# An Estimated Model of Informality with Constrained Entrepreneurship*

Pierre E. Nguimkeu†

Georgia State University

February 2015

**Abstract**

This paper provides an econometric analysis of the role of individual and institutional factors on entrepreneurship and informality in developing countries. Emphasizing the link between micro-foundations and the data, I present a simple model of occupational choice in which heterogeneous agents who differ with their skills and initial wealth endowment choose between formal entrepreneurship, informal entrepreneurship and non-entrepreneurial work. Using reduced-form estimation techniques and specification analysis I test the validity of the model and its main implications with data from Cameroon, an economy where 90% of the labor force works in the informal sector and 31% of the gross national product comes from informal activities. I also provide structural estimates of the theoretical model which explicitly accounts for the presence of registration sunk costs, imperfect credit markets, and limited law enforcement. The estimated model is then used in counterfactual policy simulations to quantify the impact of enterprise registration reforms, taxation and enforcement policies on the size of informality, the allocation of skills across occupations and aggregate income.

*Keywords:* Entrepreneurship, Informality, Regulation, Maximum likelihood, Counterfactual analysis.

*JEL Classification:* O12, O17, C51, C52, C54.

*I thank conference participants of the World Economic Congress 2014 and the Southern Economic Association meeting 2014

†Department of Economics, Andrew Young School of Policy Studies, Georgia State University, 14 Marietta Street NW, Suite 524, Atlanta, GA 30303, USA; Email: nnguimkeu@gsu.edu.

# 1 Introduction

It is difficult to overemphasize the importance of entrepreneurial activity for both developed and developing economies. They are universally considered a crucial ingredient in promoting and sustaining economic growth because of their potential for creating jobs, delivering innovation and raising productivity. In the context of developing countries the existence of a large informal sector (60-90% of the overall workforce) that coexists with the traditional formal sector brings additional challenges to policy-makers' objective to understand and promote entrepreneurship. On the one hand, formal enterprises have higher prospects to grow and a higher productivity and income potential (60% of GDP in Africa, Schneider and Enste 2000), and they generate tax revenues that can be used by the state to improve the provision of public goods and services. On the other hand, informal entrepreneurship is generally low-scale, is largely untaxed and creates important shortfalls in government revenues. But at the same time it represents an important avenue of job creation (more than 90% of jobs created in Africa between 1990-1994, Kuchta-Helbling 2000), and is considered an incubator for business potential and a stepping stone for accessibility and graduation to the formal economy (ILO, 2002).

This paper contributes to the existing literature that endeavours to understand the connections between entrepreneurship, informality and economic growth by providing some answers to the following key questions: (i) What are the drivers of entrepreneurship and enterprise performance in developing countries and their relationship to formality? (ii) What are the relevant policies and programs that could be implemented to promote entrepreneurial growth and improve output while harnessing informality? To answer these questions, I formulate a model of occupational choice where heterogeneous agents choose to become workers, informal entrepreneurs or formal entrepreneurs, and their decision-making process depends on both personal and institutional characteristics. I test the implications of the model using reduced form techniques and provide structural estimates using data from Cameroon, an economy where 90% of the labor force operates in the informal sector. The estimated model is then used in counterfactual policy simulations to examine and quantify the effects of several policies including registration reforms, taxation, enforcement, and business training on entrepreneurship, informality and aggregate income.

The model discussed and estimated is an amended version of several models that have appeared in the literature (e.g. Amaral and Quentin 2006, Antunes 2007, 2008, Prado 2011, 2014, Ordoñez 2014). As discussed below, I present an extension that improves in fitting the context of many developing countries as evidenced by recent research work and survey data (e.g. *Doing Business* 2005, Djankov et al. 2002, Auriol and Walters 2005). Specifically, the model economy assumes an institutional environment that accounts for the presence of registration sunk costs, imperfect credit markets and a tax collection policy with limited enforcement. Individuals differ in this economy by their entrepreneurial skills and their initial wealth endowment, and each face a discrete occupational choice:

whether to be a formal entrepreneur, an informal entrepreneur or a non-entrepreneurial worker. A worker receives a fixed income while an entrepreneur establishes a firm with capital investment and hired labor and realizes profits from a decreasing-returns-to scale technology. To become formal, an entrepreneur needs to pay a registration fee and once formal the entrepreneur also pays taxes but enjoys a better access to credit. In contrast, informal entrepreneurs do not pay taxes but are more likely to face borrowing constraints while also facing a probability of being caught and getting their profit forfeited. The main implications of the model are as follows. First, entrepreneurship is positively related to skills and initial wealth, which is used as collateral to obtain loans from financial institutions. Second, entrepreneurs with low productivity choose informality whereas the most productive ones choose to operate in the formal sector. In fact, since formality is costly through both taxation and high registration costs, only entrepreneurs whose firms promise a sufficiently high return and prospects to grow are willing to afford it. Because low productivity entrepreneurs naturally choose lower scale, the high costs of registration acts as an implicit exclusion mechanism to enterprises with low productivity. This induces a low entry decision of entrepreneurs with low ability, which in turn increases the number of unproductive enterprises in the informal sector, even when the tax rate is reasonably small.

To structurally estimate the model, I build a likelihood function by matching the expected probability of each occupation generated by the theoretical model with the corresponding occupational status observed in the data. I also test the model predictions using reduced-form discrete choice estimation (probit regressions) that make use of a large set of explanatory variables. This allows to check the consistency of the structural findings as well as the role of several other controls that are not explicitly part of the structural model. Evidence is provided using a cross-sectional sample of an original dataset from Cameroon. These data stem from the 2005 National Survey on Employment and the Informal Sector (EESI) and gather a comprehensive set of information on households characteristics as well as the characteristics of their economic activities.[1]The results confirm that observable talent such as education is a key determinant of entrepreneurial choice. There is however a U-shaped relationship between entrepreneurs' education and their decision to formalize. In other words, less educated entrepreneurs find it more profitable to remain informal as their level of education increases. It is only above a certain educational threshold that formality becomes increasingly attractive to informal entrepreneurs with increasing levels of education. I also found that parents occupations play a key role on entrepreneurial choice and explain why more than 40% of formal entrepreneurs in the data are sons or daughters of entrepreneurs. This means that informal business training received at home has an effective impact on entrepreneurial skills and enterprise success and suggests that business training may be useful to foster entrepreneurship. Finally, consistently with previous results, the estimations confirm that informal entrepreneurs are credit constrained and that individual wealth is a significant driver of entrepreneurship as it plays a key role in obtaining loans from financial

[1]Available in the Cameroon National Institute of Statistics website at www.statistics-cameroon.org.

institutions.

To empirically assess and quantify the role of institutional factors such as registration costs, tax rates, the degree of law enforcement and financing constraints, I perform a set of counterfactual policy simulations using the estimated model. I found that a 50% decrease in registration costs (e.g. through government subsidies or substantial reduction of administrative steps) doubles the proportion of formal enterprises (through both formalization of informal firms and new formal enterprise creation), increases employment by 10% and aggregate income by 15%. Interestingly, total tax revenues net of the foregone registration fees (or cost of subsidies) increase by more than 50%. Counterfactual results on tax policy shows that there is an optimal tax rate, estimated at 23% (i.e. about half of the current tax rate) that induces the maximum tax revenue gains at 30% above the current level. At the same time this optimal tax rate induces twice as much formal enterprises, 10% more employment and a 20% increase in aggregate income. By contrast, while an increase in the enforcement of formality status slightly increases the fraction of formal firms and the associated tax revenues, it has an overall perverse effect on the economy in terms of enterprise and job destruction, as well as aggregate income loss. I also examine the impact of other policy such as business training programs as recently reviewed by McKenzie and Woodruff (2013). The results indicate that a small increase in entrepreneurial skills resulting from business training has a large impact on entrepreneurship - mostly through informal enterprise creation - and generate important tax revenues and aggregate income gains. Put together, these results show that economic growth and important welfare gains can obtain from reducing the cost of registration and optimally choosing the tax rate while fostering entrepreneurial skills and enterprise creation through business training and better access to credit.

This paper relates to the active literature of entrepreneurship in countries with a sizeable informal sector. The closest to the model discussed in this paper is the one recently presented by Ordoñez (2014). A distinctive difference, though, is that while Ordoñez (2014) assumed away the entry cost into the formal sector, I formally include it in my model to reflect the original intuition of De Soto (1989) and the recent empirical findings of Djankov et al (2002) who both emphasize that start-up costs generate large informal sectors and may discourage formal entrepreneurship. Another key difference is that unlike Ordoñez (2014) who assumed that capital is freely optimized in the informal sector I allow informal entrepreneurs to face credit constraints, an assumption consistent with empirical work that have examined micro-enterprise financing in developing countries (e.g. Paulson et al 2006 for Thailand) and has formally been tested in Nguimkeu (2014) with the Cameroon data. In Antunes and Cavalcanti (2007), informality choice is an alternative that comes at no risk, while in my model, informal entrepreneurs face a nonzero probability of getting caught and having their output forfeited. Other papers such as Prado (2011), focus only on firms' decision to formalize and do not investigate the decision to become an entrepreneur which is an important focus of my paper. Another originality of this work, which to the best of my knowledge has not been attempted

in the literature, is that it uses a structural econometric approach to connect data from a typical developing country to a comprehensive theoretical model of informality with entrepreneurial choice. Thus, important relationships can therefore be estimated, consequences of informality, credit constraints, availability of skills and regulation costs can be studied, and counterfactual policy simulations can be performed to quantify the impact of relevant policies on entrepreneurship.

The rest of the paper is organized as follows. Section 2 presents a theoretical model of occupational choice and derives the main implications. Section 3 presents the data, descriptive statistics and reduced form results. Section 4 presents structural estimates and robustness checks. Policy simulations are presented in Section 5 and concluding remarks are given in Section 6. The Appendix gathers other technical material.

# 2 Model Description

The economy is populated with individuals that differ with their skills $\theta$ and their initial wealth $z$. If an agent is a wageworker or subsistence worker, its earning is a given amount $w$. This income is compared to the profits the agent could receive if they start a firm. Regardless of whether they are formal or informal, an entrepreneur with skill $\theta$ uses capital $k$ and hire $l$ units of labor to produce goods according to the technology

$$y = \theta k^{\alpha} l^{\beta} \epsilon \tag{1}$$

where $\alpha, \beta \in (0, 1)$ are the elasticities of output with respect to capital and labor respectively, and we have $\gamma = \alpha + \beta < 1$, implying diminishing returns to scale in variable factors at the establishment level (Lucas 1978). The term $\epsilon$ is a productivity shock, independent from $\theta$ and $z$, with positive support and unit mean.

