Determinants of the Allocation of Resources and their Efficiency Implications in Less Developed Countries

by

Jonathan Goyette

B.Eng., McGill University, 2000
M.Sc., HEC Montreal, 2004

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE STUDIES
(Economics)

The University of British Columbia
(Vancouver)

November 2010

© Jonathan Goyette, 2010
Abstract

Less developed countries typically exhibit lower output per worker and too few medium firms compared to developed economies. The purpose of this thesis is to isolate the distortions driving this misallocation of resources and examine their efficiency implications. Using firm-level data, we show that the probability of being audited for taxes increases significantly around a size threshold of 30 employees in Uganda. This results in a break in the density of firm size around this size threshold and in significantly higher [lower] capital-labour ratios [growth rate] for firms below the size threshold. We argue that entrepreneurs on the verge of having a medium firm hide below the size threshold to avoid a rise in expected regulation costs. They can then substitute capital for labor to scale-up production and wait for a productivity shock that will offset the cost of growing. Other explanations such as a differential in technology or in factor prices due to inefficient credit markets do not reconcile the patterns observed in the data. Based on these empirical evidence we extend Hopenhayn [1992a]'s model with heterogeneous firms to include the tax and credit environment observed in Uganda. Under some simplifying assumptions we solve the model analytically and derive comparative statics for the parameters of interest. We finally revert to the full-blown version of the model and estimate it by Indirect Inference. The calibration strategy consists in choosing a set of structural parameters such that the model generates endogenous patterns in capital-labor ratios, growth of output and firm size density similar to the data. We show that the model does a reasonable job at explaining the data along several dimensions using a set of over-identifying restrictions. We conduct a positive analysis by comparing the effect of the uneven auditing scheme and credit constraints. Finally, we suggest policy interventions to improve efficiency in Uganda.
# Table of Contents

Abstract ...................................................... ii

Table of Contents ........................................ iii

List of Tables ........................................ vi

List of Figures ........................................ viii

Glossary ................................................... ix

Acknowledgments .......................................... x

1 Introduction .............................................. 1

2 Related Work ............................................ 8
   2.1 Cross-Country Differences in Economic Performance ........ 8
      2.1.1 Aggregate Estimates ............................. 9
      2.1.2 Reallocation of Resources across Productive Units .... 10
   2.2 The Missing Middle .................................. 14
      2.2.1 Fiscal Environment ............................. 14
      2.2.2 Credit Market Imperfections .................... 16
      2.2.3 Other Institutional Factors ...................... 18
   2.3 Corruption .......................................... 19
      2.3.1 Wage/Enforcement Structure .................... 19
      2.3.2 Efficiency Implications of Corruption .......... 20
   2.4 Modelling Entrepreneurs’ Behavior ................... 20
      2.4.1 Models based on Lucas and Prescott (1971) ........ 20
5 Numerical Analysis .............................................. 116
  5.1 Numerical Methodology ..................................... 116
    5.1.1 Indirect Inference .................................... 116
    5.1.2 Simulated Annealing .................................. 117
    5.1.3 Value Function Iteration ............................... 118
  5.2 Parametrization ............................................ 120
  5.3 Model Validation ........................................... 126
    5.3.1 A Test of the Model ................................. 126
    5.3.2 Non-Targeted Statistics ............................. 127
  5.4 Positive Analysis ......................................... 133
    5.5 Normative Analysis: Policy Experiments .............. 138
      5.5.1 Smooth Tax Schedule ............................... 138
      5.5.2 Honest Bureaucracy ............................... 140

6 Conclusion .................................................... 143

Bibliography ..................................................... 150

A ................................................................. 159
  A.1 Empirics .................................................. 159
  A.2 Theory ................................................... 161
    A.2.1 Effect of $l^c$ on ic .............................. 161
    A.2.2 Effect of $r$ on ic ............................... 162
  A.3 Numerical Results ...................................... 164
List of Tables

3.1 Summary Statistics .................................................. 29
3.2 Summary Statistics by Size Bins .................................... 30
3.3 Audit Intensity, Number of Employees and Capital: OLS Estimates 42
3.4 Audit Intensity, Nb. of Emp. and Capital: OLS Estimates by Size Groups .................................................. 43
3.5 Differences across Groups .............................................. 46
3.6 Audit Intensity, Nb. of Emp., Sales and Taxes: OLS Estimates 47
3.7 Audit Coverage in the U.S. in 2008 (%) ............................. 49
3.8 Audit Intensity in Uganda in 1997 (%) ............................... 49
3.9 Comparison of Size-Thresholds for the Density of Firm Size .... 51
3.10 Comparison of Bandwidth for the Density of Firm Size .......... 52
3.11 Capital-Labor Ratio and Number of Employees: OLS Results 54
3.12 Growth of Output and Number of Employees: OLS Estimates 56
3.13 Growth of Output and Number of Employees: OLS Estimates by Size Groups ................................................. 57
3.14 Bribes and Number of Employees: OLS Estimates ............... 59
3.15 Taxes and Number of Employees: OLS Estimates ............... 60
3.16 Correlation Coefficients between Sources of Financing and Firm Size .................................................. 62
3.17 Differences across Small Firms ...................................... 65
3.18 Capital and Labor Shares .............................................. 69

4.1 Comparative Statics .................................................... 98

5.1 Parameters ............................................................ 120
5.2 Bargaining Power and Enforcement Costs .......................... 122
List of Figures

3.1 Size Distribution of Firms in Uganda in 1997 .................. 33
3.2 Kernel Density Estimates for Full Sample (left) and Firms with 100
   Employees or Less (right) .................................. 33
3.3 Share of Medium Firms v.s. GDP per Capita .................. 34
3.4 Share of Small and Large Firms v.s. GDP per Capita .......... 34
3.5 Skewness v.s. GDP per Capita ............................... 35
3.6 Audit Intensity v.s. Firm Size in Uganda in 1997 .......... 40
3.7 Step Function: Audit Intensity v.s. Number of Employees ... 41
3.8 Audit Intensity v.s. Capital per Size Groups ................. 45
3.9 Density of Firm Size ....................................... 50
3.10 Capital-Labor Ratios v.s. Firm Size ......................... 53
3.11 Bribes and Taxes per Employee v.s. Firm Size .............. 58

5.1 Size Distribution of Firms: Benchmark v.s. Data ........... 129
5.2 Size Distribution of Firms (Logs): Benchmark v.s. Data .... 129
5.3 Capital-Labor Ratios v.s. Firm Size ........................ 131
5.4 Capital-Labor Ratios v.s. Firm Size (Case 2) ............... 135
5.5 Capital-Labor Ratios v.s. Firm Size (Case 3) ............... 136
5.6 Productivity versus Interest Rate ........................... 138

A.1 Audit Intensity v.s. Firm Size in Uganda in 1997 ........... 159
A.2 Density of Firm Size with a Bandwidth of 5 Employees ...... 160
A.3 Capital-Labor Ratio v.s. Firm Size in Uganda in 1997 ...... 160
Glossary

**ADRES**  Association for the Development of Research in Economics

**BETA**  Bureau d’economie theorique et appliquee

**CSAE**  Centre for the Study of African Economies

**DGP**  Data Generating Process

**GDP**  Gross Domestic Product

**IRS**  Internal Revenue Service

**LDCS**  Less Developed Countries

**NEUDC**  North Eastern Universities Developement Consortium

**RES**  Royal Economic Society

**SSHRC**  Social Sciences and Humanities Research Council

**TFP**  Total Factor Productivity

**UPSES**  Uganda Private Sector Enterprise Survey

**WBES**  World Bank Enterprise Surveys
Acknowledgments

This thesis would never have seen the day without the love and the unconditional support of my parents and close friends. I give them thanks and praises for making my life balanced and extraordinary throughout these years of hard work. My mother and father, who are an inspiration to live a good and meaningful life. My brother, who’s hardwork in arts has inspired me to do the same in my endeavors. Estelle, who’s big heart has been a source of support, comfort and strength during difficult times and who’s warm laugh has been the catalyst of so many other happy times. Vincent, for being a true friend and for his thirst of new experiences. Sue, for being so attentive and supportive and exemplifying with her own life style what it truly means to be a balanced, happy and positive person.

Accomplishing this project would not have been possible without the help and support of my advisor Giovanni Gallipoli who was readily available and has generously offered is time to answer all of my questions and to discuss more philosophical concerns inherent to the doctoral process.

A special thank to Paul Beaudry and Patrick Francois whom, as the members of my committee, have provided me with enlightening advices and encouragements throughout this project. I also wish to thank Mick Devereux, Viktoria Hnatkovska, Vadim Marmer, Amartya Lahiri and Henry Siu who provided insightful comments at some point or another during the completion of this endeavor. The comments of Dirk Krueger, Rodolfo Manelli and Makoto Nakajima have also been helpful.

I thank my companions at UBC, Takuro Miyamoto and Cristian Troncoso-Valverde for their support during the course work of the first 2 years and afterward, Marcos Agurto and German Pupato who’s advices and encouragements allowed keeping my eyes on the big picture during the job market.
I thank participants at the Small Open Economy in a Globalized World Conference held in June 2008 at Wilfrid Laurier University, the 2008 Canadian Economic Association conference held at the University of British Columbia in June 2008, the 2009 conference of the Centre for the Study of African Economies (CSAE) at the University of Oxford, the 2009 North Eastern Universities Development Consortium (NEUDC) conference held in Tufts University, Boston, the 2009 Workshop on Dynamic Macroeconomics held in Soutomaíor, the Royal Economic Society (RES) meetings held in City University London in January 2010, the Bureau d’économie théorique et appliquée (BETA) seminar held at the University of Strasbourg in January 2010 and the Association for the Development of Research in Economics (ADRES) meetings held in University Jean Moulin Lyon 3 in January 2010 for helpful discussions and comments.

I acknowledge the financial support of the Social Sciences and Humanities Research Council (SSHRC) for the Joseph-Armand Bombardier CGS Doctoral Scholarships, my advisor Giovanni Gallipoli for my presenting at NEUDC conference and the Workshop on Dynamic Macroeconomics in 2009, Professor Brian Copeland for my presenting at the CSAE conference in 2009, my advisor and the AMS Financial Commission for my presenting at the Royal Economic Society Meetings in London and the ADRES Meetings in Lyon in January 2010.
Chapter 1

Introduction

Cross-country differences in levels and growth rates of output are large and persistent (Hall and Jones, 1999; Klenow and Rodriguez-Clare, 1997). Understanding why this is the case and what could be done about it is an important challenge for less developed countries (LDCS), especially in the light of the United Nations’ millenium development goals which are coming to terms in 2015. Differences in institutions are often suggested as a reason for the variation in productivity across countries (North et al., 1981; Acemoglu et al., 2001). According to Prescott [1998] and Easterly and Levine [2001], much of this variation is attributable to heterogeneity in total factor productivity (TFP). Recent efforts in theoretical modelling and empirical research have been directed toward both a broader and deeper theory of TFP. Increasing attention has thus been devoted to the effect of institutional distortions on the reallocation process of resources across productive units and efficiency (Restuccia and Rogerson, 2008; Alfaro et al., 2007). However, little is known about mechanisms mapping specific distortions into the allocation of resources and the ensuing consequences in terms of efficiency. Restuccia and Rogerson [2008] emphasize the necessity to restrict the magnitude and types of distortions studied, to obtain better measures of specific distortions and to evaluate their aggregate consequences. This is exactly the purpose of this thesis. We use micro-level data on Ugandan firms to identify two distortions which are typical of LDCS. These distortions are: 1) the burden from regulation falls disproportionately on firms of medium size in LDCS and 2) credit markets are usually inefficient or absent in LDCS. We
then use this information to rationalize, with a theoretical model, the productive choices of entrepreneurs that we observe in the data and to investigate, in a numerical exercise, the extent to which these distortions result in a mis-allocation of resources and loss in measured productivity.

The basic idea for this thesis derives from anecdotal evidence collected during field work in Cameroon. In the summer of 2004, the author of this thesis interviewed 80 entrepreneurs in Douala and Yaounde with respect to the constraints they were facing in their business environment. When it came to discuss issues about informality and corruption, many entrepreneurs readily admitted staying small to avoid dealing with tax officials and other bureaucrats. A review of the literature corroborates this observation as it is often argued that tax administrations in LDCS target larger firms more intensively because tax revenues are expected to be greater than enforcement costs (de Soto 1989, Gauthier and Gersovitz 1997, Gauthier and Reinikka 2006). One consequence of this predatory behavior is that entrepreneurs have incentives to mis-allocate resources to minimize efficiency losses. This gives rise to size distributions of firms that differ from those observed in the developed world. We observe too few medium firms in LDCS as compared to developed economies, a phenomenon that has been labeled the missing middle in the development literature. Now, what is the cause of this missing middle and what does it entail in terms of efficiency are the main questions of this thesis.

Chapter 2 reviews the literature on: 1) cross-country differences in productivity, 2) the missing middle in LDCS and its causes, 3) corruption and tax evasion and 4) recent theoretical and empirical contributions on firms’ dynamics.

The objective of Chapter 3 is to identify the driving force behind the allocation of resources in LDCS. We show that the missing middle is a recurrent fact across many LDCS. We then examine alternative explanations for the occurrence of the missing middle in Uganda using firm-level data on capital, labor, production, factor prices, taxation, bribes and the probability of being audited for taxes.

More particularly, we show that the probability of being audited by tax officials in Uganda rises sharply with firm size, around a size threshold corresponding to about 30 employees. Along with the rise in audit probability, we find evidence that regulation costs such as official and effective tax payments as well as bribes increase significantly for medium firms (between 31 and 75 employees). This trans-
lates in a significant break in the density of firm size around 30 employees. The data also suggests that entrepreneurs, who would otherwise grow their firm beyond that threshold, choose ex-ante suboptimal capital-labor allocations and that smaller firms exhibit lower growth rates of output than their larger counterparts. We argue that only relatively large productivity shocks outweigh the costs of increasing employment beyond 30 employees: in response to a positive productivity shock a large proportion of entrepreneurs choose to scale up production by substituting capital for labor so as to avoid an increase in regulation costs. This results in reduced output growth at lower size levels, a clustering of firms around the 30 employees threshold and a significant discontinuity in the level of capital-labor ratios of firms above or below the threshold. The data allows examining two other potential explanations to reconcile the aforementioned patterns. First, we show that there is a price differential across factors and across firms. Small firms are credit constrained but pay lower wages while larger firms have access to credit but pay higher wages. Second, we find evidence that technology is more labor-intensive for smaller firms. However, these evidence go against the pattern in capital-labor ratios that we observe for small and medium firms.

Chapter 4 builds on Hopenhayn [1992a]’s growth model of heterogeneous firms with entry and exit. We adapt the model to include the salient features of the Ugandan environment described in Chapter 3. The objectives of Chapter 4 are to: 1) develop a framework that will be useful for both theoretical and numerical analysis of the causes and implications of the missing middle and 2) find the effects of a change in tax rates, the audit probability, the size-threshold, the expected fine faced by corrupt bureaucrats, output price or input prices on the hiring of capital and labor, bribes and the size distribution of firms.

In the model, firms produce a homogeneous good using capital and labor. In each period there is a mass of entrepreneurs deciding whether to enter or exit the market. Firms are heterogeneous along two dimensions: transient productivity shocks and current capital stock. 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. Capital stock becomes 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.\footnote{We verify the robustness of our findings by removing this assumption in a counterfactual experiment in Chapter 5.} Consistent with empirical evidence we assume that entrepreneurs evade part of their tax liabilities. Tax officials are corruptible and, with some probability, they accept bribes when auditing entrepreneurs. We introduce a frequency of tax audits that is increasing with size.

In this thesis, we are only concerned with stationary equilibria.\footnote{As opposed to transitional dynamics.} As is standard in the literature, heterogeneity across individual entrepreneurs is best described by a probability measure which summarizes the state of the industry at a given point in time. This allows describing the evolution of the distribution of firms across time and defining a stationary equilibrium. We then derived the optimal decisions of entrepreneurs with respect to labor and capital. The optimal demand for labor is standard apart from the fact that it varies with size due to the auditing scheme. We also derive the Euler equation that governs each entrepreneur’s choice of capital investment for the next period. A particular implication of the Euler equation is that entrepreneurs will invest at a decreasing rate due to the auditing scheme.

Given the interaction between the uneven auditing scheme and missing credit markets it is of interest to examine each distortion in turn. We thus describe an economy without distortions. In this case, there are no taxes and credit markets are perfect. There are 3 implications stemming from this First-Best case. First, the first-best capital-labor ratio is constant across firms. Second, the size-dependent auditing scheme acts as a progressive tax: it decreases the amount of labor employed at any given firm but relatively more in larger firms. Third, when firms must use their own cash-flow to finance their investment, there is a non-degenerate distribution of resources across firms within a productivity class (a given level of productivity shock).

In the last section of Chapter 4, we introduce some simplifying assumptions to solve the model analytically and derive comparative statics for the parameters of interest. We show that an increase in the output [input] price increases [decreases] allocations of capital and labor in all sectors as well as the level of bribes paid, and
it alleviates [exacerbates] the clustering of firms below the size-threshold where the audit probability increases. A change in firms’ bargaining power over the surplus from tax evasion affects proportionally the allocations of capital and labor at firms dealing with tax officials. However, such a change has an ambiguous effect on the value of bribes paid as it depends on the magnitude of the elasticities of the inputs with respect to the bargaining power. We document similar patterns for the effect of a change in the tax rate. A change in the size-threshold where the audit probability increases does not affect how firms choose their inputs except for those firms right at the size-threshold. Since the bribe is a function of capital and labor it is not affected by such a change. An increase in the audit probability reduces the allocations of capital and labor as well as the size of the bribes paid at firms involved with tax officials and it increases the clustering of firms at the size-threshold where the audit probability increases. Finally, the expected fine faced by tax officials who get caught receiving a bribe does not affect the allocations of capital and labor. However, an increase in the expected fine increases the level of bribes paid which in turn leads firms to cluster at the size-threshold where the audit probability increases.

In Chapter 5 we revert to the full-blown version of the model and estimate it. The objectives in this chapter are to: 1) reproduce the 3 empirical facts observed in Chapter 3 i.e., a break in the density of firm size at 30 employees, higher capital-labor ratios and lower growth rates of output for small firms; 2) validate the numerical model with a set of over-identifying restrictions; 3) compare in terms of efficiency the 2 distortions included in the benchmark economy, i.e., the fact that larger firms are targeted more intensively by the tax administration and the fact that entrepreneurs must finance their investment with their own cash-flow; 4) suggest policy interventions in order to improve efficiency in Uganda.

The parameters of 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 needs not be the exact data generating process (DGP). 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 rejects the null that the moments from the model are equal to the moments from the data for any probability level equal to and above 2%. Low et al. [2010] develop a model of the effects of disability risk and insurance on consumption and labor decisions in the U.S. The test of their model also marginally fails to reject the null for a probability level of 2.25%, using a similar formal test of overidentifying restrictions with 23 degrees of freedom. Is this solid grounds to reject their work? Probably not given that their model replicates a wide and interesting array of features that are observed in the U.S. data. The model presented in Chapter 5 does a reasonable job at reproducing some features of the data that were not explicitly targeted such as patterns by firm size of capital-labor ratios, firms’ growth, age and shares of sales. We also observe in 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. These choices generate a gap in the size distribution of firms and lead to the missing middle in the size distribution of firms.

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 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 regulation distortions (taxes and uneven audit schedule) would close the gap in output per worker between Uganda and a First-Best economy 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 undistorted economy. The missing middle generated by the institutional distortion does not seem to have strong implications for aggregate efficiency. Firms adjust their input mix optimally to avoid productivity losses in the presence of institutional distortions and the resulting firm size distribution is a by-product of their optimal responses. In comparison, it seems that inefficiencies associated to credit markets generate a much greater efficiency loss and attention should somewhat be redirected toward financial frictions in trying to increase TFP.

Having established the ability of the model to rationalize some relevant fea-
tured 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 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 as 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, a wage that makes auditors indifferent between being honest or corrupt 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 the undistorted economy. Flattening the distribution of audit probabilities would only reduce this gap by 0.5% while the no-corruption wage would reduce this gap by 27%. In order to study the pure reallocation effect of the policies we hold aggregate labor supply constant throughout all our simulations. Therefore part of the productivity variation from each policy change is captured by the variation in the adjusted workers’ wage. In each case, the aggregate output gain must be interpreted as a lower bound on the potential effect of a policy.
Chapter 2

Related Work

The allocation of resources in LDCS presents singular differences from developed economies. Indeed, firm size distributions in LDCS typically exhibit too few medium firms as compared to developed economies and this phenomenon has been labelled the missing middle in the development literature. In light of the recent efforts to examine the effect of the reallocation of resources on efficiency this could explain part of the observed discrepancy in output per worker across countries. However, little is known about mechanisms mapping institutional and market distortions into the allocation of resources and the ensuing consequences in terms of efficiency. Given that the goal of this thesis is to identify empirically the cause and the efficiency implications of the missing middle in LDCS we describe in this chapter the need for such an endeavor. The rest of this chapter is organized as follows. In the first section, we review the literature on cross-country differences in output per worker. In the second section, we review the literature on the missing middle. In the third section we review the literature on corruption. The last section of this chapter presents the literature used to model the benchmark economy of Chapter 4 and Chapter 5.

