Distortions, Efficiency and the Size Distribution of Firms*

Giovanni Gallipoli
University of British Columbia

Jonathan Goyette
University of British Columbia

November 16, 2009

Abstract

Differences in factor productivity across countries are both large and persistent. In this paper we investigate how weak institutions, limited accountability in the tax collection system and missing credit markets may be reflected into ex-ante suboptimal production decisions, reduced output and lower productivity. Using information about an institutional distortion, typical of many developing countries, we investigate the determinants of the equilibrium distribution of firms' sizes and quantify the extent of resource mis-allocation. We consider a heterogeneous firms model where the tax environment acts as a selection mechanism restricting the growth of all but the most productive firms. We calibrate the model by Indirect Inference using firm-level data from Uganda. We test the fit of the model through a set of over-identifying restrictions and provide further validation by showing that it can match some unusual data patterns in capital-labor ratios, a dimension which is not explicitly targeted in our calibration. Through counterfactual experiments we quantify the efficiency losses associated to the institutional distortion and credit market imperfections and compare the effectiveness of alternative policy changes.

*This paper was previously circulated under the title 'Corruption and the Size Distribution of Firms'. We thank participants at the SOEGW-Waterloo 2008, CEA 2008, CSAE-Oxford 2009, NEUDC-Tufts 2009 and the 2009 Workshop on Dynamic Macroeconomics in Soutomaior for helpful discussions. Paul Beaudry, Mick Devereux, Patrick Francois, Viktoria Hnatkovska, Dirk Krueger, Henry Siu, Rodolfo Manuelli and Makoto Nakajima provided insightful comments. All errors are ours.

1 Introduction

Much of the measured variation in output per worker across countries is attributable to heterogeneity in total factor productivity (see Prescott,1998; Hall and Jones, 1999; Klenow and Rodriguez-Clare, 1997). Differences in institutions and lack of government accountability and enforcement are often suggested as possible reasons for this variation (North et al., 1981; Acemoglu et al., 2001). However, little is known about mechanisms mapping observable institutional and market distortions into the allocation of resources and the ensuing consequences in terms of efficiency.[1] This paper uses information about an institutional distortion, deriving from limited accountability of government officials in many developing countries, to rationalize observed productive choices of entrepreneurs and to investigate the extent to which this distortion results in a mis-allocation of resources and loss in measured productivity. Building on empirical results from Goyette (2009), who uses firm-level data to study the determinants of the size distribution of firms in Uganda, this paper develops and estimates a model of firms' growth and inspects the ensuing equilibrium allocation of productive resources and distribution of firms' sizes.

It is often argued that tax administrations in less developed countries (LDCs) target larger firms more intensively because tax collection is expected to be greater than enforcement costs (de Soto,1989; Gauthier and Gersovitz,1997; Gauthier and Reinikka,2006). Goyette (2009) shows that the probability of being audited by tax officials in Uganda rises sharply with firm size, especially beyond a certain size threshold corresponding to about 30 employees. This translates in a significant 'break' in the density distribution of firm sizes around that size threshold. The data suggests that entrepreneurs, who would otherwise grow their firms beyond that threshold, choose ex-ante suboptimal capital-labor allocations in order to avoid steep increases in regulation costs. Only relatively large productivity shocks outweigh the costs of increasing employment beyond the 'formality' threshold: in response to a positive productivity shock a large proportion of entrepreneurs choose to scale up production by substituting capital for labor. This results in the observed clustering of many firms at or below the 30-employee threshold and to a significant discontinuity in the level of capital-labor ratios of firms above or below the threshold.

The proposed model extends the one by Hopenhayn and Rogerson (1993)'s, HR hereafter, featuring entry and exit of heterogeneous firms. We allow for a tax distortion due to uneven auditing and for tax-evasion and bribing of officials.

In the model, firms produce a homogeneous good using capital and labor. They decide whether to enter and exit the market and are heterogeneous along two dimensions: transient

[1]Recent work by Banerjee and Duflo (2008) and Buera et al. (2009) highlight how credit market imperfections and heterogeneity in financing needs generate distortions in the allocation of resources in developing countries and result in efficiency and productivity losses.

productivity shocks and current capital stock.[2] Apart from entry and exit, entrepreneurs make two choices in each period. First, they choose labor based on their current productivity shock, current capital stock and the probability to meet with a tax official in the current period. Second, they choose capital investment for the following period based on their expectations of future productivity. Consistent with empirical evidence we assume that entrepreneurs evade part of their tax liabilities and that tax officials are corruptible and, with some probability, they accept bribes when auditing entrepreneurs. The frequency of tax audits is size-dependent.

Under some simplifying assumptions we solve the model analytically and derive comparative statics for the parameters of interest. We show that an increase in aggregate demand [input prices] alleviates [exacerbates] the clustering of firms below the size-threshold set by the tax administration. We also show that a change in firms' bargaining power over the surplus from tax evasion has an ambiguous effect on the value of bribes paid, and that this depends on the magnitude of the elasticities of the inputs with respect to the bargaining power. We document a similar pattern for the effect of the tax rate on bribes.

We then revert to the full-blown version of the model and calibrate it. The parameters for the regulation environment are calibrated using firm-level information from Ugandan data. We calibrate the technology parameters based on the existing literature. Finally, we assess the values of the parameters for the processes of firms survival and transition of shocks using an Indirect Inference approach (Gourieroux et al., 1993; Smith Jr, 1993). This simulation-based method relies on the specification of an auxiliary model which is different from the exact data generating process. If there are more auxiliary parameters than structural parameters one can use this set of over-identifying restrictions to test the fit of the model. The test of over-identifying restrictions suggests that our model cannot be rejected.[3] Moreover, the model does a reasonable job at reproducing some features of the data that were not explicitly targeted such as capital-labor ratios patterns by firm size, firms' growth, age and shares of sales. We also observe both in the Ugandan data and the simulated data a clustering of small firms just before the exogenous size threshold. These small firms substitute capital for labor while waiting for a productivity shock that will offset the cost of growing. We note that these choices generate a gap in the size distribution of firms and lead to the so-called 'missing middle' in the size distribution of firms, typical of many LDCs.[4]

[2]Capital stocks become a state variable under the assumption that the market for the rental of capital is not functioning and entrepreneurs finance capital purchases through cash flow. This assumption seems realistic for developing countries, especially Uganda, as discussed in Goyette (2009). We verify the robustness of our findings by removing this assumption in some of the counterfactual experiments.

[3]In particular, it seems to perform well in fitting firms' average number of employees, average age, average growth rate, average capital-output ratio and average capital-labor ratio. The statistical power of this type of test is, however, unclear.

[4]Many papers in the development literature argue that LDCs exhibit too few medium firms as compared to

We examine the theoretical implications of the model by sequentially removing the regulation distortion and the missing credit market assumption. In this way we show that the increase in audit probability, as identified in Goyette (2009), is key in generating the clustering of firms at the size-threshold, and thus the 'missing middle' in the size distribution of firms. Moreover we find that removing all regulation distortions would close the gap in output per worker between Uganda and an undistorted economy (like, to a certain extent, the US) by 11%. However introducing perfect capital markets generates much larger gains and explains roughly 92% of the output difference between our benchmark economy and the first-best. It is apparent that the 'missing middle' generated by the institutional distortions does not have strong implications for aggregate efficiency. Firms adjust their input mix optimally to avoid large productivity losses in the presence of institutional distortions and the resulting firms size distribution is a by-product of their optimal responses.

Having established the ability of the model to rationalize some relevant features of the data, we design and implement two counterfactual experiments. First, we show that only small efficiency gains are associated to a smoothing of the tax schedule (through an even distribution of audit probabilities across firms). This result is obtained in the context of missing credit markets and suggests that, in the presence of credit constraints, having a tilted, size-dependent distribution of audit probabilities might in fact have some advantages for smaller firms, which have a harder time financing activities through their cash flow. In a second experiment, we change wage compensation of auditors so to make them indifferent between being honest or corrupt. We argue that bribing may serve as a useful second-best mechanism in that it buys entrepreneurs a tax rebate. However, this higher wage brings about larger efficiency gains than corruption. Finally, we examine how these policy changes may affect the gap in output per worker between Uganda and an undistorted economy with working credit markets. Flattening the distribution of audit probabilities would only reduce this gap by a mere 0.5% while the higher (no corruption) wage would reduce this gap by 27%. In order to study the 'pure' reallocation effects of the policies we hold aggregate labor supply constant throughout all our simulations. Therefore much of the productivity variation from the policy changes is captured by the variation in the adjusted workers' wage and the aggregate output gains are lower bounds on the potential effect of the policy.

Related Literature Recently, increasing attention has been devoted to the effects of institutional constraints on the allocation of productive resources. Using data from 79 developed and developing countries Alfaro et al. (2007) show that differences in the allocation of resources

developed economies like the US (Tybout, 2000; Liedholm and Mead, 1987; Little et al., 1987; Steel, 1993;Steel and Webster, 1992;Sleuwaegen and Goedhuys, 2002).

across heterogeneous plants are an important determinant of cross-country income differences per worker. Similarly Restuccia and Rogerson (2008) use a calibrated version of the neoclassical growth model to argue that differences in allocation of resources across heterogeneous establishments may account for significant cross-country differences in output per capita. In a related paper Guner et al. (2008) examine the costs of size-dependent policies distorting production scale.[5]

Our model is based on early contributions on firms' dynamics starting with Lucas Jr and Prescott (1971), who show that the competitive equilibrium for an industry can be solved by maximizing the net aggregate surplus of an economy. In their paper, the production technology exhibits constant returns to scale, all shocks are aggregate and there is no entry and exit. Jovanovic (1982) provides a similar result in a context where shocks are experienced at firm level and follow a non-stationary process. However, his model does not allow for entry and exit in the limit. Hopenhayn (1992) extends the above models to allow entry and exit in a stationary equilibrium, and a more general process for firms' idiosyncratic shocks. HR further extend these results to a general equilibrium framework. They calibrate their model to U.S. data and examine the impact of labor regulations on job creation and destruction. HR's model has provided the basic framework for several other papers. Just to name a few, Veracierto (2001) examines the welfare effect of firing taxes by introducing a flexible form of capital in HR's model. Atkeson and Kehoe (2005) also amend HR's model to explore the process by which firms accumulate plant-specific knowledge. Melitz (2003) uses a reduced-form of Hopenhayn (1992) model to explain plant export activities.

A somewhat different theoretical framework was originally developed by Lucas (1978) in his span-of-control model where entrepreneurs differ in ability. Rauch (1991) extends Lucas (1978) model to investigate why in LDCs tax enforcement seems to be enforced mostly for large firms. The aforementioned models by Alfaro et al. (2007) and Guner et al. (2008) are also based on Lucas (1978).

Also of interest, Ericson and Pakes (1995) develop a model of industry dynamics that allows for entry, exit and firm-specific uncertainty where R&D investment serves as a selection mechanism. Sleuwaegen and Goedhuys (2002) describe the size distribution of firms in Côte d'Ivoire with a growth model where firms exploit scale enlargements and learning to make efficiency gains. Rossi-Hansberg and Wright (2007) develop a model where the scale dependence of firm size dynamics (growth rate, exit rate, etc.) relies on the response of production decisions to the allocation and accumulation of industry specific human capital. This scale dependence is one potential reason in their view for observing thicker tails in the size distribution of firms of LDCs.[6]

[5] Our work departs from this study in two dimensions: i) we focus on an observable distortion and use data from a LDC; ii) they use Lucas (1978) span-of-control framework with no entry and exit.

[6] There is a huge body of literature on financial constraints and heterogeneous firms. Albuquerque and Hopen-

The rest of the paper is structured as follows. Section 2 summarizes the empirical evidence originally presented in Goyette (2009) which motivates our work. Based on these empirical facts, we develop a dynamic model with heterogeneous firms in section 3. We solve an analytical version of this model, based on some simplifying assumptions, and derive comparative statics for parameters of interest. Section 4 presents the calibration strategy, the numerical implementation of the model with capital investment and some policy experiments. Section 5 concludes.[7]

2 Some Facts

This section describes some interesting facts from a cross-section of Ugandan firms. Data for Uganda are taken from the Ugandan Enterprise Survey (UES) initiated in 1998 by the World Bank and by the Uganda Private Sector Foundation. Firms were randomly selected and are a representative sample of Ugandan firms. 243 firms were interviewed about their activities between 1995 and 1997. The sample covers businesses from five economic sectors: commercial agriculture, agro-processing, manufacturing, tourism and construction, and five geographical areas: Kampala, Jinja–Iganga, Mbale–Tororo, Mukono and Mbarara. The survey focuses on firms' activities including investment, finance, regulation, infrastructure, taxation, corruption and labor market.[8]

Goyette (2009) classifies as small firms those with 30 employees or less, as medium firms those between 31 and 75 employees and as large firms those with 76 employees or more. That paper documents some simple facts: 1) the audit probability depends mainly on firm size and not on capital, and increases sharply around a size-threshold of 30 employees; 2) the density of firm size is discontinuous around the size threshold; 3) small firms clustered below the size threshold substitute capital for labor; 4) tax savings 'per-bribe' are higher for smaller firms than for medium firms; and 5) small firms exhibit lower growth rates than their larger counterparts. We report some of those results below.

---

hayn (2004) develop a model to characterize constrained efficient lending contracts and study their implications for firm survival, growth, equity shares and dividend distribution policies. Mata and Cabral (2003) show that the size distribution of Portuguese manufacturing firms is right-skewed and tends toward a lognormal distribution. They argue that selection is not what matters in explaining the evolution of this industry but financial constraints. Arellano et al. (mimeo) compare how the degree of contract enforcement affects debt and growth patterns across countries. They show that, with lower financial development, the difference in growth rate between smaller and larger firms increases and that debt financing between small and large firms decreases. This footnote is by no means an exhaustive review of this literature.

