To Formalize or Not to Formalize? Comparisons of Microenterprise Data from Southern and East Africa
Alan Gelb, Taye Mengistae, Vijaya Ramachandran, and Manju Kedia Shah

Abstract

Why do firms choose to locate in the informal sector? Researchers often argue that the high cost of regulation prevents informal firms from becoming formal and productive. Our results point to a more nuanced story.

Using data from surveys of microenterprises in South Africa, Namibia, Botswana, Kenya, Uganda, Tanzania, and Rwanda, we find that the labor productivity of informal firms is virtually indistinguishable from that of formal firms in East Africa, but very different in Southern Africa. We provide a theoretical model to explain this result, based on the key assumption that firms may evade taxes subject to a cost (or concealment cost) that is increasing and convex in the firm's employment size. Consequently, the productivity distributions reflect the differences in concealment costs and the opportunity cost of formality. Greater enforcement of laws and better provision of services such as finance and electricity to formally registered firms in Southern Africa means that firms are more likely to register; those that do not are likely to be operating as "survivalist" firms. But in East Africa, weak enforcement of tax payment and no significant difference in access to services between formal and informal firms means that these variables do not explain the allocation of firms across the informal-formal divide.

We conclude that in countries with weak business environments, informal firms are just as likely as formal firms to increase their productivity as they grow. Thus, interventions to increase productivity and lower the cost of formality may be helpful. But in countries with strong business environments such as those in Southern Africa, owners of informal firms are likely to be better off entering the labor market as wage labor. In the latter case, investment in education or vocational training is probably more important.

www.cgdev.org

To Formalize or Not to Formalize?
Comparisons of Microenterprise Data from Southern and East Africa *

Alan Gelb
Taye Mengistae
Vijaya Ramachandran
Manju Kedia Shah

Revised July 2009

Introduction

Informal firms account for a large share of employment in most developing countries, often a dominant share. What drives the choice of whether a firm is informal or formal? How is that decision impacted by existing Government regulations and policies towards the MSME sector, and by the characteristics of the firms themselves?   How should informal firms be considered – as a mainspring of future growth and employment generation or as a "survivalist" employment alternative? Are there implications for policymaking?

This paper considers these questions for a set of countries in Southern and Eastern Africa, using data from the World Bank's Enterprise Surveys carried out in these countries in recent years.   These surveys have been completed for formal businesses in some 20 Sub-Saharan African (African) countries since 2005, and have been complemented by surveys for informal businesses in 10 of these countries. The content of the surveys has not been the same of course; far more detail is asked of formal firms in areas relating to production, costs, productivity and profitability. However, the surveys contain a number of similar questions, making it possible to derive some conclusions about the nature of formalization. In this paper, we use data from 7 of the countries in which surveys have been undertaken.

To provide a framework for the empirical analysis, Section 2 of the paper discusses a model of the determinants of the formalization decision. Like other decisions, many researchers view this decision as reflecting the balance of benefits and costs of formalization--for example, access to credit versus tax liability. This balance will depend on the formal regulatory framework but also on the capability and effort made by the state to enforce regulations. If tax inspectors can be bribed and firms not registered for taxes can also settle their problems by bribing inspectors, there may be little real difference whatever the formal tax rates. In any given country, the balance can also depend on the characteristics of the firm (or the entrepreneur). Some entrepreneurs may face different opportunity costs of being in business; some might also be better placed than others to achieve the potential access and productivity gains that come with formalization. We formulate a model that allows us to determine how firms "sort" into the formal vs. the informal sector.

Section 3 picks up on the empirical analysis. One question, on which studies are not unanimous, is how to define "formal". We use a uniform definition based on registration for taxes. We present some descriptive statistics for three sets of firms: formal micro-enterprises (5 employees or less), informal micro-enterprises (also 5 employees or less), and formal small firms with 5 to 10 employees.   These are compared across the seven countries: a Southern group consisting of South Africa, Namibia and Botswana and an Eastern group, Uganda, Tanzania and Kenya, and Rwanda.

Each country is unique, and there are important differences between countries in each group. Nevertheless, there appear to be systematic differences between the two groups, and these suggest very different stories in Southern and Eastern Africa. In particular,

informal firms seem to be more "survivalist" in the former region, and more likely to be potential sources of growth and employment generation in the latter one where the main differences are between smaller and larger firms rather than formal and informal firms. *Indeed, the decision of whether to formally register or not may be quite an idiosyncratic one in the East African group, in contrast to the Southern Africa group where informal firms are truly less productive than other types of firms.* Why this should be so is an interesting question, and we hypothesize that it reflects a combination of weak delivery of the services that are supposed to flow from formalization combined with weak ability to enforce formalization. Rwanda displays some distinctive characteristics relative to these two groups. Interestingly, some of the differences implied by the micro-data appear to be consistent with country-level characteristics, including from "Doing Business" and the Worldwide Governance Indicators.

Section 4 looks at the role of human capital and the business environment in the sorting of firms into formality and informality. We consider the costs of informality (avoidance of taxes and laws) and the benefits of formality (access to electricity, water, sewage, and finance) across the seven countries in our sample.

Section 5 summarizes, and draws some implications for policy. These differ between the Southern and Eastern groups. In particular, we argue that the differences between formal and informal firms in the Southern group reflect the characteristics of the entrepreneurs and labor force. Most of the survivalist firms in the informal sector (which is not to be confused with the small-scale sector) are unlikely to become formal firms. While certain initiatives may ease the situation of these firms, the approach towards improving the living standards of those engaged in this sector are most likely to be found in increasing access to employment in the formal sector, whether through temporary wage subsidies or other mechanisms. On the other hand, improvements in the business environment in East Africa are potentially more valuable in changing the balance of benefits and costs from formalization, and so encouraging small firms to formalize and grow.

2. A comparison of productivity in the informal and formal sector in seven African countries

Enterprise surveys covering the formal sector conducted for African countries cover only the formal sector and firms with a minimum employment level of 5 employees, but surveys in several countries have included a separate survey of microenterprises-those with less than five employees.   For various reasons, notably the small size of establishments, their expected high rate of turnover, the high level of "informality" of establishments in many activities and consequently the difficulty of obtaining trustworthy information from official sources, an aerial sampling approach is used to estimate the population of establishments and select the sample in this stratum for all regions of the survey. Table 1 describes the sample of firms used in this analysis.[1] Appendix 1 describes the sampling methodology.

Table 1: Sample Sizes

	Botswana	Namibia	South Africa	Kenya	Tanzania	Uganda	Rwanda
Informal micro	54	72	53	90	28	46	23
Formal micro	48	28	67	34	37	54	105
Total micro	**102**	**100**	**120**	**124**	**65**	**100**	**128**
Total formal	**342**	**329**	**937**	**657**	**419**	**563**	**212**

While enterprises in very small size classes are commonly referred to as "informal", a part of the firms with less than 5 employees in our sample is indeed classified as formal. This raises the question of how the boundary between formal and informal firms should be defined. Enterprises in the micro size class were asked about their registration status. Firms were asked whether they had:

- registered name with the Office of the Registrar or other government institutions responsible for approving company names,

- registered with the Office of the Registrar, the local courts, or other government institutions responsible for commercial registration,

- an operating or trade license or otherwise registered for a general business license with any municipal agency,

---

[1] We know that there will be some firms with more than 5 employees that are not registered for tax purposes in our small-firm sample.   But the scope of this paper is to look at the determinants of formality amongst microenterprises; thus the bulk of the analysis is confined to this subset of firms.

- obtained a tax identification number from the tax administration or other agency responsible for tax registration.

In all the countries in our sample, only a subset of firms in the first three categories are registered for tax purposes. We define firms are registered for taxes to be formal microenterprises, they are defined as informal otherwise. This distinction of microenterprises into these two categories follows earlier studies. It is clear that to grow into larger size classes, a firm must first formalize its operations by registering with tax authorities.

As a first step towards looking at the differences between formal and informal firms, we look at a simple measure of productivity—value added per worker—of informal microenterprises, formal microenterprises and formal small firms in each of the seven countries in our sample.

Figure 1: Labor productivity of firms in the sample (median)
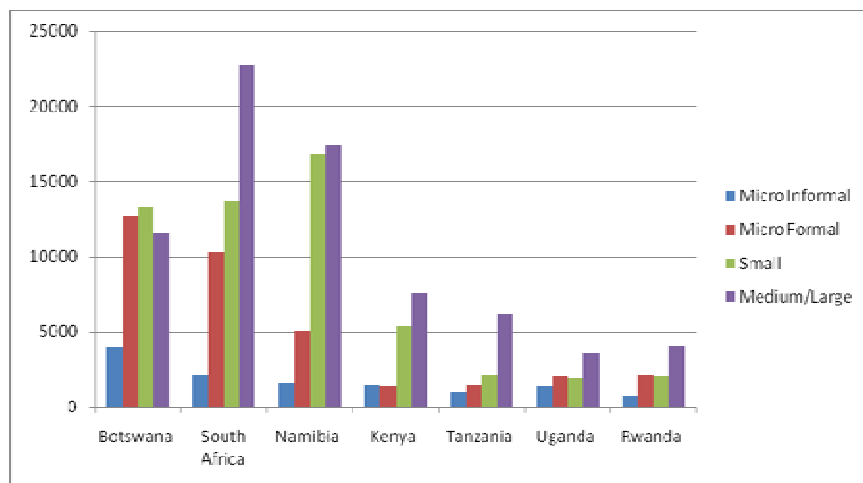
We begin our analysis by examining productivity differences across firm groups within countries. Figure 1 presents the median productivity between microenterprises and formal sector small firms in our sample. We see that, between Southern and Eastern Africa, differences in formal sector productivity are much greater than differences in microenterprise productivity. There are other large gaps as well--formal microenterprises are more productive in Southern Africa compared to their informal counterparts, and much more productive than formal microenterprises in Eastern Africa.

However, median values present only part of the picture. The dispersions around the median provide further information on the nature of enterprise performance, and competitiveness of the sector. Figures 2-8 show kernel density estimations for labor productivity across the three types of firms in South Africa, Botswana and Namibia (Southern Africa) and in Kenya, Tanzania and Uganda (East Africa) and Rwanda.

A striking feature of the figures below is that in some countries the probability density of labor productivity is bimodal along the formal vs. informal firms divide but not in others. Labor productivity is lower in informal firms on average in every country sample, and the density of labor productivity in those firms also significantly overlaps the density for formal firms. But the degree of overlap also varies a great deal from country sample to country sample, being much smaller in samples from Southern Africa than it is in country samples from East Africa. It is possible that this has something to do with differences in survey sample designs, but we also know that the sample design is reasonably standardized across countries. It is therefore likely that the smaller extent of overlap in labor productivity across the formal-informal divide in Southern Africa reflects differences in the structure and workings of the economies of the two regions. Specifically we hypothesize that it may indicate that governments in Southern Africa enforce business tax laws and codes of regulation more than their counterparts do in Eastern Africa and that the formal environment provides better services.