## 2.1 Informal Entrepreneurship

Informal entrepreneurship implies that the entrepreneur do not pay taxes to the government. Tax avoidance comes with a risk of being caught, in which case the informal entrepreneur's profit is forfeited. Moreover, when operating in the informal sector the entrepreneur has limited access to the credit market. In order to get loan from the banks, they need to provide a collateral as a guarantee. They can therefore only borrow up to a fixed multiple, $\lambda \geq 1$, of their initial wealth, $z \geq 0$, that they use as collateral. Denoting by $p$ the probability of getting caught, the informal entrepreneur's optimal investment capital and labor then solves for his expected profit maximization problem

$$\max_{k,l} \left\{ (1-p) \left[ \theta k^{\alpha} l^{\beta} - wl - rk \right] \right\} \qquad \text{s.t.} \qquad 0 \leq k \leq \lambda z, \quad l \geq 0$$

For $\lambda = \infty$ the credit market is perfect and $\lambda = 1$ corresponds to financial autarky, where all capital is self-financed by the informal entrepreneurs. This specification captures the

common prediction from models of limited contract enforcement typical of the informal economy where the amount of credit is limited by individuals wealth.[2]

The optimization constraint on capital then gives rise to two types of informal entrepreneurs. Those who are financially unconstrained, i.e. those whose optimal investment capital is an interior solution of the above optimization problem, and those who are financially constrained, i.e., those whose capital constraint is binding. The interior solutions of the entrepreneurs maximization problem are

$$k^* = \theta^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{r}\right)^{\frac{1-\beta}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} \quad \text{and} \quad l^* = \theta^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{1-\alpha}{1-\gamma}} \tag{2}$$

This solution is feasible only if $k^*$ is lower than $\lambda z$, or equivalently

$$\theta \leq (\lambda z)^{1-\gamma} \left(\frac{r}{\alpha}\right)^{1-\beta} \left(\frac{w}{\beta}\right)^{\beta} \equiv \theta_c(z) \tag{3}$$

When the constraint is binding the investment capital and optimized labor are given by

$$k^* = \lambda z \qquad \text{and} \qquad l^* = \theta^{\frac{1}{1-\beta}} \left(\frac{\beta}{w}\right)^{\frac{1}{1-\beta}} (\lambda z)^{\frac{\alpha}{1-\beta}}.$$

The optimal informal entrepreneur's profits can therefore be expressed as follows:

$$\pi^I(z,\theta) = \begin{cases} \pi_l^I = (1-p)(1-\gamma)\theta^{\frac{1}{1-\gamma}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} & \text{if } \theta \leq \theta_c(z) \\ \pi_u^I = (1-p)\left[(1-\beta)\theta^{\frac{1}{1-\beta}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} (\lambda z)^{\frac{\alpha}{1-\beta}} - \lambda r z\right] & \text{otherwise.} \end{cases} \tag{4}$$

The informal entrepreneurs profit takes two possible values according to whether he is unconstrained, that is, $\theta < \theta_c$ or he is constrained, that is, $\theta > \theta_c$. The main difference between this profit function and the one derived by Nguimkeu (2014) is the inclusion of paid labor in the entrepreneur's objective. While this might be of relatively low importance for informal entrepreneurs of whom very few hire paid labor outside of the household in the data hired labor is a key characteristic of formal firms.

[2]The collateral constraint can be derived from a limited liability problem where at the beginning of the period an individual deposits his wealth $z \geq 0$ in the financial intermediary to rents $k$ units of capital. If this individual can abscond with this capital with probability $1/\lambda$ without any other form of punishment than loosing his collateral $z$, then in the equilibrium the financial intermediary will rent capital only to the extent that no individual will renege on the rental contract, implying $k \leq \lambda z$.

## 2.2 Formal Entrepreneurship

In order to formalize, entrepreneurs need to pay a registration cost, $c$, that represents both the license fees and the amount of time spent to get it. Moreover, at the end of the period, the government levies a tax $\tau$ on the profit of formal enterprises. Their formal status however provides them with a better access to credit from financial institutions, which we formalize by assuming that capital is optimized freely by formal entrepreneurs (see Bruhn 2011, Jansson and Chalmers 2001). Hence, the formal entrepreneurs' optimal investment capital and labor solves for the optimization problem.

$$\max_{k\geq 0, l\geq 0} \left\{(1-\tau)\left[\theta k^{\alpha} l^{\beta} - wl - rk\right] - rc\right\}$$

The optimal capital and labor are given by Equations (2) above and yield the following expression for the optimal formal entrepreneur's profit

$$\pi^{F}(z,\theta) = (1-\tau)(1-\gamma)\theta^{\frac{1}{1-\gamma}}\left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}}\left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} - rc \tag{5}$$

The closest to the model discussed in this section is a general equilibrium calibration model of occupational choice recently presented by Ordoñez (2014), but both differ in several distinctive ways. First, unlike Ordoñez (2014) who assumed away entry sunk costs into the formal sector, I formally include it in my model consistently with recent empirical studies that found that these costs are relatively and absolutely much higher in developing countries than in developed countries and are a serious constraint to formal entrepreneurship (see Djankov et al. 2002, Auriol and Walters 2005, Woodruff 2013, and the descriptive Table 4 below for the Cameroon data). Second, Ordoñez (2014) assumed that there is no financial constraints in the informal sector and that capital is freely optimized by informal entrepreneurs. These entrepreneurs may however choose to operate with less than an exogenously fixed amount, but only as a strategy to avoid detection. In contrast, I assume financial constraints in the informal sector and allow for this constraint to vary across individuals to account for heterogeneity in their initial wealth endowments. This assumption is consistent with empirical works that have examined micro-enterprise financing in developing countries (e.g., Paulson et al 2006), and has been empirically tested by Nguimkeu (2014) with the Cameroon data. The model in this paper is also more flexible than the one discussed by Prado (2011) and Mendicino and Prado (2014). Unless the probability of detection is higher than the tax rate, the Prado's model would deliver a corner solution equilibrium where all entrepreneurs prefer the informal sector even with a zero entry cost in the formal sector. This is inconsistent with data from developing countries where enforcement is weak and the probability of detection is usually lower than the tax rate, yet both formal and informal sectors coexist.

## 2.3 Occupational Choice

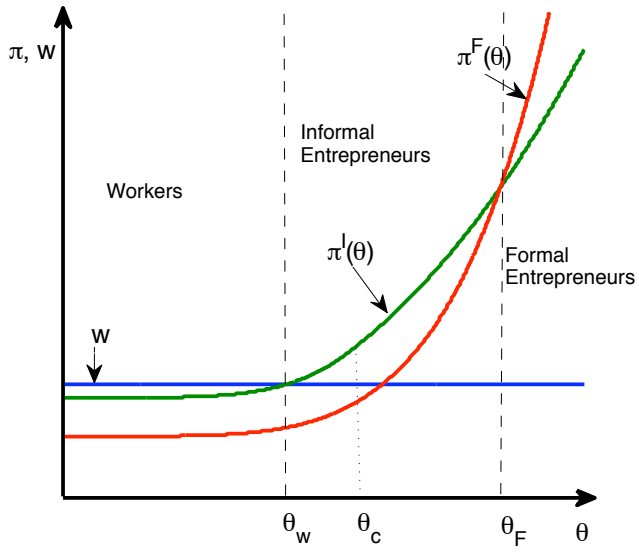
Each individual knows his personal characteristics $\theta$ and $z$, market characteristics $r$ and $w$ and institutional characteristics $\tau$, $p$ and $c$, where it is assumed that $\tau > p$ as

commonly observed in the data. Given these factors the agent chooses the occupation that would give him the maximum earning. In other words, the expected profit function of an agent with characteristics $(z, \theta)$ from the three categories of occupations can be written as

$$\pi(z,\theta) = \max\left\{w, \pi^I(z,\theta), \pi^F(z,\theta)\right\}$$

The earnings functions and related occupational choice are illustrated in Figure 1. The agent's decision is characterized by three thresholds, $\theta_W, \theta_F$ and $\theta_c$, which summarize the occupational decisions of the agents and whether the capital choices of informal entrepreneurs are constrained or unconstrained. The threshold $\theta_c$ is the one given in Equation (4). For $\theta_W$, notice that entrepreneurs profits are increasing with $\theta$ while workers earnings do not vary with $\theta$. It follows that there exists an ability threshold $\theta_W$ such that $w = \max\left\{\pi^I(z,\theta_W), \pi^F(z,\theta_W)\right\}$. Hence, all agents with $\theta < \theta_W$ become workers and the rest become entrepreneurs. Finally, note that informal entrepreneurs with initial wealth $z$ can not operate with capital above $\lambda z$. This constraint is more costly for higher skilled entrepreneurs as they would prefer larger scale firms given their high productivity. Hence, there exists an ability threshold $\theta_F$ such that $\pi^I(z,\theta_F) = \pi^F(z,\theta_F)$, above which all entrepreneurs prefer the formal sector and the rest the informal sector. These results are illustrated in Figure 1 and summarized in the following proposition.

Figure 1: Earning Functions and Occupations




**Proposition 1.** *Consider an agent with characteristics $\theta$ and $z$. There exist three critical ability thresholds $\theta_W(z)$, $\theta_c(z)$ and $\theta_F(z)$, with $\theta_W(z), \theta_c(z) < \theta_F(z)$, such that*

*(i) If $\theta < \theta_W(z)$ the agent chooses to be a wageworker*

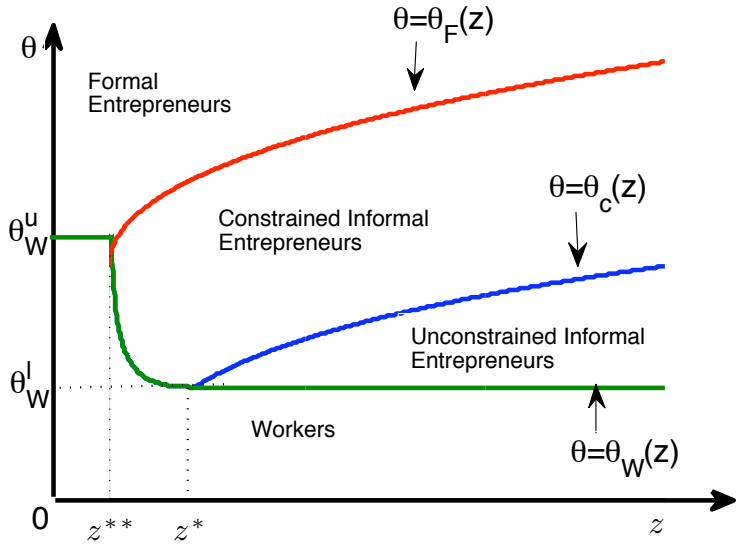
*(ii) If $\theta_W(z) \leq \theta < \theta_F(z)$ the agent is an informal entrepreneur. In particular, when $\theta_W(z) \leq \theta < \theta_c(z)$ he is an unconstrained informal entrepreneur, and when $\theta_c(z) \leq \theta < \theta_F(z)$ he is a constrained entrepreneur.*

*(iii) If $\theta \geq \theta_F(z)$ the agent is a formal entrepreneur.*

*Proof.* See Appendix. □

Whether or not an informal entrepreneur is constrained is a key determinant for his prospects to eventually formalize. Clearly, as long as the tax rate $\tau$ is higher than the enforcement $p$, unconstrained informal entrepreneurs will not formalize, regardless of the amount of the registration sunk cost $c$. Hence, inducing formality from registration reforms would mainly target only constrained entrepreneurs. In contrast, entrepreneurial training might induce both types of informal entrepreneurs to formalize. The nature of the selection to different types of occupations (i.e. workers, constrained informal entrepreneurs, unconstrained informal entrepreneurs and formal entrepreneurs) is depicted in Figure 2. This figure depicts a partition of the wealth-ability space that forms the basis of the structural estimation discussed in Section 4.

Figure 2: Characterization of Occupations




Note again that the implications derived here significantly differ from Ordoñez (2014). In particular, consistently with empirical evidence both constrained and unconstrained informal entrepreneurs always coexist in my model (see Figure 2). In contrast, Ordoñez (2014) model implies the existence of a corner solution equilibrium where all informal entrepreneurs are constrained.