[7] A theoretical appendix, containing proofs of Proposition 1 and Claim 1, as well as the expressions for aggregate variables and the equations for the comparative statics, is available online at http://grad.econ.ubc.ca/jgoyette/

[8] For a detailed discussion about the data and summary statistics see Goyette (2009).

2.1 Empirical Methodology

Tools from the Regression Discontinuity Design McCrary (2008) suggests a formal test to verify whether the running variable of a regression discontinuity design (RDD) is manipulated and discontinuous around the cutoff where the treatment takes effect.[9] In what follows, this test is used to show that entrepreneurs with a firm at the lower boundary of the medium-size bin reduce their labor input and self-select into the small-size bin. The algorithm developed in McCrary (2008) allows to back out the density of the running variable. This algorithm provides a detailed distribution of the running variable which can then be plotted to give suggestive graphical evidence of the discontinuity around the cutoff.[10] Define $\xi(s)$ as the density of running variable s, and c as the cutoff where the discontinuity potentially occurs. Define $\xi^+ = \lim_{s \to c^+} \xi(s)$ and $\xi^- = \lim_{s \to c^-} \xi(s)$. If the running variable is manipulated, i.e. if there is self-selection, one should observe:

$$\xi^+ \neq \xi^-$$

The parameter of interest in such an analysis is:

$$\vartheta = \ln \xi^+ - \ln \xi^- \tag{1}$$

This parameter is the difference of the intercepts in log at the discontinuity and its standard error is obtained following McCrary (2008).

The reduced form analysis in Goyette (2009) uses a regression equation in the spirit of the RDD, keeping in mind that self-selection prevents any causal inference in the context at hand.[11] Consider the regression model:

$$Y_i = \gamma + \tau D_i X_i + \beta Z_i + \varepsilon_i \tag{2}$$

where Y_i is the dependent variable, Z_i and X_i are observable characteristics of the individuals and D_i is a binary treatment equal to zero if s $\leq$ c and 1 if s $>$ c.

[9] A discontinuity in the density of the running variable indicates that individuals self-select into (or out) of the treatment which the Regression Discontinuity Design (RDD) tries to identify. Self-selection violates the identification requirement of continuity which is needed to implement the RDD. McCrary (2008) and Lee and Lemieux (2009) use the terminology 'running variable' and 'forcing variable' to define the variable with respect to which treatment is assigned.

[10] The algorithm is the following. In a first step, one obtains an histogram of the running variable. In the second step, this histogram is used to estimate a fourth order polynomial on each side of the cutoff. The midpoints of the histogram are treated as regressors and the number of observations per size bins as the outcome variable in a weighted regression where more weight is given to the bins closest to the discontinuity point (i.e. using a triangle kernel). The density function is then estimated by looping over evaluation points of the running variable.

[11] Except for the audit probability which is the treatment itself. More on this below.

This framework is useful to investigate the statistical significance of any discontinuity at the size-threshold in the variables of interest, i.e. capital-labor ratio, firms' growth[12], bribes and taxes.

2.2 Empirical Estimates

Audit probability One key assumption in both theoretical model and numerical exercise is that the audit probability depends on firm size. Table (E1, panel A) reports estimates for equation 2 where we set $D_i = 1$ for all s, i.e. we run a standard OLS of the audit probability on the log of firm size and/or capital with no interaction with the treatment dummy. For the purposes of this paper we consider the audit probability to be the 'treatment' itself and interpret size or capital as having an exogenous effect on the dependent variable.[13] In column 1, we regress the audit probability on size only. The coefficient is significant and the Wald statistic F1 rejects the null that coefficient on size is equal to zero. In column 2, we regress the audit probability on capital only. The coefficient on capital is significant and we reject the null that it is equal to zero (F1=4.45). Columns 3 and 4 present the results when the audit probability is regressed on both size and capital without and with controls for sector and location, respectively. In both regressions size is significant but capital is not. However, the Wald statistic F2 does not allow rejecting the null that the coefficients on labor and capital are the same.

Examining in more details the determinants of the audit probability, table (E1, panel B) reports results from different regressions of the audit probability on the interaction of the binary treatment D_i with labor and capital, i.e. $X_i' = [\ln(l_i), \ln(k_i)]$ in equation 2 where $c = 30$.
Column 1 reports results from a regression of the audit probability on labor, restricting our sample to small firms only: the coefficient is highly significant and so is the Wald statistic F1. In column 2 we run the same regression but we restrict our sample to firms with more than 30 employees. The coefficient on labor is not significant in this case. Columns 3 and 4 report results for a similar exercise with capital stock for small and larger firms in turn. The coefficient for capital is never significant. Column 5 shows results from the regression of the audit probability on size interacted with the binary treatment D_i where $c = 30$. Coefficients for both terms are positive and significant with the coefficient on small firms half of what it is in column 1. The

[12]Firms' growth refers to the growth of sales between 1996 and 1997

[13]Potential endogeneity of the size threshold can be associated to the decisions of tax administrators, which could depend on the size distribution of firms. Anecdotal evidence supports the view that the choice of a threshold is rather arbitrary. In a report from the Article IV Consultation for Uganda made by the IMF in 1998 it is noted that the Ordinary Audit Section has a 'target...to carry out 75 audits a month, 50% of which would be on the largest taxpayers'. The report provides no grounds for this particular assignment nor does it define precisely these largest taxpayers. Also, given the assessment of tax evasion and the low returns on auditing (less than 15% of evaded taxes) it appears that the tax administration's auditing strategy does not maximize tax revenues.

Table (E1, panel A): OLS results				
Dependent Variable: Audit Probability				
	1	2	3	4
labor	.089***		.077**	.075**
	(.021)		(.035)	(.037)
Capital		.043**	.007	.006
		(.02)	(.027)	(.027)
Constant	.516***	-.045	.260	.493
	(.143)	(.442)	(.482)	(.511)
R-squared	.14	.13	.07	.16
Industry-Location	Yes	Yes	No	Yes
F1	17.9***	4.45**	5.22***	4.32**
p-value	[0.0]	[.037]	[0.0]	[.02]
F2			1.47	1.37
p-value			[.23]	[.24]
Nb. Obs.	242	150	150	150

Note:* significant at 10%, ** significant at 5%, *** significant at 1%. Std. Dev. in parenthesis, F-value in brackets; all regressors are in logs; F1: Wald test stat with H_0: coefficients are jointly zero; F2: Wald test stat with H_0: the coefficients on labor and capital are equal.

Table (E1, panel B): Regression Discontinuity results							
Dependent Variable: Audit Probability							
	1	2	3	4	5	6	7
labor	.277***				.131**		.259**
small	(.071)				(.054)		(.122)
labor		.003			.105***		-.002
large		(.042)			(.028)		(.060)
Capital			.067			.017	-.014
small			(.045)			(.024)	(.029)
Capital				-.014		.025	.028
Large				(.027)		(.022)	(.029)
R_squared	.22	.08	.22	.11	.14	.16	.18
F1	15.08***	.01	2.25	.28	9.31***	4.13**	3.22**
p-value	[0.0]	[.94]	[.14]	[.60]	[0.0]	[.02]	[.02]
F2					.73	3.72*	1.84
p-value					[.39]	[.06]	[.14]
Nb. Obs.	115	127	62	88	242	150	150

Note: * significant at 10%, ** significant at 5%, *** significant at 1%. Std. Dev. in parenthesis, F-value in brackets; all regressors are in logs; F1: Wald test stat with H_0: coefficients are jointly zero; F2: Wald test stat with H_0: all coefficients on labor and capital are equal. All regressions include dummies for industry-location.

Wald statistic F1 rejects the null that these coefficients are jointly equal to zero. However these coefficients are not statistically different, as can be inferred from the Wald statistic F2. In column 6 we do the same exercise with capital interacted with the binary treatment D_i where $c = 30$. The coefficients are not significant. Note however that the Wald statistic F2 is significant. This implies that even though the coefficient on capital for each size group have very large variances we reject the null that they are identical. Figure 1 below offers a suggestive graphical representation of this result. The figure presents the correlation between audit probability and capital for small and larger firms. Firms are, or are not, audited and the audit variable thus takes value 0 or 1 as represented by the dots on the graph. We note that the quadratic curve for larger firms lies above the seemingly linear curve for small firms over the whole range of capital values. This graph shows clearly that larger firms are audited more intensively than their smaller counterparts.[14]

[14]The relationship between the audit probability and sales or taxes exhibit somewhat similar trends as for the relationship between labor and capital (graphs available upon request). Note that the correlation coefficient

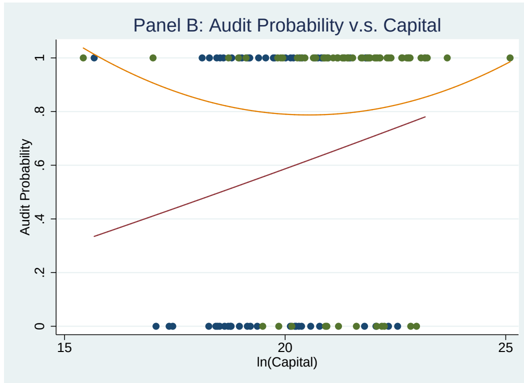

Figure 1: Audit Probability v.s. Capital per Size Groups

Finally, we introduce all covariates in column 7 of table 1, panel B. Only the coefficient on small firms enters significant and is back to a magnitude comparable to column 1. The Wald statistic F1 is significant and implies that taken jointly all 4 coefficients on labor and capital are different from zero. The Wald statistic F2 does not allow to reject that all 4 coefficients are equal. However, a Wald test for each combination of a pair of coefficients rejects the null that the two coefficents are equal except for the pair of coefficient labor-large and capital-large.[15] As a further robustness check Goyette (2009) reports the results of a t-test for the difference in audit probability across size bins and documents a significant 34% increase in the audit probability between small and medium firms. This evidence suggests that the audit probability depends on size, especially around 30 employees. Evidence of any dependence of audit probability on capital seems rather mild once we control for firm size. In the next sections we study the implications of differences in auditing probability on firms' input choices.[16]

Density of firm sizes We use McCrary (2008) local density estimator to test whether entrepreneurs with a firm marginally close to becoming medium size self-select in the smaller size bin. Figure 2 shows the firm size density estimated with the algorithm described in section 2.1 for size bins of 5 employees and a bandwidth of 50 employees. A clear discontinuity at 30 employees is present (vertical line in Figure 2). The difference in the intercepts at the discontinuity, i.e. ϑ, is equal to -.99 and its standard deviation is 0.298 (p-value=0.0). We interpret this as evidence

between labor and capital is 0.68, between labor and sales, 0.79, and between labor and taxes paid, 0.64.

[15]For the pair labor-small and labor-large F2 equals 3.65 with a p-value of 0.06; for the pair capital-small and capital-large, F2 equals 4.10 and the p-value is 0.05; for the pair labor-small and capital-small F2 equals 3.83 and the p-value is 0.05; for the pair labor-large and capital-large F2 equals 0.14 and p-value is 0.71.

[16]Internal Revenue Service data on filled and examined tax reports for U.S. corporations suggest that the rate of examination of reports also increases with size. However, most increases seem to occur for very large firms, above the 90th percentile.

that entrepreneurs' labor choices are distorted by the tax administration targeting of firms above a certain size.[17]

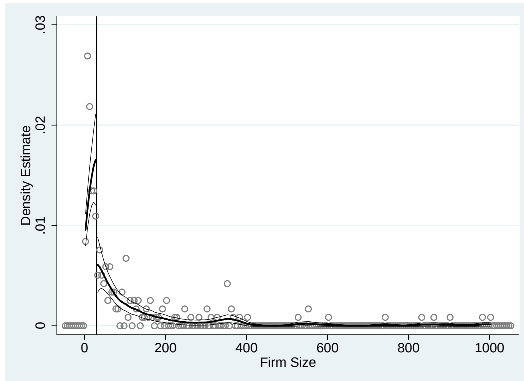

Figure 2: Density of Firm Size

Capital–labor ratios An implication of the size-dependence of audit probabilities is that entrepreneurs may respond to positive productivity shocks by scaling up production through 'ex-ante suboptimal' capital-labor choices. We use the framework of equation 2 to study the distribution of capital intensities by firm sizes: table (E1, panel C) reports results from a regression of the resale value of capital (per employee) on firm size, interacted with D_i where $c = 30$.

Column 1 presents the results of a regression with no constant and covariates. The coefficients on small and larger firms are positive and highly significant. Wald tests suggest that coefficients are not jointly different from zero (F1) and not significantly different from each other (F2). In Column 2 we control for sector, location, firm's age and whether the survey respondent is the firm's owner. The use of the last covariate requires some explanation. The variable 'respondent is the firm's owner' and 'resale value of capital' have a correlation coefficient of 0.2 and we interpret this as evidence that owners might not completely disclose the true value of their firms to their employees. This could bias the amount of reported capital downward. Introducing these covariates reduces the coefficients of interest but does not alter the results of the Wald tests. One could also be worried that the higher coefficients for smaller firms are driven by an initial investment in capital to start up a business. We redo these regressions introducing a constant in our analysis in columns 3 and 4. The coefficients on small and larger firms are still significant but the Wald statistic F2 does not reject the null that these coefficients are the same. Given our focus

[17]The automatic selector in McCrary (2008) suggests the use of size bins of 22 employees and a bandwidth of 45 employees. Such large size bins do not convey a good graphical representation of the density around the threshold and we use smaller size bins. The value of the coefficient ϑ and its standard deviation are robust to the use of the automatic selector.