Figure 2: South Africa
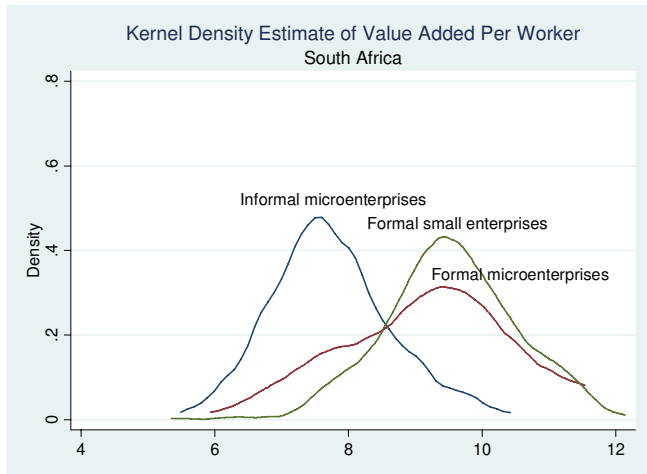
Figure 3: Botswana
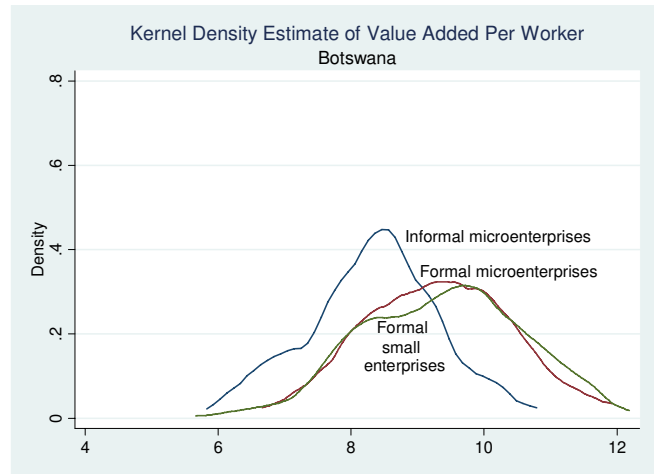
Figure 4: Namibia

Figure 5: Kenya

Kernel Density Estimate of Value Added Per Worker — Namibia

Kernel Density Estimate of Value Added Per Worker — Kenya

Figure 6: Tanzania
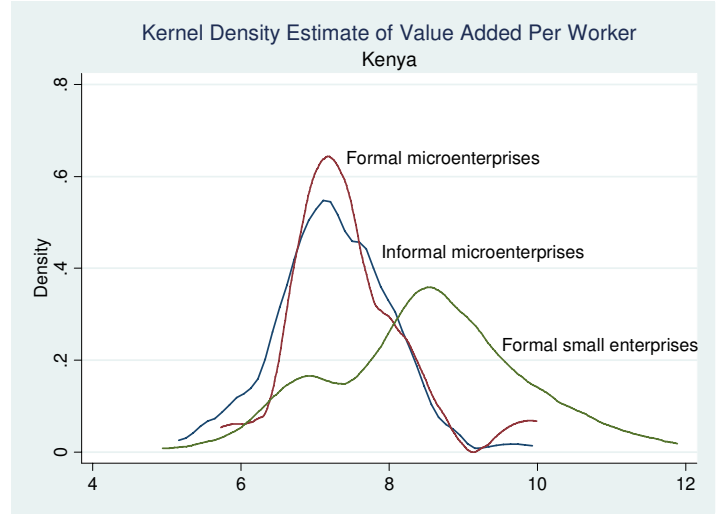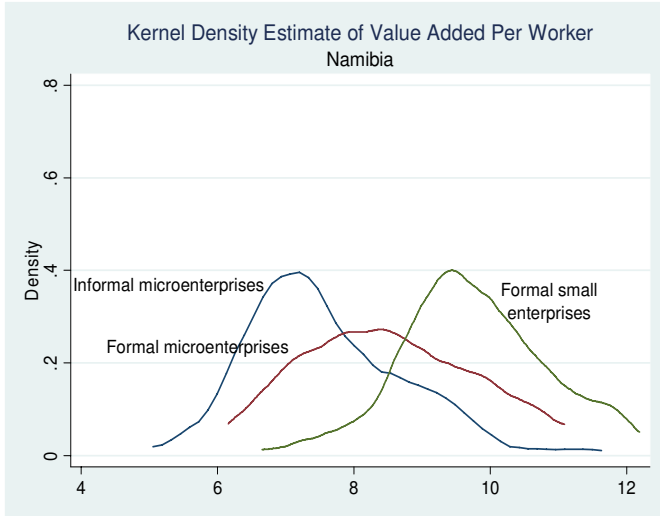
Kernel Density Estimate of Value Added Per Worker — Tanzania

Figure 7: Uganda
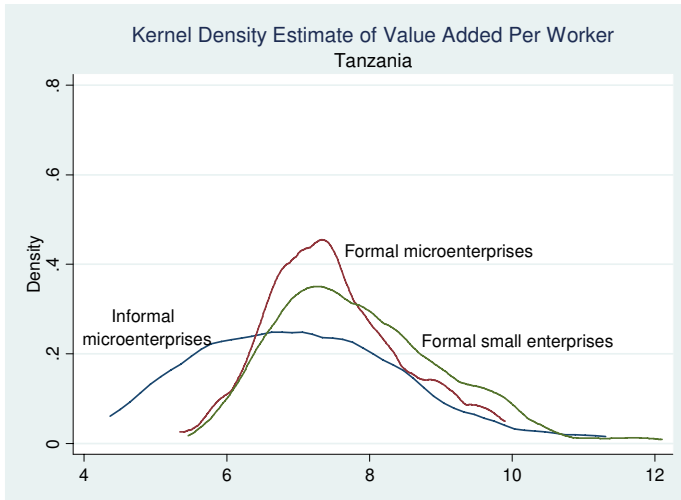
Kernel Density Estimate of Value Added Per Worker — Uganda
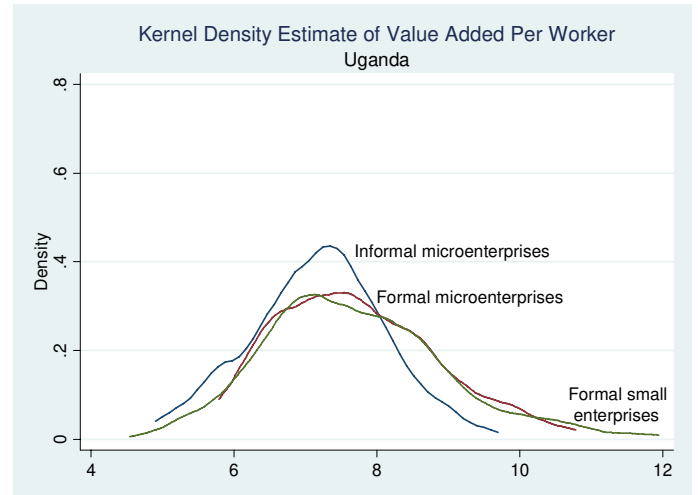
For the three countries in Southern Africa, the difference in labor productivity between formal and informal micro-firms is especially stark for South Africa. There the density functions barely overlap, and productivity in the modal formal firm is several times higher than in the modal informal firm. Indeed, in South Africa there almost no difference between the productivity density functions for micro formal firms and small formal firms. In Botswana and Namibia, the picture is muddier, yet with some similarities.

In contrast, the East African pattern is one with a large productivity overlap between firms that are informal versus those that are formally registered microenterprises. In Kenya there is some difference between formal and informal micro densities, but there is a great deal of overlap in Uganda and Tanzania; indeed, there is little or no difference between the densities for formal and informal micro-firms in these countries. In Uganda

the overlap is almost total across all three types of firms. In Tanzania and particularly Kenya, small firms tend to be somewhat more productive than micro firms.

Figure 8: Rwanda

Kernel Density Estimate of Value Added Per Worker
Rwanda
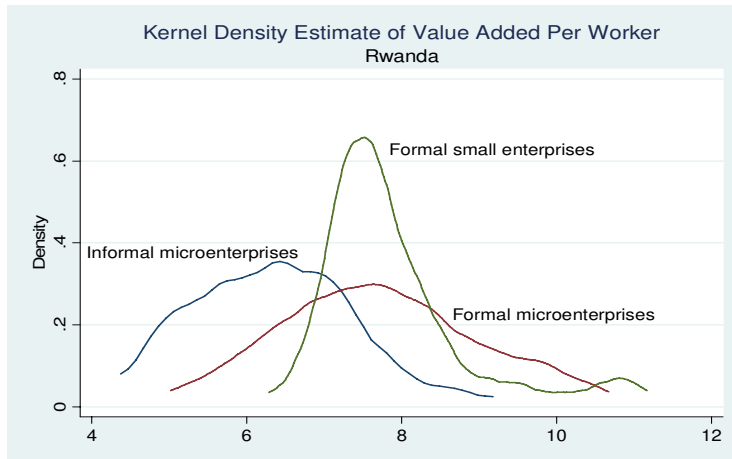
Rwanda is clearly a special case (Figure 8). As in Southern Africa, productivity of informal micro-firms is lower than for the formal micro and small firms. However, because of the stringent enforcement of registration requirements and punitive penalties for non-registration, there are almost no real informal micro firms in Rwanda. The survey found only 23 such firms, all very small and engaged in fringe activities.   Alone among the countries, in this case the question of whether or not to register cannot be regarded as a practical choice; the question appears to be more one of whether or not to establish a business.

Figures 9 and 10 show the kernel density estimations aggregated across East and Southern Africa. .[2] The picture is clear—in East Africa, informal and formal firms have more or less the same level of productivity; in Southern Africa, the picture is starkly different. It is also interesting to note that informal firms in Southern Africa look very similar to informal firms in East Africa, but formal firms are far more productive in Southern Africa than in East Africa.

Figure 9   Figure 10

---

[2] It is not strictly correct to pool data across the countries because of variations in price parity. One possibility is that the distributions overlap due to very high costs of informality that lead to the exit of the least efficient firms in the informal sector. But as we show below, the costs (and opportunity costs) of informality are much lower in East Africa than in South Africa.

