. Série de recherche 2. Centre d'Economie du Développement. Université Montesquieu Bordeaux IV.

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

Loayza, N., A.M. Oviedo et L. Serven (2006).The Impact of Regulation on Growth and Informality Cross-Country Evidence, In B. Guha-Khasnobis, R. Kanbur, and E. Ostrom (Eds.), Linking the Formal and Informal Economy, Concepts and Policies, Oxford: Oxford University Press

Millan, J. M., Congregado, E. et Roman, C. Determinants of Self-Employment Dynamics and their Implications on Entrepreneurial Policy Effectiveness. *Lect. Econ.* [online]. 2010, n.72, pp. 45-76. ISSN 0120-2596

NguetséTegoum (2009) Estimating the returns to education in Cameroon Informal sector. Cameroon Ministry of Economy, Planning and Regional Development (MINEPAT, Yaoundé)

Ouédraogo M.I. (1996) The Urban Informal Sector in Burkina Faso. Thèse de Ph.D University of Tennessee, Knoxville, U.S.A.

Perry, G. (2007).Informality: Exit and Exclusion, World Bank Latin American and Caribbean Studies

Rankin, Sandefur et Teal (2010) Learning &Earning in Africa: where are the returns to education high? CSAE WPS/2010-02

Rees H. et Shah A. (1986) An empirical analysis of self-employment in the UK. *Journal of Applied Econometrics*, 1:101_108

Schneider, F. et D.H. Enste, (2000) Shadow Economies: Size, Causes, and Consequences. *Journal of Economic Literature,* Vol. 38.

Yamada G. (1996) Urban informal employment and self-employment in developing countries: theory and evidence, *Economic Development and Cultural Change*, 44, 289-314.

Zerbo, A. (2006b) Marché du travail et pauvreté en Afrique Subsaharienne : un modèle d'analyse . Document de Travail n°129 ; Centre d'économie du développement, Institut fédératif de recherche sur les dynamiques économiques, Université Montesquieu Bordeaux IV.

Zerbo, A. (2001). Les économies urbaines subsahariennes et le problème de la mobilisation de ressources via le secteur informel ; Document de Travail n°62, Centre d'économie de développement, Université Montesquieu - Bordeaux IV.