Table (E1, panel C): Regression Discontinuity results					
Dependent Variable: ln(Resale Value of Capital per employee)					
	1	2	3	4	5
labor	4.54***	2.92***	1.18***	1.17***	1.35***
small firms	(.107)	(.189)	(.212)	(.216)	(.255)
labor	2.81***	1.90***	1.03***	.99***	1.11***
larger firms	(.047)	(.103)	(.111)	(.113)	(.145)
Constant	No	No	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	Yes
R-squared	.97	.98	.49	.49	.44
F1	2707.2***	169.5***	73.8***	61.0***	45.43***
p-value	[0.0]	[0.0]	[0.0]	[0.0]	[0.0]
F2	218.7***	72.0***	1.61	2.11	3.27*
p-value	[0.0]	[0.0]	[.21]	[.15]	[.07]
Nb. Obs.	157	157	157	157	145

Note: * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis, F-value in brackets; Regressors are in logs and covariates are dummies for industry-location, Age and a dummy for the respondent being the owner of the firm; F1: Statistic for a Wald test with H0: coeffficients are jointly zero and F2: Statistic for a Wald test with H0: coefficients on labor-small and labor-large are equal; Column 5: larger firms: between 31 and 500 employees.

on growth dynamics of small and medium firms, in column 5 we restrict the sample to include only firms with 500 employees or less. For this sub-sample the Wald statistic (F2) rejects the null that the coefficients for small and larger firms are equal. Overall this cross-sectional evidence supports the view that small firms tend to substitute capital for labor at the size-threshold.[18]

Next, we show some graphical evidence on the atypical capital-labor ratios in Uganda. Figure 3 (panel A) uses the resale value of capital per employee as a proxy for capital intensity and plots it against firm size for small firms (less or equal than 30 employees) and larger firms (more than 30 and less than 201 employees). Capital-labor ratios of small firms appear to be higher,

[18] As a robustness check Goyette (2009) examines t-test differences across size bins for different proxies of the capital-labor ratios and shows that the difference between small and medium firms is significant in all cases. Goyette (2009) presents also a lowess estimation based on Cleveland (1979) of the patterns in capital-labor ratios in the data.

on average, than those of larger firms. We compare this to Figure 3 (panel B), which presents similar evidence for the U.S. using data from the NBER-CES Manufacturing Industry Database. Clearly, capital-labor ratios increase with firm size in the US.[19] Similar evidence for LDCs are also available. Soderbom and Teal (2004) show that capital-labor ratios are increasing with firm size in Ghana. We note that even if the coefficients on small and larger firms in column 5 of table E1 (panel C) were equal the case of Uganda would still be atypical compared to the US, or even to a LDC like Ghana.

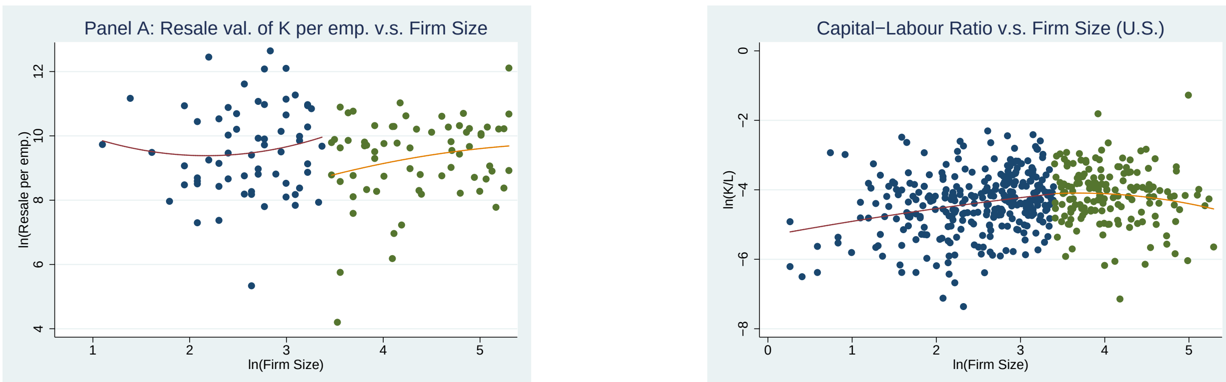

Figure 3: Capital-labor Ratio v.s. Firm Size

Bribes and taxes If larger firms are targeted more often then bribes and taxes paid should also be higher at these firms. We show here that this is indeed the case. We start with some graphical evidence. Figures 4 panels A and B plot bribes and taxes paid against firm size for small firms (less than 31 emp.) and larger firms (more than 30 emp. and less than 201 emp.). Both graphs document that larger firms face higher payments than small firms.

Turning to regression estimates, table (E1, panel D) reports results of a regression of bribes on firm size interacted with D_i where $c = 30$. Column 1 presents the results of a regression without constant or covariates. The coefficients on small and larger firms are positive and highly significant and the Wald statistic F2 rejects the null that the coefficients are equal. However, the coefficient for small firms is unexpectedly higher than for larger firms. In column 2 we control for sectors and locations. The coefficient for small firms remains higher than for larger firms but the Wald statistic F2 is no longer significant: we cannot reject the null that these two coefficients are the same. In columns 3 and 4 we introduce a constant in equation 2. The coefficient on small firms becomes not significant and the constant seems to pick up the amount of bribes paid by small firms. Goyette (2009) shows that the smallest of the small firms seem to pay higher bribes for renegotiating licenses and permits due to the ambiguity of their legal status. This

[19] The capital-labor ratio in this case is given by the ratio of total real capital stock and the number of employees.

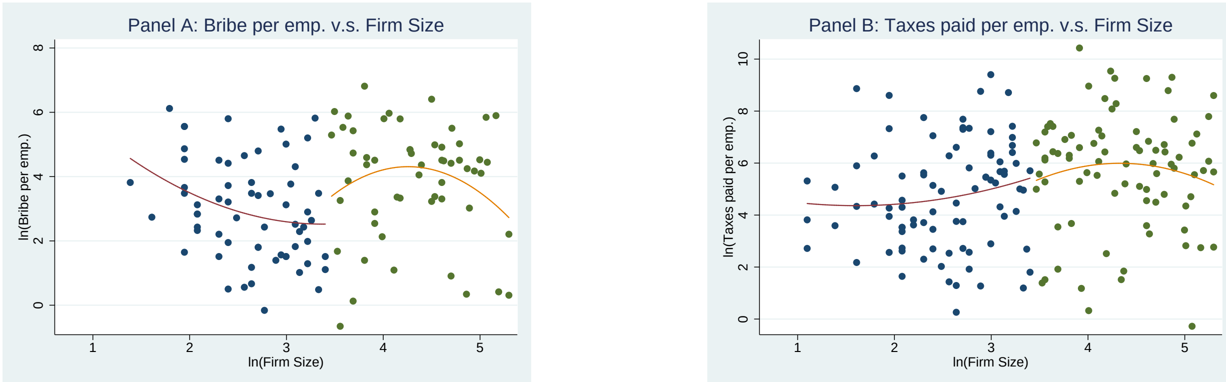

Figure 4: Bribes and Taxes per emp. v.s. Firm Size

helps to understand the observation of a negative slope in figure E1, panel D, for small firms. The coefficient for larger firms in table E1, panel D, columns 3 and 4, remains highly significant and is now higher than for small firms. The Wald statistic F2 is significant again and rejects the equality of the coefficients.

Table E1, panel E, presents the results for a regression of taxes paid on size interacted with D_i where $c = 30$. In this case the coefficient on small firms remains significant in all specifications. Columns 1 and 2 show estimates for a regression without a constant. The coefficients are higher for small firms than larger firms with a Wald statistic F2 rejecting the null that the coefficients are equal. However, in columns 3 and 4 we introduce a constant in our estimating equation and this yields coefficients that are similar across small and larger firms and Wald statistics for test of equality are no longer significant.

Goyette (2009) presents t-test differences across size-bins for tax obligations per employee, taxes paid per employee, bribes per employee and tax savings through bribes. These tests show that: 1) medium firms owe three times the amount of tax liabilities per employee owed by smaller firms and this difference is significant; 2) medium firms pay about three times what smaller firms pay in taxes and bribes per worker and these differences are also significant; and 3) tax savings made though bribes are significantly higher (about 6 times) for small and large firms than for medium firms.

Firms' growth rates Goyette (2009) reports figures for average sales growth by size group. Sales growth rates are given by:

$$FirmsGrowth = ln(Sales1997) - ln(Sales1996) \tag{3}$$

Average sales growth is 0.02 for small firms, 0.13 for medium firms and 0.17 for large firms. The results of a regression of firms growth rate on size are shown in table E1, panel F.

Table (E1, panel D): Regression Discontinuity results				
Dependent Variable: ln(Bribes)				
	1	2	3	4
labor	2.05***	1.46***	.08	.03
small firms	(.09)	(.2)	(.27)	(.27)
labor	1.72***	1.38***	.65***	.65***
larger firms	(.04)	(.12)	(.14)	(.14)
Constant	No	Yes	No	Yes
Covariates	No	No	Yes	Yes
R-squared	.94	.94	.49	.52
F1	1039.9***	83.45***	65.98***	67.23***
p-value	[0.0]	[0.0]	[0.0]	[0.0]
F2	11.19***	.38	15.77***	18.28***
p-value	[0.0]	[.54]	[0.0]	[0.0]
Nb. Obs.	142	142	142	142

Note: * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis, F-value in brackets; Regressors are in logs; F1: Statistic for a Wald test with H0: coefficients are jointly zero and F2: Statistic for a Wald test with H0: coefficients on labor-small and labor-large are equal; Smaller firms have less than 31 employees and larger firms have more than 30 employees.

Table (E1, panel E): Regression Discontinuity results				
Dependent Variable: ln(Taxes)				
	1	2	3	4
labor	5.44***	3.13***	1.1***	1.09***
small firms	(.14)	(.25)	(.29)	(.28)
labor	3.55***	2.23***	1.23***	1.2***
larger firms	(.06)	(.14)	(.15)	(.15)
Constant	No	Yes	No	Yes
Covariates	No	No	Yes	Yes
R-squared	.96	.98	.42	.49
F1	2320.5***	121.01***	69.53***	64.29***
p-value	[0.0]	[0.0]	[0.0]	[0.0]
F2	158.95***	33.07***	.67	.54
p-value	[0.0]	[0.0]	[.42]	[.46]
Nb. Obs.	196	196	196	196

Note: * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis, F-value in brackets; Regressors are in logs; F1: Statistic for a Wald test with H0: coefficients are jointly zero and F2: Statistic for a Wald test with H0: coefficients on labor-small and labor-large are equal; Smaller firms: less than 31 emp.; larger firms: more than 30 emp.

Table (E1, panel F): Regression Discontinuity results						
Dependent Variable: Growth Rate						
	1	2	3	4	5	6
labor	-.04		.066	-.054*		-.072
small firms	(.028)		(.053)	(.031)		(.055)
labor		.031**	.061**		.039**	.074**
larger firms		(.015)	(.029)		(.017)	(.031)
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	No	No	Yes	Yes	Yes
R-squared	.009	.02	.03	.07	.09	.09
F1	2.11	4.07**	2.7*	3.02*	5.36**	3.49*
p-value	[.148]	[.045]	[.069]	[.084]	[.022]	[.033]
F2			.02			0.0
p-value			[.977]			[.96]
Nb. Obs.	208	208	208	208	208	208

Note: * significant at 10%; ** significant at 5%; ***significant at 1%; Std. Dev. in parenthesis and F-Value in brackets; Covariates are dummies for sector and location and Age; F1 Wald test with H0: coefficients on Small and Larger firms are zero jointly; F2: Wald test with H0: coefficients on labor-small and labor-large are equal; Smaller firms: less than 31 emp. and larger firms: more than 30 emp. Results without a constant: significant (not shown). Results not significantly affected by gross growth rate.

Columns 1 to 3 present the regression results of the growth rates on size with no other covariates. In columns 1 and 3 the coefficient on small firms is not significant. In column 4, where we control for industry and location, the coefficient on small firms is weakly significant and negative, but this effect vanishes in column 6 where the results from a regression of growth rates on both size variables are reported. The coefficient on larger firms remains significant and positive throughout. The evidence on firms' growth presented here provide some evidence that small firms restrict employment to avoid heavier regulation costs.[20]

Robustness: other potential explanations Goyette (2009) examines also other potential

[20]Growth patterns observed in Ugandan data contrast with findings for the US. For example Evans (1987) and Hall (1987) argue that small firms usually exhibit higher growth rates than larger firms.

explanations for the observed patterns in capital-labor ratios, firms' growth and size distribution of firms in Uganda. In particular, Goyette (2009) investigates the effect of a factor price differential, differences in technology across firms, and compositional effects across industry, sector or location and shows that none of these explanations reconcile the evidence on the density of firm size, patterns in capital-labor ratios and regulations costs observed in the data.

3 Theory

3.1 The benchmark economy

In this section we describe the basic elements of the model. The economy is made of entrepreneurs, workers and bureaucrats. We assume that these agents are ex-ante different and that their career choice is exogenous.[21]

3.1.1 Entrepreneurs

There is a measure 1 of entrepreneurs. Entrepreneurs are risk neutral and set up a firm to produce a homogeneous consumption good. In each period, potential entrants decide whether to enter or not. Incumbents and new entrepreneurs make hiring and production decisions based on their current productivity shock and capital stock, as well as the probability to be audited by a tax official. All entrepreneurs on the market then decide their level of investment for the next period based on their expectations over productivity shocks. We assume that the market for renting capital is not functioning because of conditions typical of a LDC, e.g. high default probabilities, usuary lending rates from private money lender, etc.[22] As a consequence, entrepreneurs can only finance their investment through cash flow and invest in their own firm.[23] This assumption is made for two reasons. First as observed in Goyette (2009) small firms (and some medium firms) have almost no access to credit. Second, capital-labor ratios are the outcome of a dynamic process of investment that is affected by the interaction of the regulation distortion and credit constraints. It is thus of interest to examine the implications of each distortion in turn as well as jointly.