Kernel Density Estimate of Value Added Per Worker
Kenya, Tanzania and Uganda

Kernel Density Estimate of Value Added Per Worker
Botswana, Namibia and South Africa
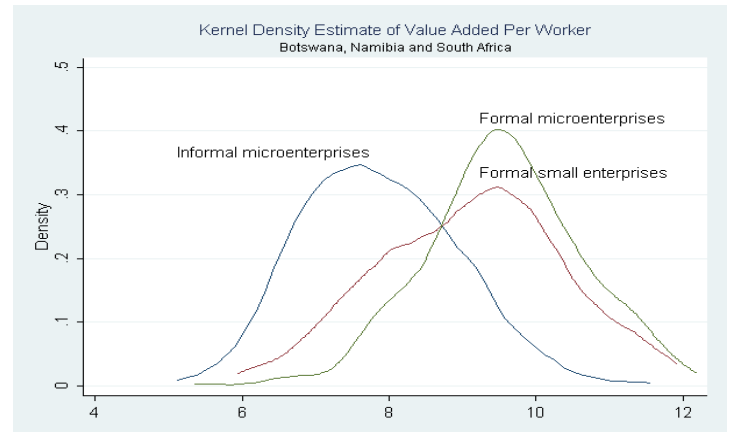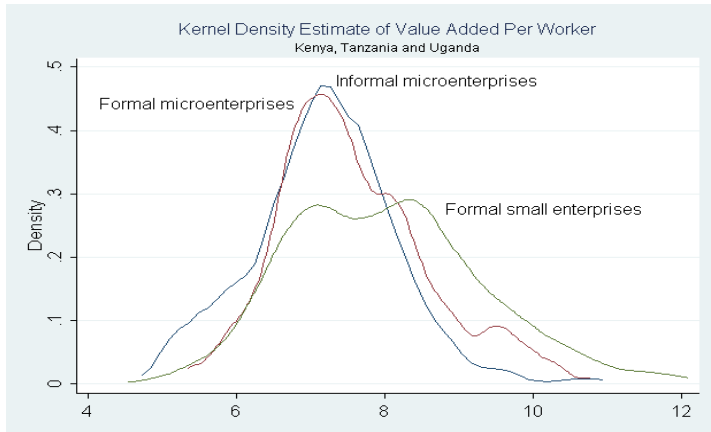
3. A Simple Framework for Understanding Productivity Differences

Perhaps the most visible proponent of the cost-benefit view of formalization is Hernando de Soto.   In his seminal analysis, *The Mystery of Capital,* he argues:

> "Extra legal businesses are taxed by the lack of good property law and continually having to hide their operations from the authorities. Because they are not incorporated, extralegal entrepreneurs cannot lure investors by selling shares; they cannot get low interest formal credit because they do not even have legal addresses. They cannot reduce risks by declaring limited liability or obtaining insurance coverage. In fact, the only 'insurance' available to them is that provided by their neighbors and the protection that local bullies or mafia are willing to sell them. Moreover, because extralegal entrepreneurs live in constant fear of government detection and extortion from corrupt officials, they are forced to split and compartmentalize their production facilities between many locations, thereby rarely achieving important economies of scale. With one eye always on the lookout for police, underground entrepreneurs cannot openly advertise to build up their clientele or make less costly bulk deliveries to customers."

The question then is why firms choose to endure this situation. DeSoto concluded that informal actors remained out of the formal economy due to the cost imposed by the regulatory structure within the country (DeSoto, 2000). Increasing unemployment, either through loss of public sector jobs or a growing labor force, has also been cited as a factor for the growth of the informal economy, which serves as a social safety net for the unemployed.[3] Whatever the mix of reasons, the informal economy in Africa is indeed large, both in terms of the number of enterprises as well as its contribution to GDP. Schneider estimates it to be about a third of total economic activity (Schneider, xx).

---

[3] . Recent unemployment estimates for the seven countries in this study range from 23% (for Botswana) to 56% (for Kenya).

The past few decades have witnessed the emergence of a large volume of literature on the informal sector. The literature is quite diverse, covering informal labor (Almeida, 2005; Chen, 2004), tax policy (Ihrig and Moe, 2004; Emram and Stiglitz, 2004), as well as corruption (Johnson et al, 1998; Marcouiller and Young, 1995). Many researchers argue that firms locate in the informal sector because the manager/entrepreneur believes that the benefits of informality outweigh the costs (Djankov et al 2002; Loayza et al 2005; Ishengoma and Kappel 2006). Djankov et al enumerate benefits and costs of informality and conclude that regulatory barriers must be lowered to encourage formalization. Gatti and Honorati look at informal firms in 49 developing countries to investigate the role of access to credit and external finance. The authors find that more tax compliance is significantly associated with more access to credit and that the link between credit and formality is stronger in "high-formality" countries (Gatti and Honorati, 2008). Arterido et al find that a weak business environment shifts downward the size distribution of firms and that poor quality finance and business regulation reduces the employment growth of micro and small firms in particular. They argue that significant reforms are needed to for micro firms to cross over into the category of SMEs (Arterido et al, 2007).

Loayza et al discuss the informal economy as a result of excessive taxes and regulation. They study the determinants and effects in an endogenous growth model and find that the size of the informal sector depends positively on the proxies for tax burden and labor restrictions and negatively on a proxy of the quality of government institutions.   This model was tested using country level data in Latin America (Loayza et al 2005). Fajnzylber et al (2005) look at the effect of improvements in the business environment on the performance of microenterprises in Mexico. Ishengoma and Kappel (2006) set up a formalization model around the costs and benefits in the business environment (formal and informal). They then assess approaches to encourage formalization in this paradigm and make recommendations for specific measures. Implicit is many of these analyses is the notion that informality impedes growth and acts as a drag on productivity.[4]

Shleifer and LaPorta argue that informal firms are neither capable of crossing over to formality when registration costs are lowered nor are they a burden on society (Shleifer and LaPorta, 2008). According to Shleifer and LaPorta, "there is very little support for the romantic view, and indeed the differences in productivity between the formal and informal firms are so large that it is very hard to believe that the registration of unregistered firms would eliminate the gap. On the other hand, there is not much support for the parasite view, either, and the evidence suggests that beefing up "enforcement" against the unofficial firms would devastate the livelihood of millions of people surviving near subsistence. The evidence rather points to the dual view, with the fairly standard

---

[4] Fafchamps (1994) discusses six factors that may govern this choice. Fafchamps argues that the observed informality is only a short-run disequilibrium phenomenon; however, given that the number of firms in each of these countries in this sector has grown rapidly, this is unlikely to be the case. High transport costs may limit a firm's market, it therefore produces on a small scale. However, this alone, cannot determine informality. Market failures, information asymmetries and management requirements play a role as well-- in each case, the demands are fewer on informal enterprises. Government policies and regulations such as registration procedures, costs, tax laws, labor regulations, worker safety laws can be avoided by informal firms. Informal firms, with flexible technologies, can adjust more easily to market demand. And finally, large scale production requires managerial skills that entrepreneurs may not have.

implication that the hope of economic development lies in the creation of large registered firms, run by educated managers and utilizing modern practices, including modern technology, marketing, and finance."

Are regulatory costs the driver of the decision to formalize? If so, why do we observe so little difference in productivity between informal and formal firms in East Africa relative to Southern Africa? In this section we spell out a partial equilibrium model of business informality that provides an explanation of the difference in labor productivity between informal and formal firms in Southern Africa and the contrast it bears to East Africa, where the overlap between the densities of the formal and informal sectors is almost complete.

The Model

We begin with a two factor model, following Lucas (1978). Factor prices are assumed to be given and fixed, along with the distribution of the workforce between employees and entrepreneurs. Our focus is on the sorting of observed entrepreneurs between the formal and the informal sectors rather than the question of selection into entrepreneurship which was the main focus in the Lucas model. In the next few paragraphs we describe our model and the key propositions that are derived from it. Appendix 2 describes each of the propositions in detail.

As in Lucas (1978), "managerial talent" level, x is drawn from a fixed distribution $\Gamma : R^+ \rightarrow [0,1]$. A firm in the economy consists of a single entrepreneur, managing n identical employees, who use k units of homogenous capital to produce y units of net output by means of a two-level production function given by $y = xg[f(n,k)]$, where $f,g : R^+ \rightarrow R^+$ and each of f and g is twice differentiable, increasing and strictly concave, with $g(0) = 0$. We also assume that the production technology f is one of constant returns to scale so that, denoting factor proportion as $r = \dfrac{k}{n}$, $f(n,k) = n\phi(r)$, where $\phi : R^+ \rightarrow R^+$ is also twice differentiable, increasing and strictly concave. Firm face an identical wage rate, w, in the labor market and a uniform rental price for capital services, u, and are assumed to maximize profits in their production and hiring decisions.

Following Fortin et al. (1997), we assume that all firms are required to pay a profit tax at the uniform rate of t_π and a payroll tax t_w. *Firms may evade both taxes subject to a cost of evasion (or concealment cost) that is increasing and convex in the firm's employment size.* The concealment cost can be described as $c = c(n)$ such that $c'(n), c''(n) > 0$. We assume that all formal firms pay both taxes while all informal firms evade both.

The degree of convexity varies with the overall strength with business environment. It is greater in countries where regulations and laws are better enforced, both because

10

detection is more difficult to avoid and because the opportunity costs of informality, in the form of access to services enabled through formalization, are higher.

 It follows from our assumptions that for any firm x, profits are given by

(1a) $\pi_f = (1-t_\pi)\left[xg(n\phi(r)) - (1+t_w)wn - uk\right]$

if the firm operates formally, and by

(1b) $\pi_i = xg(n\phi(r)) - c(n) - wn - uk$

if x is choose to be informal instead. If x is in the formal sector, its input demands would maximize (1a), for which the first order conditions are

(2a) $xg'[n\phi(r)][\phi(r) - r\phi'(r)] = (1+t_w)w$
 and
(3) $xg'[n\phi(r)]\phi'(r) = u$,

and from which it follows that

(4a) $\dfrac{\phi(r) - r\phi'(r)}{\phi'(r)} = \dfrac{(1+t_w)w}{u}$

 Equation (4a) defines the factor proportion, r as an implicit function of factor prices and shows that formal sector firms should exhibit uniform capital intensity under our technological assumptions.

The firm's labor demand function is implicitly defined by each of equations (2a) and (3), and is of the form $n_f = n_f(x, w, t_w, u)$. Because of the uniformity of r, the profit maximizing demand for capital services is $k_f = k_f(x, w, t_w, u) = rn_f(x, w, t_w, u)$. Employment size is therefore a reliable indicator of the scale of output of formal firms.