2.1 Cross-Country Differences in Economic Performance

Lucas et al. [1988] observe that average annual growth rates over several decades vary substantially across countries. According to Hall and Jones [1997] small dif-
ferences in such growth rates can add up over time to enormous differences in levels of income. Indeed, these authors argue that the ratio of gross domestic product (GDP) per worker in the fifth-richest country to that in the fifth-poorest country in 1988 was 29 while this ratio was 26 in 1960. Understanding what drives these differences is thus of importance. In what follows, we look in turn to explanations related to aggregate factor accumulation and the reallocation of resources across productive units.

2.1.1 Aggregate Estimates

Romer [1990] and Grossman and Helpman [1990] suggest that endogenizing a country’s technology may explain these differences in productivity across countries. Hall and Jones [1997] argue that cross-country differences in economic levels are driven by institutional and government policies. Societies which encourage the redistribution of output instead of its production produce much less output per worker. This echoes the conclusions of Murphy et al. [1991] who find that countries with a higher proportion of engineering majors grow faster than countries with a higher proportion of law degrees. In the same vein, Mauro [1995] finds that corruption lowers investment which, in turn, lowers growth.

A seminal contribution to the study of economic growth is Solow [1956] who suggests to use a standard neoclassical production function with decreasing returns to capital. He showed that, because saving and population growth rates vary across countries, different countries reach different steady states. Barro [1991] corroborates this result using a cross-section of 98 countries and provides empirical evidence that the growth rate of real per capita GDP is positively related to initial human capital and negatively related to the initial level of real per capita GDP. Mankiw et al. [1992] show that the Solow model augmented to include human capital explains 78% of the cross-country variance of output per worker in 1985. Klenow and Rodriguez-Clare [1997] also offer evidence on the relative importance of productivity and capital (physical and human) in explaining cross-country differences in levels and growth rates of output. They focus on the measurement of human capital, using primary school attainment rather than secondary school attainment as in Mankiw et al. [1992] and argue that their measure of human capital
fares better at explaining international productivity differences. Prescott [1998] argues that differences in physical and human capital cannot account for the large international income differences nor does growth in the stock of technical knowledge. He claims that a theory of cross-country differences in TFP is needed which does not rely on technical change. Easterly and Levine [2001] also argue in favor of such an endeavor. A recent attempt in that direction is Acemoglu and Zilibotti [2001] who argue and provide empirical evidence that the mismatch between the supply of skills in LDCS and the requirement of technologies imported from OECD countries lead to sizeable TFP differences. The next section presents other attempts at developing a broader theory of TFP.

2.1.2 Reallocation of Resources across Productive Units

Until recently the literature on cross-country differences in economic performance was concerned with examining the effects of differences in relative prices on aggregate accumulation. However, this abstracts from heterogeneity across production units. As such, recent advances on firms’ dynamics modelling show that resources’ reallocation across productive units also affects cross-country differences in output per worker.

Alfaro et al. [2007] perform a development accounting exercise using data for more than 20 million establishments from 79 developed and developing countries. Their model follows Melitz [2003] in terms of the modelling of heterogeneous production units, Restuccia and Rogerson [2008] in terms of the modelling of policy distortions as idiosyncratic heterogeneous prices faced by individual entrepreneurs and Hsieh and Klenow [2007] in terms of the modelling of constant returns to scale technologies and monopolistic competition. They calibrate their model so that a tax and subsidy schedule for each country matches this country’s plant-size distribution, using the United States as an undistorted benchmark economy. The authors calculate for each country how much aggregate output is wasted due to the misallocation of resources. Using these estimates from the calibrated model, they show that this misallocation of resources is a powerful explanatory factor of cross-country differences in income. The model explains 58% of the log variance of cross-country income dispersion while the usual model which considers physical
and human capital only has a success rate of 42%.\(^1\)

Bartelsman et al. [2006] execute a similar type of accounting exercise as in Alfaro et al. [2007] but use Olley and Pakes [1996]'s empirical decomposition of the level of industry productivity. They find that increasing distortions to allocation of resources tends to decrease the Olley and Pakes cross term. However, they argue that distortions can negatively affect alternative channels of resource allocation that may not change the Olley and Pakes cross term. For example, when potential entrants must pay an entry cost before learning their productivity, distortions will affect the proportion of entrants, the mix of firms and the scale of activity (average firm size and the capital-labor ratio) without necessarily affecting the Olley and Pakes cross term. The authors note that caution must be used in interpreting the patterns of empirical decompositions of productivity.

Restuccia and Rogerson [2008] use a calibrated version of the growth model developed by Hopenhayn [1992a] and Hopenhayn and Rogerson [1993] to argue that differences in allocation of resources across heterogeneous establishments due to a generic and abstract form of policy distortions may account for significant cross-country differences in output per capita. The production technology exhibits decreasing returns to scale as this allows pinning down the size of each individual production units. The authors consider idiosyncratic policies that alter prices faced by individual producers. These distortions should be broadly interpreted as policies that could generate such an effect from the vantage point of producers such as non-competitive banking systems, product and labor market regulations, corruption, trade restrictions, etc. In the simulations, these distortions affect the reallocation of resources across plants but not aggregate accumulation nor aggregate relative prices. Nevertheless, Restuccia and Rogerson [2008] show that such policies lead to a decrease in output in the range of 30 to 50 percent. In the undistorted economy, all plants with the same productivity are of the same size whereas in the distorted case there is a non-degenerate distribution of plant size within a productivity class. They show that negative effects on output and productivity result whether low productivity or high productivity plants are targeted by subsidies: in both cases the reallocation of resources entail an efficiency loss.

\(^1\)Note however that this explains less than what has been accomplished by Mankiw et al. [1992] who augmented the Solow model with human capital.
The framework of Restuccia and Rogerson [2008] encompasses a few papers that have focused on specific distortions. In a paper closely related to ours, Guner et al. [2008] examine the costs of size-dependent policies distorting production scale. They find that the effects of this class of policies are large: policies that reduce the average size of establishments lead to sizeable reductions in output per establishment as well as a large increase in the number of establishments. Our work departs from this study in two dimensions: i) we focus on an observable distortion and use data from a LDC whereas they use information on regulations in specific OECD countries; ii) they use Lucas [1978] span-of-control framework with no entry and exit. Hsieh and Klenow [2007] provide quantitative evidence on the impact of resource misallocation on efficiency in India and China versus the United States. Their analysis is based on a standard model of monopolistic competition with heterogeneous firms similar to Melitz [2003] without international trade. Distortions in their model generate a differential between the marginal products of capital and labor across firms and this, in turn, lowers efficiency. They find that moving to US efficiency would increase TFP by 30-50% in China and 40-60% in India. They argue that output gains would be roughly twice as large if capital accumulated in response to aggregate TFP gains. These effects are of the same order of magnitude as those described in Restuccia and Rogerson [2008]. Our findings in Chapter 5 corroborate these results: removing all distortions from the Ugandan benchmark economy also entails a doubling of output per worker.

There are a few other papers which focus on a specific distortion. Parente and Prescott [1999] argue that monopoly in determining the use of inefficient technologies is an important cause of cross-country variation in output per worker. Schmitz Jr. [2001] examines a similar channel but where government policies support inefficient public enterprises. Veracierto [2001] and Lagos [2006] follow Hopenhayn and Rogerson [1993] in examining the effect of labor market regulations on efficiency. We also review a large literature that examines the effect of financial frictions on efficiency in Section 2.2.2.

Our work is a natural extension of Restuccia and Rogerson [2008] who emphasize the necessity to restrict the magnitude and types of policy distortions studied, obtain better measures of specific distortions and evaluate their aggregate consequences. This is exactly what we do in this thesis: we identify the most important
cause of the missing middle in Uganda and then evaluate its efficiency implications. However, finding a natural experiment or an identification strategy that would allow observing sizeable effects of a policy change on the reallocation of resources across productive units is a difficult task.\footnote{Banerjee and Duflo [2008] is a successful attempt in that direction.} Instead, we build a structural model based on patterns observed in the Ugandan data. Such a structural approach presents the advantage of allowing for a quantitative analysis of the potential impact of resource misallocation on aggregate productivity. In that sense our work is closely related to Hsieh and Klenow [2007] who argue that distortions driving wedges between the marginal product of labor and capital across firms lower aggregate TFP. A distinct contribution of our study is the fact that we identify such distortions, namely a rise in audit intensity at a certain size-threshold and the absence of capital market for small and medium firms. Moreover, as argued by Acemoglu [2010], economic theory should play a central role in formulating a model so that its estimates can be used for counterfactual and policy analysis. As such, our structural model permits to conduct policy experiments with respect to the distortions of interest.

We also estimate some parameters by Indirect Inference (Gourieroux et al., 1993; Smith Jr, 1993). This simulation based-method presents two advantages. First, it is useful when the coefficients one tries to target during the estimation procedure do not have analytical expressions. This is indeed the case here. We introduce a non-convexity in the auditing scheme and this prevents us from calculating a single theoretical expression for each aggregate moment. For example, we use the average number of employees as one of our targets during the calibration procedure. However, firms face different audit intensities depending on their size and this implies that they also choose labor based on a different theoretical expression (see Equation 4.38). Thus we cannot define this moment with a single theoretical expression for the aggregate. As mentionned in Low et al. [2010], unlike other methods based on simulations (like the Simulated Method of Moments), Indirect Inference only requires the specification of an auxiliary model which needs not be the correct DGP. One then tries to minimize the distance between the parameters of the auxiliary model estimated from the observed data and the parameters of the
auxiliary model estimated from the simulated data.

Second, if the number of coefficients targeted is smaller than the number of parameters to be estimated a set of overidentifying restrictions can be used to formally test the fit of the model to the data. We believe to be among the first studies with Low et al. [2010] using such a formal test of the goodness of fit for a structural model.

The next section reviews the literature on the missing middle and its potential causes. This serves as reference point in Chapter 3 to identify the driving force behind the missing middle in Uganda.

2.2 The Missing Middle

Firms size distributions tend to be highly skewed and this observation has been robust over time [Axtell, 2001]. Based on a large body of evidence it seems that LDCS present firm size distributions that are either more skewed to the right or to the left than developed economies. Many papers document the fact that firms size distribution in LDCS exhibit too few medium firms as compared to developed economies, a phenomenon which has come to be denominated as the missing middle. Liedholm and Mead [1987] report evidence for many developing countries in the seventies and eighties, including Colombia, Ghana, Honduras, India, Indonesia, Jamaica, Kenya, Korea, Philippines, Nigeria, Sierra Leone, Tanzania and Thailand. Little et al. [1987] observe this in South Korea in the seventies. Steel [1993] examines the case of Indonesia and Steel and Webster [1992] that of Ghana in the eighties. Sleuwaegen and Goedhuys [2002] provide evidence for Ivory Coast. We also provide similar evidence for Uganda in Chapter 3 as well as for 42 other LDCS using data from the World Bank. The next sections review some of the potential causes behind the missing middle.

2.2.1 Fiscal Environment

One particular channel we wish to examine in this thesis is the misallocation of resources due to a lack of accountability in the fiscal environment of a country. de Soto [1989] argues that many Peruvian entrepreneurs stay small to avoid regu-
lation. Gauthier and Gersovitz [1997] show that small Cameroonian firms refrain from growing to avoid taxes. Gauthier and Reinikka [2006] argue that "smaller firms may find it easier to slip out of the tax collector’s net as enforcement costs easily exceed tax revenues collected" in Uganda. We find supporting evidence of this claim in a report from the Article IV Consultation for Uganda by the IMF where 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."

Johnson et al. [2000] find that managers in Russia and Ukraine have more incentives to hide their activities as compared to managers from Poland, Slovakia and Romania because they face higher tax rates, worse bureaucratic corruption, a greater incidence of mafia protection and have less faith in their court system. Ingram et al. [2007] show that formality is negatively correlated with the rate of taxation and corruption using firm-level data from the World Bank from six African countries which include Uganda. Dabla-Norris et al. [2008] show with data from the World Business Environment Survey (WBES) that informality is negatively related to firm size and that, although heavier regulations may conduct to higher informality, this needs not be so in countries with a strong rule of law.3

Our work in this thesis is closely related to the following two papers. Keen and Mintz [2004] examine the choice and efficiency implications of an optimal threshold for a value-added tax. Their model provides an explanation for the gap in the size distribution of firms observed in LDCS. However, capital is absent from their model and they do not examine the implications of tax evasion and corruption. As we show in Chapter 5 bribes buy entrepreneurs a tax rebate and this diminishes the impact of a stark threshold on the magnitude of the missing middle, thus attenuating Keen and Mintz [2004]'s results. Fortin et al. [1997] develop a model of heterogeneous firms in which a formal sector arises endogenously due to taxation and wage controls. They offer an explanation for the missing middle based on the fact that the marginal cost of tax and regulation evasion increases with size. Again, capital is absent from their model and this does not allow entrepreneurs to substitute capital for labor when facing a distortion on labor and, in turn, to minimize efficiency losses.

3 As the modelling of corruption is an important part of our work we devote a full section to this subject below.
2.2.2 Credit Market Imperfections

Recent works highlight how credit market imperfections and heterogeneity in financing needs generate distortions in the allocation of resources in developing countries and result in productivity losses. Banerjee and Duflo [2008] use variation in access to a lending program to evaluate the extent of credit constraints faced by Indian firms. They argue that constrained firms use credit to expand production while unconstrained firms use credit as a substitute for other borrowing. Using a difference-in-difference-in-difference approach they show that many firms are credit constrained and their marginal rate of return to capital is very high. Banerjee and Duflo [2005] explain that even if there are numerous firms of smaller size being credit constrained in an economy the share of capital in these firms is too small to explain cross-country differences in output per worker. We find a similar result as most of the production is carried out by the largest firms in our model.

However, Banerjee and Duflo [2005] argue that if firms of medium size are credit constrained this could go as far as explaining the entire productivity gap between India and the U.S. One reason for their result is that the production function of firms in their model uses only one input: capital. They argue that credit constraints might explain the misallocation of capital across firms. This observation is correct, however, firms in the model of Chapter 4 and Chapter 5 can use two inputs and substitute them (albeit imperfectly). Therefore if credit constraints adversely affect the intertemporal decision of entrepreneurs to invest capital for the next period while labor can be hired on the spot then it might the case that entrepreneurs substitute labor for capital to minimize their efficiency losses. In such a case, we would probably observe an impact of credit constraints on medium firms of a smaller magnitude than the one advocated by Banerjee and Duflo [2005]. In Chapter 3, we find that in response to a local distortion such as a steep rise in the probability of being audited for taxes at a threshold of a certain number of employees small firms substitute capital for labor to minimize their efficiency losses. Using the same logic to the situation of constrained medium firms, it could well be the case that medium firms substitute labor for capital to attenuate the effect of credit constraints or inefficient capital markets.\(^4\)

\(^4\)This is something that prompts further research and is on our research agenda.
Buera and Shin [2008] quantitatively evaluate the impact of resource misallocation and financial frictions on the dynamics of economic development. They show that, with misallocation of initial resources, financial frictions delay the efficient reallocation of these resources and reduces production and investment. They find that an economy with financial frictions and initial misallocation of resources would take four times as long to cover half the distance to the steady state as the neoclassical benchmark. According to the authors, financial frictions are responsible for about a quarter to a half of the total misallocation in LDCS. Arellano et al. [2008] also examine how financial development influences firms’ financing and growth. The authors develop and parametrize a quantitative model to show that high credit costs limit debt disproportionately for small firms thus making their scale inefficient. They argue that financial development can rationalize the empirical facts that small firms have higher growth rates than larger firms especially in LDCS and that the difference in debt financing of small and large firms is smaller in less financially developed economies. Buera et al. [2009] provide evidence that poor countries are unproductive in tradable and investment goods sectors. They develop a quantitative framework to explain these cross-country patterns. The key to their model is that sectors with larger scales of operations such as tradable and investment goods sectors have more financing needs and are disproportionately affected by financial frictions. They test their model with a set of over-identifying restrictions. As in Restuccia and Rogerson [2008] they find that variation in financial development can explain a two-fold difference in output per worker across countries.\footnote{The removal of all distortions in our distorted benchmark economy also generates an increase in output per worker of similar magnitude.}

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. Cooley and Quadrini [2001] develop a theory based on financial market frictions to account simultaneously for the dependence of firm dynamics on size and age. By introducing financial market frictions along persistent shocks, the authors are able to generate the regularities found in the data, controlling for size in the case of age and vice-versa. Albuquerque and Hopenhayn [2004]
develop a model to characterize constrained efficient lending contracts and study their implications for firm survival, growth, equity shares and dividend distribution policies. Clementi and Hopenhayn [2006] model a multiperiod borrowing/lending relationship with asymmetric information and show that, in agreement with the empirical evidence, mean and variance of growth decrease while firm survival increases when age and size increase.

2.2.3 Other Institutional Factors

There are several other institutional distortions apart those stemming from the fiscal environment that might also affect size distributions of firms. Brunetti et al. [1998] in a worldwide survey of managers attitude find that firms in LDCS typically face institutional obstacles due to price controls, foreign currency and trade regulations, policy instability and uncertainty about costs of regulations. Even when policies do not favor certain firms, there are other features of LDCS which might lead to the misallocation of resources. Investment incentives might be set according to a certain scale of production. Special tax breaks may favor specific firms [Gauthier and Gersovitz, 1997]. Firms with political connections might have an easier time gaining control over the market through lobbying games [Bliss and Tella, 1997]. Poor road networks make it difficult to reach small pockets of consumers spread out in the countryside. The lack of infrastructure also increases sunk costs faced by entrepreneurs who must then provide themselves with their own electricity generator, roads, supply of water, etc. Engel effects bias production toward basic subsistence goods instead of more sophisticated manufactured goods. The access to inputs being limited in LDCS, entrepreneurs must either use imperfect substitutes or buy their inputs at an extra-cost from abroad. This in turn might induce some of the larger firms to vertically integrate their production which reduces the demand for intermediate firms. Human capital also plays a role in the shape of the size distribution of firms as the scarcity of high-skilled individuals affects the mix of goods manufactured in LDCS. Also the abundance of low-skilled individuals allow entrepreneurs to economize on capital by using labor-intensive technologies [Tybout, 2000]. Wage rigidities forces some workers into the informal sector due to job rationing. These individuals revert to organisation of last resort to survive
which will typically be small and inefficient. Price volatility and political instability are typical of several LDCS and reduce the incentives to incur the sunk costs from capital investment.

2.3 Corruption

2.3.1 Wage/Enforcement Structure

Our work is also related to the literature on corruption and evasion. We are primarily concerned with the wage/enforcement structure in bureaucracies of LDCS. We use the parameters of the wage/enforcement structure in the comparative analysis of Chapter 4 and the normative analysis of Chapter 5 to examine what could potentially be done to increase efficiency in Uganda.

Gang et al. [1988] ask: "If public workers [in LDCS] suffer discrimination by wage, why is it that demand for such jobs is high?". Lindauer et al. [1988] offer a potential answer. They describe the case of the Ugandan administration in the eighties where "civil servants either had to diminish their ethic standards or perish in uprightness". Svensson [2003] addresses the question indirectly using the Uganda Private Sector Enterprise Survey (UPSES) which contains data on the amounts of bribes paid by firms. The author examines which firms pay bribes and how much they pay given their bargaining power when dealing with public officials. His results suggest that public officials act as price discriminators and that prices of public services are set in order to extract bribes. This answers at least partially the question raised by Gang et al. [1988]: people are aware of the extra revenues that these jobs generate through redtape and bribing.

How should the wage/enforcement structure be set in a bureaucracy? Besley and McLaren [1993] show that, when penalties or supervision efforts are low, even infinite wages are insufficient to deter corruption completely in an administration. Di Tella and Schargrodsky [2003] find supporting evidence along these lines using auditing variation during a crackdown on corruption in Buenos Aires’ hospitals in the late nineties. Basu et al. [1992] argue that raising the level of the fine is less costly to a government than increasing the probability of detection of corrupt indi-

---

6Bardhan [1997], Tanzi [1998] and Aidt [2003] provide extensive reviews on this literature.
viduals. However, they note that the latter is more efficient at reducing corruption as its effect propagates through the hierarchy while the former has a similar effect at each stage of the administration. Hindriks et al. [1999] develop a model of tax evasion to examine the distributional effects of tax collection and corruption. They argue that a trade-off between equity and efficiency exists and suggest that tax officials could be paid commissions on high income reports.

2.3.2 Efficiency Implications of Corruption

Corruption implications for efficiency and growth have been noted in many papers since Mauro [1995]. However, most of these papers use indexes on the perception of corruption. Using micro-level data McArthur and Teal [2003] examine the effect of corruption on the performance of African firms with data from the Africa Competitiveness Report 2000/1. They find that, at the firm level, firms paying bribes see their output per worker reduced by 20% while at a national level, firms in countries with higher corruption are 70% less efficient. Fisman and Svensson [2007] show with the UPSES that an increase of 1% in the bribe rate undermines firms’ growth by 3%, an effect which is three times that of taxation.