Those who make no investment exit at the end of the period. Entrepreneurs who meet with a bureaucrat bargain over a bribe to keep part of the surplus from evasion. Every period a fraction

[21] We are thus studying the optimal decisions of entrepreneurs given the constraints from the regulation environment. An interesting extension to this work is to examine how career choices affect, and are affected by, regulation constraints. Lucas (1978)'s span-of-control model provides theoretical grounds for such an analysis.

[22] See Goyette (2009) for a description of the factor markets in Uganda.

[23] When the market for the rental of capital is functioning adequately, entrepreneurs are able to diversify their idiosyncratic risks by investing their profits also in other firms.

of firms exit and are replaced by new entrants in the next period.

Since we are only concerned with a stationary equilibrium, we assume that entrepreneurs are rational and that the prices they anticipate are the prices that realize in equilibrium.[24] The price of output is denoted by p_t and the wage is w_t. In each time period t, entrepreneur i chooses to allocate his profits, π_t^i, between consumption, c_t^i, and capital investment, k_{t+1}^i.

The objective function of entrepreneur i is:

$$\max_{c_t^i, k_{t+1}^i} E_0 \sum_{t=0}^{\infty} (\varphi\beta)^t u(c_t^i) \tag{4}$$

s.t.

$$p_t(c_t^i + k_{t+1}^i) = \pi_t^i \tag{5}$$

$$c_t^i \geq 0 \tag{6}$$

$$k_{t+1}^i > 0 \tag{7}$$

$$k_0 \text{ given} \tag{8}$$

where φ is an exogenous probability of survival, β is the discount rate and π_t^i are profits (defined below).

Entrepreneurs are heterogeneous across two dimensions. Upon entry in the market each entrepreneur is endowed with a productivity parameter, $z_t^i \in [z^{low}, z^{up}]$, drawn from an initial distribution function ν. Productivity shocks have a transition function given by $g(z_{t+1}, z_t)$. Capital stocks $k_t^i \in (0, k^{up}]$ also differ across entrepreneurs because of previous investment history. For the sake of simplicity we drop superscript i assuming it is understood that variables and equations that refer to entrepreneurs are specific to a particular individual except for the initial distribution of shocks ν and the transition function g which are the same for all entrepreneurs.

Given her current shock z_t and current level of capital stock k_t an entrepreneur decides to produce either in the informal or formal sector. These sectors are defined by an exogenous size-threshold, l^{ic}. Firms with a size greater than l^{ic} are audited more intensively. There is a proportion iu of 'informal-unconstrained' entrepreneurs who face low productivity shocks and hire the 'ex-ante optimal' amount of labor below l^{ic}. Some entrepreneurs on the verge of entering the formal sector prefer to constrain their labor at $l_t = l^{ic}$: they constitute a proportion ic of the total mass of entrepreneurs. The notation 'ic' stands for 'informal-constrained', meaning

[24] This assumption is common in this literature see for example section 4 in Lucas and Prescott (1971), Hopenhayn and Rogerson (1993) and Restuccia and Rogerson (2008).

that these entrepreneurs are constrained in the number of workers they can hire, conditional on staying informal, and they choose capital and labor combinations which would not be optimal in the absence of the conditional labor constraint. We let z_t^{ic} be the shock-threshold at which the labor constraint starts binding, for a given k_t.

Finally, there is a proportion f of 'formal' entrepreneurs who enter the formal sector where they choose any amount of labor above l^{ic}. For a given k_t, let z_t^f be the endogenous shock-threshold for which an entrepreneur decides to become formal with size $l^f = l(z_t^f, k_t)$.[25] We define an incumbent firm's current status as $s = \{iu, ic, f\}$.

Profits are given by:

$$\pi_t = p_t F(z_t, k_t, l_t) - w_t l_t + p_t(1-\delta)k_t - RC_t^s - c_F \tag{9}$$

where δ is the depreciation rate of capital and c_F is a fixed cost of carrying business.[26] $F(z,k,l)$ is a production function which exhibits decreasing returns to scale in labor and capital. We assume that:

$$F(z,k,l) = zk^\gamma l^\eta \tag{10}$$

This choice of functional form allows pinning down the size of a firm. RC_t^s are regulation costs for a given status of the firm. These costs are given by:

$$RC_t^s = p_a(l)\left\{qB_t + (1-q)(T_t^e + FA)\right\} + T_t^p \tag{11}$$

where $p_a(l)$ is the probability of being audited by a tax official which varies with firm size, q the probability that the tax official is corrupt and accepts a bribe B_t, $(1-q)$ the probability that the tax official is honest and forces the entrepreneur to pay a fine FA on top of the amount of taxes evaded T_t^e, and T_t^p is the amount of taxes initially paid by an entrepreneur. Note that $T_t^e + T_t^p = T_t^0$ where T_t^0 is the amount of taxes officially owed by a firm. In the problem at hand, we assume that taxes are on revenues and write $T_t^p = \tau_p p_t F(z_t, k_t, l_t)$ and $T_t^e = (\tau_0 - \tau_p) p_t F(z_t, k_t, l_t)$ where τ_0 and τ_p are official and effective tax rates, respectively.

3.1.2 Workers

There is a measure of workers sufficient to satisfy labor demand at the equilibrium wage in the benchmark economy. Workers are risk-averse with preferences given by:

[25] In the problem at hand, we look at cases where $l^{ic} < l^f$ (otherwise all firms would be formal).

[26] Without this fixed cost, some firms stay idle and do not have any reason to exit (Hopenhayn and Rogerson, 1993).

$$\sum_{t=1}^{\infty} \beta^t \left[u(c_t) - d(n_t)\right] \tag{12}$$

s.t.

$$p_t^F c_t \leq w_t n_t \tag{13}$$

where c_t is consumption and n_t is labor supply which is either zero or one. Following HR, we assume that workers choose employment lotteries and have access to markets to diversify idiosyncratic risks. In the context of an LDC this amounts to assume some degree of cross-insurance among workers, possibly through the use of extensive family ties or some form of micro-credit. Workers can then be jointly represented through one representative agent with preferences given by:

$$\sum_{t=1}^{\infty} \beta^t \left[u(c_t) - DN_t\right] \tag{14}$$

where N_t is the fraction of employed workers and D is a disutility parameter. We assume throughout that $u(c_t) = \log(c_t)$.

3.1.3 Government and tax officials

There is a small measure (less than 1) of risk neutral tax officials in this economy, which is sufficient to generate the observed auditing frequency observed from data. All tax officials receive exogenous wage w_b and are potentially corruptible. There is a costless technology which imperfectly detects those who have accepted a bribe with probability $\psi \in [0, 1]$. A tax official caught receiving a bribe pays a fine $A \geq 0$.

The government expenditures are enforcement costs, i.e. wages to its tax auditors, bw_b; revenues come from taxes collected, $T_t^{gvt} = T_t^p + p(1-q)T_t^e$, fines collected from entrepreneurs caught evading, $p(1-q)FA$, and fines to corrupt tax officials, $bq\psi A$. The governmental budget constraint is:

$$bw_b \leq T_t^{gvt} + p(1-q)FA + bq\psi A \tag{15}$$

3.2 Stationary Equilibrium

3.2.1 Notation

First we look at the problem of an incumbent entrepreneur, taking entry as given. In our setup the labor decision is static and employment is always set so that the marginal product of labor equals the equilibrium wage. We can thus solve for labor in terms of capital, and write the

Bellman equation of an entrepreneur conditional on his status $s = \{iu, ic, f\}$ as a simple choice over capital investment:

$$V^s(k,z) = \max_{k'} \{u(c) + \varphi\beta V^s(k',z')\} \tag{16}$$

subject to the constraints 5, 6, 7 and 8. The next proposition states that a solution to each of these fully conditional value functions exist.

Proposition 1:[27] *(i) the fully conditional value functions, V^s, are concave, continuous, differentiable, monotonically increasing in k, k' and z; (ii) the operators TV^s are contraction mappings with modulus $\varphi\beta$ where T is defined in the appendix; (iii) let l^* be the optimal labor policy and I^*, the optimal investment policy.*

Proof: See the online Appendix.

We next write the problem of an entrepreneur that can move from one status to another, depending on her productivity shock. We call this problem the semi-conditional value function of an entrepreneur: [28]

$$V^s(k,z) = \max_{k'} \{u(c) + \varphi\beta CV(k',z')\} \tag{17}$$

where the continuation value is given by:

$$CV(k',z') = \int_{z'} \max\left[V^{iu}, V^{ic}, V^f\right] dg(z',z) \tag{18}$$

A brief remark about existence of a solution for the semi-conditional functions is necessary. The jump in regulation at l^{ic} introduces a non-convexity in the entrepreneur's decision problem. However, it readily follows from Proposition 1 that a solution for the semi-conditional function must also exists: to a specific capital-shock pair (k,z) corresponds one and only one status $s = \{iu, ic, f\}$ and we have shown in Proposition 1 that a solution for each status taken alone exists.

At the end of each period before making their investment decision, entrepreneurs decide whether to stay in the market or exit forever. Exit occurs when the continuation value CV is less or equal to zero. The exit set, ex, is defined by a pair (k^{ex}, z^{ex}) such that:

$$ex = \{k^{ex}, z^{ex} : CV(k',z') \leq 0\} \tag{19}$$

[27]The specification used to prove Proposition 1 in the Appendix differs slightly from the one proposed in the text. The main differences are: 1) the proof is kept as general as possible using an abstract transition function Q to specify how shocks evolve from one period to another; 2) the discount factor is specified by $\frac{1}{R}$; 4) the continuation value does not include the possibility of exit as this introduces a non-convexity in the problem.

[28]In the quantitative model we iterate over the semi-conditionals until joint convergence.

The unconditional value function is the upper envelope of the conditional value functions of an entrepreneur and it is given by:

$$V(k,z) = \max_{s \in \{ex,u,ic,f\}} V^s(k,z) \tag{20}$$

Only when an entrepreneur exits the market there is an opportunity for an entrant in the next period. This assumption is made for two reasons. First, it ensures that our algorithm remains tractable in the computational section. Second, in order to get a stationary equilibrium, one requirement is to equalize the mass of entrants, en, to the mass of entrepreneurs exiting the market, ex. Taking entry as given, the entrepreneur's decision problem produces three decision rules: one for the optimal choice of labor, one for the optimal choice of future capital and the other for the optimal exit decision. We denote these three optimal decision rules as L^d, K and EX.

Central to our analysis is the possibility for an entrepreneur to evade tax liabilities by offering a bribe to the tax official inspecting her books. Entrepreneurs bargaining power is denoted by θ and stems from an imperfect technology to tell on corrupt tax officials and get them prosecuted. We assume that entrepreneurs are aware that this bargaining will take place and choose investment accordingly before they meet with a tax official.[29]

The outside option for an entrepreneur is to repay what they have evaded plus a fine:

$$V^\tau(k,z) = p_F F(z,k,l) - wl - p_F(k' - (1-\delta)k) - c_f - T_0 - FA + \varphi\beta CV(k',z') \tag{21}$$

When a bribe is paid the value of a firm is:

$$V^B(k,z) = p_F F(z,k,l) - wl - p_F(k' - (1-\delta)k) - c_f - B - T_p + \varphi\beta CV(k',z') \tag{22}$$

The outside option for a tax official is to be honest and force the entrepreneur to pay her taxes in full, plus a fine. In this case, the tax official only gets her wage. A corrupt tax official may accept a bribe from an entrepreneur. In this case, the bribe is added to her wage but she faces a probability of paying a fine.

The Nash-bargaining problem is:

$$B = \arg\max \left[V^B(k,z) - V^\tau(k,z)\right]^\theta \left[B - \psi A\right]^{(1-\theta)} \tag{23}$$

Solving for the equilibrium bribe yields:

[29] This simplifies our analysis because continuation values cancel out when solving for the equilibrium bribe. It also saves computational time in the numerical exercise.

$$B = (1-\theta)(\tau_0^s - \tau_p^s)pzk^\gamma l^\eta + \theta\psi A \tag{24}$$

We use the equation for the equilibrium bribe as an estimating equation to back out from the data the value of firms' bargaining power parameter, as well as the expected fine faced by tax officials (more on this below). The representative worker's problem is a static optimization problem in a stationary equilibrium:

$$\max u(c) - DN \tag{25}$$

subject to 13. We denote the solution to this problem as L^s.

3.2.2 Definition of an equilibrium

A concise way to describe heterogeneity across individuals is to use a probability measure defined on subsets of the individual state space (see for example Lucas and Prescott, 1971; Hopenhayn, 1992). Let $\mu(k,z)$ be such a probability measure, summarizing the state of the industry in a given period of time.[30] Let the mass of entrants at the beginning of next period be *en*. Furthermore, let $G(k,z;k',z')$ be a transition function mapping current states into future states. This summarizes the optimal decisions of incumbents after exit decisions have been made. The evolution of the distribution of firms is then given by:

$$\mu'(k',z') = en + \int_{z^{ex}}^{z^{up}} G(k,z;k',z')\mu(dz) \tag{26}$$

The equation above states that next period's measure of firms is given by the number of new entrants and the number of incumbents transiting from their current state (k,z) to a future state (k', z'). Aggregate inputs are obtained by adding individual inputs across all firms in the economy.[31] We now have all elements to define a stationary equilibrium for this economy.