Since the cost of concealment is a function of firm size, for informal firms (3) will still be one of the first order conditions for its input demands, but equation (2a) will not be. The other first order condition will be

(2b) $xg'[n\phi(r)][\phi(r) - r\phi'(r)] = w + c'(n)$

 Upon division by (3) this leads to

(4b) $\dfrac{\phi(r) - r\phi'(r)}{\phi'(r)} = \dfrac{w + c'(n)}{u}$

11

which shows that r is not uniform for informal firms since it depends on n and, hence, on x. Still, (2b) and (3) together implicitly define the two factor demand functions with $x, w, \ u$ as the sole arguments when x is an informal firm. The labor demand function is now of form $n_i = n_i(x, w,, u)$

to which corresponds an capital input demand function of the form $k_i = k_i(x, w,, u)$

The labor demands n_f and n_i have the standard property of strictly decreasing in w and u. They are also both strictly increasing in x. The result in Lucas (1978) that more talented entrepreneurs run larger businesses thus holds in this setting too, applying in the informal sector as well as in the formal sector. [5]

Substituting for input demands into equation (1) gives entrepreneurial rent in formal sector as $\pi_f = \pi_f(x, w, u, t_w, t_\pi)$. This is a standard profit function except that managerial talent x (scaled down by profit taxes) plays the role that output price would in an ordinary profit function, as we are using output as the numeraire. Profits are therefore strictly increasing in x by the envelope theorem and strictly decreasing in both factor prices and the two tax rates. Thus, not only do more talented managers run larger businesses, they also earn more entrepreneurial rent than those who run smaller firms.

Obviously these properties also carry over to the entrepreneurial rent function of the informal sector. This is obtained by substituting for labor and capital input $n_i(x, w, u)$ and $k_i(x, w, u)$, respectively, in equation (1b), which leads to $\pi_i = \pi_i(x, w, u)$

Each of π_f and π_i is defined for all entrepreneurs $x : x \geq x_0$ for each of which obviously only one of the two variables is realized. We assume that the other is nonetheless known to the entrepreneur as a notional value and that x is in the formal sector if $\pi_f(x, w, u, t_w, t_\pi) \geq \pi_i(x, w, u)$ and in the informal sector otherwise.

---

[5] There may be an additional sorting mechanism relating to labor. In the Lucas model (and also the other papers) the workforce is homogeneous with respect to productivity as an employee. The only personal differences are those relating to management. In the Lucas model end, the better managers get sucked into management and they also therefore earn more than they would as employees. However, in the Rauch model it is not so simple. There, the workforce is still homogeneous in terms of labor productivity, but some informal sector managers earn less than the minimum wage in the formal sector and would like to be formal sector employees. In practice however, we know that there is an important difference in the real world. Especially in the education dimension, the same factors that will make a manager more productive will also increase employee productivity and wages, especially with the large wage gradient in the formal sector. There will therefore be another selection process operating on managers, but through the labor market, if an informal business cannot generate enough surplus to compensate an educated person. This might be considered as another opportunity cost of informalization, and the convexity will depend on the degree of wage premium for an educated person. In East Africa, it may be the case that the gradient is not as steep or the formal sector is not big enough to absorb all educated applicants so that many of these become informal entrepreneurs. For example, Fox and Oviedo show that the wage differences between unskilled and skilled production workers is much higher in Namibia and Botswana versus Tanzania and Uganda (Fox and Oviedo, 2008).

This is the sorting rule motivating the econometric estimations that are presented in Section 4.

The key propositions of our model are as follows:

Proposition 1: *The entrepreneurs with the greatest managerial talent are always in the formal sector.*

Proposition 2 : *Higher taxes would push more talented entrepreneurs into the informal sector.*

Proposition 3: *Informal firms are always smaller than formal firms.*

Proposition 4: *There is discontinuity between the size distribution of formal firms and informal firms.*

Proposition 5: *Informal entrepreneurs earn less than formal entrepreneurs.*

Proposition 6: *Informal firms are less productive than formal firms.*

Proposition 7: *The distribution of labor productivity of firms in the formal and informal sectors overlap with each other.*

Proposition 8: *The extent of this overlap decreases in the convexity of concealment costs in the informal sector.*

Our model may not capture all the nuances of the decision to formalize. Other factors that drive this decision may include the following:

- Different business climates will have different costs and benefits of formality/informality. This may in turn give rise to differential probabilities of access to services such as electricity, or finance. Also, the quality of the service even if firms have access may vary across firms and regions. Furthermore, protection from predatory officials and the ability to evade taxes may vary.

- In the limit, there may be no clear difference when both regimes are very uncertain and arbitrary (Emery, 2005). In this case, it is not clear what formal or informal means. We would not expect a sorting effect with better managers choosing formality in the limit case.

- This points us towards a directly stochastic model, where self-selection probabilities depend on multiple expected differences and firm characteristics.

13

- In our model, c(.) can be expanded to include a wider range of opportunity costs of informality (generally higher in better business climates and especially for more talented managers) but the function might look very different.

- Finally, we note that the opportunity cost of informality is also higher for skilled persons (whether potential managers or employees) if there is a steep pay gradient in the formal sector. This raises the hurdle for the informal sector to retain skilled individuals.

We now turn to empirical tests of the propositions identified by our model.

4. Sorting by Human Capital, and by the Benefits and Costs of Formality

Human Capital

First, let us look at whether there is sorting by differences in the level of managerial talent, as proxied by educational attainment. Figure 11 compares the educational attainments of the entrepreneurs in the informal and formal sectors across the six countries. (Appendix 3 shows educational attainment in all categories for informal and formal microenterprises). We are far more likely to see owners with university degrees operating firms in the informal sector in East Africa than in Southern Africa, where firms are clearly sorted according to managerial talent.

Figure 11: Pct of Firm Owners with Primary or University Education
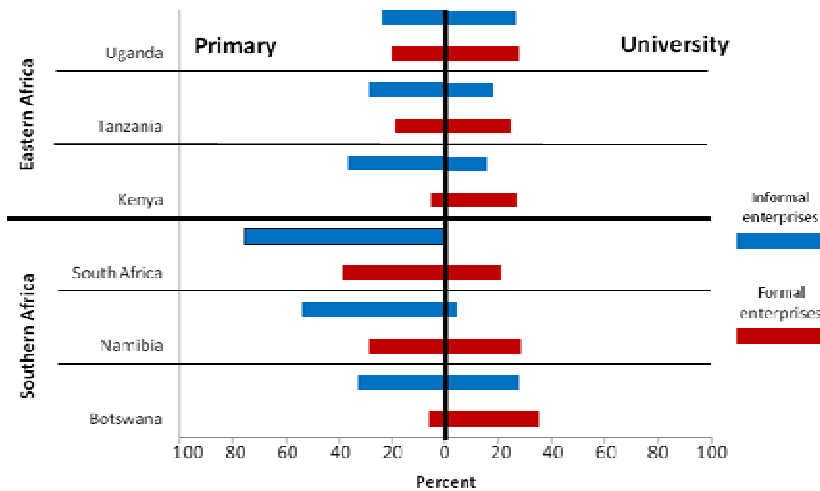
Informal entrepreneurs have lower levels of education than those in formal firms in all countries except Tanzania and Uganda where the distinction is insignificant. The sharpest difference by far is in South Africa; it is astonishing that there, despite the relatively strong human capital base of the country, informal entrepreneurs have lower levels of education than comparable businessmen in any other country. One possible factor is that the opportunity cost of being an entrepreneur in the informal sector, in the form of a high

14

formal sector wage, is particularly high in the South where skill-based wage gradients are relatively steep. Namibia comes second; there is also a sizeable gap in Botswana and Kenya.

Race is not a direct sorting mechanism in our analysis because it will proxy for the capability-related variables, in particular education. The vast majority of microenterprises our sample are Black-owned (Table 2). South Africa has the smallest share of micro formals (67 percent) that are Black-owned while Uganda has the highest at 100 percent. Histories of race discrimination will have fed into the education system and may affect the nature of the business environment (for example, access to finance or other services) across or within the countries in our sample--our model does allow us to investigate the effects of these variables on the decision to formalize.

Table 2: Share of Black-Owned Firms in the Sample

	Informal micro	Formal micro	Formal small
Botswana	0.96	0.90	0.76
Kenya	0.99	0.79	0.63
Rwanda	1.00	0.96	0.86
South Africa	0.92	0.67	0.39
Tanzania	0.96	0.97	0.77
Uganda	0.96	1.00	0.87
Namibia	0.99	0.93	0.74

From other studies across Africa, education is a key driver of rates of firm growth, except for certain "networked" firms owned and operated by ethnic minorities (Ramachandran and Shah, 2006). As discussed further below, there is also a strong relationship between educational attainment and productivity in Southern Africa, as well as between both of these variables and the likelihood of being formally registered. In contrast, there is little difference in education levels between formal and informal firms in Tanzania and Uganda, and also little productivity gap.

The model described in the previous section assumes that concealment costs are convex in employment size, so that the average as well as the marginal concealment cost will increase as a firm gets larger. The degree of convexity will depend on the overall strength of the business environment, as reflected in the enforcement of business tax laws and

codes of regulation as well as in the higher opportunity cost of formality (the provision of public services such as electricity etc). To the extent that enforcement and the provision of public services are characterized by a high degree of arbitrariness and variability, the concealment function will be less sharply convex. Therefore, it would pay for an entrepreneur in one country to remain informal while a similar entrepreneur in another country would be better off by operating formally.

The productivity and education comparisons suggest that the decision to register formally in East Africa might be a somewhat random or idiosyncratic one, as costs and opportunity costs of formality remain low across the size distribution of firms. The benefits of formal registration might not be realized because of poor delivery of financial or other services, or because the main concern of businesses—such as reliable power—is similar for formal and informal firms. Conversely, the costs of non-formalization might not be so high either, perhaps because inspectors are equally likely to seek bribes whether a firm is formal or informal. Other work (Emery, 2003) has drawn attention to a high degree of arbitrariness in the business climate in many African countries. In contrast, if the business climates of the Southern African countries are relatively strong (taking into account enforcement as well as the *de jure* constraints tabulated in *Doing Business);* individuals with the capabilities to benefit from the opportunities opened up by formalization will tend to formalize.

Benefits of Formality.

The benefits of formality lie in the access to public services and business support services, access to formal banking sector or micro finance, and availability of basic infrastructure facilities such as electricity, telephone, public sewage, as well as premises. In countries where these services are well developed, we would expect the informal sector to be smaller, and a sharper difference between the performance characteristics of informal versus formal microenterprises, *ceteris paribus*-only firms that have very low productivity/profitability will choose informality.

Examining reported infrastructure access (electricity, water and sewage), presented in Figures 12-16 we see that for firms in all six countries, infrastructure constraints are in fact greater for informal firms, compared to formal enterprises. Infrastructure access could provide incentives for firms to formalize. The benefits to formalization are clearly higher in terms of infrastructure for Southern African firms but much less so in Uganda and Tanzania. In Kenya, the difference in access to infrastructure is fairly large.