2.4 Modelling Entrepreneurs’ Behavior

2.4.1 Models based on Lucas and Prescott (1971)

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 the firm level and follow a non-stationary process. However, his model does not allow for entry and exit in the limit. Hopenhayn [1992a] extends the above models to allow entry and exit in a stationary equilibrium, and a more general process for firms’ idiosyncratic shocks. Hopenhayn and Rogerson [1993], HR hereafter, further extend these results to a general equilibrium framework. They calibrate their model to U.S. data and ex-

2.4.2 Other Models of Firms’ Dynamics

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 in 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 in LDCS.
Firm size distributions in LDCS typically exhibit too few medium firms as compared to developed economies. More particularly, we show that the missing middle in Uganda arises alongside a large number of small firms and a handful of large firms. Indeed, the Ugandan size distribution of firms is more skewed to the right than in more developed economies. The purpose of this chapter is to compare competing explanations for this observation and identify the driving force behind the size distribution of firms in Uganda. We focus on three explanations available in the data: institutional constraints stemming from the tax environment, heterogeneous technology across firm size and factor price differentials.

The empirical evidence suggest that the taxation environment generates a distortion which favors small firms in Uganda. McCrary [2008]’s density estimator allows showing that there is a break in the density of firm size around a threshold of 30 employees. One potential explanation for this break in density is a significant rise in the audit intensity for firms which is also observed around 30 employees. We also note that medium firms pay more bribes and more taxes than their smaller counterparts. As a result, medium firms face lower tax savings per bribe than small firms. Since we cannot infer from the data whether the bulk of these payments are made before or after the audit takes place, we assume that the effect of the regulation costs adds up to the effect of the rise in the audit intensity around the size
threshold. The break in the density of firm size could thus be due to entrepreneurs self-selecting just below the size threshold in order to hide from tax officials and avoid the rise in regulation costs that a larger size entails.

As a consequence of this we observe significantly greater capital-labor ratios for small firms than medium firms. Entrepreneurs seem to substitute capital for labor while hiding and waiting for a productivity shock that is large enough to outweigh the costs of growing beyond the size threshold. Another consequence that goes hand-in-hand with the capital-labor substitution is that smaller firms exhibit lower growth rates than their larger counterparts. This observation is at odd with previous findings for developed economies (Evans 1987; Hall 1987).

The data allows examining two other potential explanations to reconcile the aforementioned patterns. First, we show that there is a price differential across factors and across firm sizes. Small firms are credit constrained but pay lower wages while larger firms have access to credit but pay higher wages. These evidence go against the patterns in capital-labor ratios for small and medium firms. Second, we find evidence that technology is more labor-intensive for smaller firms. This also goes against the patterns in capital-labor ratios observed in the data for small and medium firms.

The relationship between the patterns we try to identify stems from a set of endogenous decisions. The best way to verify the hypothesis brought forward in this chapter is to develop a structural model based on the empirical evidence. This is done in Chapter 4 and Chapter 5 where we develop a model of firms’ growth and show that a rise in audit intensity rationalizes the size distribution of firms and patterns in growth and capital-labor ratios observed in the Ugandan data.

This chapter is structured as follows. The first section presents the data. The second section shows that the missing middle is a stylized fact across many LDCS and why we should find a distortion that favors small firms in Uganda. The third section describes the empirical approach. The last section presents the empirical results.
3.1 Data

3.1.1 Variables

Data for Uganda are taken from the Ugandan Private Sector Enterprise Survey (UPSES) 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 on their activities from 1995 to 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.

Taxation  Taxation data were collected on the main taxes paid by Ugandan businesses, in particular the corporate income tax (CIT), the sales tax/value added tax (VAT), the National Social Security Fund (NSSF) levy, the excise tax and the withholding tax. Information was obtained on special tax reduction and exemption programs available to firms within the Ugandan tax system. One of the main sources of tax exemptions is the 1991 Investment Code, which provides exemptions to large investors. The Minister of Finance also grants tax exemptions on a case-by-case basis from CIT, import duties, and domestic sales taxes.

Bribes  Data on corruption were collected in different parts of the questionnaire. Information on bribe amounts was obtained indirectly as respondents were asked to estimate the typical bribe payment a firm in their line of business would pay each year to deal with public officials, in customs, taxes, licences, regulation, etc. The questionnaire has data on the frequency at which bribes were asked.

---

1 Svensson [2003], Gauthier and Reinikka [2006] and Fisman and Svensson [2007] provide specific details about this dataset.

2 See Gauthier and Reinikka [2006] for the official rates and exemption conditions for each of these taxes. Other taxes include import duties, the presumptive tax on small businesses and the local property tax.

3 Reinikka and Svensson [2006] show that with appropriate survey methods and interview techniques, quantitative micro-level data on corruption can be obtained.

4 The question was: “Many business people have told us that firms are often required to make informal payments to public officials to deal with customs, taxes, licenses, regulations, services, etc.
when dealing with tax officials. We compute a probability of meeting a corrupt bureaucrat using this information. The dataset provides also information on senior management’s time spent each month dealing with officials and on auditing probabilities for the CIT and the VAT.

**Evasion Surplus** In a related paper, Azam et al. [2004] construct a proxy for tax obligations using the Ugandan Tax Code and firms’ information on sales, size and exemptions. This constructed variable represents the amount of taxes officially owed by firms as opposed to what firms actually pay for the main taxes enumerated above. Subtracting taxes paid from tax obligations, we obtain an evasion surplus for each firm. Finally, taking the ratio of the evasion surplus to bribes we get an amount of tax savings per US dollar worth of bribing.

**Labor and Wage Bill** The dataset details the number of employees as being hired on a full-time, contractual or casual basis. This is the measure of firm size we will use throughout this thesis unless specified otherwise. In what follows, small firms have 30 employees or less, medium firms have between 31 and 75 employees, and large firms have 76 employees or more. Tybout [2000], Gauthier and Reinikka [2006], Soderbom and Teal [2004] and many other papers in the development literature use a similar classification of firm size. We are thus making this choice of categorization to be consistent with the rest of the literature. We then verify whether the data presents interesting patterns around that threshold. It does indeed as shown in Section 3.4.2.

Total costs net of interest payments and taxes are recorded in the dataset as well as the percentage of these costs that accounts for wages and benefits. Using this information we compute a wagebill for each firm.

**Capital** There is no direct measure of capital in the UPSES. Calculating the quantity of capital owned by firms seems to be a daunting task for economists when there are no direct measures available. As Pakes and Griliches [1984] put it:

———

Can you estimate what a firm in your line of business and of similar size and characteristics typically pays each year?”  

The choice of answers is qualitative and we assign increasing values between zero and one to each answer. The assignment of values was: never = 0; seldom =0.25; occasionally=0.5; often =0.75; always=1.
In general, there is no obvious unique right way to construct “capital”; it all depends on the purpose for which such a variable is to be constructed.

Attempts at estimating capital usually rely on perpetual inventory methods and specific, thus limiting, frameworks. Baily et al. [1981] use a model of vintage capital to show that the aggregate measure of capital services in an economy is equal to the perfect foresight market value of the capital stock. In such a model, old capital is equivalent to a smaller quantity of new capital. So, in order to find this equivalence, a simple system where rents generated by capital act as weights works just right. This would not be true in a more general model. Similarly, Hall [2001] develops a quantity revelation theorem where the quantity of capital and its shadow price are backed out using a model with adjustment costs. Given we do not have access to firms’ initial level of capital stock in our sample, any Perpetual Inventory Method like the one proposed by Baily et al. [1981] or in Soderbom and Teal [2004] can not be implemented. Also, there is no way we can trace back the market value of the firms in our sample from the Uganda Securities Exchange given the confidential nature of the survey. This rules out Hall [2001]’s approach or other methodologies that rely on Tobin’s q (Hayashi, 1982; Hopenhayn, 1992b; Chung and Pruitt, 1994). This is also true of other methodologies used in finance to back out the value of a firm or of its capital, including the Free Cash Flow Method (Kaplan and Ruback, 1995), Capital Cash Flow method (Ruback, 2002), Adjusted Present Value method (Myers, 1974).6 Finally, Griliches [1963] reviews different concepts and measurements of capital stock.

In what follows, we suggest an alternative approach based on the information available in our dataset on interest payments, the replacement value of machinery and equipment, and the resale value of plant, land, equipment, machinery and other assets. We approximate the rate of rental of capital \( r \) using interest payments \( \text{int} \) and the resale (scrap) value of capital \( V^0 \) with the following relationship:

\[
\frac{\text{int}}{V^0} = \frac{r(p_0k)}{p_0k} = r
\]

where \( k \) is scrap capital and \( p_0 \) its price. Note that this ratio gives us the cost of servicing existing capital. Rates of rental of capital calculated for each firms

6This is by no means an exhaustive list of the subject.
are thus likely to be lower than interest rates on outstanding debt. Furthermore, if we assume that the price of capital is the discounted sum of the returns (rate of rental minus depreciation) on a unit of consumption invested as capital, we have that:

\[ p_0 = \sum_{t=0}^{\infty} \left( \frac{r - \delta}{1 + r} \right)^t = \frac{1 + r}{1 + \delta} \]  

(3.2)

where \( \delta \) is the rate of depreciation. Each firm’s own price of capital can be used to identify its own stock of capital with the following relationship:

\[ \ln(k) = \ln(V^0) - \ln(p_0) \]  

(3.3)

In Table 3.5, we examine the capital-labor ratios by size-bins. To construct these capital-labor ratios we use the methodology described above to assess the level of capital stock at a given firm. Obviously any mis-measurement of the estimated capital might alter the results found below. We thus examine capital-labor ratios where capital is approximated in 3 different ways. First, we use the resale value of capital as a proxy for the stock of capital. Second, the quantity of capital is derived from the resale value of capital using each firm’s own price of capital as calculated with the aforementioned methodology. Third, the quantity of capital is again derived from the resale value of capital but, this time, using the price of capital for a given size-bin (small, medium or large firms) using the same methodology as above. Table 3.5 summarizes the results using these proxies for capital labor ratio. We have also conducted robustness checks using the information available in the dataset on the replacement value of capital, on investment and borrowing. We discuss these results in the last paragraph of Section 3.4.3.

3.1.2 Summary Statistics

Table 3.1 presents some basic characteristics of the sample based on data from 1997. The variable Age gives the average number of years of existence of Ugandan firms in the sample. The variable Workers gives the average number of full-time and temporary employees in the sample. The average wage bill per worker is listed next. We report values for sales, replacement and resale values of capital as well
as tax obligations, taxes paid and bribes in thousands of US dollars (USD). The cost of servicing existing capital is computed based on the methodology described above. Tax obligations, taxes paid and bribes are also presented as a per employee ratio in thousands of USD. We report the value of tax savings (the difference between tax obligations and taxes actually paid) per bribe as described above. The variable Audit Intensity exhibit the average of self-reported values on whether an entrepreneur has been audited or not in the three years preceding 1998. We also presents the probability that an entrepreneur has been asked for a bribe. Finally, we report the growth rate of output. See Table 3.2 for a summary of these statistics by size bins.
### Table 3.1: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Observations</th>
<th>Mean</th>
<th>Std Dev</th>
<th>Median</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>242</td>
<td>13.9</td>
<td>12.5</td>
<td>10</td>
<td>1</td>
<td>74</td>
</tr>
<tr>
<td>Workers&lt;sup&gt;b&lt;/sup&gt;</td>
<td>243</td>
<td>124</td>
<td>259</td>
<td>35</td>
<td>0</td>
<td>2000</td>
</tr>
<tr>
<td>Wage/wkr&lt;sup&gt;c&lt;/sup&gt;</td>
<td>222</td>
<td>1.096</td>
<td>1.954</td>
<td>0.488</td>
<td>0</td>
<td>16.538</td>
</tr>
<tr>
<td>Sales&lt;sup&gt;c&lt;/sup&gt;</td>
<td>225</td>
<td>2486</td>
<td>9499</td>
<td>173</td>
<td>0.8</td>
<td>89100</td>
</tr>
<tr>
<td>Replacement Value of Capital&lt;sup&gt;c,d&lt;/sup&gt;</td>
<td>221</td>
<td>2606</td>
<td>10600</td>
<td>236</td>
<td>0.227</td>
<td>90900</td>
</tr>
<tr>
<td>Resale Value of Capital&lt;sup&gt;c,e&lt;/sup&gt;</td>
<td>155</td>
<td>2713</td>
<td>10200</td>
<td>500</td>
<td>2.0</td>
<td>103000</td>
</tr>
<tr>
<td>Cost of servicing capital</td>
<td>150</td>
<td>0.023</td>
<td>0.055</td>
<td>0</td>
<td>0</td>
<td>0.390</td>
</tr>
<tr>
<td>Tax obligations&lt;sup&gt;c,f&lt;/sup&gt;</td>
<td>230</td>
<td>488</td>
<td>2147</td>
<td>30</td>
<td>0</td>
<td>23100</td>
</tr>
<tr>
<td>Tax obl/worker&lt;sup&gt;c,f&lt;/sup&gt;</td>
<td>230</td>
<td>2.686</td>
<td>5.547</td>
<td>0.637</td>
<td>0</td>
<td>41.227</td>
</tr>
<tr>
<td>Taxes paid&lt;sup&gt;c,f&lt;/sup&gt;</td>
<td>231</td>
<td>407</td>
<td>3081</td>
<td>5</td>
<td>0</td>
<td>39300</td>
</tr>
<tr>
<td>Taxes paid/worker&lt;sup&gt;c,f&lt;/sup&gt;</td>
<td>230</td>
<td>1.498</td>
<td>5.668</td>
<td>0.150</td>
<td>0</td>
<td>65.065</td>
</tr>
<tr>
<td>Bribe&lt;sup&gt;c&lt;/sup&gt;</td>
<td>176</td>
<td>7</td>
<td>17</td>
<td>0.455</td>
<td>0</td>
<td>163</td>
</tr>
<tr>
<td>Bribe/worker&lt;sup&gt;c&lt;/sup&gt;</td>
<td>175</td>
<td>0.071</td>
<td>0.126</td>
<td>0.018</td>
<td>0</td>
<td>0.909</td>
</tr>
<tr>
<td>Tax savings/ Bribe&lt;sup&gt;c,f&lt;/sup&gt;</td>
<td>135</td>
<td>0.097</td>
<td>0.333</td>
<td>0.005</td>
<td>-0.515</td>
<td>2.411</td>
</tr>
<tr>
<td>Audit Intensity</td>
<td>250</td>
<td>0.62</td>
<td>0.485</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Probability being asked for a bribe</td>
<td>235</td>
<td>0.33</td>
<td>0.30</td>
<td>0.25</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Growth Rate of Output</td>
<td>208</td>
<td>0.09</td>
<td>0.55</td>
<td>0.07</td>
<td>-1.57</td>
<td>3.09</td>
</tr>
</tbody>
</table>

---

<sup>a</sup> Age is in years at the end of 1997.  
<sup>b</sup> Workers include permanent and temporary workers.  
<sup>c</sup> Sales, Replacement and Resale values, amounts of Tax obligations, Taxes and Bribes, Wage per worker, Tax obligations per worker, Taxes per worker and Bribes per worker, are in thousands of US dollars.  
<sup>d</sup> Replacement value of capital includes value of machinery and equipment.  
<sup>e</sup> Resale value of capital includes value of building, land, machinery, equipment and other assets.  
<sup>f</sup> Tax includes company income tax, sales tax VAT and NSSF.
Table 3.2: Summary Statistics by Size Bins

<table>
<thead>
<tr>
<th>Variable</th>
<th>≤ 30</th>
<th>30 &lt; x ≤ 75</th>
<th>&gt; 75</th>
<th>Variable</th>
<th>≤ 30</th>
<th>30 &lt; x ≤ 75</th>
<th>&gt; 75</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observations</td>
<td>116</td>
<td>89</td>
<td>78</td>
<td>Observations</td>
<td>116</td>
<td>89</td>
<td>78</td>
</tr>
<tr>
<td>Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>12.0</td>
<td>15.4</td>
<td>15.9</td>
<td>Tax obl./worker&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
<td>1.5</td>
<td>4.5</td>
<td>3.3</td>
</tr>
<tr>
<td></td>
<td>(10.0)</td>
<td>(13.3)</td>
<td>(14.9)</td>
<td>(3.3)</td>
<td>(8.1)</td>
<td>(6.1)</td>
<td></td>
</tr>
<tr>
<td>Workers&lt;sup&gt;b&lt;/sup&gt;</td>
<td>14</td>
<td>50</td>
<td>334</td>
<td>Taxes&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
<td>9</td>
<td>123</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>(7.57)</td>
<td>(12.52)</td>
<td>(379.3)</td>
<td>(29)</td>
<td>(306)</td>
<td>(5380)</td>
<td></td>
</tr>
<tr>
<td>Wage/worker&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.82</td>
<td>1.44</td>
<td>1.33</td>
<td>Taxes/worker&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.58</td>
<td>2.17</td>
<td>2.47</td>
</tr>
<tr>
<td></td>
<td>(1774.7)</td>
<td>(1825.6)</td>
<td>(2251.8)</td>
<td>(1.65)</td>
<td>(5.58)</td>
<td>(8.68)</td>
<td></td>
</tr>
<tr>
<td>Sales&lt;sup&gt;c&lt;/sup&gt;</td>
<td>121</td>
<td>1845</td>
<td>6664</td>
<td>Bribes&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.62</td>
<td>6.11</td>
<td>14.94</td>
</tr>
<tr>
<td></td>
<td>(234.4)</td>
<td>(6136.6)</td>
<td>(15600)</td>
<td>(1.35)</td>
<td>(8.89)</td>
<td>(25.9)</td>
<td></td>
</tr>
<tr>
<td>Replacement Value of Capital&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;d&lt;/sup&gt;</td>
<td>191</td>
<td>517</td>
<td>7944</td>
<td>Bribe/worker&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.05</td>
<td>0.13</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(369)</td>
<td>(832)</td>
<td>(18200)</td>
<td>(0.09)</td>
<td>(0.19)</td>
<td>(0.12)</td>
<td></td>
</tr>
<tr>
<td>Resale Value of Capital&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;e&lt;/sup&gt;</td>
<td>96</td>
<td>289</td>
<td>2673</td>
<td>Tax savings/bribe&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
<td>.12</td>
<td>.02</td>
<td>.12</td>
</tr>
<tr>
<td></td>
<td>(209)</td>
<td>(537)</td>
<td>(7139)</td>
<td>(.38)</td>
<td>(.14)</td>
<td>(.35)</td>
<td></td>
</tr>
<tr>
<td>Cost of servicing capital</td>
<td>.008</td>
<td>.030</td>
<td>.036</td>
<td>Audit Intensity</td>
<td>.53</td>
<td>.71</td>
<td>.77</td>
</tr>
<tr>
<td></td>
<td>(.032)</td>
<td>(.055)</td>
<td>(.069)</td>
<td>(.50)</td>
<td>(.46)</td>
<td>(.42)</td>
<td></td>
</tr>
<tr>
<td>Tax obligations&lt;sup&gt;c&lt;/sup&gt;,&lt;sup&gt;f&lt;/sup&gt;</td>
<td>25</td>
<td>247</td>
<td>1331</td>
<td>Probability asked for a bribe</td>
<td>.29</td>
<td>.30</td>
<td>.40</td>
</tr>
<tr>
<td></td>
<td>(55)</td>
<td>(489)</td>
<td>(3638)</td>
<td>(.30)</td>
<td>(.27)</td>
<td>(.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Growth Rate of Output</td>
<td>.02</td>
<td>.13</td>
<td>.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(.04)</td>
<td>(.09)</td>
<td>(.09)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Age is in years at the end of 1997.  
<sup>b</sup> Workers include permanent and temporary workers.  
<sup>c</sup> Sales, Replacement and Resale values, amounts of Tax obligations, Taxes and Bribes, Wage per worker, Tax obligations per worker, Taxes per worker and Bribes per worker, are in thousands of US dollars.  
<sup>d</sup> Replacement value of capital includes value of machinery and equipment.  
<sup>e</sup> Resale value of capital includes value of building, land, machinery, equipment and other assets.  
<sup>f</sup> Tax includes company income tax, sales tax VAT and NSSF.
First, we note from Table 3.1 a mean of 124 workers per firm and a median of 35 workers. The mode (not shown) at 15 workers is to the right of the median. These results are typical of a size distribution that is skewed to the right. Given our definition of the size bins small firms (30 employees or less) account for about 50% of the sample. Moreover, small firms are likely to be underrepresented in this dataset as a large informal sector exists in Uganda which is not captured by the UPSES.

The mean annual wage per employee is US $1096 for all firms, US $822 for small firms, US $1435 for medium firms and US $1328 for large firms. In 1997, the cost of servicing existing capital was approximately 2.3% for all firms, 0.8% for small firms, 3% for medium firms and 3.6% for large firms. Tax obligations represent 14.4% of sales in 1997. Tax obligations per employee are higher for medium firms which owe an average of US $4526 as opposed to US $1548 for small firms and US $3313 for larger firms. Taxes paid represent 7.7% of sales value. Overall, about half of the tax obligations disappear through evasion. In terms of taxes paid per worker, small firms pay a third of what they owe, medium firms a bit less than half of their tax obligations, and large firms evade a fifth of their tax liabilities after exemptions. Roughly 72% of the surveyed firms report paying bribes. Of these, 48% say they did so to reduce tax obligations [Azam et al., 2004]. The average bribe per employee is US $71 and bribes represent 2.4% of sales value. Again, medium firms are paying more bribes per employee (US $127) than other firms (US $47 for small firms and US $73 for large firms). The probability of being asked to pay a bribe is 0.33 on average, 0.29 for small firms, 0.30 for medium firms and 0.40 for larger firms. Finally, the probability of being audited is 62% of the sample reports being audited for CIT or VAT in one of the three years prior to 1998, 53% for small firms, 71% for medium firms and 77% for large firms.