Definition 1: *A stationary equilibrium is a set of prices, (p^*, w^*), a set of decision rules, $\{z^{ic*}\}$, $\{z^{f*}\}$, $\{z^{ex*}\}$, $\{k'^*\}$, $\{l^{d*}\}$, and $\{l^{s*}\}$, and a distribution of firms $\{\mu^*\}$ such that:*

1. Decision rules are optimal:

    - *(a) given k, there exists a z^{ic*} such that $V^{iu} = V^{ic}$ for each entrepreneur;*
    - *(b) given k, there exists a z^{f*} such that $V^{ic} = V^f$ for each entrepreneur;*

[30] As is standard in the literature on firms dynamics we define an industry as a continuum of firms which produce a homogeneous good (Hopenhayn,1992).

[31] See the appendix for a definition of aggregate variables.

- *(c) given k, there exists a z^{ex*} such that the exit problem is satisfied for each entrepreneur;*
- *(d) l^{d*} and k'^{*} satisfy each entrepreneur's labor and capital problems;*
- *(e) l^{s*} satisfy workers' problem;*

2. μ^* is defined recursively by 26 and $\mu^{*'} = \mu^*$;
3. The government budget balances;
4. Prices are market clearing: $L^{s*} = L^{d*}$.

3.2.3 Optimal Decisions

Next we examine the optimal decisions of an entrepreneur. Taking the first order conditions with respect to labor, the optimal demand for labor is:

$$l = \left[\lambda^s \left(\frac{\eta}{w}\right) z k^{\gamma}\right]^{\frac{1}{1-\eta}} \tag{27}$$

where $\lambda^s = 1 - \tau_p - p_a^s(1 - q\theta)(\tau_0 - \tau_p)$. This is standard except for λ^s which is the burden from taxation.

In the case of capital, we obtain an Euler equation which is also standard but for the fact that expectations over future states depend on capital stock:[32]

$$\frac{1}{\varphi\beta} - (1-\delta) = \int_{z^{ex'}}^{z^{ic'}} \left[\gamma z' (k')^{\gamma} (l')^{\eta} - RC_k^{iu'}\right] dg(z', z) \tag{28}$$

$$+ \int_{z^{ic'}}^{z^{f'}} \left[\gamma z' (k')^{\gamma} \left(l^{ic}\right)^{\eta} - RC_k^{ic'}\right] dg(z', z)$$

$$+ \int_{z^{f'}}^{z^{up}} \left[\gamma z' (k')^{\gamma} (l')^{\eta} - RC_k^{f'}\right] dg(z', z)$$

In the appendix we show that investment is lower in this benchmark economy than in an economy with no distortions. Furthermore, the greater the capital stock at a firm the less will the entrepreneur invest, when facing a progressive tax schedule. This can be seen from equation 28: since shock-thresholds z^{ic} and z^f are decreasing with the level of capital accumulated, more weight is put on the last terms of the Euler Equation as a firm accumulates capital and this entails an increasing negative impact from the taxation burden on investment behaviour.

[32]This means that the lower and upper bounds of the integrals in 28 vary with capital stock.

3.2.4 First best

In our model regulation distortions interact with firm's investment. For comparison we examine an economy without distortions and where the rental market for capital is fully functioning. The next definition describes the first best in this economy:

Definition 2: *The first-best framework is defined as an economy with no regulation distortions and with a market for the rental of capital that functions properly.*

Entrepreneurs are able to rent capital in the market after observing the realization of their shock. The first-best delivers expressions for factor demand that are function of parameters only. Below we present an expression for the first-best capital-labor ratio:

$$\frac{k}{l} = \frac{\gamma w}{\eta r} \tag{29}$$

The optimal mix is constant across firms and it stays constant even if we introduce a uniform or a progressive tax on revenues. Obviously a tax on only one of the inputs would distort this ratio. It is not possible to obtain an analytical expression of this sort for the benchmark because the investment decision is made prior to the realization of next period's shock. However, the quantitative model generates capital-labor ratios that we can compare to those generated in the first-best.

In the first-best the relative demand for labor between any two firms is given by:[33]

$$\frac{l_i}{l_j} = \left[\frac{z_i}{z_j}\right]^{\frac{1}{1-\gamma-\eta}} \tag{30}$$

It is easy to see that a flat tax does not affect this ratio. Next we examine the relative demand for labor in the distorted economy:

$$\frac{l_i}{l_j} = \left[\frac{\lambda_i z_i}{\lambda_j z_j}\left(\frac{k_i}{k_j}\right)^{\gamma}\right]^{\frac{1}{1-\eta}} \tag{31}$$

We derive two important implications from equation 31. First, a progressive tax on revenues decreases the amount of labor employed in any given firm but relatively more in larger firms.[34] Second, there is a non-degenerate distribution of resources across firms *within* a productivity class due to each firm's specific capital accumulation process.[35] However, the Euler equation from the benchmark shows that a firm decreases its investment as it grows larger. Thus, the

[33] A similar expression can be derived for the relative demand for capital.

[34] See appendix.

[35] It can easily be shown that the dispersion of the distribution is higher under constant returns to scale than decreasing returns to scale. This is in line with findings by Ringstad(1971) who argues that constant returns to scale rationalize the large dispersion of establishment size observed in given industries.

interaction of investment and the tax distortion might reduce the dispersion of resources *across* productivity classes. The quantitative exercise is helpful to clarify the impact and the direction of this interaction.

3.3 Analytical example

In this section we present some analytical results for our model. First we state the assumptions needed to obtain an analytical solution. Then we discuss the main implications of this solution by summarizing some comparative statics' results.

We solve for a partial equilibrium version of the model. The restrictions are as follows:

1. The production technology is $F = z^{1-\gamma-\eta}k^{\gamma}l^{\eta}$ with $\gamma + \eta < 1$;

2. Capital depreciates fully after one period and there is no investment;

3. Markets for labor and capital are complete and prices are defined by $p^F = D(Q)$ for output Q, $r = r(K)$ for capital and $w = w(L)$ for labor where D is strictly increasing in Q, and r and w are strictly decreasing in K and L respectively;

4. Productivity shocks are independent and uniformly distributed between 0 and 1;

5. Enforcement is weak, i.e. $FA = 0$ and ψA is such that all tax officials always accept a bribe and thus $q = 1$;

6. Only formal firms are audited and thus $p_a^s = 0$, for $s = iu, ic$. The probability to be audited is equal to the share of tax officials in this economy, $p_a^f = b$;

7. There is no entry and no exit and firms can stay idle: $c_f = 0$;

8. Wages paid to tax officials are negligible, i.e. $w_b = 0$.

These assumptions make expressions for shock-thresholds analytically tractable.[36] The appendix contains analytical expressions for the variables of interest, namely capital and labor decision rules for all statuses, shocks' thresholds z^{ic} and z^f, as well as the equilibrium bribe.

If the distribution of firms remains invariant, i.e. $\mu_{t+1} = \mu_t = \mu$ for all t, then equilibrium prices also remain constant and we have a stationary equilibrium indexed at a particular price vector as stated in the next claim.

Claim 1: *If the above assumptions (1 to 9) are satisfied then there exists a unique stationary equilibrium associated with each price vector.*

[36] Relaxing assumptions 2, 3 or 4 requires computational methods to solve the model.

Proof: See the online Appendix.

Based on the analytical solution and the previous claim, we derive comparative statics for the parameters of interest. The results of the analysis are summarized in table T1.[37]

Table (T1): Comparative Statics

	k^{iu}	l^{iu}	k^{ic}	k^f	l^f	z^{ic}	z^f	B	iu	ic	f
p_F	+	+	+	+	+	−	−	+	−	+/−	+
w	−	−	0	−	−	+	+	−	+	+/−	−
r	−	−	−	−	−	+	+	−	+	+/−	−
θ	0	0	0	+	+	0	+/−	+/−	0	+/−	+/−
l^{ic}	0	0	+	0	0	+	+	0	+	+/−	−
b	0	0	0	−	−	0	+	−	0	+	−
ψA	0	0	0	0	0	0	+	+	0	+	−
τ	0	0	0	−	−	0	+	+/−	0	+	−

Note: The table shows the effect of an increase in the parameters of column 1 on variables of row 1

A change in aggregate demand If the price of output [input] increases [decreases], allocations of capital and labor in all sectors and bribes increase [decrease].[38] The formal sector expands [shrinks] and the informal unconstrained sector shrinks [expands]. However, this has an ambiguous effect on the informal constrained sector. The change in the ic-proportion depends on how many firms enter from the iu-sector and how many leave for the f-sector. We show analytically (see appendix) that, under a uniform distribution of shocks, the clustering effect at the size-threshold is alleviated when the price of output goes up: there are more firms leaving the ic-sector for the formal sector than there are firms entering from the iu-sector. The converse is observed when prices of inputs go up.

A change in bargaining power An increase in bargaining power as no effect on capital and labor choices in the informal sector. It increases the amount of capital and labor hired in the formal sector. However, it has ambiguous effects on the size of the bribe. The overall effect depends on the elasticities of the inputs with respect to the bargaining power. If the elasticities are large, it could be the case that firms increase their allocation of capital and labor by so much that they pay a bigger bribe than if their bargaining power had not been raised. Conversely, if

[37] Details on the comparative analysis are in the online appendix.

[38] A change in wage does not affect ic-capital.

the elasticities are small, a formal firm does not change its allocation of inputs by much when its bargaining power is increased but it has more leverage to negotiate a lower bribe. Finally, an increase in bargaining power increases the proportion of formal firms in this economy, otherwise parameters are such that no bribes are exchanged anyway.

A change in the audit probability An increase in the number of bureaucrats does not affect the allocations (capital and labor) in the informal sector, it decreases the allocations in the formal sector, it reduces the level of the bribe (since firms are hiring less capital and labor and thus producing less). The formal sector shrinks, the ic-proportion increases, and the iu-proportion is left unchanged.

A change in expected fine for bureaucrats Allocations in all three sectors are unaffected by a change in expected fine. However, the bribe increases with stricter monitoring since bureaucrats pass on the burden of the expected fine to entrepreneurs. As a consequence, the formal sector shrinks.

A change in the tax rate A change in the tax rate leaves allocations of capital and labor in the informal sector unchanged. Formal sector allocations are reduced following a tax hike and the formal sector shrinks. However, the effect of a change in the tax rate on the bribe is ambiguous as it depends on the magnitude of the elasticities of the inputs with respect to the tax rate. Suppose these elasticities are small then the level of the bribe increases with the tax rate: firms do not change much their allocation in response to a tax increase but a higher tax rate allows bureaucrats to negotiate bigger bribes. Suppose these elasticities are large: firms hire less inputs, which leaves less profits to be taxed, and in turn, smaller bribes to give away.

4 Quantitative model

In this section we revert to the original, dynamic model with capital investment described in section 3.

4.1 Calibration Strategy

4.1.1 Methodology

The procedure used to choose some of the structural parameters relies on a simulation-based method called Indirect Inference (Gourieroux et al., 1993; Smith, 1993). Indirect inference is useful when the theoretical moments that one wants to match to the data do not have neat analytical expressions. This is the case in our model due to the non-convexity arising from the discrete change in audit probability. The idea is to use an auxiliary model which needs not be an accurate description of the data generating process and choose parameters of the structural model so that the auxiliary parameters from the data and the simulations are the same.

We choose a u-vector of structural parameters, ϕ, so that a v-vector of auxiliary parameters from the model, m_s, matches a v-vector of auxiliary parameters from the data, m_d. The indirect inference criterion is given by:

$$\widehat{\phi} = \arg\min_{\phi} \, (m_s(\phi) - m_d)' \, W(m_s(\phi) - m_d) \tag{32}$$

where W is the variance-covariance matrix from the data. The procedure is based on repeating the following steps until a minimum is reached:

1. Pick a vector ϕ of parameters to be assigned;
2. Iterate over the conditional value functions until joint convergence of the optimal functions;
3. Simulate shocks for 10000 entrepreneurs over 100 periods to attenuate the effect of initial conditions;
4. Calculate simulated moments m_s in the exact same way as they are in the data;
5. Compare the simulated moments with those from data m_d;
6. Update the vector ϕ if a minimum has not been reached.

Equation 32 is analytically intractable and we cannot rely on standard optimization tools which use gradients to search for a minimum due to the non-convexity involved in the problem at hand. Instead, we use 'simulated annealing', a minimization algorithm that proceeds by random search. Also, obtaining a solution for 32 requires iterating over the conditional Bellman equations each time a parameter is updated. In order to save computational time, we constrain the set of parameters to be identified by indirect inference and calibrate regulation parameters directly from the Ugandan data and some technology parameters from the literature.

4.1.2 Details about parametrization

Table C1 presents the value and origin of all parameters used in the benchmark calibration.

Table (C1): Parameters		
Parameter	Value	Origin
Preferences		
β	.97	Discount factor: Assumption
φ	.93	Probability of survival: Firms average age in data
D	5.80	Disutility of work: Workers' problem
Regulation		
l^{ic}	30	Size-threshold into formality: data
τ_0	.144	Official tax rate: data
τ_p	$\tau_0/2$	Effective tax rate: data
θ	.98	Bargaining power: data
ψA	71\$/yr/emp.	Expected fine: data
p_a^s	.53	Probability of audit, small: data
p_a^m	.71	Probability of audit, medium: data
p_a^l	.77	Probability of audit, large: data
q	1	Probability of being asked for a bribe: assumption
w_b	1800\$/yr	Tax Officials' wage: IMF Staff Country Report 1998
FA	0	Fine for evasion: IMF Staff Country Report 1998
Technology		
δ	.07	Depreciation Rate: Assumption
γ	.283	Basu and Fernald (1997); Restuccia and Rogerson (2008)
η	.567	Basu and Fernald (1997); Restuccia and Rogerson (2008)
w	1096\$/yr	Workers' wage: data
c_f	0	Fixed Cost of Production: Assumption
p_{zz}^l	.85	Probability to stay in current state for 10 smallest shocks: average employees in data

Preferences We set the length of a time period in our model to be one year in the data. We assume that all agents in the model discount the future at the rate $\beta = 0.97$.[39] The disutility parameter, D, is calculated in the following way. A solution for $\widehat{\phi}$ implies an equilibrium value for L^d. We then pick D to ensure that the market for labor clears using the first order condition from the household problem:[40]

$$\frac{1}{N} = D \tag{33}$$

where N is the fraction of employed workers.[41]

Regulation environment The size-threshold into formality, l^{ic}, is based on the empirical analysis from Goyette (2009) and set at 30 employees.[42] We use the ratio of tax obligations per sale found in the data to set a flat official tax rate $\tau_0 = 0.14$. Based on our estimates we assume that the effective rate of taxes paid,τ_p, is half the official rate. Audit probabilities for small, medium and large firms are also taken from the data and set to 0.53, 0.71 and 0.77, respectively.