Figure 12: Pct with electrical connections   Figure 13: Pct with water connections
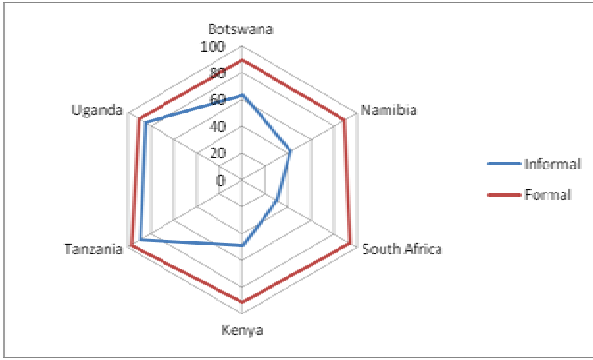
Figure 14: Pct with public sewage facilities
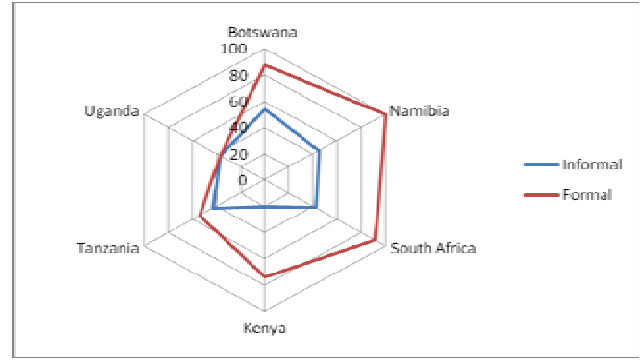
Figure 15: Percentage with externally-financed working capital
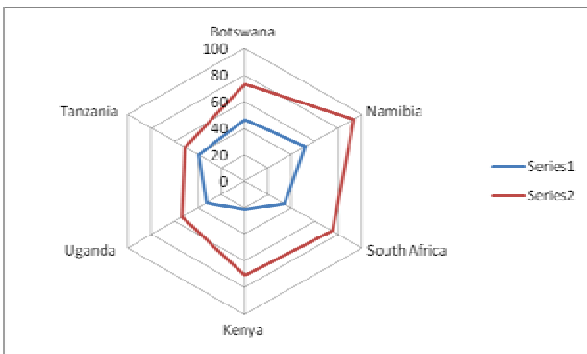
Figure 16: Percentage with a bank account
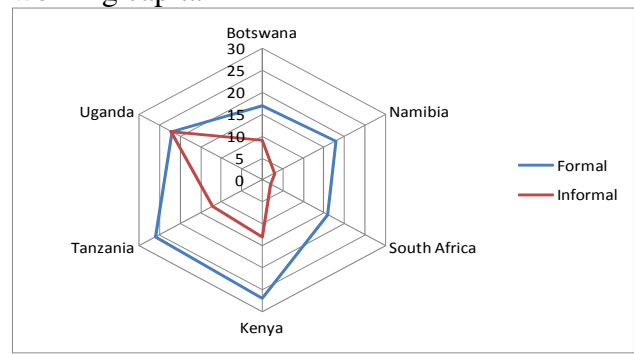
The data on access to finance also show some differences between the two sets of countries. In all of the East African countries, access to a loan is actually higher for informal firms than for such firms in any of the Southern African countries. Figure 15 shows that access to external sources of finance for working capital is higher for informal

17

firms in East Africa than in Southern Africa. Figure 16 shows that a large percentage of firms in Uganda and Tanzania have bank accounts even though they are in the informal sector. Finally, Table 3 shows that informal microenterprises finance a much higher share of working capital externally in East Africa than in Southern Africa.

Table 3: External sources of working capital, as share of all working capital
Percentages

	Formal microenterprises	Formal small enterprises	Informal microenterprises
Botswana	17	32	9
Namibia	18	39	3
South Africa	16	32	2
Kenya	27	36	13
Tanzania	26	26	12
Uganda	22	24	22

Costs of Formality

Another factor driving the decision to become formal is the cost of formality. A firm with a given level of productivity may choose informality in a country which has higher tax rates and much greater compliance costs, compared to countries where such costs are low. Figure 17 shows that a very high share of both formal and informal firms in Kenya and Uganda pay bribes and a lower but equal share of firms in the formal and informal sector in Tanzania pay bribes as well. In South Africa, the share of firms making bribe payments in the informal sector is higher than in the formal sector, suggesting that the rule of law provides a degree of protection to formal firms.[6]

Figure 17: Percentage reporting that bribe payments are required
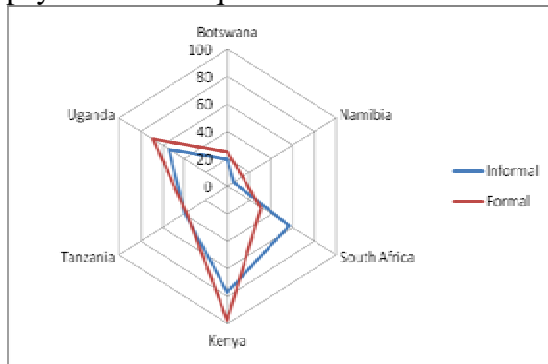
Figure 18: Percentage visited by tax inspectors

---
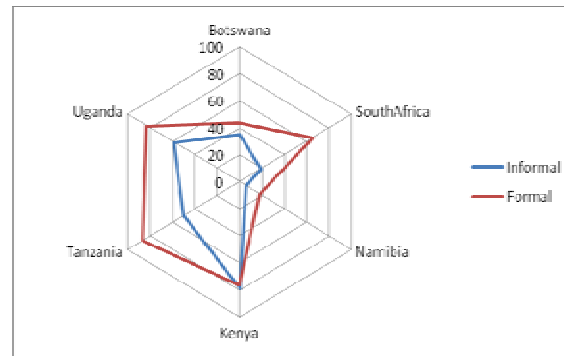
[6] Kaplinsky 1995 and Manning and Mashigo 1993 note that both small informal firms and owner-operators are particularly likely to suffer from violence and insecurity in South Africa.

Figure 18 indicates that visits from tax inspectors are likely to be higher for formal firms in every country but the share of firms visited in East Africa is much higher than in Southern Africa. Kenyan firms are equally likely to be visited, whether or not they are in the informal sector. In Southern Africa, harassment by tax inspectors is a lot lower, especially for firms in the informal sector.

Country indicators are consistent our findings that countries within the Southern and Eastern African block differ considerably in their business environments. Botswana, Namibia and South Africa are middle income countries, with relatively good governance and regulatory environments, while countries in East Africa are lower income, with mixed regulatory and governance conditions. Table 4 shows a number of indicators that might be relevant to understanding the business environment across the seven countries in our sample. Overall, these external sources of data corroborate the country-level differences that we identify in our descriptive charts above.

Table 4: Indicators of the Business Environment

	GDP per Capita	Doing Business Rankings (2007)	Institutional Investors Credit Rating	Voice and Accountability (2007)	Government Effectiveness (2007)	Control of Corruption (2007)
Botswana	5180	48	63.9	0.49	0.70	0.90
Kenya	530	83	27.0	-0.06	-0.59	-0.94
Namibia	2990	42	46.9	0.58	0.17	0.19
Rwanda	230	158	15.8	-1.24	-0.37	-0.09
South Africa	4960	32	63.3	0.74	0.72	0.32
Tanzania	340	142	25.3	0.15	-0.42	-0.45
Uganda	280	107	25.5	-0.47	-0.40	-0.76

Sources: World Bank Doing Business Indicators, Kaufmann, Kraay and Mastruzzi Governance Indicators, Institutional Investor.

South Africa, Botswana and Namibia consistently score well in terms of governance dimensions relative to the East African group. Rwanda emerges as a distinctive case— high scores relative to East Africa in dimensions related to government effectiveness and control of corruption but low in the area of voice and accountability – a pattern very consistent with the tight control over informal business noted above. Similarly, the *Doing Business* indicators suggest a higher quality business environment in the Southern countries.

5. The probability of formalizing

In this section, we summarize the results of Probit models that estimate the likelihood of formalization (the full regressions are presented in Appendix 4).[7] Table 5 shows the significance of size, education, and measures related to the business environment in determining formality. Overall, the empirical results are consistent with the predictions of the model. The probability of being formal is correlated with firm size in every country except Rwanda. Managerial talent (as proxied by educational attainment) is significant in determining likelihood of registering, *ceteris paribus*, for firms in Southern Africa and Kenya. But it does not play a role in determining formality in Tanzania, Uganda or Rwanda. Female ownership is correlated with informality in only two countries—Namibia and Rwanda—in all others, it is not a significant driver of firm status.

The benefits of formality (access to public services, external financing) are clearly driving the decision to formalize in Southern Africa and Kenya, while in Tanzania and Uganda, they are not likely to drive the decision. Interestingly, in Namibia, Uganda, Kenya, and Rwanda the likelihood of bribe payments increases the chances of formalization suggesting that formalization does not necessarily increase access to the protection of the law. In South Africa, however, the opposite is true. The regressions also confirm that access to finance is likely to be a significant driver of the decision to formalize, except in Tanzania and Uganda.

---

[7] Inequality (6) from section 3 provides a structure for the Probits, with the difference
$$\pi_D = \pi_f(x, w, u, t_w, t_\pi) - \pi_i(x, w, u)$$
providing the latent variable underlying the formality dummy D such that $D = 1$ for firms for which $\pi_D \geq 0$ and $D = 0$ if $\pi_D < 0$, where we assume that we can express π_D a stochastic function of a vector of observable determinants of or proxies, Z, of management talent, x, factor prices, w and u, and tax rates, t_w and t_π such that $\pi_{Di} = \gamma' Z + \varepsilon_i$, where ε_i a random error tem orthogonal to Z and summing up a host of other unobservable determinants of π_D.

Table 5: Determinants of formalization
(X denotes statistical significance)

	Botswana	Namibia	South Africa	Tanzania	Uganda	Kenya	Rwanda
Firm size	X	X	X	X	X	X	
Inspector visits			X	X	X		
Access to public services	X	X	X			X	
Owner has university education	X	X	X			X	
Owner is female		X(-)					X(-)
External financing	X	X	XX			X	XX
Bank account	X	X				X	X
Pays bribes		X	X(-)		X	X	X

Note: X notes significance at either the 5 percent or the 1 percent level.   XX notes cases where very few informal firms have access to external financing and a coefficient cannot therefore be estimated.