---

7See Gauthier and Reinikka [2006] for further details on tax evasion and exemption in Uganda.
8Using senior management’s time spent dealing with officials, we get 15% for all firms, 10% for small firms, 16% for medium firms and 21% for large firms. The correlation coefficient of the probability of being asked for a bribe and the variable being audited or not is 0.18, of the probability of being asked for a bribe and senior’s time is 0.32 and being audited or not and senior’s time is 0.19.
3.2 The Missing Middle in LDCS and Uganda

There is a huge body of evidence on the existence of a missing middle in LDCS. Table 1 of Tybout [2000] summarizes these results for many countries across many years. In these studies, the missing middle is usually defined by comparing the number of middle sized firms in a less developed country to the number of middle sized firms in the U.S. The U.S. size distribution of firms is assumed to be undistorted, and thus, a relatively good benchmark for the sake of this comparison. Figure 3.1 is tailored to be comparable to these studies. It is a simple histogram of 3 size bins, based on the categorization described above, i.e. small firms have 30 employees or less, medium firms have between 31 and 75 employees, and large firms have 76 employees or more. Tybout [2000], Gauthier and Reinikka [2006], Soderbom and Teal [2004] and many other papers in the development literature use a similar classification of firm size.

Using the number of employees as our measure of firm size, Figure 3.1 shows the number of firms per size-bins based on our definition of small, medium and large firms. We stress again that values reported for small firms are likely to be a lower bound. A large informal sector exists in Uganda which is not captured in our dataset, therefore the reported data most likely underestimate the actual share of small firms.

In Figure 3.2 we present the density of firm size for our sample where size is taken to be the number of employees. We have used a kernel density estimator with an Epanechnikov kernel. The left panel exhibits the density estimates for the whole size range while the right panel focuses on firms with 100 employees or less.\footnote{In both panels, the bandwidth, $h$, is given by: $h = \frac{0.9m}{n^{1/5}}$ where $m = \min(\sqrt{\text{variance}_{size}}, \frac{\text{interquartilerange}_{size}}{1.349})$.}

We clearly see from the right panel of Figure 3.2 that there is a large mass of firms between 10 and 30 employees, which accounts for about 30% of the Ugandan sample. The density estimates taper off beyond 30 employees.

The depiction of the missing middle in Figure 3.1 and Figure 3.2 is rather simplistic. Moreover, most studies have documented this fact for a restricted number of countries. To the best of our knowledge, no study has examined the relationship between the share of medium firms and the level of development using a large sam...
Figure 3.1: Size Distribution of Firms in Uganda in 1997

Figure 3.2: Kernel Density Estimates for Full Sample (left) and Firms with 100 Employees or Less (right)

ple of countries. This is what we document in Figure 3.3 where we plot the share of medium firms against GDP per capita for 43 emerging and developing countries based on data from the World Bank Enterprise Surveys (WBES).

On Figure 3.3 the slope is 0.16 and significant at the 5% level.\(^{10}\) Obviously, missing medium firms must fall either in the smaller or the larger size bin of the

\(^{10}\)The correlation is also significant with a coefficient of 0.38.
size distribution of firms, depending on the policy environment. In an environment favoring small [large] establishments entrepreneurs have an incentive to reduce [increase] their labor force. In the panels of Figure 3.4 we show the share of small and large firms against GDP per capita. It seems that poorer countries have smaller firms while richer countries have larger firms. One would tend to think that policies in LDCS inhibit the growth of firms.

**Figure 3.3:** Share of Medium Firms v.s. GDP per Capita

**Figure 3.4:** Share of Small and Large Firms v.s. GDP per Capita
Another point made along these lines is Alfaro et al. [2007] who present empirical evidence suggesting that size distributions of firms are less skewed to the right in LDCS than in developed economies. We corroborate this result in Figure 3.5 where skewness of firm size is plotted against GDP per capita using data from the WBES.

**Figure 3.5: Skewness v.s. GDP per Capita**

On Figure 3.5 the slope is 0.09 and is significant at the 10% level. The correlation between skewness and GDP per capita is positive with a coefficient of 0.25. This implies that size distributions of firms in developed countries have longer right tails and thus exhibit more larger firms. One could thus infer that in most LDCS policies not only inhibit the growth of firms but limit also the growth of those firms with the highest potential. Taking a closer look to Figure 3.5 we note that Uganda is among the most skewed countries in the sample and among the poorest. However, referring back to figures 3.3 and 3.4 we note that Uganda has fewer [more] medium

---

11 The graphical representation of kurtosis of firm size against GDP per capita is similar to Figure 3.5 and is available upon request. This implies that more of the variance in firm size of developed countries is due to infrequent extreme observations, i.e., the occurrence of very large firms is more likely in developed economies than in LDCS.
and large [small] firms than more developed countries. This implies that Uganda exhibits a large level of skewness to the right like more developed economies not because its right tail is longer like in more developed economies but because the bulk of the mass in the size distribution is concentrated to the left, i.e. there is a large preponderance of small firms. As we will show in the next section, the fiscal environment in Uganda favors small firms and this rationalizes the observed size distribution in this country. It thus seems that a redesigning of the Ugandan policies could bring about efficiency gains through the reallocation of resources toward more productive and larger units.

### 3.3 Empirical Methodology

#### 3.3.1 Density Estimator

McCrary (2008) suggests a formal test to verify whether a variable is discontinuous around a specific threshold. The test is based on an estimator for the discontinuity at the cutoff and it is implemented as a Wald test with a null hypothesis that the discontinuity is zero. The algorithm developed in McCrary (2008) allows backing out the density of the size variable and it is useful for applications where a discontinuous density function is itself of interest. McCrary (2008) cites for example Saez (1999) who shows that tax payers in the U.S. bunch at the first kink point of the U.S. income tax schedule.\footnote{Otso and Xu (2010) amend binning and local likelihood methods to estimate the discontinuity of a density function. They also suggest an empirical likelihood test and confidence sets for the discontinuity. It is on our agenda to use the test proposed by Otso and Xu (2010) and compare these results with those using McCrary (2008)’s methodology.}

The estimator is an extension of the local linear density estimator. In a first step, we obtain a finely gridded histogram of the size variable (number of employees). In the second step, we smooth this histogram using a local linear regression separately on each side of the cutoff. The midpoints of the histogram are treated as regressors and the normalized counts of the number of observations per size bins are treated as the outcome variable. We use a weighted regression where more weight is given to the bins nearest where we are trying to estimate the density, i.e. using a triangle kernel. The density function is then estimated by looping over evaluation points of
the size variable. This algorithm provides a detailed distribution of the size variable which can then be plotted to give suggestive graphical evidence of the discontinuity around the threshold.

The estimator suggested by McCrary [2008] has the advantage over traditional histogram techniques or kernel density estimates of allowing for point estimation or inference. Moreover, if one were to use a kernel density function separately for points to the left and right of the point of discontinuity, the density estimator suggested by McCrary [2008] is still better suited for this task as it allows correcting for boundary bias at the right and left of the cutoff (Marron and Ruppert, 1994). Moreover, among nonparametric methods showing good performance at boundaries, the local linear density estimator is the simplest (McCrary, 2008).

Define $\xi(s)$ as the density of size variable $s$ and $c$ as the threshold where the discontinuity potentially occurs. Define $\xi^+ = \lim_{s \to c^+} \xi(s)$ and $\xi^- = \lim_{s \to c^-} \xi(s)$. If there is a discontinuity in the density of the size variable one should observe:

$$\xi^+ \neq \xi^- \quad (3.4)$$

The parameter of interest in such an analysis is:

$$\vartheta = \ln \xi^+ - \ln \xi^- \quad (3.5)$$

This parameter is the log difference of the intercepts at the discontinuity. The procedure described above generates an estimate for $\vartheta$. Under standard nonparametric regularity assumptions, $\vartheta$ is consistent and asymptotically normal (McCrary, 2008). Its standard error is given by:

$$\sigma_{\vartheta} = \sqrt{\frac{1}{nh} \frac{24}{5} \left( \frac{1}{\xi^+} + \frac{1}{\xi^-} \right)} \quad (3.6)$$

Finally, we use the estimated parameter $\vartheta$ and its standard error to conduct a t-test of the null hypothesis of continuity at the threshold.

**Bandwidth and bin size**

McCrary [2008] claims that the estimator is robust to different choices of bin size, $b$, provided that the bandwidth, $h$, is such that $h/b > 10$. However, one has
to be careful when choosing a bandwidth. A large bandwidth provides an over-smoothed estimate of the data which may hide spurious features in the data that would otherwise be apparent with a smaller bandwidth. Below, we conduct a sensitivity analysis of the t-statistics associated to the log difference in intercepts at the size threshold using different bandwidths. McCrary [2008] also suggests an automatic procedure for the choice of the bandwidth which can be used as a basis for a subjective choice. The automatic bandwidth selection procedure is as follows (see McCrary [2008] section 3.2 for more details):

1. The first-step histogram is computed using the binsize \( \hat{b} = 2\hat{\sigma}n^{-1/2} \), where \( \hat{\sigma} \) is the sample standard deviation of the running variable.

2. Based on the first-step histogram, one estimates a global 4th order polynomial separately on each side of \( c \). For each side, one computes \( \kappa \left[ \hat{\sigma} (b - a) / \sum \hat{f}''(X_j) \right]^{1/5} \), and sets \( \hat{h} \) equal to the average of these two quantities, where \( \kappa = 3.348 \), \( \hat{\sigma} \) is the mean-squared error of the regression, \( (b - a) \) equals \( X_J - c \) for the righthand regression and \( c - X_1 \) for the left-hand regression, and \( \hat{f}''(X_j) \) is the estimated second derivative implied by the global polynomial model and the \( X_j \)'s are the grid points each \( \hat{b} \) units apart covering the support of the size variable.

### 3.3.2 Regression Model

We also consider the regression model:

\[
Y_i = \gamma_0 + \gamma_1 D_{\text{small}} X_i + \gamma_2 D_{\text{larger}} X_i + \beta Z_i + \epsilon_i \quad (3.7)
\]

where \( Y_i \) is an outcome variable, \( Z_i \) and \( X_i \) are observable characteristics of individual firms and \( D_{\text{small}} \) is a binary variable equal to one if \( s \leq c \) and zero if \( s > c \) and \( D_{\text{larger}} \) is a binary variable equal to zero if \( s \leq c \) and 1 if \( s > c \) where \( c = 30 \) employees.

This framework is useful to investigate the statistical significance of the correlation after and before the size-threshold between firm size and outcome variables such as the audit intensity, capital-labor ratios, firms’ growth of sales, bribes and taxes.

Note moreover that in our analysis below we will conduct a Wald test of the equality of the coefficients on the interaction of the variable of interest and a
dummy equal to one if below the threshold, \( c \), and zero otherwise, and the interaction of that same variable of interest and a dummy equal to one if above the threshold, \( c \), and zero otherwise. The result of this test will be reported as statistic \( F_2 \) in the relevant tables. This is equivalent to a test devised in Chow [1960] which allows verifying whether the coefficients in two linear regressions on different data sets are equal. The Chow test is most commonly used to test for the presence of a structural break.

### 3.3.3 T-Test

As a robustness check we also report t-test differences across size bins for key variables. Given the way we have defined size bins for small, medium and large firms, this amounts to compare samples of unequal size and variance\(^\text{13}\). The t-statistic for this test is given by:

\[
t = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}}
\]

where

\[
\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}
\]

and \( \bar{X} \), \( s_i^2 \) and \( n_i \) are the mean, unbiased estimator of the variance and number of observations of sample \( i \) respectively. The degrees of freedom are calculated according to:

\[
dof = \frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \left( \frac{1}{(n_1-1)} + \frac{1}{(n_2-1)} \right)
\]

\(^\text{13}\)This test seems the best compromise to carry this analysis. We have also used Mood’s median test (results not reported). However, Freidlin and Gastwirth [2000] argue that the median test performs poorly with small samples and that other tests should be preferred to it. Note that based on the median test, we do not find evidence against the results of the t-tests.
3.4 Empirical Results

3.4.1 The Audit Intensity

One key assumption in both theoretical model and numerical exercise is that the audit intensity increases with the number of employees. Figure 3.6 provides suggestive graphical evidence that entrepreneurs with small firms are audited less intensively than their larger counterparts. The graph is generated using a local smoother with an Epanechnikov kernel, a polynomial of degree 3 and a bandwidth of 26.5 employees. In the appendix, we examine the relationship between the audit intensity and firm size using a local polynomial smoother with no restriction on the polynomial (see Figure A.1). We clearly see on Figure 3.6 that the audit intensity increases sharply for small firms and then remains at the same level for medium firms and increases again for larger firms.

![Figure 3.6: Audit Intensity v.s. Firm Size in Uganda in 1997](image)

In chapter 4 and 5, we approximate the polynomial observed on Figure 3.6 using a step function. We use the size bins classification defined above, i.e. small firms have 30 employees or less, medium firms have between 31 and 75 employees,
and large firms have 76 employees or more.[14]

Figure 3.7: Step Function: Audit Intensity v.s. Number of Employees

On Figure 3.7, the first step corresponds to the average audit intensity faced by small firms which is 0.53, the second step, to the average intensity faced by medium firms which is 0.71 and the last step, to the average intensity faced by large firms which is 0.77. According to the results of the t-tests tabulated in Table 3.5 below the difference in average audit intensity between small and medium firms is significant.

Table 3.3 reports estimates for a standard OLS of the audit intensity on the log of firm size, as measured by the number of employees, and/or capital.[15] In column 1, we regress audit intensity on size only. The coefficient is significant and the

---

[14] It is on our research agenda to compare the results of Chapter 5 using a smooth polynomial instead of a step function.

[15] 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 what are 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.
Wald statistic F1 rejects the null that the coefficient on size is equal to zero. In column 2, we regress the audit intensity 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 intensity 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 size and capital are the same.

**Table 3.3:** Audit Intensity, Number of Employees and Capital: OLS Estimates

<table>
<thead>
<tr>
<th>Dependent Variable: Audit intensity</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
<td>.089***</td>
<td>.077**</td>
<td>.075**</td>
<td></td>
</tr>
<tr>
<td>Capital</td>
<td>.043**</td>
<td>.007</td>
<td>.006</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.516***</td>
<td>-.045</td>
<td>.260</td>
<td>.493</td>
</tr>
<tr>
<td>R-squared</td>
<td>.14</td>
<td>.13</td>
<td>.07</td>
<td>.16</td>
</tr>
<tr>
<td>Industry-Location</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>F1</td>
<td>17.9***</td>
<td>4.45**</td>
<td>5.22***</td>
<td>4.32**</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.037]</td>
<td>[0.0]</td>
<td>[0.02]</td>
</tr>
<tr>
<td>F2</td>
<td>1.47</td>
<td>1.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>[.23]</td>
<td>[.24]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nb. Obs.</td>
<td>242</td>
<td>150</td>
<td>150</td>
<td>150</td>
</tr>
</tbody>
</table>

*a* significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis; p-value in brackets; all regressors are in logs; b We use dummies for industry-location. c F1: Wald test stat with $H_0$: coefficients are jointly zero; d F2: Wald test stat with $H_0$: the coefficients on size and capital are equal.

Examining in more details the determinants of the audit intensity, Table 3.4 reports results from different regressions of the audit intensity on the interaction of D, the binary variable used to determine each size group, with the log of firm size
and capital, i.e. $X'_i = [\ln(l_i), \ln(k_i)]^{16}$

**Table 3.4:** Audit Intensity, Nb. of Emp. and Capital: OLS Estimates by Size Groups

<table>
<thead>
<tr>
<th>Dependent Variable: Audit intensity$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>------</td>
</tr>
<tr>
<td>Nb. of Emp. small$^e$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Nb. of Emp. large$^e$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Capital small$^e$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Capital large$^e$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>F1$^b$</td>
</tr>
<tr>
<td>p-value</td>
</tr>
<tr>
<td>F2$^c$</td>
</tr>
<tr>
<td>p-value</td>
</tr>
</tbody>
</table>

| Nb. Obs. | 115 | 127 | 62 | 88 | 242 | 150 | 150 |

$^a$ * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis; p-value in brackets; $^b$ All regressors are in logs; All regressions include dummies for industry-location; $^c$ F1: Wald test stat with $H_0$: coefficients are jointly zero; $^d$ F2: Wald test stat with $H_0$: the coefficients on size and capital are equal; $^e$ Smaller firms: less than 31 emp. and larger firms: more than 30 emp.

Column 1 in Table 3.4 reports results from a regression of the audit intensity on firm size, restricting the 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 the sample to firms with more than 30 employees. The coefficient on the number of employees 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 intensity on size interacted with the binary variable D for each size

These results are robust to the use of a probit specification.
groups. Coefficients for both terms are positive and significant with the coefficient on small firms half of what it is in column 1. The 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. As mentioned earlier, the statistic F2 is equivalent to a Chow test to examine whether there is a structural break in the data, i.e., whether the coefficients in two linear regressions on different data sets are equal. It does not seem to be the case in column 5. In column 6 we do the same exercise with capital interacted with the binary variable for each size groups. 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 3.8 below offers a suggestive graphical representation of this result. The figure presents the correlation between audit intensity and capital for small and larger firms. Firms are, or are not, audited and the audit variable thus takes value 1 or 0, respectively, 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 and this, no matter what is their level of capital.¹⁷

Finally, we introduce all 4 covariates in column 7 of Table 3.4. 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 size and capital are different from zero. The Wald statistic F2 accepts the null that all 4 coefficients are equal. However, a Wald test for each combination of two coefficients rejects the null that the two coefficients are equal except for the pair size-large and capital-large.¹⁸

¹⁷The relationship between the audit probability and sales or taxes exhibit somewhat similar trends as the relationship between the audit intensity and capital when we control for the size groups (graphs available upon request). Note that the correlation coefficient between size and capital is 0.68, between size and sales, 0.79, and between size and taxes paid, 0.64. These coefficients are significant at the 5% level.

¹⁸For the pair size-small and size-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 size-small and capital-small F2 equals 3.83 and the p-value is 0.05; for the pair size-large and capital-large F2 equals 0.14 and p-value is 0.71.
As a further robustness check we report the results of a t-test for the difference in audit intensity across size bins in Table 3.5 column 1. Rows 1 through 3 exhibit the mean, the standard deviation and the number of observations of the audit intensity in each size bin. The last two rows of Table 3.5 present t-test differences between subsequent size bins, assuming samples of unequal size and variance. As noted from row 4, the audit probability increases by 34% between small and medium firms and this difference is significant\(^{19}\).

These results suggest that the audit intensity is highly correlated to size, especially around the size threshold and that the correlation of the audit intensity on capital seems rather mild once we control for firm size.

\(^{19}\)We have examined different categorizations of firms by size and the effect on the difference in average audit intensity. The biggest difference in audit intensity occurs between small firms with 10 employees or less and firms with more than 10 employees. This difference equals 0.41 and significant at the 1% level according to a t-test with samples of unequal size and variance. The gap in average audit probability tapers off for small firms’ categorization of 15, 20 and 25 employees. All differences are highly significant. Note however that this is as expected since the audit probability schedule is increasing at a decreasing rate with the number of employees as seen on Figure 3.6. However, what is important is the effect of the gap in average audit intensity on the density of firm size. This effect is largest at 30 employees as shown in Section 3.4.2.
Table 3.5: Differences across Groups

<table>
<thead>
<tr>
<th>Variable</th>
<th>Capital-Labor Ratio</th>
<th>Average Wage</th>
<th>Cost of capital</th>
<th>Audit Intensity</th>
<th>Tax Obl. per emp.</th>
<th>Taxes paid per emp.</th>
<th>Bribe per emp.</th>
<th>Tax savings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for</td>
<td>25.2 42.2 40.7 71.6</td>
<td>.822</td>
<td>.008</td>
<td>.53</td>
<td>1548.4</td>
<td>578.6</td>
<td>46.1</td>
<td>117.6</td>
</tr>
<tr>
<td>Small Firms</td>
<td>(47.4) (68.8) (66.6) (116.9)</td>
<td>(1.775) (.032) (.05)</td>
<td>(3321.6) (1648.3)</td>
<td>(87.8) (380.5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(62) (62) (63) (62)</td>
<td>{111} {62} {116}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for</td>
<td>13.7 20.3 19.2 34.5</td>
<td>1.436</td>
<td>.030</td>
<td>.71</td>
<td>4526.2</td>
<td>2169.2</td>
<td>127.7</td>
<td>15.2</td>
</tr>
<tr>
<td>Medium Firms</td>
<td>(14.3) (18.5) (18.2) (31.4)</td>
<td>(1.826) (.055) (.07)</td>
<td>(8072.8) (5575.5)</td>
<td>(189.4) (141.8)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(34) (32) (34) (32)</td>
<td>{43} {32} {49}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average for</td>
<td>19.8 26.8 27.8 45.1</td>
<td>1.328</td>
<td>.036</td>
<td>.77</td>
<td>3313.4</td>
<td>2467.5</td>
<td>72.7</td>
<td>119.1</td>
</tr>
<tr>
<td>Large Firms</td>
<td>(40.7) (51.7) (50.5) (31.4)</td>
<td>(2.252) (.069) (.07)</td>
<td>(6077.9) (8684.6)</td>
<td>(117.7) (345.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(58) (56) (60) (54)</td>
<td>{78} {56} {78}</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difference</td>
<td>11.5* 21.8** 21.6** 37.1**</td>
<td>-.61*</td>
<td>-0.23**</td>
<td>-0.19**</td>
<td>-2977.8**</td>
<td>-1590.6**</td>
<td>-81.6**</td>
<td>102.3*</td>
</tr>
<tr>
<td>Small/medium</td>
<td>[1.76] [2.34] [2.41] [2.34]</td>
<td>[-1.88]</td>
<td>[-2.13]</td>
<td>[-2.35]</td>
<td>[-2.37]</td>
<td>[-1.88]</td>
<td>[-2.40]</td>
<td>[1.83]</td>
</tr>
<tr>
<td>Difference</td>
<td>-6.1 -6.4 -8.6 -10.6</td>
<td>.108</td>
<td>-.006</td>
<td>-.05</td>
<td>1212.7</td>
<td>-298.3</td>
<td>55.0</td>
<td>-103.8*</td>
</tr>
</tbody>
</table>

*a std. dev. in parenthesis; number of firms in curly brackets; t-statistics in brackets; *significant at 10%; **significant at 5%; ***significant at 1%;  
* Capital is obtained from resale values of building, land, machinery, equipment and other assets, and is in millions of consumption units per employee; Capital is calculated based on: resale value of capital column 1, each firm’s own price of capital in column 2 and 4, price of capital for each size-bin in column 3; Note that in column 4, we use estimated capital over wage bills divided by average wage per size-bin as a proxy for the capital-labor ratio;  
* Average wage includes overheads and is in thousand of U.S. dollars;  
* Cost of servicing capital is calculated based on $\delta = 7\%$ but significance is not affected in the range $\delta = 0$ to 100%.
Other Potential Determinants of the Audit intensity

Columns 1 and 2 in Table 3.6 report OLS estimates for a regression of the audit intensity on firm size and other potential determinants of the audit intensity: sales and tax obligations, respectively.