The bargaining power of formal entrepreneurs and the yearly expected fine for tax officials are obtained by estimating equation 24. As can be inferred from equation 24, the coefficient on the surplus from evasion delivers θ. The expected fine per employee is then obtained using θ and the value of the constant in the regression. Table C2 presents the results for an OLS of bribes on a constant, the surplus from evasion, dummies for industry and location and other covariates, including age, age squared, price of capital and the average wage per employee.[43] Given the small size of bribes compared to the amount of evasion in the data, it is not surprising to find such high values for θ in Table C2.[44] Interestingly, we note from rows 1 to 4 that expected fines

[39] Arellano, Bai and Zhang (2007 and 2008) use $\beta = 0.96$ for Ecuador and $\beta = 0.94$ for Bulgaria. Many papers based on US data use a discount rate around 0.98 (Atkeson and Kehoe, 2005). We make a middle of the range choice by setting $\beta = 0.97$. Results are not too sensitive to changes in β.

[40] Note that this procedure imposes some restrictions in the way we can conduct policy experiments later. Fixing D means that also aggregate labor stays fixed. We must thus adjust wages to keep aggregate labor as in the benchmark in all general equilibrium experiments below. Under the P.E. assumption of constant prices we instead assume that labor supply can always satisfy labor demand at all.

[41] The fraction of workers employed is given by the ratio of aggregate labor demanded generated in the benchmark to total labor force in the private sector in Uganda in 1997 (7,164,744 employees) based on data from the UN Statistics Division.

[42] This value also corresponds to the maximum number of employees in small firms used in other works. See for example, Steel and Webster (1992), Tybout (2000) and Gauthier and Reinikka (2006).

[43] Bribes in Uganda are not only paid to evade taxes. This is why we have included other covariates in the estimating equation. Dummies for industry and location control for composition effects. Age accounts for the fact that older firms might establish advantageous ties with the tax administration. Factor prices control for the advantages/disadvantages to be in the informal/formal market.

[44] Note that the average bribe per worker is approximately one month of workers' wage and half a month of a

Table (C2): Bargaining power and Enforcement costs				
	Bargaining Power (θ)	Expected Fine/emp (ψA)	R_squared (nb.obs.)	Average Bribe/emp.
Small	.995	32.12	.14	47.16
(Industry-Location)	[.95]	[.93]	(77)	(77)
Medium	.986**	210.39*	.50	120.81
(Industry-Location)	[2.40]	[1.97]	(33)	(36)
Large	.988**	72.87	.20	72.66
(Industry-Location)	[2.87]	[.86]	(58)	(62)
All	.988**	99.75**	.26	71.35
(Industry-Location)	[2.18]	[2.49]	(167)	(175)
All	.983***	380.98***	.38	71.35
(covariates)	[3.41]	[3.27]	(122)	(175)
All	.978***	476.61***	.58	71.35
(covariates + p_a)	[5.73]	[4.53]	(85)	(175)

Note: * significant at 10%, ** significant at 5%, *** significant at 1%; nb. of firms in parenthesis, t-stat in brackets; First two columns are imputed from a regression based on the Bribe equation and column 4 is the average bribe per employee; All regressions include dummies for industry-location and other covariates include Age, Age_squared, price of capital, average wage per employee; Results are not significantly affected if we use total amounts instead of amounts per employee.

(column 3) are almost equal to average bribes (column 4). The level of the bargaining power is not significantly affected by the introduction of covariates. However, the expected fine increases with the introduction of covariates. From different calibration exercises we find that an estimated value of $\theta = 0.98$ and the use of the average bribe from the data for ψA bring about the best fit of model to data.

From IMF Staff Country Report 1998 we know that, for the period going from July to December 1997, only 16% of the assessments were refunded through ordinary tax audits. We thus assume that all tax officials are potentially corruptible in the benchmark and set $q = 1$. The report also states that compliance to fill tax reports was low. We thus set the fine paid by entrepreneurs, FA, to zero . Finally we use the IMF Staff Country Report 1998 to assess the average public

tax official wage. This entails a bribe per firm (for the average firm) which is about 4 times the yearly wage for tax officials.

wage in Uganda in 1997 which is approximatively $w_b = US\$1800$.[45]

Technology Parameterizing technology presents some challenges. Goyette (2009) estimates shares into production for capital and labor. Given that output, capital and labor respond to productivity shocks at the same time, Goyette(2009) uses a two-steps estimation procedures. The results are reported in table C3.

Table (C3): Capital and labor Shares

	All		Small		Medium		Large	
	γ	η	γ_s	η_s	γ_m	η_m	γ_l	η_l
All	.43***	.68***	.14	.81***	.43**	.50***	.58***	.30*
	[.09]	[.09]	[.14]	[.15]	[.17]	[.13]	[.13]	[.16]
	(103)		(42)		(22)		(39)	
Manufacturing	.36***	.79***	.17	.81***	.51***	.43**	.51***	.60***
	[.10]	[.10]	[.15]	[.19]	[.21]	[.15]	[.14]	[.22]
	(78)		(34)		(16)		(28)	
Kampala	.55***	.59***	.23	.58	.54	.52***	.64***	.26
	[.12]	[.12]	[.24]	[.34]	[.29]	[.23]	[.15]	[.16]
	(53)		(16)		(14)		(23)	

Note: * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in Brackets, number of firms in parenthesis; Capital is estimated based on resale value of machinery, equipment and other assets and a depreciation rate of 0.07; labor is proxied by wage bills divided by average wage per size-bin; All variables are expressed in logarithms which implies that the input coefficients are input elasticities; all regressions account for sector fixed effects.

For the three size bins total returns to scale in capital and labor seem to be decreasing. However, the reported number of observations are low and, given the nature of the data, it is difficult to argue that these estimates represent aggregate technology in Uganda. Using US data on 34 industries Basu and Fernald (1997) have estimated total returns to scale for a typical industry, $\gamma + \eta$, to lie between 0.8 and 0.9. We make a conservative choice and follow Restuccia and Rogerson (2008) who use $\gamma + \eta = .85$ for the US and attribute parameter values according to capital and labor shares of income,i.e., 1/3 and 2/3, respectively. Note that Soderbom and Teal (2004) find a very similar assignment of income shares for Ghana. We assume a rate of

[45] The total public wage bill for 1997 in Uganda was 227 billions of Ugandan Shillings. The public work force was 124,664 employees. Note that the bribe per capita is US$71 which is about half a month of salary for tax officials.

depreciation of $\delta = 0.07$. Workers' wage is taken from the data and is equal to $US\$1,096$ per year. We set the fixed cost of production to zero. [46]
As in HR, there is an identification issue in that the price of output, productivity shock and capital stock enter multiplicatively in the production function: the effect of a high price is hard to distinguish from a high value of the current shock, or capital stock. A few assumptions are thus in order. First, we normalize the price of output to one and choose other parameters consistently. Second, we use the relative demand for labor, equation 31, to pin down the range of productivity shocks. In order to do so, we use the range of firm sizes between 1 and 2,000 employees in Uganda. We use a capital grid with 250 log-spaced points where we normalize the lower bound to one and conduct a sensitivity analysis to locate the upper bound.[47] This gives us a support for shocks that ranges from $z^{low} = 1$ to $z^{up} = 3.8$. We follow Hopenhayn and Rogerson (1993) and impose a grid for productivity shocks with $n_z = 20$ evenly spaced points. We assume that the initial distribution of shocks, ν, is a uniform on $[z^{low}, z^{up}]$.

4.1.3 Indirect inference

This leaves us with the calibration of the survival rate φ and the transition process $g(z', z)$. We have information on firms age but not on firms' turnover in the data. Roberts and Tybout (1997) estimate that the 1-year exit rate for Chile, Colombia and Morocco are 8.5%, 11.9% and 9.5%, respectively. We calibrate φ to match firms' average age in Uganda and verify that the 1-year exit rate is in an acceptable range of values based on Roberts and Tybout (1997). We keep the transition process $g(z', z)$ as simple as possible.[48] Let the probability that next period productivity shock, z', equals the current period productivity shock, z, be given by:

$$p_{zz} = \xi(z' = z)\,\Pr(I = 0) + \Pr(I = 1) \tag{34}$$

where ξ is a uniform distribution and I is a dichotomous random variable drawn from a binomial

[46] The fixed cost is an important feature of this class of models à la Hopenhayn and has several implications. First, the fixed cost acts as a selection mechanism in that it affects the productivity-threshold for exit. However, the effect is ambiguous and depends on the relationship between the shock process and the elasticity of profits with respect to the fixed cost (Hopenhayn, 1992). Since we already have one selection mechanism (the change in audit probability), it seems reasonable to set the fixed cost to zero and concentrate on the effect of the mechanism of interest. Second, we do not have empirical evidence to calibrate the fixed cost based on data. Finally, calibrating fixed costs with Indirect Inference would yield one more degree of freedom to fit model to data. We prefer to impose greater discipline on the model and use a smaller set of structural parameters for the calibration.

[47] Our criterion in the sensitivity analysis is that no firms should cluster at the upper bound on the capital grid. This provides us with the smallest possible capital grid which economizes on computational time. Note also that the grid is as in Hopenhayn and Rogerson (1993)

[48] One reason for this is that we want the model to generate a size distribution of firms similar to the data without explicitly targeting it.

equal to one if next period productivity shock, z', is the same as the current period productivity shock, z. Clearly, if the current state is z and $y_1, ..., y_{n_z-1}$, are all *other* possible productivity levels on the grid for shocks then $p_{zz} + \sum_{j=1}^{n_z-1} p_{zy_j} = 1$. Thus assuming that the probability to move from the current level of the productivity shock z to any *other* productivity level is the same for all y_j where $j = 1, ..., n_z - 1$, we have:

$$p_{zy} = \frac{1 - p_{zz}}{n_z - 1}, \forall y_j \tag{35}$$

We calibrate p_{zz} for the ten lowest shocks on the productivity grid (the same value for all of these shocks) and match the average firm size and the average growth rate of sales in the data. The magnitude of p_{zz} is a degree of freedom that is used to match one of the two moments. The choice of the number of grid-points on the diagonal constitutes another degree of freedom to match the other moment. From 35 above p_{zy} readily follows for the ten lowest shocks. For the ten highest shocks on the productivity grid transition probabilities across any state are simply drawn from a uniform and $p_{zz} = p_{zy}$, $\forall y_j$ where $j = 1, ..., n_z - 1$.

Notice that, if $u < v$, the model generates a set of over-identifying restrictions which can be used to test it.[49] We thus supplement the auxiliary model with two other moments that describe aggregate data: average capital-labor ratio and average capital-output ratio.

4.2 Model validation

4.2.1 A test of the model

We provide a formal test using the over-identifying restrictions from the Indirect Inference procedure. Table C4, panel A, contrasts the simulated moments obtained from the calibration against the targeted moments in the data.[50] The statistic generated by the quantitative model is $\chi^2_2 = 5.14$ and lies below a critical value of 5.99 for a χ^2-distribution with two degrees of freedom and 5% type I error probability. Thus we can't reject null hypothesis that the moments generated by the model are equal to the moments from the data, at least along the targeted dimensions.

4.2.2 Non-Targeted Statistics

In this section we examine some statistics generated by the model that were not targeted in the calibration. Table C4, panel B, summarizes results.

Firm Size Distribution The model generates an equilibrium size distribution of firms that compares well to the data. Examining the number of firms for the three size bins we note that

[49] If $u = v$ the method is equivalent to the Simulated Method of Moments.

[50] Note that we have normalized the capital-labor ratio to make it comparable to the data.

Table (C4, panel A): Targets and Wald Test		
Variable	Benchmark	Data
Av. Nb. Emp.	132	124
Av. Age	13.9	13.9
Capital/Output	.869	1.34
Capital/labor	59.3	59.3
Av. Growth	.124	.08
χ^2	5.144	—

Note: Critical Value based on a χ^2 distribution with 2 degrees of freedom and a right-tail area of 0.05 is 5.991.

the model generates a share of small firms in line with the data. The model tends to overestimate the share of large firms to the expense of medium firms. Figure 5 panels A and B compare the size distribution from the data and the benchmark using bins of 30 employees. We clearly see that the model underestimates the share of firms in the second bin (30 to 60 employees) and overestimate the share of firms in the third bin (60 to 90 employees).[51] This is due to the fact that we introduce a very stark change in the audit probability at the size-threshold whereas in the data this change occurs within a size range. Beyond 100 employees, the size distributions from the benchmark and the data lie almost exactly together. The benchmark produces average sizes for small and medium firms that are close to the data. The size of large firms is underestimated due to the fact that the model generates too many large firms.[52]

Growth and Shares of Output A peculiar aspect of the Ugandan data is that patterns in firms' growth rates actually go against some findings in the literature on firms dynamics (Evans, 1987; Hall, 1987). Small Ugandan firms exhibit lower growth rates than their larger counterparts. Interestingly, the quantitative model delivers patterns in growth rates that fit the data reasonably well. Small firms appear to refrain from growing because of the rise in audit probability and regulation costs, and larger firms can experience higher growth rates taking advantage of the high capital-labor ratios from the earlier phases of their life.