Finally, the model also provides a ground to examine transition across different occupations as a function of skills. Three comparative payoff functions can be considered for the entrepreneur's decision-making process. The transition between wage work and informal entrepreneurship is characterized by the comparative payoff function $V^{W,I}(z,\theta) = \pi^I(z,\theta) - w$; the transition between wage work and formal entrepreneurship characterized

by $V^{W,F}(z,\theta) = \pi^F(z,\theta) - w$; and the transition between informal entrepreneurship and formal entrepreneurship characterized by $V^{I,F}(z,\theta) = \pi^F(z,\theta) - \pi^I(z,\theta)$. The properties of these transition functions are illustrated in Figure 3 and summarized in the following proposition.

**Proposition 2.** *Consider agents with characteristics $\theta$ and $z$. Then*

*(i) The functions $V^{W,I}(\theta,z)$ and $V^{W,F}(\theta,z)$ are increasing in $\theta$.*

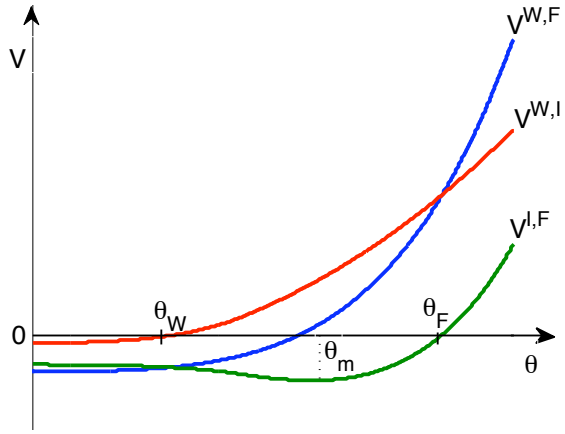
*(ii) The function $V^{I,F}(\theta,z)$ is U-shaped in $\theta$. That is, there exists $\theta_m > 0$ such that*

$$\frac{\partial V^{I,F}(\theta,z)}{\partial \theta} < 0 \quad \forall \theta \in [0,\theta_m), \qquad \text{and} \qquad \frac{\partial V^{I,F}(\theta,z)}{\partial \theta} \geq 0 \quad \forall \theta \in [\theta_m,\infty)$$

*Proof.* See Appendix. □

This result states that while the transition from wage-work to entrepreneurship is increasing in talent, the transition from informal entrepreneurship to formal entrepreneurship is non-monotonic. In fact, this transition function is U-shaped in entrepreneurial skills (see Figure 3). It implies that low ability entrepreneurs find it more profitable to remain informal when their level of skills slightly increases. This is especially true for unconstrained informal entrepreneurs since they have no incentive to formalize. However, above a certain ability threshold formality becomes increasingly attractive for increasing levels of skills, because entrepreneurs with such skills would want to operate at a large scale but would be constrained in the informal sector.

Figure 3: Transition between Occupations




Consistently with the literature on labor market, this framework assumes that labor is free to flow between sectors and that individual's expected wage is the same, regardless of the sector. This means that workers have no intrinsic preferences for the

sector they work for and only derive their utility from the wages they receive at work. While evidence show that on average formal wages are higher than informal wages, direct empirical tests of the premise that similar workers would expect higher wages in the formal sector compare to informal sector yield inconclusive results. For example, Magnac (1991), Maloney (2004), Pratap and Quentin (2006) do not find compelling evidence of labor market segmentation between formal and informal sector using data from Colombia, Mexico and Argentina, respectively.

It is important to note that while the functional forms used in this framework are standard in the literature and are useful to obtain parameter values that have meaningful economic interpretation, these functional forms are not driving the implications. In fact, the model can be fully identified non parametrically. To see this, notice that from any production function with standard Inada conditions, one can obtain the four types of occupations: wageworkers, Formal entrepreneurs, Informal unconstrained entrepreneurs, and Informal constrained entrepreneurs. To identify these occupations nonparametrically in the data it suffices to find exclusions restrictions for each occupation (see Buera 2006, Heckman and Honoré 1990). In this framework, the payoff of the worker is a constant that does not depend on its skills or its initial wealth, regardless of the technology or sector of activity. On the other hand, the payoff of formal entrepreneurs and informal unconstrained entrepreneurs depend only on entrepreneurial skills (since they are maximized via interior solutions). Hence, a proxy for entrepreneurial skills can be used to identify those agents. Since Formality status is observable, formal entrepreneurs and informal constrained entrepreneurs can be distinguished using informality dummy. Finally, the payoff of informal constrained entrepreneurs depends on both their skills and their initial wealth endowment (since their budget constraint is binding and they are using initial wealth as a collateral in the credit market).

# 3 Data and Reduced Form Evidence

This section describes some important features of the data used for this study and assesses the empirical relevance of some of the model predictions for the Cameroon Informal Sector. I give the background and provide descriptive statistics of these data and then use them to test some implications of the model through reduced-form Probit and Ordered Probit regressions.

## 3.1 Data and Background

The data used for the empirical analysis is a cross-sectional sample of households of Cameroon stemming from the Survey on Employment and the Informal Sector (EESI) conducted in 2005 by the National Institute of Statistics of Cameroon in partnership with the World Bank. This was a nationwide operation with two phases. The first phase collects sociodemographic and employment data while the second phase inter-

views a representative subsample of informal production units identified during the first phase. The methodology of the EESI is therefore similar to that of Phases 1 and 2 of the well known "1-2-3 surveys" in Central and West Africa (see www.afristat.org). This paper uses only data from Phase 1. The sampling design used for the survey was the result of the mapping of the third General Population and Housing Census conducted in 2005. A sample of 8540 households was drawn and interviewed following a two-degrees stratified sampling where the stratification was done according to the ten Cameroon provinces and areas of residence (i.e. urban, semi-urban or rural). For more details about the methodology used to collect these data, see INS 2005. For the analysis I restrict the sample to households whose heads are active and are aged 15 and above, representing a total of 6112 observations.

The definition of informality used for the EESI survey is the one adopted by the 1993 System of National Accounts, a set of international standards chosen to establish a framework for the production of national accounts statistics. The distinction between formal and informal sectors is made at the enterprise level, on the basis of administrative records and on whether or not the business keeps formal accounts. Accordingly, informal enterprises are defined as "production units that do not have written formal accounts and/or are not registered with the tax authorities." Informal sector workers are therefore persons exercising their main economic activities in informal establishments. The informal sector accounts for the vast majority of activities and employs about 89.5% of the Cameroon workforce aged 15 and above (INS 2005). The sample used consists of 6112 households where 4337 households belong to the Informal sector and 1775 households belong to the formal sector (based on the above definition of informality). The formality status of a household is categorized as formal if the household's head works in the formal sector.

### 3.2 Descriptive Statistics

Table 1 presents the characteristics of household according to their geographical areas.The average age of household heads is 36.2 years, with 37.8 in the formal sector against 35.4 in the informal sector. Women represent about 40% of the workforce and more than half of informal sector agents (54.3%). This percentage is a little more higher in rural areas (43.8%) compared to urban areas (40.0%). Conversely, in the formal sector, only one worker out of four is a woman (24.4%).

The empirical analysis requires to distinguish entrepreneurs from non-entrepreneurs according to their activity. While this distinction is clearer in the formal sector, it is not so obvious in the informal sector. For the formal sector, I follow the classic literature of entrepreneurial choice and consider self-employment or business-ownership as formal entrepreneurship (see e.g., Evans and Jovanovic, 1989; Blanchflower and Oswald 1998; Hurst and Lusardi, 2004). However, in the informal sector where the majority of people are self-employed, using the same definition in this context would be seriously misleading. In fact, self-employment in the informal sector includes both the actual informal

Table 1: Household Characteristics by Geographical Areas

| Characteristics | Geographical Areas | | |
|---|---|---|---|
| | Whole | Urban | Rural |
| Number of obs. | 6 112 | 4 295 | 1 817 |
| % of sample | 100% | 70.3% | 29.7% |
| % of women | 41.1% | 40.0% | 43.8.% |
| Av. household size | 6.1 | 6.2 | 6.0 |
| Av. age of head | 36.2 | 36.3 | 36.0 |
| Years of schooling | | | |
| 0-6 years | 48.1% | 36.5% | 56.8% |
| 6-10 years | 32.2% | 49.6% | 37.0% |
| 11+years | 25.5% | 28.8% | 17.7% |
| Av. income* | 82.2 | 91.8 | 59.3 |
| Av. wealth* | 3 374.4 | 3 749.6 | 2 487.4 |

*In thousands of local currency (CFA); $1,000$CFA $\sim$ \$2US Source: Own calculations

entrepreneurs as well as a wide category of subsisters. To distinguish between these activities, I follow Nguimkeu (2014) and define as informal entrepreneur a household that owns a business and employs others (see also Mondragón-Vélez and Peña 2010, for a similar definition). This definition particularly excludes purely self-employed (i.e. those who work just by themselves) many of which are subsisters.[3] Examples of informal entrepreneurs include taxi-drivers, grocers, tailors, carpenters, car mechanics who own shops, etc. I define as Workers the non-entrepreneurial households, that is, all other types of households including wageworkers employed by formal firms, informal firms or households as well as all categories of subsisters.[4]

Table 2 summarizes the main characteristics of households in the sample according to their occupation (formal entrepreneurs, informal entrepreneurs, workers). The sample consists of 92% workers, 7% of informal entrepreneurs and 1% of formal entrepreneurs. While it is possible that some households engage in two or more forms of activity at the

[3]Exceptions are some activities that use high capital but can only be operated by a single individual, such as taxi-drivers who own their taxi. This definition is obviously still not perfect. On the one hand, it would be inadequate to consider as entrepreneurs some employers who operate at a very low scale. On the other hand, given new technology and business models, even a purely self-employed with low physical capital could produce innovative products or services such that considering him as subsister be misleading. Such cases are however infrequent in these data as evidenced by the sensitivity analyses performed in Nguimkeu (2014).

[4]I do not distinguish between formal wageworkers and informal wageworkers since they usually do not select their formality status. While this keeps the analysis simple and the model tractable, it dilutes full extent of job losses due to enforcement in the counterfactual simulations of Section 4.

same time, I take the more relevant activity as their primary employment. In particular a household is consider as entrepreneurial if at least one member is entrepreneur in the above sense.The average number of paid employees (i.e. hired outside of family) per informal enterprise is about 1.3 against more than 50 for formal enterprises. While entrepreneurs are on average older than workers, formal entrepreneurs are on average much older than others.