**Table 3.6: Audit Intensity, Nb. of Emp., Sales and Taxes: OLS Estimates**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Dependent Variable: Audit intensity</th>
<th>Nb. of Emp. small(e)</th>
<th>Nb. of Emp. large(e)</th>
<th>Sales (e) small</th>
<th>Sales (e) large</th>
<th>Taxes owed (e) small</th>
<th>Taxes owed (e) large</th>
<th>(R^2)</th>
<th>F1(c)</th>
<th>p-value (c)</th>
<th>F2(d)</th>
<th>p-value (d)</th>
<th>Nb. Obs.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>.24***</td>
<td>.197***</td>
<td>(.074)</td>
<td>(.062)</td>
<td>(.051)</td>
<td>(.032)</td>
<td>.23</td>
<td>9.75***</td>
<td>[0.0]</td>
<td>3.45***</td>
<td>[0.02]</td>
<td>225</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(.049)</td>
<td>.068**</td>
<td>(.051)</td>
<td>(.032)</td>
<td>(.023)</td>
<td>(.021)</td>
<td>.22</td>
<td>8.91***</td>
<td>[0.0]</td>
<td>2.30*</td>
<td>[0.08]</td>
<td>224</td>
</tr>
</tbody>
</table>

\(a\) *std. dev. in parenthesis; number of firms in curly brackets; \(t\)-statistics in brackets; *significant at 10%; **significant at 5%; ***significant at 1%;

\(b\) All regressors are in logs; All regressions include dummies for industry-location;

\(c\) F1: Wald test stat with \(H_0\): coefficients are jointly zero;

\(d\) F2: Wald test stat with \(H_0\): all coefficients on size and capital are equal.

\(e\) Smaller firms: less than 31 emp. and larger firms: more than 30 emp.

In both cases, size for small firms remains positive and significant and sales
or taxes for larger firms is also positive and significant. These results corroborate what we found earlier using capital as a determinant of the Audit intensity: the number of employees is what matters for smaller firms when it comes to determine the Audit intensity whereas capital, sales and taxes can be used interchangeably to determine the Audit intensity of larger firms.  

Comparing Auditing Scheme in the United States and Uganda

It seems natural that even governments of developed countries should target the largest firms because tax revenues will be higher than enforcement costs there. Table 3.7 and Table 3.8 allow comparing the auditing scheme of the Internal Revenue Service (IRS) in the U.S. to the auditing scheme in Uganda, respectively. The first row in each table gives the size range according to which the auditing scheme varies, where size is measured in terms of million of U.S. dollars. The reason for using value in million of U.S. dollars for size is based on data availability from the IRS. We adopt this size classification for Uganda in this section only. The size bounds in million of U.S. dollars for Uganda in Table 3.8 are set to correspond approximatively to the share of firms per size-bin for the U.S. The third row in each table gives the shares of firms per size bins which differ slightly due to the small sample size in Uganda. The key comparison here is between the second row of Table 3.7 which presents the audit coverage in the U.S., defined as the ratio of examined versus filled reports, and the second row of Table 3.8 which presents the self-reported audit intensity in Uganda.

First, from Table 3.7 and Table 3.8 we note that the audit intensity in Uganda is extremely high as compared to the audit coverage in the U.S. for the smallest size bin which comprises about 97% of each sample. Second, the increase in audit coverage in the US is gradual and mild in the three subsequent size bins whereas in Uganda the audit intensity increases steeply. Finally, there is an important increase in audit coverage for the largest size bin in the US but this only concerns 0.4% of the US firms. These evidence imply that the auditing scheme in Uganda inhibits the growth and expansion of firms by targeting firms at an early stage of their

Note that we are using tax obligations as opposed to taxes paid. The reason for this is that we cannot assess whether the amount of taxes paid takes into account what has been repaid due to a successful audit.
Table 3.7: Audit Coverage in the U.S. in 2008 (%)

<table>
<thead>
<tr>
<th>Size a</th>
<th>&lt; 10</th>
<th>10 &lt; 50</th>
<th>50 &lt; 100</th>
<th>100 &lt; 250</th>
<th>&gt; 250</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit Coverage b</td>
<td>0.95</td>
<td>11.7</td>
<td>11.7</td>
<td>12.8</td>
<td>27.4</td>
</tr>
<tr>
<td>Share of firms c</td>
<td>97.2</td>
<td>1.5</td>
<td>0.3</td>
<td>0.4</td>
<td>0.6</td>
</tr>
</tbody>
</table>

a Size is in million of U.S. dollars; b Audit coverage in the US is defined as the ratio of examined versus filled reports; c The share of firms corresponding to the size categories is used in the following table to assess the size bounds for Uganda.

Table 3.8: Audit Intensity in Uganda in 1997 (%)

<table>
<thead>
<tr>
<th>Size a</th>
<th>&lt; 18</th>
<th>18 &lt; 43</th>
<th>43 &lt; 73</th>
<th>73 &lt; 88</th>
<th>88 &lt; 89</th>
</tr>
</thead>
<tbody>
<tr>
<td>Audit intensity b</td>
<td>66</td>
<td>67</td>
<td>100</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>Share of firms c</td>
<td>97.3</td>
<td>1</td>
<td>0.4</td>
<td>0.4</td>
<td>0.4</td>
</tr>
</tbody>
</table>

a Size is in million of U.S. dollars; b Audit intensity in Uganda is self-reported by entrepreneurs; c The bounds for size categories are set to match the share of firms as computed from Table 3.7.

development.

The next section reports evidence of the gap in the size distribution of firms in Uganda as a consequence of the uneven auditing schedule.

3.4.2 Density of Firm Size

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 3.9 shows the firm size density estimated with the algorithm described previously for size bins of 5 employees and a bandwidth of 50 employees.

A clear discontinuity at 30 employees is present (vertical line in Figure 3.9). The log difference in the intercepts at the discontinuity, i.e. $\vartheta$, is equal to -0.99 and its standard deviation is 0.298. A t-test rejects the null hypothesis of continuity. We interpret this as evidence that entrepreneurs’ labor choices are distorted by the tax administration targeting of firms above 30 employees.
Why 30 employees? Apart from consistency with the previous literature in development, we explain in what follows why we have chosen a size-threshold of 30 employees. The audit intensity schedule is not as stark as shown on Figure 3.7. One consequence of this is that the discontinuity in the density of firm size most likely does not occur exactly at 30 employees. Table 3.9 reports a sensitivity analysis on the log difference in intercepts using different cutoff values between 10 and 35 employees. We clearly see that the largest and most significant log difference between intercepts occurs at 30 employees. Moreover, the result remains unchanged if we use the bandwidth selected automatically for a cutoff of 30 employees, i.e. a bandwidth of 48 employees, for all other cutoff values. The largest difference still occurs at 30 employees (results not shown and available upon request).

The fact that we observe a significant discontinuity from 15 to 30 employees using McCrary (2008)’s methodology indicates that there is a preponderant mass of firms in this range. Due to the uncertainty associated with the audit intensity (it does not jump off right at 30 employees but increases starkly around that threshold), firms do not bunch right before 30 employees. So it should come as no sur-
Table 3.9: Comparison of Size-Thresholds for the Density of Firm Size

<table>
<thead>
<tr>
<th>Size-threshold (Nb. of Emp.)</th>
<th>Bandwidth (Nb. of Emp.)</th>
<th>Log Difference in intercepts</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>29.0</td>
<td>-0.13</td>
<td>0.24</td>
</tr>
<tr>
<td>15</td>
<td>32.2</td>
<td>-0.62</td>
<td>0.25</td>
</tr>
<tr>
<td>20</td>
<td>35.7</td>
<td>-0.53</td>
<td>0.26</td>
</tr>
<tr>
<td>25</td>
<td>41.2</td>
<td>-0.73</td>
<td>0.28</td>
</tr>
<tr>
<td>30</td>
<td>47.6</td>
<td>-0.97</td>
<td>0.31</td>
</tr>
<tr>
<td>35</td>
<td>51.7</td>
<td>-0.51</td>
<td>0.31</td>
</tr>
</tbody>
</table>

All estimates are based on size-bins of 5 employees.

prise that, as mentioned earlier, the mode of the size distribution is at 15 employees and not just before 30 employees.\(^{21}\)

The automatic selector from McCrary (2008) suggests the use of a binsize of 22 employees and a bandwidth of 45 employees. As we have noted above, the estimator is robust to different choices of binsize provided that \(h/b > 10\). Given that a binsize of 22 employees does not convey a good graphical representation of the density around the threshold, we have use a binsize of 5 employees. The results are robust to choices of binsize between 5 and 50 employees (results not shown). The choice of bandwidth however does affect the performance of the estimator. Table 3.10 reports the log difference in intercepts and the associated standard error for a bandwidth ranging between 25 to 50 employees. As can be inferred from Table 3.10, the results are robust to bandwidth choices of 35 employees and above. We provide in the appendix a graph of the density of firm size using a bandwidth of 5 employees. By comparing Figure 3.9 to Figure A.2, we note that the potentially oversmoothed graphical representations that we present in this section, using Figure 3.9 are not hiding spurious features in the data.

\(^{21}\) There are 73 firms between 10 and 30 employees, 47 between 15-30, 31 between 20-30 and 15 between 25-30.
Table 3.10: Comparison of Bandwidth for the Density of Firm Size

<table>
<thead>
<tr>
<th>Bandwidth (Nb. of Emp.)</th>
<th>Log Difference in intercepts</th>
<th>Standard Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>25</td>
<td>-.36</td>
<td>.48</td>
</tr>
<tr>
<td>30</td>
<td>-.43</td>
<td>.42</td>
</tr>
<tr>
<td>35</td>
<td>-.64</td>
<td>.38</td>
</tr>
<tr>
<td>40</td>
<td>-.82</td>
<td>.34</td>
</tr>
<tr>
<td>45</td>
<td>-.93</td>
<td>.32</td>
</tr>
<tr>
<td>50</td>
<td>-.99</td>
<td>.30</td>
</tr>
</tbody>
</table>

All estimates are based on a binsize of 5 employees.

3.4.3 Capital-Labor Ratios

Another implication of the size-dependence of audit probabilities is that entrepreneurs may respond to positive productivity shocks by substituting capital for labor at the early stage of their development. We first present some suggestive graphical evidence that this is indeed the case in Uganda. We present in Figure 3.10 the correlation between capital-labor ratios and firm size for a size range between 1 and 150 employees for the sake of clarity. The results hold if we take all firms with more than 75 employees. The left panel of Figure 3.10 shows that capital-labor ratios decrease with size over the whole size range.\(^{22}\) We show in Chapter 4 that this is due to the fact that entrepreneurs must finance their investment with their own savings because of missing credit markets in Uganda.\(^{23}\) We note in the right panel of Figure 3.10 that capital-labor ratios are increasing before each size threshold, that is, just before an increase in the step function of the audit intensity. We note that the average capital-labor ratio for the first size group (blue dots) is larger than for medium firms (green dots).\(^{24}\) We observe the same patterns between medium

\(^{22}\) Is this atypical? Soderbom and Teal [2004] show that capital-labor ratios are increasing with firm size in Ghana and that this is due to a differential in factor prices.

\(^{23}\) Indeed, based on the euler equation\(^{4.39}\) we show that firms decrease their investment in capital as they grow.

\(^{24}\) More on this below.
firms and firms with 76 to 150 employees.

**Figure 3.10: Capital-Labor Ratios v.s. Firm Size**

We use Equation 3.7 to study the distribution of capital intensities by firm size: Table 3.11 reports results from a regression of the Capital-Labor ratio on size for a group of small firms (less or equal to 30 employees) and a group of larger firms (more than 30 employees).

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 jointly different from zero (F1) and 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 (significant at 1%) 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 on growth dynamics of small and medium firms, in column 5 we restrict the sample to include only firms with 500 employees or less.
Table 3.11: Capital-Labor Ratio and Number of Employees: OLS Results

<table>
<thead>
<tr>
<th>Dependent Variable: ln(Resale Value of Capital per employee)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. of Emp. small(c)</td>
<td>4.54***</td>
<td>2.92***</td>
<td>1.18***</td>
<td>1.17***</td>
<td>1.35***</td>
</tr>
<tr>
<td>(small)</td>
<td>(.107)</td>
<td>(.189)</td>
<td>(.212)</td>
<td>(.216)</td>
<td>(.255)</td>
</tr>
<tr>
<td>Nb. of Emp. large(e)</td>
<td>2.81***</td>
<td>1.90***</td>
<td>1.03***</td>
<td>.99***</td>
<td>1.11***</td>
</tr>
<tr>
<td>(large)</td>
<td>(.047)</td>
<td>(.103)</td>
<td>(.111)</td>
<td>(.113)</td>
<td>(.145)</td>
</tr>
<tr>
<td>Constant</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.97</td>
<td>0.98</td>
<td>0.49</td>
<td>0.49</td>
<td>0.44</td>
</tr>
<tr>
<td>F1c</td>
<td>2707.2***</td>
<td>169.5***</td>
<td>73.8***</td>
<td>61.0***</td>
<td>45.43***</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
</tr>
<tr>
<td>F2d</td>
<td>218.7***</td>
<td>72.0***</td>
<td>1.61</td>
<td>2.11</td>
<td>3.27*</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[.21]</td>
<td>[.15]</td>
<td>[.07]</td>
</tr>
<tr>
<td>Nb. Obs.</td>
<td>157</td>
<td>157</td>
<td>157</td>
<td>157</td>
<td>145</td>
</tr>
</tbody>
</table>

\(\text{a}^*\text{ significant at 10\%, ** significant at 5\%, *** significant at 1\%; Std. Dev. in parenthesis, p-value in brackets;}

\(b\) Regressors are in logs and covariates are dummies for industry-location, Age and a dummy for the respondent being the owner of the firm;

\(c\) F1: Wald test stat with \(H_0\): coefficients are jointly zero;

\(d\) F2: Wald test stat with \(H_0\): all coefficients on size and capital are equal.

\(e\) Smaller firms: less than 31 emp. and larger firms: more than 30 emp.

\(f\) Column 5: larger firms: between 31 and 500 employees.

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 support the view that small firms tend to substitute capital for labor at the size-threshold\(^{25}\).

As a robustness check we examine the t-test differences across size bins for different proxies of the capital-labor ratios in columns 1 to 4 of Table 3.5. Rows 1 through 3 exhibit the mean, the standard deviation and the number of observations in each size bin. The last two rows of Table 3.5 present t-test differences between subsequent size bins, assuming samples of unequal size and variance. Column 1...

\(^{25}\text{We uncover similar patterns in capital-labor ratios using a local polynomial smoother (See Figure A.3 in the Appendix).}\)
uses resale value of capital per employee as a proxy for the capital-labor ratio. The difference between small and medium firms is statistically significant at the 10% level. Small firms exhibit higher resale values per employee than medium firms. In column 2 capital is estimated with the methodology described in Section 3.1 using each firm’s idiosyncratic price of capital while in column 3 the average price of capital per size-bin is used. In both cases a depreciation rate of 7% is assumed. In column 2, the difference between small and medium firms is significant at the 5% level. In column 3, we compute an average cost of servicing capital for each size bins to identify capital, using positive costs only. Note that this should go against our results. Indeed, capital stocks in column 3 are lower than in column 2. However, the difference between small and medium firms remains significant at the 5% level. Finally in column 4 we calculate capital-labor ratios using capital estimated from the methodology of Section 3.1 and wage bills to proxy for labor. The difference between small and medium is still significant at the 5% level. As further robustness checks (results not shown) we have used the replacement value of capital as a proxy for the capital stock. We still observe larger capital labor ratios for small firms but this difference is not significant anymore. We have also examined capital labor ratios using investment in capital in 1997 as a proxy for capital. We find that investment intensities are actually increasing with size for small and medium firms. The results using borrowing for investment and working capital lend to the same conclusion: borrowing intensities increase with the number of employees.

3.4.4 Firm’s Growth of Output

Table 3.2 reports figures for average output growth by size group. Output growth rates is given by:

\[ \text{FirmsGrowth} = \ln(Sales_{1997}) - \ln(Sales_{1996}) \]  

This is surprising, especially given the credit constraints faced by small firms as observed in column 6 of Table 3.5. We discuss below the particularities of the Ugandan factors market. Results are not qualitatively affected by a rate of depreciation varying between 0 and 1.

If we discard firms with a zero cost of servicing capital (result not shown and see below for a description of the factors market), the difference is not significant anymore because the number of observations drops such that the statistical power of the t-test diminishes dramatically.
Average output 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 of output on size are shown in Table 3.12 where the coefficient on the number of employees is positive and significant. A simple calculation tells us that a 1% increase in the number of employees (equivalent to 20 employees in our dataset) generates a 0.7% increase in firms’ growth rate of output.

**Table 3.12:** Growth of Output and Number of Employees: OLS Estimates

<table>
<thead>
<tr>
<th>Dependent Variable: Growth rate of Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Employees</td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td>(1.4e-4)</td>
</tr>
<tr>
<td>$R^2$</td>
</tr>
<tr>
<td>Observations</td>
</tr>
<tr>
<td>F-test$^b$</td>
</tr>
<tr>
<td>p-value</td>
</tr>
</tbody>
</table>

$^a$ *significant at 10%, **significant at 5%, ***significant at 1%; Std. Dev. in parenthesis, p-value in brackets;  
$^b$ F-test with H0: coefficient of the number of employees is zero;

In Table 3.13 we regress the growth of output on the two size groups, small or large. Columns 1 to 3 of Table 3.13 present the regression results of the growth of output 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 suggest that small firms restrict their growth to avoid heavier regulation costs.$^{29}$

$^{29}$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.
Table 3.13: Growth of Output and Number of Employees: OLS Estimates by Size Groups

<table>
<thead>
<tr>
<th>Dependent Variable: Growth Rate of Output</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. of Emp. small( ^e )</td>
<td>-.04</td>
<td>.066</td>
<td>-.054( ^* )</td>
<td>-.072</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.028)</td>
<td>(.053)</td>
<td>(.031)</td>
<td>(.055)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nb. of Emp. Large( ^e )</td>
<td>.031( ^** )</td>
<td>.061( ^** )</td>
<td>.039( ^** )</td>
<td>.074( ^** )</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.015)</td>
<td>(.029)</td>
<td>(.017)</td>
<td>(.031)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Covariates | No | No | No | Yes | Yes | Yes |
| R-squared  | .009 | .02 | .03 | .07 | .09 | .09 |

| F1\( ^c \) | 2.11 | 4.07\( ^** \) | 2.7\( ^* \) | 3.02\( ^* \) | 5.36\( ^** \) | 3.49\( ^* \) |

| F2\( ^d \) | .02 | 0.0 |
| p-value     | [.977] | [.96] |

| Nb. Obs. | 208 | 208 | 208 | 208 | 208 | 208 |

\( ^a \) * significant at 10\%, \( ^* \) significant at 5\%, \( ^** \) significant at 1\%; Std. Dev. in parenthesis, p-value in brackets; \( ^b \) Regressors are in logs and covariates are dummies for industry-location and Age; \( ^c \) F1: Wald test stat with \( H_0 \): coefficients are jointly zero; \( ^d \) F2: Wald test stat with \( H_0 \): all coefficients on size and capital are equal. \( ^e \) Smaller firms: less than 31 emp. and larger firms: more than 30 emp. All results presented here are with a constant. Results without a constant are also significant (not shown). Results not significantly affected if we use gross growth rate.