6 Conclusion

This paper has shown that the relative characteristics of informal and formal firms can differ across countries, by showing the distinct patterns that emerge from an analysis covering East and Southern Africa. Many studies simply assign firms to formal and informal categories based on their size. Here we are able to assign micro firms (those with less than 5 employees) into formal and informal categories on the basis of tax registration, and also to compare them with small formal firms with between 5 and 10 employees.

In Southern Africa, and particularly in South Africa, there is a sharp bifurcation of the micro firms.   Formal micro-firms have high labor productivity, at very similar levels to the small formal firms. This suggests that some sorting mechanism is at work, creating incentives for some entrepreneurs and their firms to operate in the formal sector. These formal micro-entrepreneurs are likely to have the capabilities to grow their firms, at least into small businesses; indeed, a survey suggests considerable upward size mobility within the South African formal sector (Clarke et al, 2007).

Informal firms in Southern Africa on the other hand have very low labor productivity, as little as one-tenth that of the formal sector. We believe that business environments are less likely to be the cause of this bifurcation. Compliance costs tend to be lower and infrastructure and financial services are more likely to be delivered in the South if firms become formal. Indeed, vulnerability to bribe-taking officials (which is sometimes cited as a disincentive to formalization) is greater among informal firms. Yet is seems that informal forms are not fully denied services; a few use banking services and other amenities.

What does this mean for the informal firms in Southern Africa? Our data suggest that, first, it is vital to distinguish micro-firms from informal firms. Many of the latter face internal barriers, in terms of management abilities and productivity, that will make it very difficult to cross over the productivity chasm that separates informal from formal firms. In South Africa for example, not one single firm in the informal sample was run by a university graduate. In addition to the other sorting factors, the opportunity cost of being a formal sector employee will be high for such individuals. This is in marked contrast to East Africa, where many informal firms are run by individuals with relatively high educational qualifications. For many firms in the informal sector in Southern Africa, it is not clear that they would be able to take advantage of the potentially expanded range of opportunities that could be expected to come from fully formal status. While there will always be exceptions, the informal sector in Southern Africa captured by the survey appears to be essentially composed of survivalist firms.

What might help these firms? Access to training programs and simple technical assistance may help to some degree. One of the areas that the South African informal firms complain about in the survey is lack of access to premises.[8] Some provisions here might ease their survival. But expanding employment in the formal sector, including in the many high-productivity micro-firms that are shown to exist by the survey, might be more useful than trying to "grow" informal firms. Whether such efforts should include entry wage subsidies or other approaches is beyond the scope of this paper. Given the low educational status of most informal entrepreneurs, it is difficult to see a route to higher-productivity employment that does not include the building of human capital, though this will be a lengthy process.

The picture is quite different in East Africa. There, especially in Uganda and Tanzania, there are indications that essentially similar firms make different decisions on whether to become fully formal. The informal sector includes many higher-productivity micro-enterprises with the capabilities to formalize and grow; the entrepreneurs running these firms are often well-educated and have access to a range of business services. Why these firms make the decisions they do plausibly relates to the business environment and its various weaknesses. Our analysis suggests that there may well be gains from improving the business environment to "crowd in" these firms into the formal sector and to widen

---

[8] Rogerson 1992 notes the legacy of slum clearance and other periodic attacks on the illegal spaces within which informal enterprises thrive. While many restrictions have been progressively lifted, there were still lingering controls and repressive bye-laws at the time of the survey.

the pool of formal micro- firms that can serve as a seedbed for small formal firms and beyond. This is not to say that efforts in these countries should not also include the building of human skills. From these efforts, we would also, however, expect duality to emerge. In a stronger business environment, the informal sector will more likely be composed of survivalist firms.

The case of Rwanda raises a distinctive question. How stringent should formalization requirements be at an early stage of development? Does a stringent registration and tax policy lead to a well-regulated and increasingly productive micro-seedbed for growth? Or, will it stultify the growth of the entire small-scale business sector? Only time will tell.

Appendices

Appendix 1: A brief note on sampling

Survey coverage

The World Bank Enterprise Surveys in Sub-Saharan Africa cover establishments located in all major regions of a country, in the following industries (according to ISIC Revision 3.1)--all manufacturing sectors (group D), construction (group F), retail and wholesale services (sub-groups 52 and 51 of group G), hotels and restaurants (group H), transport, storage, and communications (group I), and computer and related activities (sub-group 72 of group K). For establishments with five or more full-time permanent paid employees, this universe was stratified according to the following categories of industry:

1. Manufacturing: Food and Beverages (Group D, sub-group 15);
2. Manufacturing: Garment (Group D, sub group 18);
3. Manufacturing: Other Manufacturing (Group D excluding sub-groups 15 and 18);
4. Retail Trade: (Group G, sub-group 52);
5. Rest of the universe of firms, including:

    - Construction (Group F);
    - Wholesale trade (Group G, sub-group 51);
    - Hotels, bars and restaurants (Group H);
    - Transportation, storage and communications (Group I);
    - Computer related activities (Group K, sub-group 72).

The survey also sampled a selection of micro establishments (establishments with less than five full-time permanent paid employees) from the targeted universe, without stratification by industry.

Sampling methodology

The sampling methodology implemented varied by strata and region, depending on the characteristics of the population. With respect to establishments with five or more employees, the following approaches were applied (according to survey region and industrial stratum):

For the three manufacturing strata, a required second level of stratification (after sector) was firm size, defined as: small (5-19 employees), medium (20-99 employees) and large (100 and more employees).

In all surveys, the micro establishment stratum covers all establishments of the targeted categories of economic activity with less than 5 employees. For many reasons including the small size of establishments, their expected high rate of turnovers, and the high level of "informality" of establishments in many activities, the surveyors selected an *aerial sampling* approach to estimate the population of establishments and select the sample in this stratum for all regions of the survey. Aerial sampling consists of indentifying neighborhoods/clusters of informal sector activity, then dividing these into blocks and surveying door to door to identify and collect data on microenterprises. Sample sizes were chosen with the purpose of getting estimates with 5 percent precision at 90 percent confidence intervals with all variables that are 0/1 (yes/no and percentages). This implies, of course, that estimates using cost data do not have the same level of precision.

Also, the blocks, once identified, are then randomly chosen for surveying. So, not all blocks are surveyed. This is the next best solution to random selection from a population list and one which is often used when no population list is available.

The final dataset contains information on sample weights for this group, which approximates the number of micro-enterprises found in the population.

Appendix 2: Propositions derived from the model of formality

The rent functions $\pi_f(x, w, u, t_w, t_\pi)$ and $\pi_i(x, w, u)$ are defined for all entrepreneurs $x : x \geq x_0$. Although for each only one of π_f or π_i is realized for any one entrepreneurs, we assume that the other is also known to the entrepreneur as a notional value, and that that x is in the formal sector if $\pi_f(x, w, u, t_w, t_\pi) \geq \pi_i(x, w, u)$ and in the informal sector otherwise. Given this sorting rule and our other assumptions, the following propositions can be proved about the size distributions, relative earnings, and TFP and labor productivity gaps of the formal and informal sectors of the economy.[9]

Proposition 1: *The most talented entrepreneurs are always in the formal sector*

For any given set of factor prices and tax rates, there is a threshold management talent level, $x = \hat{x}$, defining the firm at the margin of formality and informality such that

(5) $\pi_f(\hat{x}, w, u, t_w, t_\pi) = \pi_i(\hat{x}, w, u)$

Moreover,

(6) $\pi_f(x, w, u, t_w, t_\pi) \geq \pi_i(x, w, u)$ if $x \geq \hat{x}$

and

$\pi_f(x, w, u, t_w, t_\pi) < \pi_i(x, w, u)$ if $x < \hat{x}$

In other words, every entrepreneur in the informal sector has less management talent than every entrepreneur in the formal sector.

Proposition 2 : *Higher taxes would push more talented entrepreneurs into the informal sector*

Equation (5) defines the talent level of the marginal formal sector entrepreneur, $x = \hat{x}$, as an implicit function of w, u, t_w, and t_π, which function can be differentiated to show that $\hat{x}$ increases in each tax rate t_w and t_π. It follows that, the higher is either of these tax rates, the higher the share of the informal in the total population of firms, and the higher is the average entrepreneurial talent in the informal sector.

Proposition 3: *Informal firms are always smaller than formal firms*

---

[9] Proofs are available from the authors.

This is in the sense that every informal firm is smaller, in employment terms, than every formal sector firm. The minimum firm size in the formal sector is $n_f(\hat{x}, w, t_w, u)$ while the largest informal sector firm has employment size $n_i(\hat{x}, w, u)$ so that $n_f(x, w, t_w, u) \geq n_f(\hat{x}, w, t_w, u)$ for every formal firm x and $n_i(x, w, u) < n_i(\hat{x}, w, u)$ for every informal firm.

Proposition 4: *The missing- middle hypothesis*

There is discontinuity between the size distribution of formal firms and the size distribution of informal firms. This is in the sense that the employment size of the marginal formal sector firm $x = \hat{x}$ is smaller than what the same firm would employ in the informal sector, that is, $n_f(\hat{x}, w, t_w, u) > n_i(\hat{x}, w, u)$. In other words, there is a gap between the minimum employment size of formal firms and the maximum employment size of informal firms.

Proposition 5: *Informal entrepreneurs earn less than owners in the formal sector*

This is in the sense that profits per firm are lower in the informal sector. For any firm $x = x_1$ observed in the informal sector and any firm $x = x_2$ observed in the formal sector, we have $\pi_f(x_2, w, u, t_w, t_\pi) > \pi_i(x_1, w, u)$. This is in fact a corollary of Proposition 1, which implies that $x_1 < \hat{x}$ and $x_2 \leq \hat{x}$ so that $\pi_f(x_2, w, u, t_w, t_\pi) \geq \pi_f(\hat{x}, w, u, t_w, t_\pi)$ and $\pi_f(\hat{x}, w, u, t_w, t_\pi) > \pi_i(x_1, w, u)$.