Table 2: **Household Characteristics by Occupations**

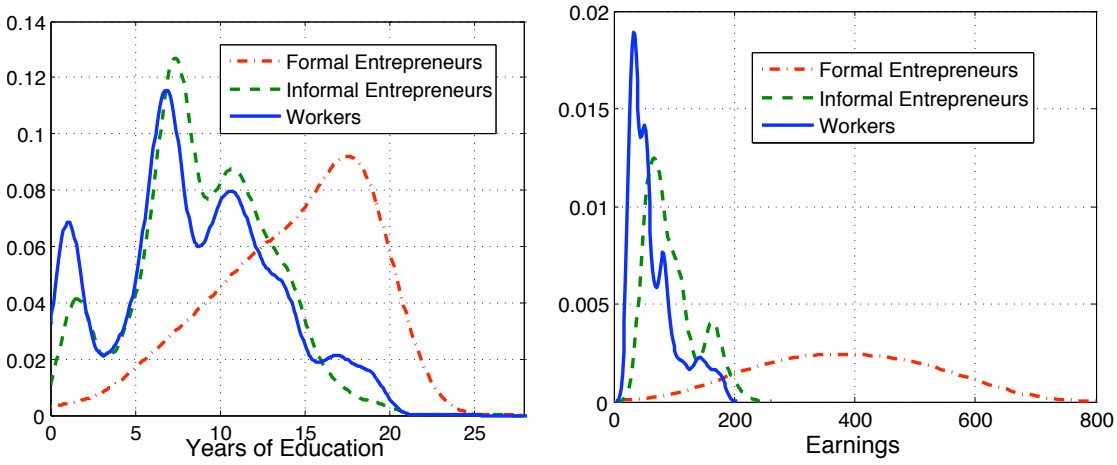
| Characteristics | Occupations | | |
|---|---|---|---|
| | Formal Entrepreneurs | Informal Entrepreneurs | Wageworkers /Subsisters |
| Num. of obs. | 65 | 424 | 5 623 |
| % of sample | 1.1% | 6.9% | 92.0% |
| % of women | 12.3% | 37.3% | 41.7% |
| Av. household size | 6.0 | 6.1 | 5.9 |
| Av. age of head | 42.4 | 37.0 | 36.1 |
| Years of schooling | | | |
| 0-6 years | 11.1% | 41.3% | 48.4% |
| 7-12 years | 31.5% | 48.6% | 36.2% |
| 13+ years | 57.4% | 10.1% | 15.4% |
| Parent Entrep. | 41.5% | 13.6% | 3.5% |
| Av. income* | 353.3 | 75.3 | 77.2 |
| Av. wealth* | 21 792.9 | 4 569.7 | 3 007.4 |

*In thousands of local currency (CFA); $1,000$ CFA $\sim$ \$2 US

Formal entrepreneurs are also more educated than informal entrepreneurs and workers, and the differences in the education composition across these occupations is sizeable. On the one hand, informal entrepreneurs and workers show similar primary education levels (0 - 6 years). The main differences between these two lies in the composition of the two top education levels; a higher proportion of informal entrepreneurs have a secondary education (41.3% against 36.2%), whereas a relatively higher proportion of workers have a college education (10.1% against 15.4%), a proportion probably driven by top ranked employees in highly-paid formal jobs. On the other hand, formal entrepreneurs are by far the most educated with 57.4% of them holding a post-secondary degree. Figure 4 provides the distribution of years of education by occupation. The data also shows that a high proportion of entrepreneurs are sons and daughters of entrepreneurs. In particular, for 41.5% of formal entrepreneurs at least one parent was an entrepreneur. This is also true for 13.7% of informal entrepreneurs. By contrast, only 3.5% of workers are children of entrepreneurs. The left pattern of Figure 4 shows the density of years of education by occupation and confirm the superior educational level of entrepreneurs over workers'

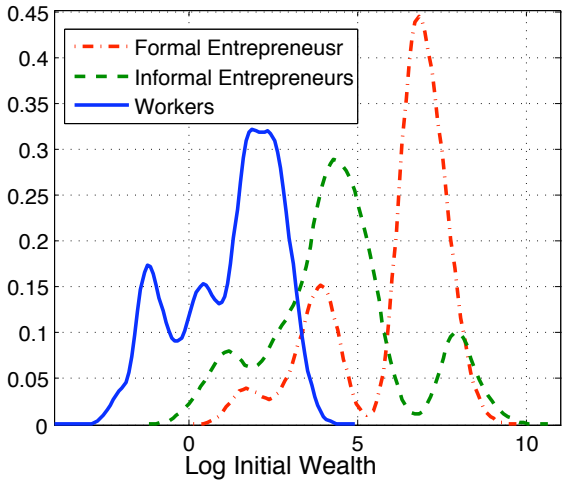
from a distributional perspective. These patterns suggest, as we assess in the structural estimation presented in Section 4, that both education and parents occupation are strongly correlated with entrepreneurial ability.

Figure 4: Distribution of Education and Earnings by Occupation




On average, entrepreneurs also earn more than workers. Consistently with the distribution of years of education given in the left panel, the right panel of Figure 4 shows that while the earning distribution of informal entrepreneurs slightly dominates the workers earning distribution, formal entrepreneurs' earnings are largely above these two.

Figure 5: Distribution of Log Initial Wealth by Occupation




As explained in the theory, initial wealth plays a key role in determining the amount

of capital that a household is able to borrow from financial institutions. While the survey do not report households ex-ante total wealth, there are retrospective questions about household belongings acquired prior to starting their activity. I use this information to build a measure of household initial wealth by computing the market value of their total initial belongings.[5] Figure 5 shows the distribution of log initial wealth by occupation. It is clear that entrepreneurs are wealthier than workers, as the bulk of their wealth distribution is more concentrate on higher values of the wealth range. For informal entrepreneurs, this shows their ability to provide collateral for credit borrowing. For formal entrepreneurs, this indicates their ability to pay the registration sunk cost upfront. Since entrepreneurs tend to be older and to possess higher initial wealth, it may be the case that households that eventually become entrepreneurs have higher past wealth because they have been saving at a higher rate or because they spent more time accumulating their savings in anticipation of starting a business in the future.

### 3.3 Regressions Analysis

In this section, I use the Cameroon data to test some of the model predictions, that is, I examine how entrepreneurship and informality is affected by financial constraints and observable business skills. In particular, I consider to what extent the likelihood of starting a formal or an informal business is related to education and parent occupation (both of which are likely to be correlated with business skills as suggested by descriptive statistics), and household wealth. To evaluate these implications I present three Probit regressions where transitions to entrepreneurship and formality are examined. These reduced-form methods have the advantage of using a large set of variables than one featured in the theoretical model and require less parsimonious or statistical assumptions than the structural estimation. They are therefore useful to check the relevance of the theory but also to examine to what extend other factors not featured in the theoretical model may influence occupational choice.

Table 3 summarizes results from probit regressions. The first panel (Informal Entr. vs. Formal Entr.) presents the transition from informal entrepreneurship to formal entrepreneurship , the second panel (Workers vs. Formal Entr) from non-entrepreneurial activity to formal entrepreneurship and the third panel (Worker vs. Informal Entr) from non-entrepreneurial activity to informal entrepreneurship. Again, unemployment is not explicitly part of the choices because in the absence of formal unemployment benefits, open unemployment is quasi inexistent as these individuals are likely to engage in some subsistence activity which are already incorporated in the workers or non-entrepreneurial category. For each set of regressions, Column (1) presents the baseline estimation while Column (2) includes a quadratic term for Education in order to assess possibly nonlinear

[5]These belongings are household durable goods including vehicles, TVs, radios, DVD/CD players, fridges, freezers, gas cookers, fans, sewing machines, mobile phones, computers, electric irons, number of houses owned by the household. A better proxy for initial wealth would have been the amount of inherited wealth as in Blanchflower and Oswald (1998). But this information is not available in the present data.

effects of observable skills as implied by the theory (see Proposition 2(ii) ).

Table 3: Probit Estimates of Transitions between Occupations

| Variable | Informal Entr. vs. Formal Entr . | | Worker vs. Formal Entr. | | Worker vs. Informal Entr. | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (1) | (2) | (1) | (2) |
| Intercept | -3.853*** | -3.129*** | -3.555*** | -3.571*** | -1.529*** | 1.789*** |
| | (0.483) | (0.564) | (0.249) | (0.354) | (0.078) | (0.094) |
| Education | 0.149*** | -0.023* | 0.046*** | 0.050 | -0.006 | 0.087 |
| | (0.029) | (0.016) | (0.015) | (0.056) | (0.006) | (0.078) |
| Education$^2$/100 | - | 0.863** | - | -0.017 | - | 0.616 |
| | - | (0.426) | | (0.256) | - | (0.412) |
| Parent | 1.770*** | 1.765*** | 0.435*** | 0.434*** | -0.713*** | -0.693*** |
| | (0.271) | (0.269) | (0.128) | (0.128) | (0.111) | (0.111) |
| Initial Wealth | 0.003*** | 0.003*** | 0.003*** | 0.003*** | 0.001*** | 0.001*** |
| | (0.001) | (0.001) | (0.000) | (0.000) | (0.000) | (0.000) |
| Experience | 0.040*** | 0.039*** | -0.003 | -0.003 | 0.0001 | 0.002 |
| | (0.015) | (0.015) | (0.008) | (0.008) | (0.003) | (0.003) |
| Female | -0.256 | -0.213 | -0.308* | -0.308* | -0.153*** | -0.169*** |
| | (0.275) | (0.273) | (0.169) | ( 0.169) | (0.053) | (0.053) |
| Married | 0.287 | 0.271 | 0.151 | 0.151 | 0.098* | 0.124** |
| | (0.262) | (0.265) | (0.156) | (0.156) | (0.053) | (0.053) |
| Urban | 0.106 | 0.098 | 0.432** | 0.433** | 0.178*** | 0.173*** |
| | (0.304) | (0.302) | (0.182) | (0.182) | (0.058) | (0.058) |
| Handicapped | 0.211 | 0.214 | 0.434* | 0.433* | -0.066 | -0.049 |
| | (0.413) | (0.414) | (0.225) | ( 0.226) | (0.139) | (0.140) |
| Number of Obs. | | | | | | |
| AIC | 194.43 | 192.54 | 506.37 | 508.37 | 3005.3 | 2974.5 |

Consistently with the theory, the results confirm that observable skills such as education are a critical requirement for formal entrepreneurship. The baseline results show that Education is positively and significantly associated to the probability of formal entrepreneurship. Interestingly, when a quadratic term for education is added in the regressions, the associated coefficient is significantly positive in the first panel while the initial coefficient becomes negative. However, this quadratic term is, as expected, insignificant in the other panels. This confirm that the transition from informal to formal entrepreneurship is U-shaped in education as suggested by the theory.

Also, the coefficient on Parent is positive and significant, implying that individuals whose parents were entrepreneurs are more likely to become entrepreneurs than others. This suggests that children of entrepreneurs may have received informal business

skills from their parents. This evidence has been also supported by a number of studies, such as Lentz and Laband (1990) and Parker and Van Praag (2012). In contrast, the association between education and the transition probability from Worker to Informal entrepreneur is insignificant. This result might arise from two conflicting effects due to the coexistence of both extremely low educated individuals and extremely highly educated individuals among Workers, compared to a more homogenous averaged educated group of Informal Entrepreneurs. The formers may therefore be driving the regression coefficient in opposite directions, resulting in an insignificant effect. However, the coefficient on Parent is still significantly positive, suggesting that those Workers whose parents were entrepreneurs are more likely to become informal entrepreneurs than others.

A key variable that allows to assess the role of financial constraints is the household initial wealth. The results show that the corresponding coefficient is significantly positive in all the regressions, although it is weaker in the second panel. This implies that there are financial constraints to entrepreneurship, especially informal entrepreneurship. As explained earlier, households that eventually become entrepreneurs may have higher initial wealth because they have been saving at a higher rate in anticipation of starting a firm in the future. Given such pattern, this result reinforces the evidence that financial constraints are important since households have to wait to start a business until they have accumulated enough wealth to finance it. This suggests that they are unable to borrow to start their business.