3.4.5 Bribes and Taxes

We show here that larger firms pay larger bribes and taxes. We start with some graphical evidence. The left and right panels of Figure 3.11 plot bribes and taxes paid against firm size for small firms (less or equal than 30 employees) and larger firms (more than 30 and less than 75 employees). Both graphs document that larger firms face on average higher regulation payments than small firms.\(^{30}\)

Turning to regression estimates, Table 3.14 reports results of a regression of

\(^{30}\)Using a quadratic fit for taxes paid against firm size we obtain a relationship between taxes and size which is smoother than the one depicted here and taxes are higher for larger firms even at the size-threshold (graph available upon request).
bribes on firm size interacted with the binary variable D for each size group. Column 1 presents the results of a regression without a 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 3.7. The coefficient on small firms is not significant anymore and the constant seems to pick up the amount of bribes paid by small firms. We show below that the smallest of the small firms pay higher bribes for renegotiating licenses and permits due to the ambiguity of their legal status. This helps to understand the observation of a negative slope in the left panel of Figure 3.11 for small firms. The coefficient for larger firms in Table 3.14 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 3.15 presents the results for a regression of taxes paid on size for two size groups: small (less or equal to 30 employees) and large (more than 30 and less or equal to 75 employees). 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,
Table 3.14: Bribes and Number of Employees: OLS Estimates

<table>
<thead>
<tr>
<th>Dependent Variable: ln(Bribes)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. of Emp. small&lt;sup&gt;e&lt;/sup&gt;</td>
<td>2.05***</td>
<td>1.46***</td>
<td>.08</td>
<td>.03</td>
</tr>
<tr>
<td></td>
<td>(.09)</td>
<td>(.2)</td>
<td>(.27)</td>
<td>(.27)</td>
</tr>
<tr>
<td>Nb. of Emp. large&lt;sup&gt;e&lt;/sup&gt;</td>
<td>1.72***</td>
<td>1.38***</td>
<td>.65***</td>
<td>.65***</td>
</tr>
<tr>
<td></td>
<td>(.04)</td>
<td>(.12)</td>
<td>(.14)</td>
<td>(.14)</td>
</tr>
<tr>
<td>Constant</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>.94</td>
<td>.94</td>
<td>.49</td>
<td>.52</td>
</tr>
<tr>
<td>F1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>1039.9***</td>
<td>83.45***</td>
<td>65.98***</td>
<td>67.23***</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
</tr>
<tr>
<td>F2&lt;sup&gt;d&lt;/sup&gt;</td>
<td>11.19***</td>
<td>.38</td>
<td>15.77***</td>
<td>18.28***</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[.54]</td>
<td>[0.0]</td>
<td>[0.0]</td>
</tr>
<tr>
<td>Nb. Obs.</td>
<td>142</td>
<td>142</td>
<td>142</td>
<td>142</td>
</tr>
</tbody>
</table>

<sup>a</sup> *significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis, p-value in brackets;  
<sup>b</sup> Regressors are in logs and covariates are dummies for industry-location;  
<sup>c</sup> F1: Wald test stat with $H_0$: coefficients are jointly zero;  
<sup>d</sup> F2: Wald test stat with $H_0$: all coefficients on size and capital are equal.  
<sup>e</sup> Smaller firms: less than 31 emp. and larger firms: more than 30 emp.

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.

Finally, Table 3.5 presents t-test differences across size-bins for tax obligations per employee, taxes paid per employee, bribes per employee and tax savings through bribes in columns 8 to 11 respectively. Rows 1 through 3 exhibit the mean, the standard deviation and the number of observations in each size bin. The last two rows of Table 3.5 present t-test differences between subsequent size bins, assuming samples of unequal size and variance. From column 8, we note that medium firms owe three times the tax liabilities per employee than smaller firms and that this
### Table 3.15: Taxes and Number of Employees: OLS Estimates

<table>
<thead>
<tr>
<th>Dependent Variable: ln(Taxes)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nb. of Emp. small(^c)</td>
<td>5.44***</td>
<td>3.13***</td>
<td>1.1***</td>
<td>1.09***</td>
</tr>
<tr>
<td>Std. Dev. (.14)</td>
<td></td>
<td>(.25)</td>
<td>(.29)</td>
<td>(.28)</td>
</tr>
<tr>
<td>Nb. of Emp. large(^c)</td>
<td>3.55***</td>
<td>2.23***</td>
<td>1.23***</td>
<td>1.2***</td>
</tr>
<tr>
<td>Std. Dev. (.06)</td>
<td></td>
<td>(.14)</td>
<td>(.15)</td>
<td>(.15)</td>
</tr>
<tr>
<td>Constant</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Covariates</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>.96</td>
<td>.98</td>
<td>.42</td>
<td>.49</td>
</tr>
<tr>
<td>F1(^c)</td>
<td>2320.5***</td>
<td>121.01***</td>
<td>69.53***</td>
<td>64.29***</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[0.0]</td>
</tr>
<tr>
<td>F2(^d)</td>
<td>158.95***</td>
<td>33.07***</td>
<td>.67</td>
<td>.54</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.0]</td>
<td>[0.0]</td>
<td>[.42]</td>
<td>[.46]</td>
</tr>
<tr>
<td>Nb. Obs.</td>
<td>196</td>
<td>196</td>
<td>196</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) * significant at 10%, ** significant at 5%, *** significant at 1%; Std. Dev. in parenthesis, p-value in brackets;  
\(^b\) Regressors are in logs and covariates are dummies for industry-location;  
\(^c\) F1: Wald test stat with \( H_0 \): coefficients are jointly zero;  
\(^d\) F2: Statistic for a Wald test with \( H_0 \): coefficients on size-small and size-large are equal;  
\(^e\) Smaller firms: less than 31 emp. and larger firms: more than 30 emp.

This difference is significant at 5%. In columns 9 and 10, medium firms pay about three times what smaller firms pay in taxes and bribes per worker. These differences are also significant. In column 11, tax savings made through bribery are significantly higher (about 6 times) for small and large firms than for medium firms.

**Bribes for the Smallest of the Small Firms**

We examine other determinants of bribes for smaller firms such as the probability of being asked for a bribe 1) to get connected to public services, 2) to get licenses and permits, 3) when dealing with taxes and tax collection, 4) to gain government...
contracts or when dealing with customs. A regression of bribes per employee on book accounting yields a highly significant and negative coefficient. This effect vanishes as we introduce each probability of being asked for a bribe in turn. The full specification with the accounting variable and all probabilities yield the best fit and only the probability of being asked for a bribe to get licenses and permits is significant and positive. This is in line with the findings of Table 3.14 above and Table 3.17 below: the smallest firms are more prone to harassment and extortions. Bureaucrats re-negotiate licenses and permits because of the marginal status of these firms.

### 3.4.6 Robustness

**Differential in Factor Prices**

A natural explanation for observed patterns in capital-labor ratios is a differential in prices across factors and/or across firm sizes. In the Ugandan sample, small firms typically pay lower wages and seem to be credit constrained. Labor being cheaper and capital more expensive for these small firms implies that we should observe capital-labor ratios that are increasing with firm size, the opposite of what we find. This means that the capital-labor ratios in the Ugandan sample are driven by other forces than a differential in factor prices. Let us examine the price of labor and capital in more details. Table 3.5 columns 5 and 6, exhibit t-test differences across size bins for wage and the cost of servicing existing capital respectively. Rows 1 through 3 exhibit the mean, the standard deviation and the number of observations in each size bin. The last two rows of Table 3.5 present t-test differences between subsequent size bins, assuming samples of unequal size and variance. Examining column 5, we note that medium firms pay higher wages per employee than small firms, while this difference is not significant between medium and large firms. In column 6, we observe a similar pattern for the cost of servicing capital. The differences in wages and the cost of servicing capital in Table 3.5 reflect the structure of the market in Uganda. Using information on firms’ legal status, we find that smaller firms are usually family enterprises while larger firms tend to be partner-

---

31 Previous results for all firms or formal firms (medium and large) are not affected by the introduction of these covariates.
ships or corporations. The wage differential explains naturally from the fact that smaller firms are hiring relatives and friends at lower wages whereas larger firms are hiring individuals through the formal labor market. For the cost of servicing capital, the dataset gives information on the origins of financing.\textsuperscript{32} Table 3.16 shows the correlation between these sources of investment and size bins.

<table>
<thead>
<tr>
<th>Source of financing</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Own profits</td>
<td>-.12</td>
</tr>
<tr>
<td>b. Personal savings</td>
<td>-.25</td>
</tr>
<tr>
<td>c. commercial banks</td>
<td>.11</td>
</tr>
<tr>
<td>d. development finance</td>
<td>.24</td>
</tr>
<tr>
<td>e. parent or holding company</td>
<td>.19</td>
</tr>
</tbody>
</table>

Other sources of financing are family and relatives, suppliers credit, money lender, sale of assets, lease finance, new partner; all other sources yield coefficients below 0.1 and are positively related to size. Number of observations is 157 for a. and b., and 156 for c., d. and e.

As seen from Table 3.16, small firms use as their main sources of financing their own profits and personal savings. Larger firms (medium and large) finance their investment through commercial banks, development finance and parent or holding companies.\textsuperscript{33} The extremely low cost of servicing existing capital for small firms reflects the fact that these firms do not make interest payments and thus do not have access to formal credit.\textsuperscript{34} Indeed, Nisbet [1967] argues that on the one hand, loans to small firms are likely to be provided by relatives, friends or other noncommercial lenders who sometimes lend at negative rates and on the other hand

\textsuperscript{32}Sources are: profits, personal savings, family, relatives and friends, commercial banks, development finance, supplier's credit, money lender, parent or holding company, sale of assets, lease finance, or new partner.

\textsuperscript{33}Below equilibrium interest rates are typical of capital markets in LDCS [Berry, 1978]. Practitioners usually favor low interest rates as observed from those charged by public credit agencies and from governmental limits placed on loans. This could explain the low rates that we observe for medium and large firms.

\textsuperscript{34}We examine below differences between small firms for whom we have calculated a rate of rental equal to zero and those with a strictly positive rate of rental.
private moneylender rates sometimes range from 20% to 100% and dissuade small entrepreneurs from borrowing at all.\textsuperscript{35} In Chapter 4 and 5, we take this as evidence of an inefficient credit market where we define the degree of inefficiency as the extent to which the actual output is driven away from a First-Best level of output by distortions or impediments such as high collateral and high lending rates.\textsuperscript{36}

The cost of renting or buying new capital is actually high for these firms. This and the fact that small firms pay on average lower wages should actually drive down the capital-labor ratio of small firms, a conclusion that goes against what we actually observe in the data. To see this examine the objective function of a firm with a decreasing returns to scale Cobb-Douglas production function:

$$\max_{k,l} kl^{\gamma} - wl - rk$$

where \(w\) and \(r\) are factor prices and \(\gamma + \eta < 1\).

The optimal mix is then given by:

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

It is easy to see that if small firms are credit constrained but pay lower wages while larger firms have access to credit but pay higher wages the capital-labor ratios should be increasing with firm size. Soderbom and Teal \textsuperscript{2004} find that log of capital intensities are increasing across firm size in Ghana. They show that this is due to a factor price differential rather than technical inefficiency.\textsuperscript{37} The costs of servicing capital that we observe in the Ugandan sample should generate capital-labor ratios in line with the findings of Soderbom and Teal \textsuperscript{2004}. However, we find the opposite pattern for small and medium firms. We take these results as evidence that some capital-labor substitution occurs for smaller firms in the sample due to some distortion that offsets the differential in factor prices.\textsuperscript{38}

\textsuperscript{35}Note that based on the 1998 IMF Country Staff Report for Uganda we calculate a commercial banks’ lending rate of 21.33\% on average in 1997.

\textsuperscript{36}Note that our definition of inefficiency is not standard. Usually, an inefficient credit market is defined as a market where people cannot borrow the present discounted value of what they can be expected to repay over their lifecycle.

\textsuperscript{37}Our results also hold if we compare the logs of our different versions of capital intensities.

\textsuperscript{38}We could not find information on the auditing scheme for taxes in Ghana. However, the results
Differences across Small Firms

We also examine differences between small firms for whom we have calculated a cost of servicing existing capital equal to zero \( (r = 0) \) and those with a strictly positive cost of servicing existing capital \( (r \neq 0) \). Table 3.17 depicts these results. There are no significant differences in terms of resale values, tax obligations per employee, taxes paid per employee, tax savings through bribes or meeting probability. However, the former \( (r = 0) \) are significantly smaller in size, pay significantly more bribes and are more likely to not keep a record of their transactions (the correlation coefficient between a nil rate of rental and keeping an accounting book is -0.22). These results confirm that firms with a zero cost of servicing existing capital do not have access to credit (at least from commercial banks) probably because they are not registered officially. As mentioned earlier, one potential reason for them paying higher bribes is that these smaller firms could be prone to greater bureaucratic harassment due to the fact that there is no book to confirm their economic activities. Such a situation leaves discretion to officials to re-negotiate a licence, a patent or a permit of exploitation even though a firm has already paid for its right to operate.

\[ \text{Reference: Soderbom and Teal [2004] make a strong case for the effect of a price differential on capital intensities. Our point here is to show that Soderbom and Teal [2004]'s story does not fit the Ugandan data.} \]
Table 3.17: Differences across Small Firms

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number of Employees</th>
<th>K/L</th>
<th>Tax Obl. per emp.</th>
<th>Tax per emp.</th>
<th>Bribe per emp.</th>
<th>Tax savings</th>
<th>Probability asked for a bribe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Firms with $r = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>(6)</td>
<td>{44}</td>
<td>334</td>
<td>1221</td>
<td>267</td>
<td>59</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(773)</td>
<td>(2720)</td>
<td>(500)</td>
<td>(107)</td>
<td>(464)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>1221</td>
<td>267</td>
<td>59</td>
<td>125</td>
<td>.29</td>
</tr>
<tr>
<td>$r \neq 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>(6)</td>
<td>{18}</td>
<td>584</td>
<td>2475</td>
<td>970</td>
<td>20</td>
<td>172</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(874)</td>
<td>(4367)</td>
<td>(2815)</td>
<td>(38)</td>
<td>(341)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2475</td>
<td>970</td>
<td>20</td>
<td>172</td>
<td>.30</td>
</tr>
<tr>
<td>Difference</td>
<td>-4**</td>
<td>-250</td>
<td>-1254</td>
<td>-703</td>
<td>39*</td>
<td>-47</td>
<td>-.01</td>
</tr>
</tbody>
</table>

*significant at 10%, ** significant at 5%, *** significant at 1%; std. dev. in parenthesis; t-statistics in brackets; number of firms in curly brackets;
bribes are significantly lower for small firms with $r=0$ than medium firms;
Small Firms Piling up Capital

Lending rates are close to nil for small firms. This implies that they either use their own or their relatives’ savings to invest in capital. If relatives were a sound source of financing entrepreneurs could pile up capital for 2 potential reasons. First, this would allow meeting the stringent collateral requirement for borrowing. Second, they could actually lend capital to larger firms (Banerjee and Duflo [2008]). In this case, the marginal product of capital of these entrepreneurs would be low and such that they do not want to borrow from the formal lending market. However, about half of the respondents in the sample did not request a loan and only 10% did so because they did not need to borrow. More particularly, only 5% of the entrepreneurs with small firms indicated that they did not need to borrow. Other entrepreneurs invoked collateral or interest rates being too high. Moreover, Banerjee and Duflo [2005] argue that the average product of capital was of the order of 27% in Uganda in the late 1990s as compared to 9% in the U.S., i.e. on average the marginal product of capital across firms was higher in Uganda than in the U.S. Based on these evidence, we rule out the possibility that the higher capital-labor ratio for small firms are the result of entrepreneurs piling up capital either for collateral or lending purposes.

Technology

Examining the share of labor and capital into production for each size bins could also shed light on the behavior of Ugandan firms. In estimating these shares, OLS results will be biased if the explanatory variables are correlated with unobserved firm specific effects. There is thus a clear endogeneity issue due to the fact that output, capital and labor all respond to productivity shocks at the same time. We thus infer the share of labor and capital into production by using a two-step estimation procedure where the set of equations is:

\[
\ln F = \alpha_0 + \alpha_1 \ln l_t + \alpha_2 \ln k_t + \delta + \epsilon
\]  
(3.12)

\[
\ln l_t = \beta_0 + \beta_1 l_{t-s} + \beta_2 k_{t-s} + \delta + \nu
\]  
(3.13)
\[ \ln k_t = \gamma_0 + \gamma_1 l_{t-s} + \gamma_2 k_{t-s} + \delta + \mu \]  

(3.14)

where \( F \) is output, \( k_t \) and \( l_t \) are instrumented variables obtained from vectors of lags of capital, \( k_{t-s} \), and of labor, \( l_{t-s} \), the vector \( \delta \) contains sector dummies to account for compositional (fixed) effects across sectors, and \( \epsilon, \nu \) and \( \mu \) are error terms.

We use sales in 1997 to proxy for output. For labor, we have that:

\[ L = \sum_i e_i h_i = i(eh) \]  

(3.15)

where \( e_i \) is the i-th worker’s efficiency, \( h_i \) is the i-th worker’s number of hours worked, \( e \) is average efficiency per worker in a specific firm and \( h \) is average number of hours worked per worker in a specific firm. We assume the same average number of hours per worker for all firms. Moreover, multiplying \( i \) by \( e \) gives us total efficiency in a specific firm: we approximate this using annual wage bills. As for capital, we back it out using the methodology described above. The results presented in Table 3.18 are based on a regression of sales on capital and wage bills in 1997, using lags of capital and wagebills and sector dummies as instruments.\(^{39}\)

Examining shares of labor and capital into production for each size bins, we note that smaller firms are more intensive in labor than other firms, whereas medium firms seem more intensive in capital than all other firms. This could explain the different cost of servicing capital observed for these firms. Small firms exhibit a low marginal product of capital due to a small share of capital into their production function. This seems typical of a cottage economy where small entrepreneurs rely mostly on labor and borrow for small amounts of physical capital at low interest rates from family and relatives. Medium firms exhibit a higher marginal product of capital due to the greater share of capital into their production function. This in turn would explain the higher cost of servicing existing capital for these firms which rely on credit markets just as larger firms.

If it was the case that some of these medium firms are hiding with small firms,\(^{39}\)

\(^{39}\)The estimates of the shares are quite sensible to the variables used in either stages of the procedure. We have examined regressions using inventory instead of sales for output, using investment instead of capital and using the number of employees instead of wage bills for labor.
these hiding firms would have more incentives to accumulate capital than their smaller counterparts due to the nature of their production function. This could explain the difference in capital intensity across small and medium firms. We have analyze whether firms at the size-threshold (30 employees) have a share of capital into production similar to that of medium firms by further disaggregating our size-bins. However, this analysis does not give interesting results as the sample of firms is too small to be statistically meaningful.
Table 3.18: Capital and Labor Shares

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>small</th>
<th>medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\gamma^b$</td>
<td>$\eta^c$</td>
<td>$\gamma^b$</td>
<td>$\eta^c$</td>
</tr>
<tr>
<td>All</td>
<td>.43***</td>
<td>.68***</td>
<td>.14</td>
<td>.81***</td>
</tr>
<tr>
<td></td>
<td>(.09)</td>
<td>(.09)</td>
<td>(.14)</td>
<td>(.15)</td>
</tr>
<tr>
<td></td>
<td>{103}</td>
<td>{42}</td>
<td>{22}</td>
<td>{39}</td>
</tr>
<tr>
<td>Manufac</td>
<td>.36***</td>
<td>.79***</td>
<td>.17</td>
<td>.81***</td>
</tr>
<tr>
<td>-turing</td>
<td>(.10)</td>
<td>(.10)</td>
<td>(.15)</td>
<td>(.19)</td>
</tr>
<tr>
<td></td>
<td>{78}</td>
<td>{34}</td>
<td>{16}</td>
<td>{28}</td>
</tr>
<tr>
<td>Kampala</td>
<td>.55***</td>
<td>.59***</td>
<td>.23</td>
<td>.58</td>
</tr>
<tr>
<td></td>
<td>(.12)</td>
<td>(.12)</td>
<td>(.24)</td>
<td>(.34)</td>
</tr>
<tr>
<td></td>
<td>{53}</td>
<td>{16}</td>
<td>{14}</td>
<td>{23}</td>
</tr>
</tbody>
</table>

- * significant at 10%, ** significant at 5%, *** significant at 1%; std. err. in parenthesis; number of firms in curly brackets;
- Share of capital is given by $\gamma$ where capital is estimated based on resale value of machinery, equipment and other assets and a depreciation rate of 0.07;
- Share of labor is given by $\eta$ where labor is proxied by wage bills divided by average wage per size-bin;
- All variables are expressed in logs which implies that the input coefficients are input elasticities;
- All regressions account for sector fixed effects.
Compositional Effect

Another potential explanation for the observed capital-labor ratios in Uganda is that a compositional effect across sectors may drive the results for the capital-labor ratio. However, reducing the sample to include only manufacturing firms (188 observations) does not affect the results significantly.\textsuperscript{40}

Extortion

Extortion is a potential explanation for observing the missing middle in LDCS. A tax official could create red tape by threatening or even forcing an entrepreneur to shut down his firm until an informal agreement involving the payment of a bribe has been reached. However, given the small amount paid in bribes compared to the amounts of taxes evaded extortion does not seem to be a factor which has as strong an effect as the rise in the Audit intensity.\textsuperscript{41}

\textsuperscript{40}Results for the manufacturing sector only are available upon request.