Capital-labor ratios An important robustness check is whether the model endogenously generates capital-labor ratios that are similar to those measured from data, especially for small and medium firms. We can see in table C4 (panel B) that the average capital-labor ratio of small

[51]Note however that we also observe a second mode in the data around 400 employees.

[52]The median size in the model (51 employees) is higher than in the data (35 employees).

Table (C4, panel B): Non-Targeted Statistics			
Variable	Size bins	Benchmark	Data
Share of firms	Small	.47	.48
	Medium	.12	.20
	Large	.41	.32
Average Size	Small	17	14
	Medium	59	50
	Large	285	334
Average Age	Small	9	12
	Medium	14	15
	Large	20	16
Maximum Age	All	97	74
Output Growth	Small	.04	.02
	Medium	.19	.13
	Large	.20	.17
Share of Output	Small	.06	.03
	Medium	.06	.14
	Large	.88	.83
Bribes per Sales	All	.02	.02
Capital/labor	Small	65.4	78.0
	Medium	60.9	38.0
	Large	52.4	50.0

Note: Scaling factor to convert a unit of simulated capital to a unit of capital in the data is 59.3; K/L are in millions of consumption unit per unit of labor.

firms is greater than the ratio for medium firms. Note that large firms have a lower capital-ratio than medium firms. This is in line with our previous discussion of equation 28: for any given shock z, an entrepreneur faces higher regulation costs the higher her capital stock. It thus seems optimal for larger firms to marginally reduce their capital stock and adjust labor after the realization of the productivity shock. [53] The increase in capital-labor ratio from medium to large firms observed in the data can be rationalized if collateral is needed to have access to credit. In

[53]This implication from the Euler equation is not the driving force behind the pattern in capital-labor ratios between small and medium firms. In the next section we will show that after removing this distortion, capital-labor ratios of small firms are actually lower than those of medium firms.

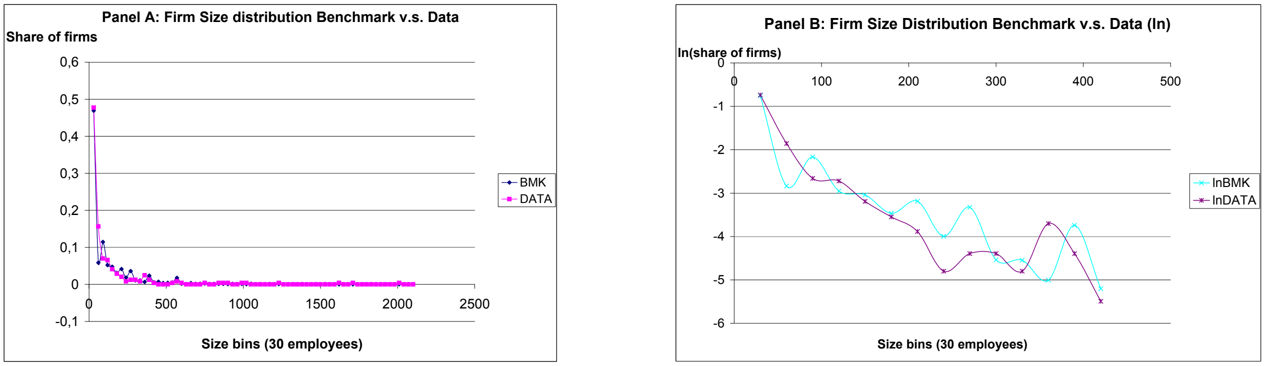

Figure 5: Size Distribution of Firms: Benchmark v.s. Data

such a situation larger firms may levy more debt if they have more capital. [54]

Figure 6 shows the correlation of the capital-labor ratio with firm size for firms with 200 employees or less in the benchmark and should be compared to Figure 3 (panel A). Figure 6 documents a clustering of firms at the size threshold ($\ln(30) = 3.4$) in the benchmark. This is where small firms substitute capital for labor.

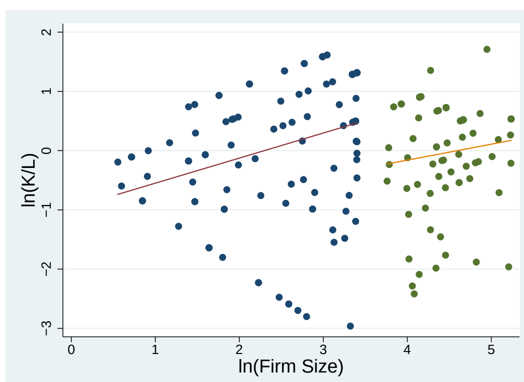

Figure 6: Capital-labor Ratios v.s. Firm Size, Benchmark

Column 1 and 2 in table (C5) present the results of discontinuity regressions using both Ugandan data and simulated data, for the group of firms with less than 31 employees and the group of firms with more than 30 employees. From column 2 we note that the coefficients for small and larger firms are significant but of smaller magnitude than those found in the real data. The Wald statistic F1 rejects the null that the coefficients are jointly zero and the Wald statistic

[54] Goyette (2009) shows that small firms are heavily credit constrained and do not levy any debt while large firms have some access to credit in Uganda.

F2 rejects the null that the coefficients are equal both for the Ugandan data and the simulated data.

Table (C5): Regression Discontinuity Results

Dependent Variable: ln(K/L)

	Data	Benchmark	Case 2	Case 3
Smaller	.80***	.36***	.198*	.0059***
firms	(.23)	(.095)	(.105)	[.0007]
Larger	.56***	.15***	.108**	.0007*
firms	(.13)	(.051)	(.052)	(.0004)
R_squared	.09	.06	.02	.25
F1	10.54***	7.39***	2.17	39.92***
	[0.0]	[0.0]	[.12]	[0.0]
F2	3.90**	14.0***	1.88	78.24***
	[.05]	[0.0]	[.17]	[0.0]
Nb. Obs.	206	243	243	243

Note: * significant at 10%; ** significant at 5%; ***significant at 1%; Std. Dev. in parenthesis and F-Value in brackets; F1 Wald test with H0: coefficients on Small and Larger firms are zero jointly; F2: Wald test with H0: coefficients on Small and Large firms are equal; smaller firms ¡ 31 employees and larger firms ¿ 30 employees. All regressions with a constant. Only firms with 500 employees or less in the regression using Uganda data.

Taxes and bribes By construction taxes per sale in the model fit the data exactly. The model does also a good job at generating the ratio of bribes to sales.[55]

Age and firms turn-over The benchmark generates a one-year exit rate of 7.2%, which is slightly below the values found in Roberts and Tybout (1997). In terms of age, small firms are younger than in the data. The model does a reasonable job at estimating the average age of medium firms while the age of large and oldest firms are over-estimated.

[55] Equation 24 contains two terms. The first term concerns the bargaining over the surplus from evasion and accounts for a very small part of a bribe. The second term is the bulk of the side-payment and is a function of the expected fine that tax officials pass on to entrepreneurs for letting them evade.

4.2.3 Regulation distortions and capital market

Next, we examine the implications of the distortions we have added to the model of HR (1993). We compare the benchmark to 3 alternative cases. Case 2 is an economy where tax on firms revenues are set to zero and the market for the rental of capital is absent. In case 3 we introduce a perfect market for the rental of capital, but there is still a regulation distortion due to taxes and bribes. Finally, case 4 is what we have described as the first best. Table (C6) summarizes the percentage change from the benchmark for key variables in cases 2 to 4, except for capital-labor ratios for which the actual values are given. In all cases we focus on the allocation of a given amount of labor input: therefore the wage is adjusted so that aggregate labor supply is kept at the level of the benchmark.

Table (C6): Model Validation
(% change from the benchmark)

Experiment	Case 2	Case 3	Case 4
Variable	%	%	%
wage	14.3	50.2	67.2
Share small	-18.7	49.7	28.2
Share medium	72.5	-22.9	59.7
Share Large	-.07	-50.1	-50.0
Agg. Capital	22.3	60.9	78.9
Agg. Output	6.1	50.5	55.1
Agg. Taxes	-100	50.5	-100
Agg. Bribe	-100	50.5	-100
Exit	0	50.1	50.1
K/L small	60.0	59.4	59.3
K/L medium	68.0	59.0	59.3
K/L large	53.9	59.0	59.3

Case 2: No market for capital and no distortions; Case 3: Perfect market for capital and distortions; Case 4: Perfect market for capital and no distortions. K/L are in millions of consumption unit per unit of labor.

Case 2 From table (C6) we note that in Case 2 the wage goes up by 14.3%. As expected small firms that were on the verge of becoming medium in the benchmark and refraining from growing due to distortions reallocate themselves to the medium size bin. Overall this entails a 6.1% increase in aggregate output. The last rows of table (C6) show that the average capital-labor ratio is lower for small firms than for medium firms. This should come as no surprise since we have removed the regulation distortions. In Figure 7 (panel A) both clustering and gap in the capital-labor distribution have vanished. This should be compared to Figure 3 panel B for US data. Examining the regression results with the simulated data from case 2 in column 3 of table (C5) we note that the coefficient on small firms is not significantly different that for medium firms. These results suggest that the 'missing middle' is due to regulation and tax distortions, rather than imperfect credit markets. They also suggest that the optimal responses of entrepreneurs to regulation distortions, resulting in the 'missing middle' in the distribution of firms' size, is associated to a relatively small loss in terms of efficiency and output per worker.

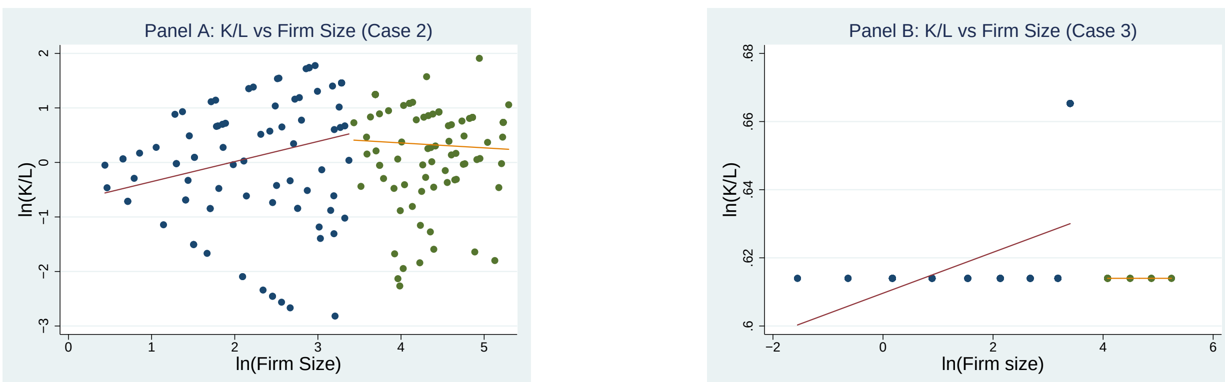

Figure 7: Capital-labor Ratios v.s. Firm Size, Case 2 and 3

Case 3 Introducing perfect capital market entails large changes in aggregate variables. Table (C6) shows that in case 3 the wage must be increased by 50.2% to keep aggregate labor at the level of the benchmark. Nevertheless aggregate capital and output go up by 60.9% and 50.5%, respectively. The average capital-labor ratio is higher for small firms than for medium firms but equal for medium and large firms: this is because large firms partly lose the incentive to accumulate extra capital. A graphical representation of the capital-labor ratios is best suited to examine other implications from the theory. First, Figure 8 (panel B) documents no dispersion *within* each productivity class. Observations for firms with similar productivity collapse at identical points and lie on top of each other in Figure 7 (panel B). However comparing the overall dispersion of firms' size in the benchmark (variance of firms size equals 42263) and Case 3 (variance of firm size equals 85755) we note that differences in employment level between large and small firms offset the effect of having a degenerate distribution of resources *within* each productivity class.[56] We notice that even with perfect capital markets small firms on the verge of becoming medium-sized substitute capital for labor at the size threshold when there is a distortion: this corresponds to the cluster of higher points at the size threshold in Figure 7 (panel B).

Case 4 In case 4, distortions are removed and the market for the rental of capital functions properly. This brings about the highest changes in aggregate variables (see table C6). The wage increases by 67.2%. Aggregate capital and output increase by 78.9% and 55.1%, respectively. The reallocation of firms towards the small size bin is not as stark as in Case 3, thanks to the removal of regulation distortions. As predicted in the theoretical section, capital-labor ratios are now similar across all firms.

In case 2 and 3 the removal of the distortions due to the tax environment and/or the lack of capital markets generates gains in output per worker that are equal to the gain in aggregate output in Table (C6). This is due to the fact that aggregate labor is kept constant by adjusting the wage in all experiments. Case 4 (first-best) is a rough proxy for the output that could be generated by a relatively undistorted economy, like the US. Using data from Hall and Jones (1999), output per worker in the US was about 31 times higher than in Uganda in 1988. The model shows that removing the regulation constraint (Case 2) would only reduce this gap by 11%. Note however that this estimate constitutes a lower bound on potential efficiency gains because of our assumption that aggregate labor must be kept constant in the experiment. The wage is thus capturing part of the productivity increase due to the change in the environment. However the introduction of a working capital market (Case 3) could go as far as reducing the gap in output per worker between the US and Uganda by 92%.