# 4 Structural Estimation

In this section, I estimate the model by maximum likelihood. What is fitted is the probability of being formal entrepreneur, informal entrepreneur or worker as a function of wealth and ability generated by the model with the actual household occupational status from the data. The goal is to produce structural estimates that allows to examine the content of the mechanisms implies by the theoretical model and to use the estimated model in counterfactual simulations to quantify the impact of relevant policies on entrepreneurship and informality.

## 4.1 Estimation Technique

While the initial wealth, $z$, is observable, the individual ability, $\theta$, is not observable by the econometrician. Following previous studies, I make the structural assumption that this ability depends on education and parent occupation.

$$\ln \theta = \delta_0 + \delta_1 s + \delta_2 P + \varepsilon \tag{6}$$

where $s = \ln(1 + S)$ are the logs of years of schooling $S$, $P$ is a dummy indicating whether at least one parent was an entrepreneur. In this baseline specification current entrepreneurial ability is assumed independent from initial wealth. But I relax this assumption in the robustness checks to allow for the entrepreneurial skills to be correlated

with wealth. There are three categories of agents observed in the data: Formal entrepreneurs, $(F = 1, I = 0, W = 0)$, Informal entrepreneurs $(F = 0, I = 1, W = 0)$ and wage-workers/subsisters $(F = 0, I = 0, W = 1)$. Denote by $X = [1 \quad s \quad P \quad z]$ the vector of observable covariates and by $\mathbf{1}[\cdot]$ an indicator function that takes the value one when its argument is true and zero otherwise.

Given the model predictions, the probability of formal entrepreneurship is[6]

$$\Pr[F = 1|X] = \Pr[\theta > \theta_F(z)] = 1 - \Phi\left\{\frac{\ln \theta_F(z) - \delta_0 - \delta_1 s - \delta_2 P}{\sigma}\right\} = H_F(\psi, X) \quad (7)$$

where $\psi$ denotes the vector of all structural parameters of the model and $\theta_F(z)$ is given by Equation (14) in the appendix.

The probability of non-entrepreneurial activity is given by

$$\begin{aligned}\Pr[W = 1|X] = \Pr\left[\theta \leq \theta_W(z)\right] &= \Pr\left[\theta \leq \theta_W^u\right] \mathbf{1}\left[z \leq z^{**}\right] + \Pr\left[\theta \leq \theta_W^l\right] \mathbf{1}\left[z \geq z^*\right] \\ &\quad + \Pr\left[\theta \leq \theta_W^m(z)\right] \mathbf{1}\left[z^{**} \leq z \leq z^*\right] \\ &= \Phi\left\{\frac{\ln \theta_W^u - \delta_0 - \delta_1 s - \delta_2 P}{\sigma}\right\} d_i^u(z) + \Phi\left\{\frac{\ln \theta_W^l - \delta_0 - \delta_1 s - \delta_2 P}{\sigma}\right\} d_i^l(z) \\ &\quad + \Phi\left\{\frac{\ln \theta_W^m(z) - \delta_0 - \delta_1 s - \delta_2 P}{\sigma}\right\}\left(1 - d_i^u(z) - d_i^l(z)\right) = H_W(\psi, X)\end{aligned} \quad (8)$$

where $d_i^u(z) = \mathbf{1}\left[z \leq z^{**}\right]$ and $d_i^l(z) = \mathbf{1}\left[z \geq z^*\right]$. The thresholds ability levels $\theta_W^u(z)$, $\theta_W^m(z)$ and $\theta_W^l(z)$ are given by Equation (12) in the appendix, while $z^*$ and $z^{**}$ are respectively given by Equations (11) and (13) in the appendix.

The probability of informal entrepreneurship is then obtained by

$$\Pr[I = 1|X] = 1 - \Pr[F = 1|X] - \Pr[W = 1|X] = 1 - H_F(\psi, X) - H_W(\psi, X) = H_I(\psi, X). \quad (9)$$

Given a sample of independent observations of size $n$, $\{(F_i, I_i, W_i, X_i), i = 1, \ldots, n\}$, the log-likelihood function of the econometric model can therefore be written as:

$$L_n(\psi) = \sum_{i=1}^{n} F_i \ln H_F(\psi, X_i) + I_i \ln H_I(\psi, X_i) + W_i \ln H_W(\psi, X_i) \quad (10)$$

where $F_i, I_i$ and $W_i$ are zero-one indicator variables for the observed occupational choice of household $i$, and $X_i$ is the vector of their observable characteristics as defined above. Both the gross interest rate $r$ and the wage rate $w$ are exogenously fixed at the observed average rates. The maximum likelihood estimation is therefore performed over the set

[6]Note that unlike in a standard probit model, the standard error $\sigma$ is identifiable because of the nonlinearity of the model in the parameters.

of parameters $\psi = [\delta_0, \delta_1, \delta_2, \alpha, \beta, \sigma, \lambda]$. These parameters correspond respectively to the constant term of the ability distribution, $\delta_0$; the interaction between education and ability, $\delta_1$; the interaction between parents occupation and ability, $\delta_2$; the productivity of capital in the production technology, $\alpha$; the productivity of labor in the production technology, $\beta$; the standard deviation of the ability distribution, $\sigma$; and the degree of financial friction, $\lambda$. With the given observations, standard maximization routines can be used to search for the maximum numerically. The standard errors of the estimated parameters can be computed by bootstrap methods using draws of the original sample with replacement. Details about the maximization algorithm that I use are presented in the Appendix.

### 4.2 Institutional Parameters

There are three exogenous institutional parameters entering the model: the tax rate $\tau$, the entry sunk cost $c$, and the probability of detection, $p$. Table 4 provides useful information that may help to set the values for some of these parameters. In particular, the total tax rate as a percentage of enterprise profit is readily available and fixed at $\tau = 48\%$ as per the measurements of *Doing Business Survey (2005)*.

Table 4: Characteristics of the Institutional Environment

| Indicator | Starting a Business | Indicator | Paying Taxes |
|---|---|---|---|
| Number of procedures | 12 | Number of payments/year | 44 |
| Number of days | 37 | Number of days | 90 |
| Registration fees (% GNI/capita) | 182.5 | Total tax rate (% profit) | 48.9 |
| Min. capital (% GNI/capita) | 232.0 | | |
| GNI per capita = \$640 $\approx$ CFA 320,000 | | *Source: Doing Business 2005* | |

For the entry sunk cost to formality, I take the registration fees estimated by Doing Business (2005) which I top up by the foregone income incurred during the days spent in the registration office for the procedures. That is,

$$c = \text{Registration Fees} + \text{Number of days} \times \text{Average daily Earnings}$$

Table 4 shows that the registration fee can be calculated at CFA 582,400, the number of days for the registration procedure is 37, and the average daily earnings in our data is measured at CFA 3,200. Hence, the parameter $c$ is exogenously fixed at $c =$ CFA 700,800 $\approx$ \$1,400 (see also Djankov et al. 2002 for a similar estimation). Note that this is still an underestimation of the actual entry cost since it does not account for the amount of bribes given to the registration officers, psychological costs due to the large and exhausting number of procedures, etc.

However, the probability of detection $p$ is not directly available from the data. In order to approximate it, I use information on the number of tax inspections and the

level of corruption in the country. Since inspections are on-site, the probability of being in trouble with tax authorities can be obtained as the ratio of the total number of tax inspections over the total number of firms. This should however be deflated by the degree of integrity of tax inspectors. Thus, I calculate the probability $p$ of getting caught and having the firm's output forfeited by

$$p = \frac{\text{Total number of tax inspections}}{\text{Total number of firms}} \times \text{Degree of Integrity of tax inspectors}$$

The Total number of tax inspections is obtained by multiplying the number of tax inspectors (proxied by the size of the tax department of the Ministry of Finance) by the number of per-period inspections (all available in the Cameroon Statistical Yearbook, at www.statistics-cameroon.org). The total number of businesses with fixed locations was measured during the 2009 General Enterprise Census (see INS 2009). Finally, I use the Corruption Perception Index (CPI) produced by Transparency International as a measure of the integrity of tax authorities in Cameroon. The CPI is a score that indicates the perceived level of public sector corruption on a scale of 0 (highly corrupt) to 100% (very clean) (see www.transparency.org for details). Given these measures, I compute the probability of detection at $p = 0.78\%$.[7]

### 4.3 Structural Results

The results of the maximum likelihood estimation of the model are presented in Table 5. The estimated correlation between entrepreneurial ability and education, $\delta_1$ is estimated at 0.4, implying that a 10 percent increase in the amount of schooling increases entrepreneurial ability by 4 percent. This suggests that education may be a reasonable indicator of entrepreneurial talent in Cameroon. Having a parent who was an entrepreneur is also an important ingredient, since the estimated coefficient that relates parent occupation and entrepreneurial ability, $\delta_2$, is positive and significant. In the second column, the estimated correlation between entrepreneurial ability and assets is statistically insignificant. Thus we can reject the hypothesis that wealth is a positive proxy for entrepreneurial ability. Moreover, the inclusion of wealth in the log talent specification significantly deteriorates the model fit (see the discussion in Section **??** below). An important finding is that there are binding capital constraints, since the the degree of financial friction, $\lambda$, is estimated at 11.4. This estimate, however, seems to be too high. It means that agents can borrow up to 10 times the value of their wealth. In practice, loan to collateral values are typically quite low and the value of the loan is very often significantly less than the value of the collateral used to secure it (i.e, $\lambda \leq 2$). On the other hand, there are also many unsecured loans in the data, based on reputation, trust, etc. (i.e., $\lambda \to \infty$). A combination of these two patterns should yield a fairly large

[7]Note that while these values of institutional parameters that I use for the structural estimation may be imperfect, the simulations performed in Section **??** allow to examine the sensitivity of the results for a wide range of different values that they may respectively possibly take.

estimated $\lambda$ as the one obtained in Table 5.

The estimates of $\alpha$ and $\beta$ mean that a 10 percent increase in the capital devoted to a business leads to a 1.2 percent increase in earnings, while a 10 percent increase in hired labor increases earnings by 4.5 percent, respectively. Since returns to capital are usually high in the informal sector (see, eg. Kremer et al. 2010, Schundeln 2004, Udry and Anagol 2006, De Mel et al. 2008) it must be that the estimated value of $\alpha$ obtained here is pinned down by firms of the formal sector where returns to capital turn out to be very low in most studies (e.g., Alby, Auriol and Nguimkeu 2013).

Table 5: Structural Maximum Likelihood Estimates of the Model

| Parameter | Name | **Estimate** | |
|---|---|---|---|
| | | (1) | (2) |
| Log ability - constant | $\delta_0$ | -2.8372 | -3.0314 |
| | | (0.0118) | (0.0980) |
| Log ability - education | $\delta_1$ | 0.4013 | 0.4270 |
| | | (0.0210) | (0.0118) |
| Log ability - parents | $\delta_2$ | 0.0241 | 0.0452 |
| | | (0.0079) | (0.0118) |
| Log ability - wealth | $\delta_3$ | ... | 0.0186 |
| | | ... | (0.0214) |
| Capital share | $\alpha$ | 0.2201 | 0.2252 |
| | | (0.059) | (0.028) |
| Labor share | $\beta$ | 0.4502 | 0.4702 |
| | | (0.092) | (0.022) |
| St. Dev for ability | $\sigma$ | 2.4610 | 2.2692 |
| | | (0.0380) | (0.0330) |
| Capital constraint | $\lambda$ | 11.417 | 10.624 |
| | | (3.410) | (3.512) |
| Log-likelihood | | -1.9602 | -1.9692 |
| Number of Obs. | | 6112 | 6112 |

Bootstrap standard errors in parenthesis

For the labor returns coefficient $\beta$, on the one hand most informal firms produce output without hiring labor out of their household (corresponding to a $\beta \to 0$); and on the other hand, formal firms output are very elastic to labor, consistently with a $\beta$ that tends to 0.7 according to stylized fact. The average therefore reasonably falls within the estimated value given in Table 5. The sum of these two coefficients is thus estimated at 0.67 and is statistically lower than unity, implying that there are diminish-

ing returns to managerial ability at the establishment level (as discussed by Lucas 1978).