\textsuperscript{41}However, the effect might be complementary as we observed relatively higher bribes for firms above the size threshold.
Chapter 4

Theoretical Model

4.1 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.\footnote{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’s (1978) span-of-control model provides theoretical grounds for such an analysis.}

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 LDCs, e.g. high default probabilities, usury lending rates from private money lender, etc.\footnote{See Chapter 3 for a description of the factor markets in Uganda.} We define the degree of inefficiency of credit or capital market as the extent to which
the actual output is driven away from a First-Best level of output by distortions or impediments such as high collateral and usury lending rates. As a consequence, entrepreneurs can only finance their investment through cash flow and invest in their own firm. This assumption is made for two reasons. First, our focus is on small and medium firms and, as observed in Chapter 3, about half of the sample of Ugandan firms, mainly small and some medium firms, has no access to credit because the collateral demanded by lending institutions is too high. Second, coding a dual market for capital for the exercise of Chapter 5 proves to be very difficult. Thus we present two extreme situations: 1) the benchmark where no firms can rent capital and 2) a situation where capital market are fully functioning. It is important to note that the Ugandan reality with its dual capital market lies somewhere in between the bounds defined by these two extreme situations.

Entrepreneurs pay part of their taxes. Those who meet with a tax auditor bargain over a bribe to keep part of the surplus from evasion.

Entrepreneurs who face a series of bad shocks and have completely run down their capital stock exit at the end of the period. Every period a fraction of firms exit and are replaced by new entrants in the next period.

Since we are only concerned with stationary equilibria, we assume that entrepreneurs are rational and that the prices they anticipate are the prices that realize in equilibrium. Entrepreneurs are heterogeneous across two dimensions. Upon entry in the market each entrepreneur is endowed with a productivity parameter, \( z_t \in [z^{low}, z^{up}] \), drawn from an initial distribution function \( \nu \). Productivity shocks have a transition function given by \( g(z_{t+1}, z_t) \) that we keep as general as possible until Chapter 5. Capital stocks \( k_t \in (0, k^{up}) \) also differ across entrepreneurs because of previous investment history. For the sake of simplicity we do not identify entrepreneurs with a superscript \( i \) assuming it is understood that variables and equations that refer to entrepreneurs are specific to a particular individual except

---

3 Based on Buera et al. (2009), we will develop a setup where entrepreneurs can either use their capital or rent it from the capital market. In this version of the model, the level of financial development of a country will dictate how much capital entrepreneurs can borrow based on an incentive compatibility constraint. We have undertaken this path of research but the results will be presented in future work.

4 This assumption is common in this literature. See for example section 4 in Lucas Jr and Prescott [1971], Hopenhayn and Rogerson [1993] and Restuccia and Rogerson [2008].
for the initial distribution of shocks $\nu$ and the transition function $g$ which are the same for all entrepreneurs.

The price of output is normalized to one and the wage is $w_t$. In each time period $t$, an entrepreneur chooses to allocate his profits, $\pi_t$, between consumption, $c_t$, and capital investment, $k_{t+1}$.

The objective function of a risk neutral entrepreneur is:

$$\max_{c_t, k_{t+1}} E_0 \sum_{t=0}^{\infty} (\phi \beta)^t c_t$$

s.t.

$$c_t + k_{t+1} = \pi_t$$

$$c_t \geq 0$$

$$k_{t+1} > 0$$

where $k_0$ is given, $\phi$ is an exogenous probability of survival, $\beta$ is the discount rate and profits are given by:

$$\pi_t = F(z_t, k_t, l_t) - w_t l_t + (1 - \delta) k_t - RC(l_t) - c_F$$

where $\delta$ is the depreciation rate of capital and $c_F$ is a fixed cost of carrying business.\(^5\) The production function exhibits decreasing returns to scale in labor and capital:

$$F(z, k, l) = zk^\gamma l^\eta$$

where $\gamma + \eta < 1$. This choice of functional form allows pinning down the size of a firm. $RC(l_t)$ are regulation costs which vary with size. These costs are given by the following general form:

$$RC(l_t) = p_a(l_t) \{ qB_t + (1 - q)(T_t^e + FA) \} + T_t^p$$

\(^5\)Without this fixed cost, some firms stay idle and do not have any reason to exit (Hopenhayn and Rogerson, 1993).
where \( p_a(l_t) \) is the probability of being audited by a tax official which varies with firm size in a step-wise fashion:

\[
p_a = \begin{cases} 
  p^s_a & \text{if small} \\
  p^m_a & \text{if medium} \\
  p^l_a & \text{if large}
\end{cases}
\]

and \( p^s_a < p^m_a < p^l_a \). There is a probability \( q \) that the tax official is corrupt and accepts a bribe \( B_t \). With probability \( (1 - q) \) the tax official is honest and forces the entrepreneur to pay a fine \( F_A \) on top of the amount of taxes evaded \( T^e_t \). \( T^p_t \) is the amount of taxes initially paid by an entrepreneur. Note that \( T^e_t + T^p_t = T^0_t \) where \( T^0_t \) is the amount of taxes officially owed by a firm. In the problem at hand, we assume that taxes are on revenues and write \( T^p_t = \tau_p F(z_t, k_t, l_t) \) and \( T^e_t = (\tau_0 - \tau_p) F(z_t, k_t, l_t) \) where \( \tau_0 \) and \( \tau_p \) are official and effective tax rates, respectively. Although we observe some variation in terms of taxes paid across firms in the data (see the section on summary statistics), we assume, for the sake of simplicity, in this chapter and the following chapter that all entrepreneurs face the same official and effective tax rates.\(^6\)

Note that in the rest of the analysis we set \( F_A = 0 \) and \( q = 1 \). 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 assume that the poor performance of the audits is due to the fact that most tax officials are potentially corruptible. We thus set in the benchmark \( q = 1 \). The report also states that compliance to fill tax reports was low. We would thus set the fine paid by entrepreneurs, \( F_A \), to zero if it was included in Equation 4.7. In what follows, Equation 4.7 thus reduces to:\(^7\)

\[
RC(l_t) = p_a(l_t)B_t + T^p_t
\]

\(^6\)In the analytical example below, we depict the situation even more starkly, allowing small firms to evade completely while larger firms pay part of their liabilities and a bribe.

\(^7\)Note that in the second policy experiment, we set \( q = 0 \) and Equation 4.7 takes the following form:

\[
RC(l_t) = p_a(l_t)T^e_t + T^p_t
\]
Given his current shock \( z_t \) and current level of capital stock \( k_t \) an entrepreneur chooses to operate according to three different status \( s = \{iu, ic, f\} \). These status are defined by the steps in the audit probability occurring at the exogenous size-threshold \( l^{ic} \). There is a proportion \( iu \) of informal-unconstrained entrepreneurs who face low productivity shocks and hire an 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 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 \( k_t \). There is a proportion \( f \) of formal entrepreneurs who 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 \( l^f = l(z_t^f, k_t) \).

**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:

\[
\sum_{t=1}^{\infty} \beta^t [u(c_t) - d(n_t)]
\]

\[\text{s.t.} \quad c_t \leq w_t n_t\]

where \( c_t \) is consumption and \( n_t \) is labor supply which is either zero or one. Following Hopenhayn and Rogerson [1993], 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 microcredit. Workers can then be jointly represented through one representative agent.

\[\text{In the problem at hand, we look at cases where } l^{ic} < l^f (\text{otherwise all firms would be formal}).\]
with preferences given by:

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

(4.10)

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

**Government and Tax Officials**

There is a small measure (less than 1) of risk neutral tax officials in this economy, given by $b$, 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 comprise enforcement costs, i.e. wages to its tax auditors, $bw_b$ while revenues come from taxes collected, $T_{t}^{evt} = T_{t}^{p} + p_a(l_t)(1 - q)T_{t}^{e}$, fines collected from entrepreneurs caught evading, $p_a(l_t)(1 - q)FA$, and fines to corrupt tax officials, $bq\psi A$. The governmental budget constraint is:

$$bw_b \leq T_{t}^{evt} + p_a(l_t)(1 - q)FA + bq\psi A$$

(4.11)

### 4.2 Stationary Equilibrium

**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:

---

9In our model, the government levies taxes which are not redistributed. These taxes can be considered as another form of rent captured by the elite of the country.
subject to the constraints 4.2, 4.3 and 4.4. The next proposition states that a solution to each of these fully conditional value functions exists.

**Proposition 1:** (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 $\varphi \beta$ and where $T$ is defined below; (iii) $l^*$ and $I^*$ are the optimal labor and investment policy, respectively.

**Proof.**

We prove this proposition by first verifying the following assumptions which are drawn from Stokey et al. [1989] (SL, in what follows). These assumptions are needed to prove part (i) of proposition 1, i.e., the properties of the fully conditional value functions, namely concavity, continuity, differentiability and monotonicity. For part (ii) of proposition 1, we then verify the modified Blackwell’s conditions to check the applicability of the Contraction Mapping Theorem. Finally, in part (iii), we state three assumptions needed to ensure that policy functions for labor and investment are optimal.

**Part 1i) Concavity, continuity, differentiability and monotonicity**

We first state the nine assumptions needed to ensure that the fully conditional value functions are concave, continuous, differentiable and monotonically increasing.

1: Assumption 9.18, p.271 in SL

Define $K$ as the set of possible values for capital stock. Let $K \subset \mathbb{R}^l$ be a closed convex cone, with its Borel sets $\mathcal{K}$ and $\sigma$-algebra $\mathcal{K}$, such that $(K, \mathcal{K})$ is a measurable space.
2: Assumption 9.5, p.271 in SL

Define $\mathcal{Z}$ as the set of possible values for exogenous shock $z$ with $\sigma$-algebra $\mathfrak{F}$, such that $(\mathcal{Z}, \mathfrak{F})$ is a measurable space. Either:

a. $\mathcal{Z}$ is a countable set and $\mathfrak{F}$ is the $\sigma$-algebra containing all subsets of $\mathcal{Z}$; or

b. $\mathcal{Z}$ is a compact, convex subset of $\mathbb{R}^l$, with its Borel subsets $\mathfrak{F}$, and the transition function $g$ on $(\mathcal{Z}, \mathfrak{F})$ has the Feller property (it maps bounded continuous functions on $\mathcal{Z}$ into itself).

3: Assumption 9.19, p.271 in SL

Define the correspondence, $\Gamma_s : K \times \mathcal{Z} \rightarrow K$, as the set of feasible values for next period’s endogenous state variable given the current state, that is:

$$\Gamma_s(k, z) = \{k' : 0 \leq k' \leq \lambda^s F(z, k, l) - wl + (1 - \delta)k - c_f - c\}$$

where $\lambda(l) = 1 - \tau_p - p_o(l)(1 - q\theta)(\tau_0 - \tau_p)$

$\Gamma_s$ is nonempty as making no investment is always an option for a firm. The fact that $K$ is closed and bounded implies that $\Gamma_s$ is compact-valued. Finally, since $F(z, k, l)$ is continuous so is $\Gamma_s$, $\forall s \in \{iu, ic, f\}$.

Finally, define $SS$ as the set of possible states for the system with $\sigma$-algebra $\mathfrak{D}$, such that $(SS, \mathfrak{D}) = (K \times \mathcal{Z}, K \times \mathfrak{F})$ is a measurable space.

4: Assumption 9.20, p. 271 in SL

Define the graph of $\Gamma_s$ as $U_s = \{(k, k', z) \in K \times K \times \mathcal{Z} : k' \in \Gamma_s\}$ and let $Q_s : U_s \rightarrow \mathbb{R}$ be the return function defined by the objective function of a firm depending on its status $s$.

Since $F(z, k, l)$ and $\Gamma_s$ are continuous, $Q_s(z, k, k')$ is also continuous, $\forall s \in \{iu, ic, f\}$. Now, $F(z, k, l)$ is homogenous of degree one or less, $\forall s \in \{iu, ic, f\}$, hence $Q_s(z, k, k')$ is homogenous of degree one or less, for $s = \{iu, ic, f\}$.\(^{10}\)

\(^{10}\)This implies we need to verify conditions for unbounded returns in theorem 9.12, p.274:

For $0 < \xi < \infty$, $|Q(z, k', l)| \leq \xi (\|k\| + \|k'\|)$ and $0 < \varphi \beta < 1$. 

78
5: Assumption 9.8, p.264 in SL

Given \((k', z)\), \(Q^s(z, \cdot, k') : U^s_{k'} \rightarrow \mathbb{R}\) is increasing in its first argument since \(F(z, k, l)\) is increasing in \(k\).

6: Assumption 9.9, p.264 in SL

Due to the fact that \(\bar{F}^s(k, z) = \lambda^s F(z, k, l^*) - w^*\) is monotonically increasing in \(k\) (where * indicates that labor was chosen optimally), we have that, if \(k \leq \hat{k}\) and \(c^s \leq \hat{c}^s\) then:

\[
\bar{F}^s(k, z) - c^s + (1 - \delta)k \leq \bar{F}^s(\hat{k}, z) - \hat{c}^s + (1 - \delta)\hat{k}
\]

which implies that \(\Gamma^s(k) \subseteq \Gamma^s(\hat{k})\) and \(\Gamma^s\) is also increasing.

7: Assumption 9.10, p.265 in SL

For a given \(z\), \(Q^s(z, \cdot, \cdot) : U^s_{k} \rightarrow \mathbb{R}\) is concave in the pair \((k, k')\). Define \(\bar{F}^s(k, z)\) as above in point 6. Consider two pairs \((k_1, k'_1) \neq (k_2, k'_2)\). Let \(k_\alpha = \alpha k_1 + (1 - \alpha) k_2\) and \(k'_\alpha = \alpha k'_1 + (1 - \alpha) k'_2\). Note that from point 8, below, we know that \(\Gamma^s\) is convex and \(k'_\alpha\) is thus feasible. We thus have:

\[
Q^s(z, k_\alpha, k'_\alpha) = F^s(k_\alpha, z) - k'_\alpha + (1 - \delta)k_\alpha
\]

\[
= F^s(\alpha k_1 + (1 - \alpha) k_2, z) - \alpha k'_1 + (1 - \alpha) k'_2 + (1 - \delta) [\alpha k_1 + (1 - \alpha) k_2]
\]

\[
\geq F^s(\alpha k_1, \alpha z) + F^s((1 - \alpha) k_2, (1 - \alpha) z) - \alpha k'_1 + (1 - \alpha) k'_2 + (1 - \delta) [\alpha k_1 + (1 - \alpha) k_2]
\]

\[
= \alpha F^s(k_1, z) + (1 - \alpha) F^s(k_2, z) - \alpha k'_1 + (1 - \alpha) k'_2 + (1 - \delta) \alpha k_1
\]

\[
+ (1 - \delta)(1 - \alpha) k_2
\]

\[
= \alpha Q^s(z, k_1, k'_1) + (1 - \alpha) Q^s(z, k_2, k'_2)
\]
where we have used the identity \( z = \alpha z + (1 - \alpha)z \), and where the inequality stems from the concavity of \( \tilde{F}^s(k, z) \).

8: Assumption 9.11, p.265 in SL

Since \( K \) is convex, then \( \Gamma^s \) is also convex, that is, if \( k' \in \Gamma^s(k) \) and \( \hat{k}' \in \Gamma^s(k) \) then:

\[
\theta k' + (1 - \theta)\hat{k}' \in \Gamma^s(\theta k' + (1 - \theta)\hat{k}')
\]

9: Assumption 9.12, p. 266 in SL

Since \( F^s(z, k, l) \) is continuously differentiable in the interior of \( U^s \), so is \( Q^s(z, k, k') \), \( \forall s \in \{iu, ic, f\} \).

These nine assumptions insure that lemmas and theorems 9.5 through 9.12 in Stokey et al. [1989] apply to our setup. To see this, first define the policy correspondence \( H^s : K \times \hat{Z} \rightarrow K \) by:

\[
H^s(k, z) = \{ k' \in \Gamma^s(k, z) : V^s(k, z) = Q^s(z, k, k') + \phi \beta V^s(k', z') \}
\]

(4.13)

Define also the space of bounded and continuous functions \( V^s \) as \( C^s(S) : \{ V^s : K \times \hat{Z} \rightarrow \mathbb{R} \} \) with norm defined as (\( C^s(S) \) is a Banach space):

\[
\|V^s\| = \sup_{z \in Z} \sup_{k \in K |k| \leq 1} |V^s(k, z)| < \infty
\]

And define the operator \( T \) as follows:

\[
T(V^s)(k, z) = \max_{k' \in \Gamma^s(k)} \left\{ Q^s(z, k, k') + \phi \beta V^s(k', z') \right\}
\]

(4.14)

Assuming that \( K, C^s(S), \Gamma^s, Q^s \) and \( \phi \beta \) satisfy points 1 through 9 the equivalent to Lemma 9.5, p.261 in SL holds. In other words, if \( Q^s : K \times \hat{Z} \rightarrow \mathbb{R} \) is continuous
and bounded, quasi-concave in its first $\ell$ arguments, and is an element of $C(S)$, then

$$(MQ^s)(k', z) = \int Q^s(z', k', k'')g(z, dz')$$

is also continuous and bounded, quasi-concave in its first $\ell$ arguments, and in $C(S)$, that is $M : C(S) \rightarrow C(S)$.

Hence, integration preserves the required properties of the integrand (boundedness, continuity, monotonocity, and concavity).

Lemma 9.5 also implies that most theorems from the non-stochastic part have equivalent analogues.

**Theorem 9.6 (p.263, SL)** holds, that is,

i) $T$ maps $C(S)$ into itself, $T : C(S) \rightarrow C(S)$

ii) $T$ has a unique fixed point $V^s \in C(S)$;

iii) For all $V^0_0 \in C(S)$, $\|TV^0_0 - V^s\| \leq (\varphi \beta)^n \|V^0_0 - V^s\|$, $n = 0, 1, 2, \ldots$

iv) $H^s(k, z)$ is non-empty, compact valued and upper hemi-continuous.

**Theorem 9.7 (p.264, SL)** holds:

If $V^s$ is the unique solution to the firm’s value function, then $V^s$ is strictly increasing.

**Theorem 9.8 (p.265, SL)** holds:

Let $V^s$ satisfy the firm’s value function and let $H^s$ satisfy Equation 4.13, then $V^s$ is strictly concave and $H^s$ is a continuous single-valued function.

**Theorem 9.9 (p.265, SL)** holds:

Let $V^s$ satisfy the firm’s value function and let $H^s$ satisfy Equation 4.13 and let $\{V^s_n, h^s_n\}$ be defined by:

$V^{s}_{n+1} = TV^s_n$, $n = 0, 1, 2, \ldots$

$h^s_n(k, z) = \arg \max_{k' \in \Gamma^s(k)} \{Q^s(z, k, k') + \varphi \beta V^s(k', z')\}$, $n = 0, 1, 2, \ldots$
then $h_n \to h$ pointwise and since $K$ and $\bar{Z}$ are both compact the convergence is uniform.

**Theorem 9.10 (p.265, SL)** holds:

The functions $V^s(\cdot, z_0)$ are differentiable in $k$ at $k_0$ if $k_0 \in \text{int}(K)$ and $h^s(k_0, z_0) \in \text{int}(\Gamma^s(k_0, z_0))$.

We have shown that the stochastic functions $V^s$ are strictly concave, continuous, differentiable and increasing and we are done with point i) of proposition 1.

**Part ii) Blackwell’s conditions and the Contraction Mapping Theorem**

We now verify Blackwell’s conditions to ensure the applicability of the Contraction Mapping theorem. Indeed, we show that $T$, as defined by Equation 4.14, is a contraction mapping by examining Blackwell’s conditions in turn.

1. **Monotonicity:**

For a given $z$, let $V^s_1(k, z) \geq V^s_2(k, z), \forall s$, imply $\max V^s_1(k, z) \geq \max V^s_2(k, z), \forall s$, and let $V^s_1 = \max \left[V^{iu}_1, V^{ic}_1, V^f_1\right]$ and similarly for $V^s_2$. Hence,

$$V^s_1(k, z) = T(V^s_1)(k, z)$$

$$= \max_{k' \in \Gamma^s(k)} \left\{ Q^s(z, k, k') + \phi \beta V^s_1(k', z') \right\}$$

$$= \max_{k' \in \Gamma^s(k)} \left\{ Q^s(z, k, k') + \phi \beta \int_{j'} \max \left[V^{iu}_1, V^{ic}_1, V^f_1\right] d j' \right\}$$

$$\geq \max_{k' \in \Gamma^s(k)} \left\{ Q^s(z, k, k') + \phi \beta \int_{j'} \max \left[V^{iu}_2, V^{ic}_2, V^f_2\right] d j' \right\}$$
\[
\max_{k' \in \Gamma^s(k)} \{ Q^s(z,k,k') + \phi \beta V^s_2(k',z') \}
\]
\[
= T(V^s_2)(k,z) = V^s_2(k,z)
\]

2. Discounting:

\[
T(V^s + a)(k,z) = Q^s(z,k,k') + \phi \beta [V^s(k',z') + a]
\]
\[
= Q^s(z,k,k') + \phi \beta \int_{j'} \max \{ V^{iu} + a, V^{ic} + a, V^f + a \} \, dj'
\]
\[
= Q^s(z,k,k') + \phi \beta \int_{j'} \max \{ V^{iu} + a, V^{ic}, V^f \} \, dj' + \phi \beta \int_{j'} a \, dj'
\]
\[
= Q^s(z,k,k') + \phi \beta V^s(k',z') + \phi \beta a
\]

Hence, \( T \) is a contraction mapping and it satisfies the conditions required to apply the analogue of the contraction mapping theorem to the stochastic case. The contraction mapping theorem guarantees that \( V^s \) has a unique fixed point and that \( T \) preserves the properties of \( V^s \), that is, monotonicity, continuity, concavity and differentiability.