4.3 Policy Analysis

4.3.1 Smooth tax schedule

In this experiment, we look at the effect of smoothing the audit probability evenly across firms. In order to exclusively focus on the reallocation of existing labor resources, we let wage adjust to keep aggregate labor employed at the same level as in the benchmark. Enforcement costs are also held constant by assuming that the number of bureaucrats stays constant and the same number of firms as in the benchmark are audited. The audit probability is set to $p_a = .65$ for all firms. The results are summarized in column 1 of Table (C7) where we tabulate percentage

[56] The variance of firm size in the data is 67038.

changes in aggregate variables relative to the benchmark. Capital-labor ratios are presented in levels.

Table (C7): Policy Experiments (% from the benchmark)		
Experiment	Smooth Auditing	Honest Bureaucracy
Variable	%	%
τ_c	—	-.0463
wage	.2	.546
Share Small	-10.0	-3.3
Share Medium	41.1	24.7
Share Large	-.8	-3.5
Agg. Capital	1.0	45.6
Agg. Output	.3	15
Agg. Taxes	.3	0
Agg. Bribe	-7.4	-100
K/L small	61.5	71.2
K/L medium	68.2	97.3
K/L large	53.3	64.9
Note: % change from the benchmark except for K/L which are in levels.		

Smoothing the regulation schedule implies a 0.2% increase in wage. About 10% of small firms reallocate into the medium size bin. Overall, this entails an increase of 41.1% in the number of medium firms. Aggregate capital and output increase by 1.0% and 0.3%, respectively. Note that the aggregate bribe decreases by 7.4% and this is due to the fact that more of the auditing is now done on smaller firms. Capital labor ratios move in the expected direction: small firms now exhibit lower capital-labor ratios than medium firms. The efficiency gains from this policy change are quite small and suggest that the optimal distribution of audit probabilities might indeed be progressive when smaller firms have a harder time financing their activities than their larger counterparts. [57]

[57]We contrast these results with the existing literature. Keen and Mintz (2004) examine the efficiency implications of the choice of an optimal value-threshold for a value added tax. They find that, depending on the regulation environment, the optimal threshold could be set either to a very low value, capturing most firms, or to a very high value, capturing only the firms which contribute the most to the tax base. However, their model does not incorporate capital accumulation nor tax evasion. In our model, setting l^{ic} to a high value is not an option because entrepreneurs hide under the threshold and substitute capital for labor. Setting l^{ic} to a low value, which

4.3.2 Honest Bureaucracy

What is the effect of corruption and bribes in this model? We answer this question by setting the probability that a bureaucrat accepts a bribe to zero, i.e. $q = 0$. Also in this experiment aggregate labor is kept as in the benchmark by adjusting the wage. We introduce a tax (rebate) on consumption to keep total tax collection at the same level as in the benchmark: we focus on changes in consumption taxes because they are a relatively common revenue-maker in developing countries. Finally, we keep tax officials' total earnings (official wage bill + aggregate bribe) at the same level as in the benchmark, so that auditors are made indifferent between being in a corrupt r non-corrupt equilibrium. The equivalent of the aggregate bribe from the benchmark is now paid to tax officials out of tax revenues.[58]

The effects of a change from a corrupt to a honest tax administration are summarized in column 2 of Table (C7). By changing q we are effectively changing the expected costs from regulation. Audited entrepreneurs now pay roughly double the amount of taxes they were paying in the benchmark, and this has two effects. First, it increases the aggregate amount of taxes collected by the government. Second, entrepreneurs scale down their operations due to greater regulation costs. The extra revenues from taxation can thus be used to subsidize consumption ($\tau_c = -.0463$) and this in turn stimulates aggregate demand and entices entrepreneurs to scale up their operations. Moreover, given that auditors are non corruptible, no size-dependent distortion is present. As a consequence of this, 3.3% of small firms become medium firms. The wage is increased to keep aggregate labor constant and a share of 3.5% of large firms become medium firms. Overall, this entails increases in aggregate capital and sales of 45.6% and 15% respectively.

Examining the capital-labor ratios across firm size we note that the increase in aggregate capital is mainly due to an increase in capital accumulated by medium firms. This implies that the increase in production should also be due to medium firms. Indeed, the share of output produced by small firms and large firms falls by 11% and 0.8% respectively while that of medium firms increases by 26%.[59]

In the benchmark bribes act as a second-best mechanism and actually help smooth the taxation schedule by buying a tax rebate to entrepreneurs. However, the extra-revenues generated by the introduction of a reservation wage that makes tax officials honest imply a greater efficiency gain than bribes. To see this note that the effect of a change in q or θ on the burden from regulation λ^s is the same.[60] If $\theta = 0$ then entrepreneurs who are audited pay double the amount

is equivalent to smoothing the regulation schedule, seems a more appropriate prescription in the context of our model.

[58]We assume that such a wage is sufficient to lead tax officials to collect tax revenues that were previously evaded and not accept bribes.

[59]Results available upon request.

[60]Indeed, the burden from regulation is: $\lambda^s = 1 - \tau_p - p_a^s(1 - q\theta)(\tau_0 - \tau_p)$

of regulation costs they were paying in the benchmark.[61] However, the extra regulation costs paid on top of what was paid in the benchmark are now paid as bribes. Clearly consumption can no longer be subsidized and this reflects in lower demand and, perhaps surprisingly, in an aggregate efficiency loss.

5 Conclusion

We analyze the implications of an institutional distortion from the fiscal environment and of missing credit markets on resource allocation and aggregate efficiency. We find that amending a growth model with entry and exit of heterogeneous firms to include audit and tax distortions and capital accumulation generates an equilibrium firm size distribution, and patterns in capital-labor ratios, similar to those observed in micro-data for Uganda. The numerical counterpart of the model fits the data reasonably well in several dimensions and is validated using over-identifying restrictions and untargeted data moments.

Using an undistorted economy with working capital markets as a comparison, we find that removing all regulation and tax distortions would close the gap in output per worker by 11%. However, introducing perfect capital markets would generate larger gains and account for a 92% reduction in the observed gap. We also find that the 'missing middle' in the firms' size distribution is a by-product of the institutional distortions in the tax system and is associated to relatively small losses in aggregate efficiency. Our results suggest that the 'missing middle' simply reflects the optimal responses of forward looking firms to the tax and auditing environment, and that such responses are quite effective in reallocating resources so to minimize the impact of the distortion.

We conduct two policy experiments and show that negligible efficiency improvements may result from an even distribution of audit probabilities across firms. This evidence suggests that a tilted, progressive auditing probability – as the one observed in our data set for Uganda – might actually have some positive effects when credit markets are missing. To the extent that small firms find it hard to finance their investment through cash flow, it might be helpful to have a reduction in their de facto tax liabilities due to a size-dependent auditing probability.

We find that bribes may act as a useful second-best mechanism in a distorted environment and may help smooth the regulation schedule; nonetheless offering a wage compensation to tax officials such that they are just as well off as they would be in the 'corruption' equilibrium – and therefore willing to forego bribes – can generate very large efficiency gains. We find that such a policy might reduce the gap in output per worker between Uganda and an undistorted economy

[61] Note that when $\theta = 0$, the second term in equation 24 drops out. From the point of view of the entrepreneurs the effect of a change in q or θ is the same.

by roughly 27%.

One key feature of the model is that the discrete jump in audit probability and regulation costs generates a cluster of firms below the exogenous size-threshold. Such a phenomenon is not specific to the environment at hand, as one would suspect similar clusters to arise wherever a sharp change in policy conflicts with agents' incentives. The threshold around which the clustering takes place could be endogenized in the context at hand, as well as in other contexts. For example the model of this paper could be extended to have individuals decide whether to become productive entrepreneurs or rent-seeking bureaucrats, based on expected payoffs; alternatively one might explicitly model the impact of remuneration incentives on the decision to become corrupt and the dynamic bargaining process between firms and auditors.

References

Acemoglu, D., Johnson, S., Robinson, J., 2001. The colonial origins of comparative development: An empirical investigation. The American Economic Review 91 (5), 1369–1401.

Albuquerque, R., Hopenhayn, H., 2004. Optimal lending contracts and firm dynamics. The Review of Economic Studies 71 (2), 285–315.

Alfaro, L., Charlton, A., Kanczuk, F., 2007. Firm-size distribution and cross-country income differences. manuscript, Harvard Business School.

Arellano, C., Bai, Y., Zhang, J., mimeo. Capital Structure and Contract Enforcement. Working Paper.

Atkeson, A., Kehoe, P., 2005. Modeling and measuring organization capital. Journal of Political Economy 113 (5), 1026–1053.

Banerjee, A., Duflo, E., 2008. Do firms want to borrow more? Testing credit constraints using a directed lending program. MIT Working Paper Series.

Basu, S., Fernald, J., 1997. Returns to scale in US production: Estimates and implications. Journal of Political Economy 105 (2), 249–283.

Buera, F., Kaboski, J., Shin, Y., 2009. Finance and development: A tale of two sectors. NBER Working Paper.

Cleveland, W., 1979. Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association 74 (368), 829–836.

de Soto, H., 1989. The Other Path: The Invisible Revolution in the Third World. London.. IB Tauris.

Ericson, R., Pakes, A., 1995. Markov-perfect industry dynamics: A framework for empirical work. The Review of Economic Studies, 53–82.

Evans, D., 1987. The relationship between firm growth, size, and age: estimates for 100 manufacturing industries. The Journal of industrial economics 35 (4), 567–581.

Gauthier, B., Gersovitz, M., 1997. Revenue erosion through exemption and evasion in Cameroon, 1993. Journal of Public economics 64 (3), 407–424.

Gauthier, B., Reinikka, R., 2006. Shifting tax burdens through exemptions and evasion: an empirical investigation of Uganda. Journal of African Economies 15 (3), 373.

Gourieroux, C., Monfort, A., Renault, E., 1993. Indirect inference. Journal of Applied Econometrics 8, 85–118.

Goyette, J., 2009. Determinants of the Size Distribution of Firms in Uganda. Working Paper, UBC.

Guner, N., Ventura, G., Xu, Y., 2008. Macroeconomic implications of size-dependent policies. Review of Economic Dynamics 11 (4), 721–744.

Hall, B., 1987. The relationship between firm size and firm growth in the US manufacturing sector. The Journal of industrial economics 35 (4), 583–606.

Hall, R., Jones, C., 1999. Why Do Some Countries Produce So Much More Output Per Worker Than Others?*. Quarterly Journal of Economics 114 (1), 83–116.

Hopenhayn, H., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica 60 (5), 1127–1150.

Hopenhayn, H., Rogerson, R., 1993. Job turnover and policy evaluation: A general equilibrium analysis. The Journal of Political Economy 101 (5), 915–938.

IMF, 1998. Uganda: Selected Issues and Statistical Appendix. IMF Country Staff Report No. 98/61.

Jovanovic, B., 1982. Selection and the Evolution of Industry. econometrica 50 (3), 649–670.

Keen, M., Mintz, J., 2004. The optimal threshold for a value-added tax. Journal of Public Economics 88 (3-4), 559–576.

Klenow, P., Rodriguez-Clare, A., 1997. The neoclassical revival in growth economics: Has it gone too far? NBER macroeconomics annual 12, 73–103.

Lee, D., Lemieux, T., 2009. Regression discontinuity designs in economics. NBER working paper.

Liedholm, C., Mead, D., 1987. Small scale industries in developing countries: empirical evidence and policy implications. International Development Papers.

Little, I., Mazumdar, D., Page, J., 1987. Small manufacturing enterprises: A comparative analysis of India and other economies. A World Bank Research Publication.

Lucas, R., 1978. On the size distribution of business firms. The Bell Journal of Economics 9 (2), 508–523.

Lucas Jr, R., Prescott, E., 1971. Investment under uncertainty. Econometrica 39 (5), 659–681.

Mata, J., Cabral, B., 2003. On the Evolution of the Firm Size Distribution: Facts and Theory. American economic review 93 (4), 1075.

McCrary, J., 2008. Manipulation of the running variable in the regression discontinuity design: A density test. Journal of Econometrics 142 (2), 698–714.

Melitz, M., 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. Econometrica 71 (6), 1695–1725.

North, D., et al., 1981. Structure and change in economic history. Norton New York.

Prescott, E., 1998. Needed: A theory of total factor productivity. International Economic Review 39 (3), 525–552.

Rauch, J., 1991. Modelling the informal sector formally. Journal of Development Economics 35 (1), 33–47.

Restuccia, D., Rogerson, R., 2008. Policy distortions and aggregate productivity with heterogeneous establishments. Review of Economic Dynamics 11 (4), 707–720.

Roberts, M., Tybout, J., 1997. Producer turnover and productivity growth in developing countries. The World Bank Research Observer 12 (1), 1.

Rossi-Hansberg, E., Wright, M., 2007. Establishment size dynamics in the aggregate economy. American Economic Review 97 (5), 1639–1666.

Sleuwaegen, L., Goedhuys, M., 2002. Growth of firms in developing countries, evidence from Cote d'Ivoire. Journal of Development Economics 68 (1), 117–135.

Smith Jr, A., 1993. Estimating nonlinear time-series models using simulated vector autoregressions. Journal of Applied Econometrics 8, 63–84.

Soderbom, M., Teal, F., 2004. Size and efficiency in African manufacturing firms: evidence from firm-level panel data. Journal of Development Economics 73 (1), 369–394.

Steel, W., 1993. Small Enterprises in Indonesia: Role. Growth, and Strategic Issues, DSP Working.

Steel, W., Webster, L., 1992. How small enterprises in Ghana have responded to adjustment. The World Bank Economic Review 6 (3), 423.

Tybout, J., 2000. Manufacturing firms in developing countries: How well do they do, and why? Journal of Economic literature 38 (1), 11–44.

Veracierto, M., 2001. Employment flows, capital mobility, and policy analysis. International Economic Review 42 (3), 571–595.