Before taking the estimated model seriously to a counterfactual policy simulation exercise, it is useful to first examine how well it fits the data. I first check how sensitive the results are to various specifications of the log entrepreneurial talent and also calculate a goodness of fit statistic for the empirical model.

### 4.4 Robustness and Goodness-of-fit Checks

A common concern in the entrepreneurship literature is the possible correlation between wealth and entrepreneurial ability. For example, to study the relationship between starting a firm and initial wealth, some authors (e.g., Holtz-Eakin et al. 1994, Blanchflower and Oswald 1998) use data on inheritances which are likely exogenous. Lacking such data I make an effort to control for endogeneity by using assets acquired by households several years prior to starting their activity as the initial wealth variable in the model. However, the baseline specification given in Equation (6) where the unobserved ability distribution is independent of wealth may still be potentially problematic. In this section I discuss the implications of relaxing this assumption by allowing correlation between ability $\theta$ and initial wealth $z$, that is, $\ln\theta = \delta_0 + \delta_1 s + \delta_2 P + \delta_3 z + \varepsilon$. The results from re-estimating the model using this specification are given in Column (2) of Table 5. The parameter $\delta_3$ provides a measure of the correlation between wealth and ability and is estimated at 0.0186 with a standard error of 0.0214. This indicates a positive but insignificant correlation between ability and the measure of wealth. However, with this new specification, the coefficient on Parent becomes negative and the overall fit deteriorates. This means that wealth is not acting as a proxy to entrepreneurial talent. Moreover, the inclusion of wealth does not significantly change the initial results.

The second concern is to assess how well the model fits the data. First, a comparison between the actual relative sizes of each occupation with their fitted counterparts shows a close proximity. While the data features 1.1% of formal entrepreneurs, 6.9% of informal entrepreneurs and 92.0% of workers, the estimated model delivers fitted probabilities of 1.3%, 6.8% and 91.9, respectively. Second, a more formal standard statistical maneuver for performing a specification test is to compare observed and expected values, since large departures between them would seemingly indicate lack of fit. I apply a simple Pearson test statistic based on the standardized residuals within each occupation. The test is defined by

$$\widehat{T} = \sum_J \frac{(n_J - n\bar{\hat{H}}_J)^2}{n\bar{\hat{H}}_J} = \sum_J \frac{\left(\sum_{i=1}^n \left[\mathbf{1}(J_i = 1) - H_J(\widehat{\psi}, X_i)\right]\right)^2}{\sum_{i=1}^n H_J(\widehat{\psi}, X_i)}, \quad J \in \{F, I, W\}$$

and has a limiting chi-squared distribution with 2 degrees of freedom under the null of correct model specification. Large values of the statistic would imply that the model is inappropriate. This statistic is computed at 1.906 for the estimated model, that is, a $p$-value of 0.3856, suggesting that the empirical model is not at odd with the data.

# 5 Counterfactual Policy Analysis

In this section, I perform a set of counterfactual experiments to evaluate the impact of some policies on entrepreneurship and informality with the Cameroon data. The estimated model is considered the current equilibrium and departures from this initial state are assessed by evaluating discrepancies that may occur from changes in policy parameters due to various possible reforms. Since the theory suggests that entry sunk costs, taxation, enforcement and entrepreneurial skills are the main drivers of entrepreneurship and formality, I examine the impact of each of these factors.

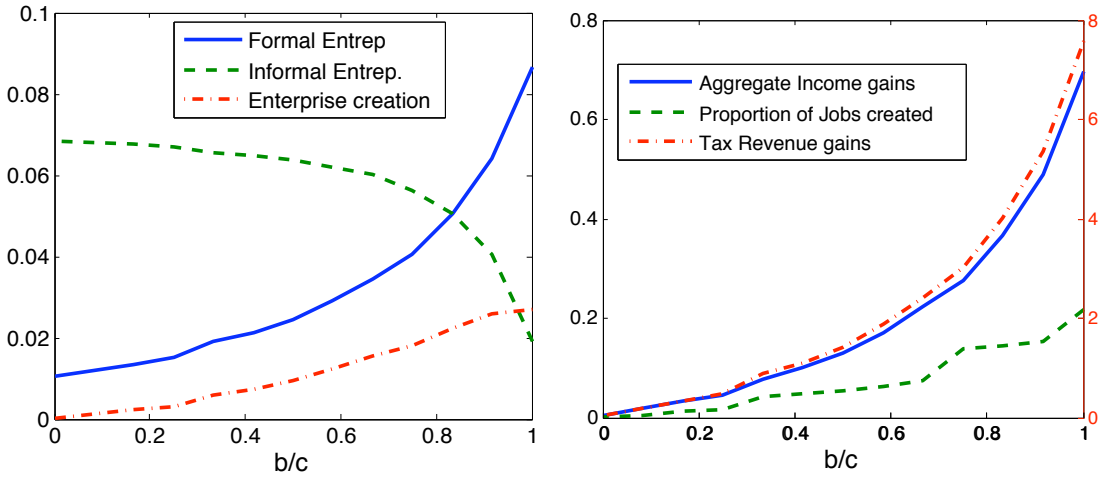
## 5.1 Registration Reforms

One policy that has proven to be quite successful, at least for the case of Mexico, is the firm registration reform. The policy consists in substantially reducing the cost of registration and the number of procedures required to start a firm (see Bruhn 2011 for more details). In the context of Cameroon, Doing Business 2005 documents that firm registration involves 12 procedures, 37 days and a registration fee that is about 1.82 times the GNI per capita (see Table 4). This represents an extremely high cost that is difficultly affordable by medium-size entrepreneurs. Since the entry sunk cost to formal entrepreneurship is captured by the institutional parameter, $c$, in the model, the counterfactual experiment consists in examining how the equilibrium would change from a reduction of $c$ from the current cost to a smaller amount. The policy is thus implemented as

$$c' = c - b, \qquad 0 \leq b < c,$$

where $b$ represents the decrement in the entry cost implied by the reform. Practically, this policy may also take the form of a government registration subsidy that encourages small firms to formalize.

Figure 6: Impact of a Registration Reform/Subsidy



The impact of this registration reform with the Cameroon data within the context of the model is depicted in Figure 6. The effect is quantified for a range of relative costs decrements, $b/c$, starting from 0, the current state as produced by the structural parameter estimates, to 1, the idealistic state where the entry cost is zero. The left panel of Figure 6 shows the variation in the fraction of formal entreprises, informal entreprises and new enterprise creation. As $b/c$ increases, the fractions of formal entrepreneurs and new enterprises increase while the fraction of informal entrepreneurs decreases. In particular, a 50% decrease in registration costs (i.e $b/c$ goes from 0 to 0.5) doubles the proportion of formal enterprises, through both formalization of informal firms and new formal enterprise creation. The right panel of Figure 6 depicts the variation in aggregate income gains, computed as the total income from all sectors, the proportion of job creation implied by formalization and enterprise creation, and the tax revenue gains, computed as the total tax revenues net from the foregone registration fees due to the reform (right vertical axis). Results show that all those three aggregates increase with decreasing entry costs. In particular, a 50% decrease in registration costs increases employment by 10%, aggregate income by 15% and total net tax revenues by more than 50%. These findings are consistent with empirical results recently found by Bruhn (2011) and Kaplan et al. (2011) about the effect of the "System of Fast Opening of Firms" (SARE) on entrepreneurship in Mexico. Other compelling empirical evidence have also been found from recent reforms in some developed countries such as Germany (Rostam-Afschar 2014) and Portugal (Branstetter et al. 2014).

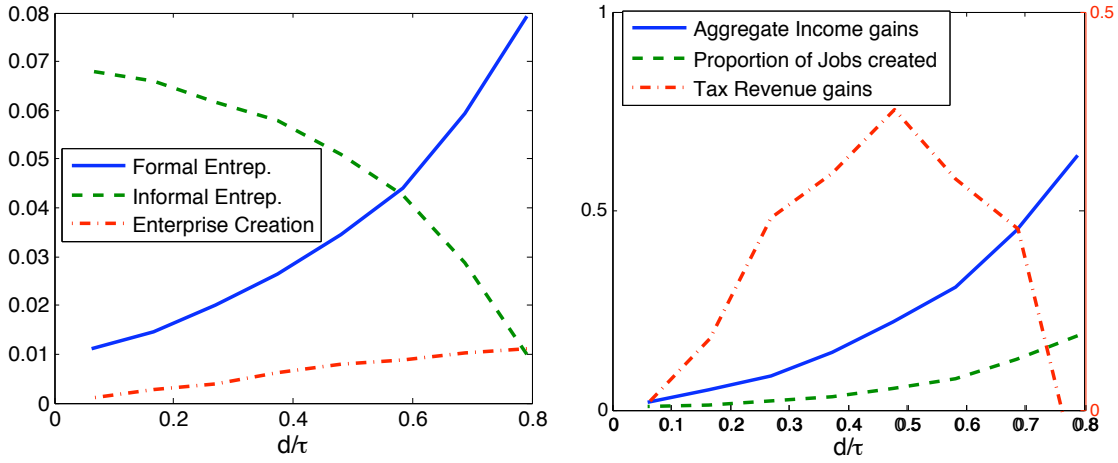
### 5.2 Tax Reforms

Given the large size of the informal sector, a natural question is whether the government is choosing the tax rate in the best possible way. As argued by Woodruff (2013), it is not clear whether high rates of taxation implies higher tax collections. In fact just the opposite may occur. While high taxation may increase tax revenues and provide the state with the resources necessary to build law enforcement capacity and the capacity to offer some of the benefits of being formal, too much taxation may as well push economic activity out of the formal economy. Here, I investigate the impact of tax reforms on the entrepreneurship and informality. Formally, I assume reductions in tax rates of magnitude $d$, such that

$$\tau' = \tau - d, \qquad 0 \leq d < \tau$$

Figure 7 depicts the effects of variations in tax rates on the economy. The relative reduction in tax rate, $d/\tau$, ranges from 0, the current state, to about 0.8, representing an 80% tax reduction. As one would expect, a decrease in taxes increases the fraction of formal entrepreneurs, both in terms of formalization of informal firms as well as new enterprise creation (see left panel of Figure 7). Only constrained entrepreneurs are affected by this policy. Unconstrained entrepreneurs are not. All the constrained move and the remaining fraction of entrepreneurs are the unconstrained. The more interesting pattern, however, is the evidence of the sub-optimality of the current tax rate in the Cameroon data as depicted in the right panel of Figure 7. In fact, the counterfactual

Figure 7: Impact of a Tax Reduction Policy




results show that taxation rates have a U-inverse effect on government revenues; there exists an optimal tax rate, estimated at 23% (i.e. about half of the current tax rate) that induces the maximum tax revenue gains at 30% above the current revenues. At the same time this optimal tax rate induces twice as much formal enterprises, 10% more employment and a 20% increase in aggregate income.