Proposition 6: *Informal firms are less productive than formal firms*

This is in the sense that total factor productivity any informal firms is lower than total factor productivity in any formal firm. This too follows form Proposition 1. The Lucas model should be interpreted as one of selection (into entrepreneurship or informality) on total factor productivity, given that management talent can only be partially observable. For example, under the assumption of Gibrat's Law (that firms grow at a uniform mean rate independent of initial size) made in much of the Lucas (1978), the production function is of the form: $y = A(x)k^{\lambda_1} n^{\lambda_2}$, where $A(x) = \alpha x$ is TFP, $\alpha > 0$ and $0 < \lambda_1, \lambda_2 < 1$. Underlying the size distribution of firms in the baseline model is therefore an inter-firm dispersion in TFP, which is increasing monotonically over the size spectrum starting from lowest observation for the marginal firm. Thus for any two firms $x = x_1$ and $x = x_2$, and a given set of factor prices w and u, their employment size ratio would be a monotonic transformation of their relative total factor productivities:

(7) $$\frac{A(x_2)}{A(x_1)} = \left(\frac{n(x_2)}{n(x_1)}\right)^{1-\lambda_1-\lambda_2}$$

27

<u>Proposition 7:</u> *The densities of labor productivity in the formal and informal sector overlap*

Total factor productivity is unambiguously lower in the informal sector than in the formal sector in the sense that the $A(x)$ of every informal firm is smaller than that of every formal firm. And in general labor productivity is higher in the formal sector for that reason. However, there is significant overlap between the probability density of labor productivity in the formal sector and the density of labor productivity in formal firms due to the convexity of concealment costs.

<u>Proposition 8:</u> *The extent of overlap in the densities of labor productivity decreases in the convexity of concealment costs in the informal sector*

The range of x (or subset of firms) over which the density of labor productivity in the formal sector overlaps with the density of labor productivity in informal firms depends on the curvature of the concealment cost function, the range being smaller the more convex is the cost of concealment, that is the greater is, $c^{"}(n(x))$. Arguably the curvature of the concealment cost function will be greater the higher is government's ability or willingness to enforce tax laws. If that is indeed the case, then the degree of overlap in the density of labor productivity between the formal and informal sectors will be smaller where the enforcement of laws is stronger.

28

Appendix 3: Educational Attainment of Firm Owners
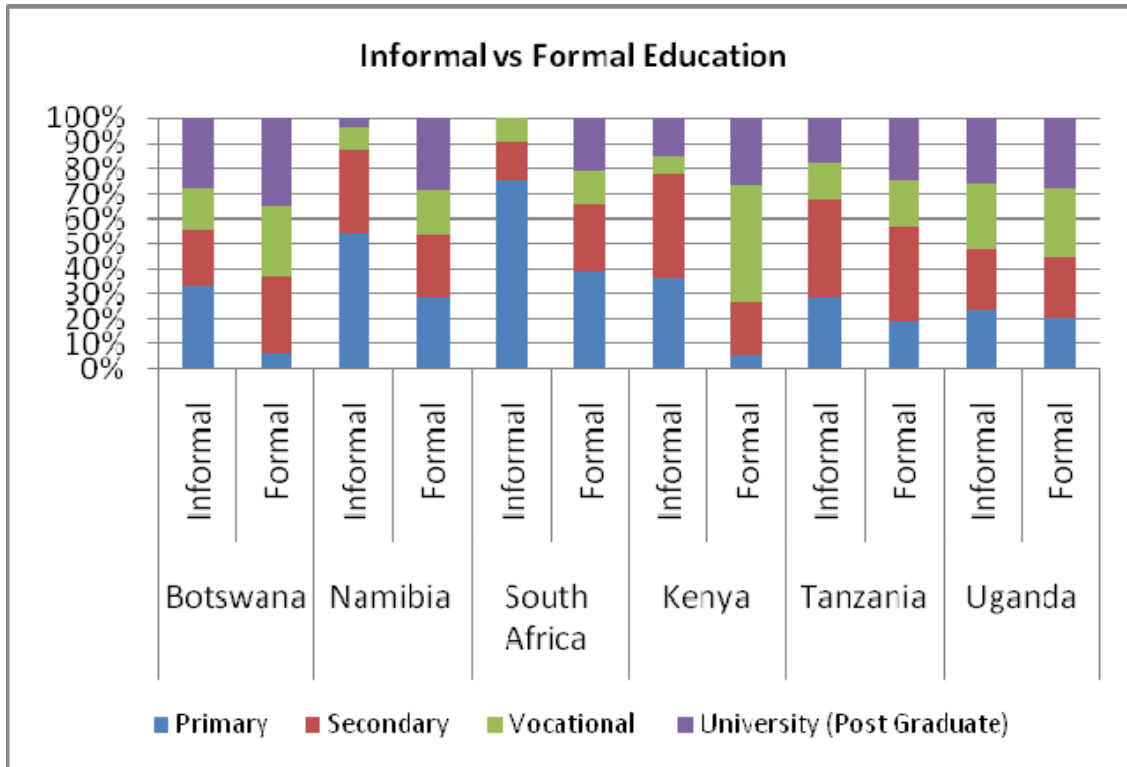
Appendix 4: Econometric Estimations of the Decision to Formalize

Model I: Impact of Education (Categorical dummies) on Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.49	-1.56***	-0.67*	-1.77***	0.12	-0.03	0.21
	(0.336)	(0.574)	(0.374)	(0.468)	(0.454)	(0.386)	(0.579)
secdary	0.41	0.64	0.18	0.41	0.85	0.01	0.85**
	(0.346)	(0.706)	(0.450)	(0.727)	(0.779)	(0.544)	(0.424)
vocat	0.99***	1.38***	0.21	0.74	0.55	0.14	0.43
	(0.382)	(0.569)	(0.398)	(0.497)	(0.481)	(0.449)	(0.381)
univ	1.46***	1.35***	1.13***	1.83***	0.56	0.12	0.56
	(0.405)	(0.542)	(0.413)	(0.485)	(0.480)	(0.403)	(0.435)
retailit	-0.12	0.63*	-0.03	-0.28	-0.64*	-0.13	0.78
	(0.305)	(0.324)	(0.298)	(0.336)	(0.332)	(0.285)	(0.587)
female	0.20	-0.36	-0.66**	0.31	-0.39	0.15	-0.62**
	(0.290)	(0.281)	(0.298)	(0.277)	(0.361)	(0.282)	(0.286)
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-72.199	-59.413	-49.592	-56.513	-40.668	-67.465	-54.924

Model 2: Impact of Firm Size on Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.83***	-1.16***	-0.63**	-0.81***	-0.09	-0.58	-0.61
	(0.332)	(0.397)	(0.300)	(0.301)	(0.411)	(0.403)	(0.734)
Log(Workers)	0.99***	0.75***	0.86***	0.32*	0.66**	0.38*	0.49*
	(0.200)	(0.228)	(0.242)	(0.184)	(0.313)	(0.215)	(0.271)
retailit	0.18	0.90***	-0.01	-0.41	-0.48	0.08	1.39**
	(0.302)	(0.332)	(0.320)	(0.329)	(0.343)	(0.303)	(0.692)
female	0.24	-0.46*	-0.89***	0.06	-0.35	0.25	-0.54*
	(0.292)	(0.277)	(0.329)	(0.266)	(0.346)	(0.289)	(0.296)
Number of Observations	120	101	85	107	65	98	101
Log Likelihood	-67.34	-59.80	-42.34	-62.65	-39.24	-65.94	-50.54

Model 3: Impact of Education (Continuous Variable-years of Education) on Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-3.42***	-2.30***	-2.43***	5.88***	-0.19	-0.59	-1.11
	(1.154)	(0.986)	(0.823)	(1.271)	(1.352)	(0.986)	(1.126)
years of educ.	0.34***	0.17**	0.20***	0.45***	0.07	0.06	0.14
	(0.111)	(0.077)	(0.071)	(0.103)	(0.113)	(0.081)	(0.095)
retailit	-0.15	0.59*	-0.19	-0.28	-0.78***	-0.08	0.78
	(0.361)	(0.325)	(0.333)	(0.337)	(0.343)	(0.292)	(0.605)
female	0.55	-0.41	-0.77***	0.25	-0.36	0.01	-0.51*
	(0.381)	(0.284)	(0.332)	(0.285)	(0.352)	(0.288)	(0.296)
Number of Observations	78	91	80	115	61	88	101
Log Likelihood	-44.49	-56.70	-40.40	-53.61	-38.25	-60.52	-48.54

Model 4: Size, Education and the Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.97***	-2.14***	-0.73*	-1.73***	-0.47	-0.67	-0.83
	(0.386)	(0.628)	(0.437)	(0.521)	(0.542)	(0.532)	(0.780)
secdary	0.20	0.84	-0.19	0.28	0.78	-0.04	0.95**
	(0.362)	(0.704)	(0.523)	(0.750)	(0.774)	(0.554)	(0.475)
vocat	0.52	1.19**	-0.05	0.72	0.51	0.18	0.52
	(0.411)	(0.579)	(0.454)	(0.522)	(0.488)	(0.454)	(0.406)
univ	0.99***	1.11**	0.79*	1.76***	0.30	0.06	0.76
	(0.434)	(0.554)	(0.468)	(0.529)	(0.503)	(0.409)	(0.461)
Log(workers)	0.83***	0.69***	0.79***	0.08	0.69**	0.39*	0.46
	(0.213)	(0.248)	(0.256)	(0.206)	(0.337)	(0.218)	(0.282)
retailit	0.04	0.82***	-0.10	-0.15	-0.47	0.08	1.36*
	(0.323)	(0.339)	(0.337)	(0.364)	(0.349)	(0.308)	(0.730)
female	0.13	-0.33	-0.85***	0.24	-(0.360)	0.28	-0.68**
	(0.306)	(0.291)	(0.342)	(0.296)	(0.369)	(0.296)	(0.323)
Number of Observations	120	99	85	107	65	98	100
Log Likelihood	-64.00	-55.38	-39.44	-52.20	-38.43	-65.80	-47.18

Model 5: Impact of Infrastructure (electricity, water and public sewage) on the Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.91***	-2.10***	-0.94***	-1.97***	0.06	-0.04	-0.06
	(0.383)	(0.633)	(0.400)	(0.551)	(0.456)	(0.388)	(0.622)
secdary	0.27	0.94	-0.35	0.40	0.87	-0.02	0.88**
	(0.372)	(0.742)	(0.520)	(0.873)	(0.781)	(0.549)	(0.426)
vocat	0.65	1.38***	-0.10	0.74	0.44	0.09	0.37
	(0.418)	(0.597)	(0.431)	(0.584)	(0.494)	(0.455)	(0.382)
univ	1.00***	1.11*	0.47	1.49***	0.42	0.07	0.49
	(0.442)	(0.573)	(0.475)	(0.567)	(0.500)	(0.408)	(0.452)
retailit	-0.03	0.80***	0.005	-0.63	-0.55	-0.12	1.06*
	(0.330)	(0.338)	(0.322)	(0.405)	(0.344)	(0.286)	(0.643)
female	0.25	-0.26	-0.94***	0.04	-0.46	0.13	-0.65**
	(0.315)	(0.290)	(0.332)	(0.315)	(0.370)	(0.283)	(0.294)
Infrastructure Services	1.16***	0.82***	1.26***	1.46***	0.37	0.24	-
	(0.267)	(0.319)	(0.363)	(0.322)	(0.379)	(0.306)	
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-62.38	-56.00	-42.94	-45.45	-40.20	-67.16	-53.41