### 5.3 Law Enforcement

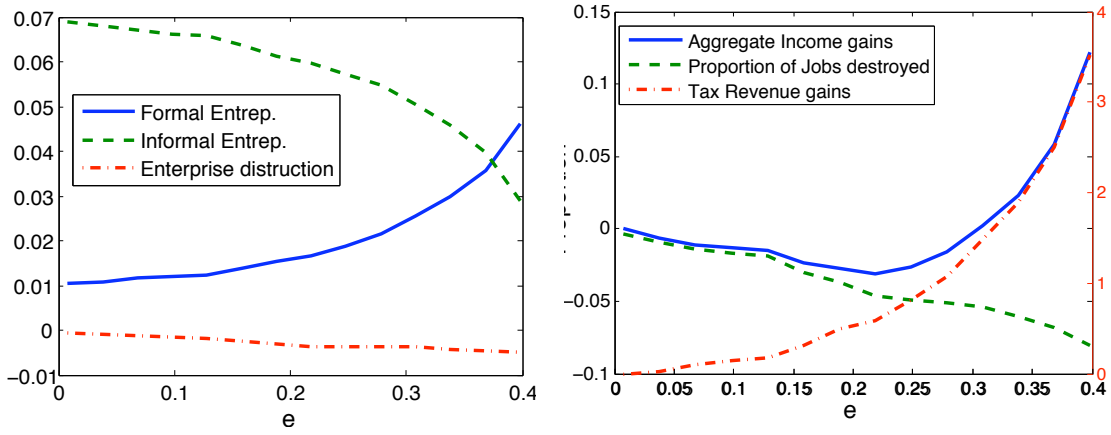
While there is an extensive theoretical literature that emphasizes the role of law enforcement as a typical cause of the existence of large informal sector, there is very little empirical work that quantitatively assesses how changes in law enforcement affect entrepreneurship and informality. This limitation is perhaps due to the difficulty of getting counterfactual data. In this study, I take advantage of the structural framework which allows to approximately quantify the effect of law enforcement through counterfactual simulations. In this experiment, increased enforcement is modelled by

$$p' = p + e, \qquad 0 \leq e < 1 - p$$

where $e$ represents increases in the probability of detection. Unlike in the previous setting where the cost of implementing the policy is accounted for in the benefits calculations, in this setting, I am abstracting from the costs incurred by enforcing the formality status. Figure 8 illustrates the effects of increased formality enforcement on the Cameroon economy. The values of the enforcement probability increments range from 0, the current state, to 0.42 corresponding to the state where an informal firm has a 50% chance of being detected and getting in trouble with the tax authorities. The left panel of Figure 8 shows the impact of enforcement on formal entrepreneurship and enterprise creation. By contrast to the pervious policies, increased enforcement shows a very different pattern. While increase enforcement increases the fraction of formal enterprises, the fraction of

new enterprises, however, decreases. This means that only informal firms are formalizing, but new formal firms are not being created. Moreover, many informal firms are being destroyed. This has some immediate consequences on jobs losses that can be seen in the right panel of Figure 8. The figure shows that for increasing enforcement, there is an increased fraction of jobs destructions, although tax revenues increase due to the formalization of informal firms.

Figure 8: Impact of Law Enforcement




On the other hand aggregate output has a U-shape in enforcement. For lower levels of enforcement, only small unconstrained firms formalize and total income goes down. However, as the probability of detection becomes larger, more productive entrepreneurs join the formal sector, are able to increase their productivity by borrowing more. This also implies higher tax revenues, hence the aggregate output goes up. The results show that an enforcement probability of 0.45 will completely eradicate informal entrepreneurship. In fact, high productivity informal enterprises will formalize while informal firms with with low productivity will shut down. The corresponding job loss is about 3%. These results may be seen as an upper bound of the loss. In reality, enforcement may imply more a lot of job losses. The simulations portrays only losses from workers who are employed in informal firms that are destroyed. It does not account for informal subsistence job losses (like street-vending) which represents about 90% of the informal activities. While an increase in the enforcement of formality status slightly increases the fraction of formal firms and the associated tax revenues, it has an overall perverse effect on the economy in terms of enterprise and job destruction, as well as aggregate income loss.

### 5.4 Business Training

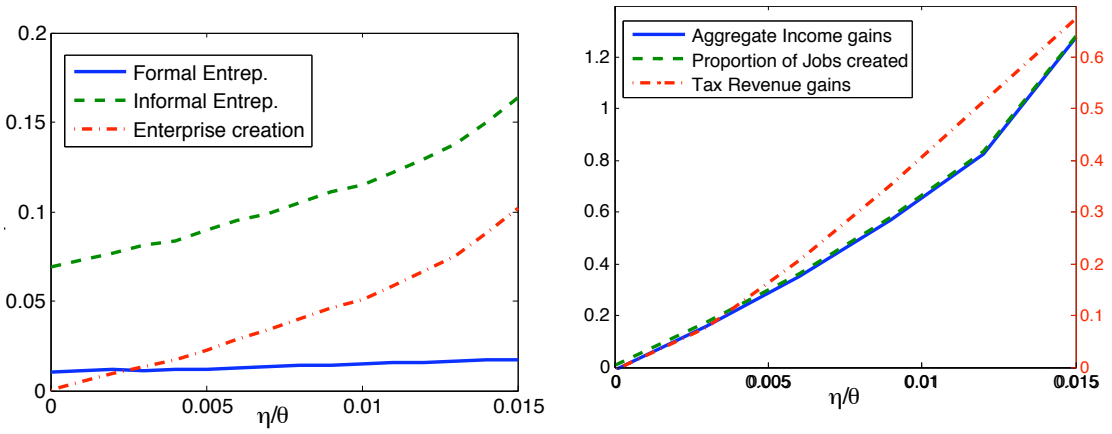
Several studies have found a positive effect of business training on business startups.However, how does such a policy affect formality in our context? Since the theory emphasizes that skills are a drivers of formality, we should expect that business training should positively

affect formality. However, it is not clear how much such a policy compares with the ones discussed above in terms of impact. In this section, I also examine the impact of business training programs as recently reviewed by McKenzie and (2013). The impact of a this policy in the model can be quantified by allowing individuals' talent to shift as a result of business training

$$\theta' = \theta + \eta, \quad 0 \leq \eta < \infty$$

where $\eta$ is the increase in ability inculcated by the training. Also, in this setting, I do not account for the costs incurred by financing and operating such a policy, and I also assume a perfect take-up rate. The impact is quantified for a range of increments $\eta$ in

Figure 9: Impact of Business Training Programs




ability due to business training starting from 0, the current state, corresponding to the situation where no program is initiated then progressively increases in terms of fractions of the initial average ability level $\eta/\theta$. The results are presented in Figure 9. They indicate that a small increase (from 0 to 1.5% of average talent) in entrepreneurial skills resulting from business training has a large impact on entrepreneurship - mostly through informal enterprise creation - and generate important tax revenues and aggregate income gains. Almost all the enterprise creation generated by a business training program are informal enterprises. As suggested by the theory, this implies the current average business skills in this economy is low and may explain why small increases in business skills only increase the likelihood of becoming informal entrepreneur (see Proposition 2).

Put together, these counterfactual policy experiments performed in this section suggest that economic growth and important welfare gains can obtain from reducing the cost of registration and tax rate while fostering entrepreneurial skills and enterprise creation through business training and better access to credit. The results also suggest, by contrast, that an enforcement policy that is not accompanied by other measures that provide incentives to formalize would have overall perverse effects in the economy, especially through enterprise destruction and job losses.

# 6 Conclusion

This paper provides an econometric analysis of the impact of individual and institutional factors on entrepreneurship and economic growth in countries with large informal sectors. I develop a simple theoretical model of occupational choice where agents differ with their initial wealth and ability, to analyze the implications of financial constraints and skills as well as taxation, law enforcement and entry sunk cost into the formal sector. The validity of the model and its main implications are first tested using reduced form estimates of transition probabilities between non-entrepreneurial work, informal and formal entrepreneurship with data from Cameroon, an economy where 90% of the labor force operates in the informal sector. The results suggest that while initial wealth and parents' entrepreneurial status are the main factors that influence the probability of becoming an informal entrepreneur, higher level of education is the additional ingredient that definitely drives formal entrepreneurship .

I also structurally estimate the theoretical model using maximum likelihood by matching the observed occupational status with the theoretical probability of each occupation, and assess the robustness of the structural estimates using sensitivity analysis and specification test. This approach is particularly useful in making explicit the contribution of institutional factors to the occupational patterns observed in the data. The results confirm that there are binding capital constraints in informal entrepreneurship and emphasize the role of education as an important proxy of entrepreneurial ability. Moreover, they show that while there are few entrepreneurs who are comfortable in the informal sector (unconstrained informal entrepreneurs), the remaining larger fraction of informal entrepreneurs have the potential to become formal, but may be discouraged by institutional factors such as the high entry sunk costs.

To further investigate this fact, the estimated model is subject to a set of counterfactual policy simulations where the impact of institutional factors are quantified. The counterfactual policy analysis supports this intuition and shows that if the state reduces the registration costs by half, they can induce twice as much formal enterprises and levy more than twice the amount of tax revenues that they currently collect. Similarly, there exists an optimal tax rate, set at half of the current rate, that would also induce twice as much formal enterprises and produce four-third of the current tax revenues. In contrast, a law enforcement policy whose objective is to increase the probability of detection would have an overall perverse effect in the economy in terms of firms shut downs and job losses.

To the best of my knowledge the structural test on the role of personal and institutional factors in entrepreneurship and informality choice that I perform in this study is new to the literature. The results show that entrepreneurship, the growth of a modern formal sector and important income gains can obtain from reducing the cost of registration and optimally choosing the tax rate while fostering entrepreneurial skills and enterprise creation through education, business training and better access to credit. The

choice to resort to a simple static model was imposed by data limitations and the computational complexity of structurally estimating dynamic models of occupational choice. Moreover, the role of risk has not been incorporated in the model. I will work to address these limitations in my future research.