Model 6: Access to Formal Banking (checking account) on the decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-1.45***	-1.85***	-1.14***	-1.76***	-0.14	-0.41	-0.04
	(0.457)	(0.650)	(0.444)	(0.470)	(0.504)	(0.395)	(0.613)
secdary	0.38	0.75	-0.06	0.23	0.85	0.03	0.76
	(0.377)	(0.722)	(0.477)	(0.755)	(0.802)	(0.545)	(0.469)
vocat	0.68	1.39***	-0.12	0.62	0.59	0.001	0.43
	(0.417)	(0.577)	(0.442)	(0.505)	(0.496)	(0.463)	(0.424)
univ	1.35***	1.35***	0.73	1.45***	0.42	0.01	0.32

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
	(0.429)	(0.550)	(0.450)	(0.529)	(0.500)	(0.411)	(0.489)
retailit	-0.22	0.60*	-0.02	-0.30	-0.61*	-0.16	0.77
	(0.331)	(0.327)	(0.311)	(0.343)	(0.338)	(0.287)	(0.626)
female	0.39	-0.29	-0.64**	0.17	-0.59	0.17	-0.81***
	(0.320)	(0.290)	(0.309)	(0.293)	(0.392)	(0.285)	(0.322)
Bank Deposit account	**1.43***	**0.36**	**1.00***	**0.56***	**0.54**	**0.32**	1.14***
	(0.332)	**(0.329)**	**(0.370)**	**(0.322)**	**(0.384)**	**(0.282)**	(0.337)
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-61.50	-58.81	-45.52	-55.01	-39.67	-66.81	-48.19

Model 7: Impact of Formal sector borrowing on decision to formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.46	-1.60***	-0.75*	-1.82***	0.13	0.24	0.20
	(0.343)	(0.574)	(0.403)	(0.470)	(0.453)	(0.432)	(0.586)
secdary	0.09	0.64	0.14	0.27	0.88	-0.25	0.96**
	(0.364)	(0.705)	(0.476)	(0.757)	(0.781)	(0.575)	(0.447)
vocat	0.68*	1.22**	0.16	0.74	0.55	-0.11	0.43
	(0.409)	(0.575)	(0.417)	(0.496)	(0.479)	(0.484)	(0.418)
univ	1.25***	1.29***	0.81*	1.66***	0.58	-0.03	0.63
	(0.421)	(0.542)	(0.461)	(0.491)	(0.483)	(0.427)	(0.458)
retailit	-0.19	0.58*	-0.12	-0.30	-0.64*	-0.20	0.59
	(0.321)	(0.328)	(0.325)	(0.340)	(0.332)	(0.289)	(0.596)
female	0.35	-0.27	-0.63*	0.27	-0.37	0.12	-0.65**
	(0.319)	(0.290)	(0.321)	(0.282)	(0.365)	(0.285)	(0.310)
Bank Overdraft or Loan		**0.94****	**2.03***	**0.61***	**-0.15**	**-0.76**	
		(0.475)	**(0.641)**	**(0.316)**	**(0.433)**	**(0.463)**	
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-60.93	-57.24	-42.37	-54.64	-40.61	-66.09	-48.47

Note: In South Africa, no Informal firms had access to banking sector loans or overdrafts

Model 8: Burden of Inspections and the decision to formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.98***	-1.62***	-0.83**	-1.77***	-0.71	-0.81	0.29
	(0.388)	(0.592)	(0.398)	(0.536)	(0.591)	(0.497)	(0.580)
secdary	0.31	0.70	0.26	0.41	1.13	0.03	0.82*
	(0.378)	(0.718)	(0.464)	(0.727)	(0.885)	(0.574)	(0.424)
vocat	0.70*	1.39***	0.33	0.74	0.58	0.33	0.44
	(0.416)	(0.576)	(0.412)	(0.499)	(0.565)	(0.478)	(0.387)
univ	1.54***	1.35***	1.13***	1.83***	0.48	0.16	0.66
	(0.440)	(0.549)	(0.420)	(0.485)	(0.567)	(0.425)	(0.448)
retailit	-0.09	0.63*	0.01	-0.28	-1.18***	-0.06	0.88
	(0.335)	(0.325)	(0.303)	(0.336)	(0.415)	(0.296)	(0.592)
female	0.15	-0.38	-0.63**	0.31	-0.90**	0.37	-0.57*
	(0.316)	(0.283)	(0.302)	(0.277)	(0.440)	(0.305)	(0.292)

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Inspector Visits	1.33***	0.15	0.64	-0.003	1.82***	0.87***	-0.32
	(0.276)	(0.279)	(0.463)	(0.317)	(0.469)	(0.307)	(0.316)
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-59.47	-59.26	-48.64	-56.51	-31.46	-63.30	-54.41

Model 9: Bribe Payments and the Decision to Formalize

	South Africa	Botswana	Namibia	Kenya	Tanzania	Uganda	Rwanda
Intercept	-0.26	-1.56***	-0.63*	2.74***	0.11	-0.32	0.26
	(0.359)	(0.574)	(0.375)	(0.702)	(0.463)	(0.423)	(0.578)
secdary	0.50	0.64	0.09	0.36	0.85	-0.17	0.98***
	(0.357)	(0.706)	(0.461)	(0.744)	(0.781)	(0.559)	(0.436)
vocat	1.09***	1.37***	0.03	0.83	0.55	0.04	0.37
	(0.394)	(0.573)	(0.415)	(0.512)	(0.482)	(0.458)	(0.384)
univ	1.38***	1.34***	1.02***	1.82***	0.56	0.07	0.36
	(0.413)	(0.547)	(0.422)	(0.497)	(0.481)	(0.408)	(0.470)
retailit	-0.14	0.63*	0.02	-0.26	-0.64*	-0.01	0.65
	(0.311)	(0.324)	(0.303)	(0.347)	(0.335)	(0.296)	(0.588)
female	0.10	-0.37	-0.78***	0.31	-0.40	0.17	-0.73***
	(0.296)	(0.285)	(0.311)	(0.283)	(0.364)	(0.287)	(0.298)
Bribe Payment	-0.54**	0.08	0.93*	1.05**	0.04	0.50*	0.84**
	(0.259)	(0.327)	(0.528)	(0.530)	(0.339)	(0.282)	(0.431)
Number of Observations	120	99	99	124	65	98	127
Log Likelihood	-70.00	-59.38	-48.02	-54.00	-40.66	-65.90	-52.72

REFERENCES

Almeida, Rita and Carneiro, Pedro Manuel. 2005. "Enforcement of Labor Regulation, Informal Labor, and Firm Performance." World Bank Policy Research Working Paper No. 3756

Aterido, Reyes, Mary Hallward-Driemeier and Carmen Pages. 2007. Investment Climate and Employment Growth: The Impact of Access to Finance, Corruption and Regulations Across Firms (November). IZA Discussion Paper No. 3138

Chen, Martha Alter. 2004. "Rethinking the Informal Economy: Linkages with the Formal Economy and the Formal Regulatory Environment".

Clarke, George and Robert Cull. "Access to Credit and Firm Ownership in South Africa."

De Paulo, Aureo, José A Scheinkman. 2007. "The Informal Sector, Second Version". PIER Working Paper No. 07-035.

De Soto, Hernando. *The Mystery of Capital*. Basic Books. 2000

Djankov, Simeon, Ira Lieberman, Joyita Mukherjee and Tatiana Nenova, 2002. "Going Informal: Benefits and Costs" in Boyan Belev, editor. *The Informal Economy in the EU Accession Countries*. Center for the Study of Democracy.

Emery, James in Nicole Ball, Nicolas van de Walle and Vijaya Ramachandran eds. 2003. *Beyond structural adjustment: the institutional context of African development*. Palgrave Macmillan.

Emran, M. Shahe and Joseph Stiglitz. 2004. "Price-neutral Tax Reform With an Informal Economy". Manuscript.

Fafchamps, Marcel. 1994. "Industrial Structure and Micro-Enterprises in Africa." *Journal of Developing Areas*, 29(1):1-30, October.

Fajnzylber, Pablo, William Maloney and Gabriel Rojas (2005). "Releasing Constraints to Growth or Pushing on a String? The Impact of Credit, Training, Business Associations and Taxes on the Performance of Mexican Micro Firms, World Bank Working Paper #2807.

Fortin, Bernard, Guy Lacroix, and Claude Montmarquette. 1997. "Are Underground Workers More Likely to be Underground Consumers?" Cahiers de Recherche, Université Laval.

Fox, Louise and Ana Maria Oviedo, 2008. "Are Skills Rewarded in Sub-Saharan Africa? Determinants of Wages and Productivity in the Manufacturing Sector," World Bank Policy Research Working Paper 4688.

Gatti, Roberta and Maddalena Honorati, 2008. "Informality among Formal Firms: Firm-level, Cross-country Evidence on Tax Compliance and Access to Credit." World Bank Policy Research Working Paper 4476.

Ihrig, J. and K.S. Moe. 2004. "Lurking in the shadows: the informal sector and government policy." *Journal of Development Economics*.

Ishengoma, Esther and Robert Kappel. 2006. "Economic Growth and Poverty: Does Formalisation of Informal Enterprises Matter?" German Institute of Global and Area Studies (GIGA) Working Paper No. 20.

Johnson, Simon, Daniel Kaufmann, Pablo Zoido. 1998. "Regulatory Discretion and the Unofficial Economy." AER, Vol. 88, No. 2.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi. 2008. "Governance Matters V: Aggregate Individual Governance Indicators for 1996-2005". World Bank Policy Research Working Paper No. 4012.

La Porta, Rafael and Andrei Shleifer, "The Unofficial Economy and Economic Development," Manuscript, November 2008.

Loayza, Norman V, Ana Maria Oviedo, and Luis Serven. 2005. "The impact of regulation on growth and informality - cross-country evidence." Policy Research Working Paper Series 3623. World Bank.

Lucas, Robert 1978. On the Size Distribution of Business Firms. *Bell Journal of Economics* Autumn, Volume: 9, Issue: 2, pp. 508-523

Marcouiller, D. and Young, L. 1995. "The Black Hole of Graft: The Predatory State and the Informal Economy." *American Economic Review*. 1995

Ramachandran, Vijaya and Manju Kedia Shah 2006. "Why Are There So Few Black-Owned Firms in Africa? Preliminary Results from Enterprise Survey Data". Working Paper 104. Center for Global Development.

Rauch, James E., 1991. "Modelling the informal sector formally," *Journal of Development Economics*. Elsevier, vol. 35(1), pages 33-47, January.