# 7 Appendix

## 7.1 Proof of Proposition 1

*Proof.* (i) The critical threshold $\theta_W(z)$ is solution to the equation $w = \max\{\pi^I(\theta, z), \pi^F(\theta, z)\}$. Denote by $\theta_W^u$ and $\theta_{IW}$ the unique solutions of the equations $w = \pi^F(\theta, z)$ and $w = \pi^I(\theta, z)$ respectively. Then solving for these equations using the expressions given in (4) and (5) yields $\theta_W^u = \left(\frac{w + rc}{(1-\gamma)(1-p)}\right)^{1-\gamma} \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \equiv \theta^{**}$ and

$$\theta_{IW}(z) = \begin{cases} \left(\frac{w}{(1-\beta)(1-p)} + \frac{\lambda r z}{1-\beta}\right)^{1-\beta} (\lambda z)^{-\alpha} \left(\frac{w}{\beta}\right)^{\beta} & \text{if} \quad z \leq z^* \\ \left(\frac{w}{(1-\gamma)(1-p)}\right)^{1-\gamma} \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \equiv \theta^* & \text{otherwise} \end{cases}$$

where

$$z^* = \frac{\alpha w}{(1-\gamma)(1-p)\lambda r} \tag{11}$$

The desired solution is then defined by $\theta_W(z) = \min\{\theta_{FW}, \theta_{IW}(z)\}$, that is,

$$\theta_W(z) = \begin{cases} \theta_W^u = \left(\frac{w + rc}{(1-\gamma)(1-p)}\right)^{1-\gamma} \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \equiv \theta^{**} & \text{if} \quad z \leq z^{**} \\ \theta_W^m = \left(\frac{w}{(1-\beta)(1-p)} + \frac{\lambda r z}{1-\beta}\right)^{1-\beta} (\lambda z)^{-\alpha} \left(\frac{w}{\beta}\right)^{\beta} & z^{**} < z \leq z^* \\ \theta_W^l = \left(\frac{w}{(1-\gamma)(1-p)}\right)^{1-\gamma} \left(\frac{r}{\alpha}\right)^{\alpha} \left(\frac{w}{\beta}\right)^{\beta} \equiv \theta^* & \text{otherwise} \end{cases} \tag{12}$$

where the cut-off $z^{**}$ is given by

$$\left(\frac{w}{(1-\beta)(1-p)} + \frac{\lambda r z^{**}}{1-\beta}\right)^{1-\beta} \left(\frac{\lambda r z^{**}}{\alpha}\right)^{-\alpha} = \left(\frac{w + rc}{(1-\gamma)(1-p)}\right)^{1-\gamma} \tag{13}$$

(ii) Consider the function $V^{I,F}(z, \theta) = \pi^F(z, \theta) - \pi^I(z, \theta)$. For $\theta \leq \theta_c(z)$, we clearly have $V^{I,F}(z, \theta) < 0$. Suppose $\theta > \theta_c(z)$; then since $\tau > p$ and $\frac{1}{1-\gamma} > \frac{1}{1-\beta}$, $V^{I,F}(z, \theta)$ is strictly increasing in $\theta$, and moreover, $\lim_{\theta \to \infty} V^{I,F}(z, \theta) = +\infty$ for any given $z$. It follows by the intermediate value theorem that there exists a unique threshold $\theta_F = \theta_F(z) \in (\theta_c(z), \ \infty)$ such that $V^{I,F}(z, \theta_F) = 0$, that is,

$$\pi^F(z, \theta_F) = \pi^I(z, \theta_F) \tag{14}$$

Note that this equation does not have a closed form solution. I provide below an algorithm that numerically solves for this solution and show how it is simultaneously used for the structural estimation. □

## 7.2 Proof of Proposition 2

*Proof.* (i) Since $w$ is independent of $\theta$ and both $\pi^I(\theta, z)$ and $\pi^F(\theta, z)$ are strictly increasing in $\theta$, then $V^{W,I}(\theta, z)$ and $V^{W,F}(\theta, z)$ are clearly also strictly increasing in $\theta$.

(ii) For $\theta \leq \theta_c(z)$,

$$\frac{\partial V^{I,F}(z,\theta)}{\partial \theta} = -(\tau - p)\theta^{\frac{\gamma}{1-\gamma}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} < 0.$$

For $\theta > \theta_c(z)$,

$$\begin{aligned} \frac{\partial V^{I,F}(z,\theta)}{\partial \theta} &= (1-\tau)\theta^{\frac{\gamma}{1-\gamma}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} - (1-p)\theta^{\frac{\beta}{1-\beta}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\beta}} (\lambda z)^{\frac{\alpha}{1-\beta}} \\ &= (1-\tau)\theta^{\frac{\beta}{1-\beta}} \left(\frac{\alpha}{r}\right)^{\frac{\alpha}{1-\gamma}} \left(\frac{\beta}{w}\right)^{\frac{\beta}{1-\gamma}} \left[\theta^{\frac{\alpha}{(1-\gamma)(1-\beta)}} - \theta_m^{\frac{\alpha}{(1-\gamma)(1-\beta)}}\right], \end{aligned}$$

where

$$\theta_m = \theta_m(z) = \left(\frac{1-p}{1-\tau}\right)^{\frac{(1-\gamma)(1-\beta)}{\alpha}} \left(\frac{r}{\alpha}\right)^{1-\beta} \left(\frac{w}{\beta}\right)^{\beta} (\lambda z)^{1-\gamma}.$$

Hence, when $\theta < \theta_m$, $\frac{\partial V^{I,F}(z,\theta)}{\partial \theta} <$, and when $\theta \geq \theta_m$, $\frac{\partial V^{I,F}(z,\theta)}{\partial \theta} \geq 0$ □

## 7.3 Algorithm for the maximum likelihood estimation of the model

The structural estimation of the model parameters involves finding the root of the non-linear equations and using it in the likelihood function as an input for the maximization problem. The following algorithm is used for this purpose.

1. Define a grid for parameters $\alpha$ and $\beta$ within $[0.001, 0.999]$, say $\alpha(i) = 0.999\frac{i}{N}$, and $\beta(j) = 0.999\frac{j}{N}$, $i, j = 1, 2, \ldots, N$, where $N = 999$. Define a grid for $\lambda$ within $[1, 99]$, $\lambda(k) = 1 + 99\frac{k}{N}$
2. For each pair $(i, j)$ such that $\alpha(i) + \beta(j) < 1$,
   (a) Solve for $\pi^F(z, \theta_F) - \pi^{Iu}(z, \theta_F) = 0$. This yields a value $\theta_F^{ijk} = \theta_F(\alpha(i), \beta(j), \lambda(k), z)$, for each $z$.
   (b) Solve the maximization problem

$$\max_{\delta_0, \delta_1, \delta_2, \sigma} L_n(\delta_0, \delta_1, \delta_2, \alpha(i), \beta(j), \sigma, \lambda(k))$$

   This yields

$$L_n^{ijk} = L_n(\delta_0(i,j,k), \delta_1(i,j,k), \delta_2(i,j,k), \alpha(i), \beta(j), \sigma(i,j,k), \lambda(k))$$

3. The maximum likelihood is given by $L_n^{i^*,j^*,k^*} = \max_{1\leq i,j,k\leq N} L_n^{i,j,k}$, and the optimal parameter values are $\delta_0(i^*,j^*,k^*)$, $\delta_1(i^*,j^*,k^*)$, $\delta_2(i^*,j^*,k^*)$, $\alpha(i^*), \beta(j^*)$, $\sigma(i^*,j^*,k^*)$, $\lambda(k^*)$), where $i^*, j^*, k^*$ are the optimal indexes form the grids.

To implement the algorithm I used MATLAB routines starting from a wide variety of predetermined guesses. The standard errors of the estimated parameters are computed by bootstrap methods using 100 draws of the original sample with replacement.

# References


[1] Alby, P., Auriol, E., and Nguimkeu, P. (2013) "Social Barriers to Entrepreneurship in Africa: The Forced Mutual Help Hypothesis", Working Paper.

[2] Amaral, P. and E. Quentin (2010): "A Competitive Model Of The Informal Sector", Journal of Monetary Economics, 53, 15411553.

[3] Buera, F. 2006. Nonparametric Identification and Testable Implications of The Roy Model. Unpublished Manuscript.

[4] Blanchflower, D. and Oswald, A. (1998). What makes an Entrepreneur? Journal of Labor Economics 16(1): 26-60.

[5] Bruhn, M., (2011). "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico", Review of Economics and Statistics 93 (1), 382-386.

[6] Cagetti, M., and M. De Nardi. 2006. Entrepreneurship, frictions, and wealth. Journal of Political Economy, 114 (5): 835-70.

[7] De Mel, S., McKenzie, D., and , C. (2008), "Returns to Capital in Microenterprises: Evidence from a Field Experiment", The Quarterly Journal of Economics, 123(4), 1329-1372.

[8] De Soto, H. (1989). *The Other Path: The Invisible Revolution in the Third World*, I B Tauris & Co Ltd.

[9] Djankov, S., R. LaPorta, F. Lopez-De Silanes, AND A. Shleifer (2002). "The Regulation of Entry", Quarterly Journal of Economics, 117, 1-37.

[10] Evans, David S. and Boyan Jovanovic (1989). "An Estimated Model of Entrepreneurial Choice under Liquidity Constraints." Journal of Political Economy 97, 808- 827.

[11] Heckman, J., Honoré, B. (1990). The Empirical Content of the Roy Model., Econometrica, 58, 1121-49.

[12] Holtz-Eakin D., D. Joulfaian and H. Rosen, 1994, "Sticking It Out: Entrepreneurial Survival and Liquidity Constraints", The Journal of Political Economy, 102(1), 53-75.

[13] Hurst, E., and A. Lusardi. (2004). Liquidity constraints, wealth accumulation and entrepreneurship. Journal of Political Economy 112 (2): 319-47.

[14] International Labour Organisation (2002a), "Decent Work and the Informal Economy".

[15] Kaplan, D., Piedra, E., Seira, E., (2007). "Entry regulation and business start-ups: evidence from Mexico", Journal of Public Economics 95 (2011) 1501-1515

[16] Kremer,M., Lee,J.N. and Robinson, J.M. (2010) "The return to capital for small retailers in Kenya: Evidence from Inventories," Mimeo, Havard University

[17] Kuchta-Helbling C. (2000), "Background paper for the workshop Barriers to Participation: The Informal Sector in Emerging Democracies", Sao Paulo, Brazil November 13, 2000.

[18] Lentz, B. F., and Laband, D.N., (1990), "Entrepreneurial Success and Occupational Inheritance among Proprietors", The Canadian Journal of Economics, 23(3) 563-579

[19] Lucas Jr., R.E., (1978). "On the size distribution of business firms." Bell Journal of Economics 9, 508523.

[20] McKenzie, D. and , C. (2013). What are we Learning from Business Training and Entrepreneurship Evaluations around the Developing World? Quarterly Journal of Economics.

[21] Mondragón-Vélez, C. and Pena, X. (2009), "Business Ownership and Self-Employment in Developing Economies: The Colombian Case", Manuscript.

[22] Nguimkeu, P. (2014), "A structural econometric analysis of the informal sector heterogeneity", Journal of Development Economics 107 (2014) 175-191.

[23] Ordoñez, J.C., (2014). "Tax collection, the informal sector, and productivity", Review of Economic Dynamics 17 (2), 262286.

[24] Parker S. and Van Praag, M. (2010) "The Entrepreneur's Mode of Entry: Business Takeover or New Venture Start", Journal of Business Venturing, 2382

[25] Paulson, A., R. Townsend and A. Karaivanov, (2006), Distinguishing Limited Commitment from Moral Hazard in a Model of Entrepreneurship, Journal of Political Economy, 144(1), pp. 100-44.

[26] Piketty, T., (1997), The Dynamics of the Wealth Distribution and the Interest Rate with Credit Rationing", Review of Economic Studies, 64(2), pp. 173-89.

[27] Prado, M., (2011). "Government policy in the formal and informal sectors", European Economic Review 55 (8), 1120-1136.

[28] Schundeln, M., (2004), "Modeling Firm Dynamics to Identify the Cost of Financing Constraints in Ghanaian Manufacturing", Mimeo, Harvard University.

[29] Schneider, F., Enste, D.H., (2000). "Shadow economies: size, causes and consequences", Journal of Economic Literature 38, 77114.

[30] Udry, C. and Anagol S. (2006) " The return to capital in Ghana" American Economic Review 96 (2) 388-393.

[31] , C. (2013), "Registering for Growth: Tax and the Informal Sector in Developing Countries", The CAGE-Chatham House Series, No. 7, July 2013.