We are done with point ii) of proposition 1.

**Part iiii) The policy functions are optimal.**

All we are left to show is that \( l^* \) and \( I^* \) are optimal policies. To see this we need to verify three more assumptions.
1: Assumption 9.1, p.243 SL

$\Gamma^s$ is non-empty (shown above) and has a measurable selection, that is, there exists measurable function $Q^s : S \rightarrow K$ such that $Q^s(z,k,k') \in \Gamma^s, \forall (k,z) \in SS$.

2: Assumption 9.2, p.244, SL

$Q^s : U \rightarrow \mathbb{R}$ is $U$-measurable, and either (a) or (b) holds:

(a) $Q^s \geq 0$ or $Q^s \leq 0$.

(b) For each $(k_0,z_0) \in SS$ and each plan $\pi \in \Pi(k,z), Q^f_r[z_t,\pi_{t-1}(z'^{-1}),\pi(t')]$ is $\mu^t(z_0,\cdot) - integrable, t = 1,2,...$, where $\mu^t$ is a probability measure such that $\mu^t(z_0,\Delta) : \mathcal{F}^t \rightarrow [0,1]$ and the limit

$$Q^s[z_0,k_0,\pi_0] + \lim_{n \rightarrow \infty} \sum_{t=1}^{\infty} \int_{\mathcal{F}^t} (\varphi \beta)^t Q^f_r[z_t,\pi_{t-1}(z'^{-1}),\pi(t')] \mu^t(z_0, dz')$$

exists.

3: Assumption 9.3, p.248, SL

If $Q^s$ takes on both signs, there is a collection of nonnegative measurable functions $L_t : S \rightarrow \mathbb{R}_+, t = 0,1,...$, such that for all plans:

a. $|Q^f(z_0,k_0,\pi_0)| \leq L_0(k_0,z_0)$

b. $|Q^f[z_t,\pi_{t-1}(z'^{-1}),\pi(t')]| \leq L_t(k_0,z_0), \forall z' \in \mathcal{Z}, t = 1,2,...$

c. $\sum_{t=1}^{\infty} (\varphi \beta)^t L_t(k_0,z_0) < \infty$

If the three assumptions enumerated above hold then theorem 9.2 (the equivalent for unbounded returns is theorem 9.12) and 9.4 hold in Stokey et al. [1989]. This implies that $V^s = V^{**}$ and any plan $(l^* and I^*)$ generated by $H^s$ attains the supremum. The partial converse is given by theorem 9.4 which states that the plans $(l^* and I^*)$ are optimal if they have been generated almost everywhere by $H^s$. 

84
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:

\[ V_s(k, z) = \max_k \{ c' + \varphi \beta CV(k', z') \} \]  

(4.15)

where the continuation value is given by:

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

(4.16)

A brief remark about existence of a solution for the semi-conditional functions is necessary. The jumps 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 each of the semi-conditional functions 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:

\[ ex(k, z) = \begin{cases} 0 & \text{if } CV(k', z') > 0 \\ 1 & \text{otherwise} \end{cases} \]  

(4.17)

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

\[ V(k, z) = \max_{s \in \{ex, iu, ic, f\}} V_s(k, z) \]  

(4.18)

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

\[ 11 \text{In the quantitative model we iterate over the semi-conditionals until joint convergence.} \]
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 his books. Entrepreneurs bargaining power is denoted by $\theta$ 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.$^{12}$

The outside option for an entrepreneur is to repay what they have evaded plus a fine.$^{13}$

$$V^\tau(k, z) = F(z, k, I) - wI - k' + (1 - \delta)k - c_f - T_0 - FA + \varphi \beta CV(k', z')$$ (4.19)

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

$$V^B(k, z) = F(z, k, I) - wI - k' + (1 - \delta)k - c_f - B - T_p + \varphi \beta CV(k', z')$$ (4.20)

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

The Nash-bargaining problem is:

$$B = \arg \max [V^B(k, z) - V^\tau(k, z)]^\theta [B - \psi A]^{(1-\theta)}$$ (4.21)

$^{12}$This simplifies our analysis because continuation values cancel out when solving for the equilibrium bribe.

$^{13}$In what follows, we drop superscripts for the sake of brevity
Solving for the equilibrium bribe yields:

\[ B = (1 - \theta)(\tau_0 - \tau_p)z^\gamma l^{\eta} + \theta \psi A \]  \hspace{1cm} (4.22)

**Claim 1:** The equilibrium bribe exhibits the following properties: \( B_l > 0 \), \( B_{ll} < 0 \), \( B_k > 0 \) and \( B_{kk} < 0 \).

**Proof.**

Effects of \( k \) on \( B \)

\[ \frac{dB}{dk} = (1 - \theta)(\tau_0 - \tau_p)\gamma z^\gamma l^{\eta - 1} > 0 \]  \hspace{1cm} (4.23)

and

\[ \frac{d^2B}{dk^2} = (1 - \theta)(\tau_0 - \tau_p)\gamma(\gamma - 1)z^\gamma l^{\eta - 2} < 0 \]  \hspace{1cm} (4.24)

Effects of \( l \) on \( B \)

\[ \frac{dB}{dl} = (1 - \theta)(\tau_0 - \tau_p)\eta z^\gamma l^{\eta - 1} > 0 \]  \hspace{1cm} (4.25)

and

\[ \frac{d^2B}{dk^2} = (1 - \theta)(\tau_0 - \tau_p)\eta(\eta - 1)z^\gamma l^{\eta - 2} < 0 \]  \hspace{1cm} (4.26)

The representative worker’s problem is a static optimization problem in a stationary equilibrium:

\[ \max u(c) - DN \]  \hspace{1cm} (4.27)

subject to (4.9). We denote the solution to this problem as \( L^* \).

---

\(^{14}\)If we were using the general expression for the regulation costs, the equilibrium bribe would take into account the fine faced by entrepreneurs if caught evading and would be:

\[ B = (1 - \theta)\left[(\tau_0 - \tau_p)z^\gamma l^{\eta} + FA\right] + \theta \psi A \]
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 Jr and Prescott [1971]; Hopenhayn [1992a]). Let \( \mu(k, z) \) be such a probability measure, summarizing the state of the industry in a given period of time.\(^\text{15} \) Let the mass of entrants at the beginning of next period be \( e_n \). 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:

\[
\mu'(k', z') = \int \xi(k_0, z') \nu dz + \varphi \int (1 - ex(k, z))G(k', z'; k, z)\mu(k, z)d(k \times z) \tag{4.28}
\]

where the measure of exiting firms is given by:

\[
\xi(k, z) = \int [\varphi ex(k, z) + (1 - \varphi)]\mu(k, z)d(k \times z) \tag{4.29}
\]

Equation 4.28 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 future state \((k', z')\).

Aggregate inputs are obtained by adding individual inputs across all firms in the economy. Aggregate employment is given by:

\[
L^{agg} = \int l(k, z)\mu(k, z)d(k \times z) \tag{4.30}
\]

and aggregate capital by:

\[
K^{agg} = \int k(z)\mu(k, z)d(k \times z) \tag{4.31}
\]

Similarly, aggregate production is then given by:

\[
F^{agg} = \int F(k, z)\mu(k, z)d(k \times z) \tag{4.32}
\]

\(^{15}\) As is standard in the literature on firms dynamics we define an industry as a continuum of firms which produce a homogeneous good [Hopenhayn 1992a].
Aggregate regulation costs are:

\[
RC^{\text{agg}} = \int \{ p_a(l) [qB(k,z) + (1-q)(T_e(k,z) + FA)] + T_p(k,z) \} \mu(k,z)d(k \times z)
\]  

(4.33)

and the aggregate bribe is:

\[
B^{\text{agg}} = \int B(k,z)\mu(k,z)d(k \times z)
\]  

(4.34)

Entrepreneurs consumption is given by:

\[
C^e = \int c(k,z)\mu(k,z)d(k \times z)
\]  

(4.35)

and aggregate consumption is given by the consumption of the workers, the bureaucrats and entrepreneurs:\[16\]

\[
C^{\text{agg}} = C^{\text{wks}} + C^b + C^e
\]  

(4.36)

Aggregate feasibility is then given by:

\[
C^{\text{agg}} + K^{\text{agg}'} - (1-\delta)K^{\text{agg}} \leq F^{\text{agg}} - wL^{\text{agg}} + wL^{\text{agg}} - RC^{\text{agg}} + B^{\text{agg}} - c_f \int \mu(k,z)d(k \times z)
\]  

(4.37)

Equation (4.37) states that aggregate consumption plus aggregate investment cannot exceed net aggregate production across all firms, the total wage bill received by workers and the bribes received by bureaucrats:\[17\]

We now have all elements to define a stationary equilibrium for this economy.

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

---

\[16\]Workers eat their wage. Tax officials eat their bribes as we are assuming they receive negligible public wages.

\[17\]Based on Gang et al. [1988] and Lindauer et al. (1988), we consider that wages paid to bureaucrats are negligible. However, they can easily be incorporated in the aggregate regulation costs and then added on the left side of Equation 4.37.
1. Decision rules are optimal:
   - (a) given \( k \), there exists a \( z_{ic}^* \) such that \( V^{iua} = V^{ic} \) for each entrepreneur;
   - (b) given \( k \), there exists a \( z_{if}^* \) such that \( V^{ic} = V^f \) for each entrepreneur;
   - (c) given \( k \), there exists a \( z_{ex}^* \) such that the exit problem is satisfied for each entrepreneur;
   - (d) \( l^{ds} \) and \( k^* \) satisfy each entrepreneur’s labor and capital problems;
   - (e) \( l^s \) satisfy workers’ problem;

2. \( \mu^* \) is defined recursively by 4.28 and \( \mu^{*'} = \mu^* \);

3. The government budget balances.

4. Prices are market clearing: \( L^{ss} = L^{ds} \).

**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:

\[
l = \left[ \lambda(l) \left( \frac{\eta}{w} \right) z k^\gamma \right]^{\frac{1}{1-\eta}} \quad (4.38)
\]

where \( \lambda(l) = 1 - \tau_p - p_o(l)(1 - q \theta)(\tau_0 - \tau_p) \). This is standard except for \( \lambda \) which is the burden from taxation.\(^{18}\)

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 the current capital stock.\(^{19}\)

**Claim 2:** The Euler Equation is

\[
\frac{1}{\varphi \beta} (1 - \delta) = \int_{z^{ex}} \gamma' (k') \gamma^{-1} (l') \varphi - R c_{ik}^{iu} \, dg(z', z) \quad (4.39)
\]

\(^{18}\)Given the functional form for the audit probability (i.e. step-function), the first order condition for labor can be solved in closed hand form.

\(^{19}\)This means that the lower and upper bounds of the integrals in Equation 4.39 vary with capital stock.
Proof.

Taking the first-order condition of the value function with respect to $k'$ gives:

$$\frac{dV}{dk'} : -1 + \varphi \beta \frac{d}{dk'} CV(k', z') = 0$$

where:

$$\frac{d}{dk'} CV(k', z') = \frac{d}{dk'} \max \left[ 0, V^{iu}(k', z'), V^{ic}(k', z'), V^{f}(k', z') \right]$$

$$= \frac{d}{dk'} \left[ \int_{z_{ic}}^{z_{ex}} V^{iu} d' + \int_{z_{ic}}^{z_{up}} V^{ic} d' + \int_{z_{ic}}^{z_{up}} V^{f} d' \right]$$

In obtaining the euler equation, we keep the integrals in the envelope condition and make use of Leibniz’s rule. Let’s evaluate in turn all three terms of the continuation value:

**Term 1:**

$$\frac{d}{dk} \left[ \int_{z_{ex}}^{z_{ic}} V^{iu} d' \right] = \frac{d}{dk} V^{iu}(z_{ic}) - \frac{d}{dk} V^{iu}(z_{ex}) + \int_{z_{ex}}^{z_{ic}} \left\{ F_k + (1 - \delta) - RC^{iu}_k \right\} dz$$

**Term 2:**

$$\frac{d}{dk} \left[ \int_{z_{ic}}^{z_{up}} V^{ic} d' \right] = \frac{d}{dk} V^{ic}(z_{up}) - \frac{d}{dk} V^{ic}(z_{ic}) + \int_{z_{ic}}^{z_{up}} \left\{ F_k^{ic} + (1 - \delta) - RC^{ic}_k \right\} dz$$

**Term 3:**

$$\frac{d}{dk} \left[ \int_{z_{ic}}^{z_{up}} V^{f} d' \right] = \frac{d}{dk} V^{f}(z_{up}) - \frac{d}{dk} V^{f}(z_{ic}) + \int_{z_{ic}}^{z_{up}} \left\{ F_k^{f} + (1 - \delta) - RC^{f}_k \right\} dz$$

The support for $z$ is continuous and we know that $V^{iu} = V^{ic}$ at $z_{ic}$ and $V^{ic} = V^{f}$ at $z_{f}$ so the first term of Term 1 cancels off with the second term of Term 2, the first term of Term 2 cancels off with the second term of Term 3. The first term of Term 3 vanishes because the upper bound on the support for shocks is fixed and exogenous. Finally, we make the assumption that at $z_{ex}$, $V^{iu} = 0$. 

$\blacksquare$
4.3 First Best

In our model regulation distortions interact with firms’ investment. For the sake of comparison we examine an economy without distortions associated with taxes and where the rental market for capital is fully functioning. We call this setup the First Best. It is important to note that the setup of the First Best is an abstract tool that is useful to examine the effects of each distortion in turn. The setup of the First Best is not a realistic description of what could be achieved in Uganda, i.e. taxes need to be levied and capital markets always present some imperfections even in the developed world.

In the First Best, there is one representative household made of infinitely many entrepreneurs [Guner et al., 2008]. The household makes the investment decision and rents capital to entrepreneurs at rental rate, $r_t$. We assume furthermore that the representative household makes no labor-leisure choice and is endowed with one unit of time each period. The household problem is:

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

s.t.

$$\sum_{t=0}^{\infty} [c_t + k_{t+1} - (1 - \delta)k_t] = \sum_{t=0}^{\infty} [r_t k_t + w_t l_t + \pi_t]$$

where profits are given by the static optimization of the firm:

$$\pi_t = \max_{k_t, l_t} z_t k_t^{\eta} l_t^{1-\eta} - r_t k - w_t l - c_f$$

One thing to note from this setup is that the introduction of a market for capital allows firms to adjust their hiring of capital after observing their shock. Moreover, an entrepreneur’s source of investment is not limited to its own profits anymore.

The household first order condition with respect to capital yields the standard euler equation:

$$-1 + \varphi \beta \int z'[r' + (1 - \delta)]dz' = 0$$

We know that in equilibrium: $r = F_k$ and $w = F_l$. We can thus rewrite Equation 4.43 as:
\[
\frac{1}{\phi \beta} = \int_{z}^{z'} \left[ F_k' + (1 - \delta) \right] dz'
\] (4.44)

From the firm’s problem we obtain neat expressions for the allocation of labor and capital:

\[
I = \left[ z \left( \frac{Y}{r} \right)^{\gamma} \left( \frac{\eta}{w} \right)^{1-\gamma} \right]^{1/(1-\eta)}
\] (4.45)

\[
k = \left[ z \left( \frac{Y}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^{\eta} \right]^{1/(1-\eta)}
\] (4.46)

The optimal capital-labor ratio thus reduces to:

\[
\frac{k}{I} = \frac{\gamma w}{\eta r}
\] (4.47)

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 in the next chapter generates capital-labor ratios that we can compare to those generated in the first-best.

By comparing the benchmark economy to the First-Best we derive two important implications in the following claim.

**Claim 3:** (i) A progressive tax on revenues decreases the amount of labor employed in any given firm but relatively more in larger firms; (ii) There is a non-degenerate distribution of resources across firms within a productivity class due to each firm’s specific capital accumulation process.

**Proof.**

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

\footnote{A similar expression can be derived for the relative demand for capital.}

93
\[
\frac{l_i}{l_j} = \left[ \frac{z_i}{z_j} \right]^{\frac{1}{1-\eta}}
\]  

(4.48)

It is easy to see that a flat tax does not affect this ratio. However, the relative demand for labor in the benchmark economy is:

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

(4.49)

where \(\lambda(l) = 1 - \tau_p - p_u(l)(1 - q\theta)(\tau_0 - \tau_p)\). Abstracting from capital accumulation, we have:

\[
\frac{l_i}{l_j} = \left[ \frac{\lambda_i z_i}{\lambda_j z_j} \right]^{\frac{1}{1-\eta}}
\]  

(4.50)

A progressive tax would be distortive for the relative demand of labor. Indeed, for a given \(z_i\) and \(z_j\), if \(z_i < z_j\) and thus \(\lambda_i > \lambda_j\) then \(\left[ l_i / l_j \right]_{dist} > \left[ l_i / l_j \right]_{opt}\). Tax on revenues decrease the amount of labor employed in any given firm but more in larger firms because larger firms are facing higher tax rate. If we remove the distortions and shut down the market for capital, the relative labor demand between any two firms is:\[21\]

\[
\frac{l_i}{l_j} = \left[ \frac{z_i \left( k_i / k_j \right)}{z_j} \right]^{\frac{1}{1-\eta}}
\]  

(4.51)

Firms in this case differ along two dimensions: their productivity shocks and the level of capital they have accumulated at \(t\). Note from equation Equation 4.51 that within a productivity class \((z_i = z_j)\) resources are distributed across firms due to each firm’s specific accumulation process.

Finally, we note from Equation 4.39 that a firm decreases its investment as it grows. To see this let us compare each of the three terms from Equation 4.39 to Equation 4.44. It is easy to see that for each term of equation Equation 4.39 to be equal to its First-Best counterpart investment must be lower in the bench-
mark due to the concavity of the production function and regulation costs. Moreover, since the expected regulation costs grow with size the effect of the progressive tax schedule on investment behaviour also grows with size. In addition to this, shock-thresholds $z^{ic}$ and $z^{f}$ are decreasing with the level of capital accumulated. This means that more weight is put on the last terms of Equation 4.39 as a firm accumulates capital. This also entails an increasing negative impact from the taxation burden on investment behaviour. The interaction of investment and the tax distortion might reduce the dispersion of resources across productivity classes. The quantitative exercise in the next chapter is helpful to clarify the impact and the direction of this interaction.

### 4.4 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 conduct a comparative analysis for the parameters of interest.

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 $\psi A$ is such that all tax officials always accept a bribe, i.e. $q = 1$;
6. Only formal firms are audited in the analytical example. The probability to be audited is equal to the share of tax officials in this economy, i.e. $b$. We thus only consider a single step in the audit probability which is given by:
\[ p_a = \begin{cases} 
0 & \text{if } l \leq l^{ic} \\
b & \text{if } l > l^{ic}
\end{cases} \]

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.\[22\]

The analytical expressions for the variables of interest, namely capital and labor decision rules for all statuses, shocks’ thresholds \( z^{ic} \) and \( z^{f} \) are given by:

\[
\begin{align*}
\text{\( l^{iu} = z \left[ p_F \left( \frac{r}{r} \right)^{\gamma} \left( \frac{\eta}{w} \right)^{1-\gamma} \right]^\frac{1}{1-\gamma} \right] \quad (4.52) \)} \\
\text{\( k^{iu} = z \left[ p_F \left( \frac{r}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^{\eta} \right]^\frac{1}{1-\gamma} \right] \quad (4.53) \)} \\
\text{\( k^{ic} = \left( \frac{\gamma p_F(z)^{1-\eta} (l^{ic})^\eta}{r} \right)^\frac{1}{1-\gamma} \quad (4.54) \)} \\
\text{\( l^{f} = z \left[ p_F \left( \frac{r}{r} \right)^{\gamma} \left( \frac{\eta}{w} \right)^{1-\gamma} \right]^\frac{1}{1-\gamma} \right] \quad (4.55) \)} \\
\text{\( k^{f} = z \left[ p_F \left( \frac{r}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^{\eta} \right]^\frac{1}{1-\gamma} \right] \quad (4.56) \)}
\end{align*}
\]

To obtain \( z^{ic} \) when the production functions is DRS, we first derive the optimal capital of \( iu \)-firms at \( z^{ic} \), then substitute this in Equation 4.54 and isolate the shock threshold:

\[
\begin{align*}
\text{\( z^{ic} = l^{ic} \left\{ \frac{1}{p_F} \left[ \frac{r}{\gamma} \left( \frac{w}{\eta} \right)^{1-\gamma} \right]^\frac{1}{1-\gamma} \right\} \quad (4.57) \)}
\end{align*}
\]

\[22\]Relaxing assumptions 2, 3 or 4 requires computational methods to solve the model.
To obtain $z^f$, we use the following relationship: $V^f = V^{ic}$ and get:

$$
p_F \left( z^f \right)^{1-\gamma-\eta} \left\{ \lambda (k^f)^\gamma (l^f)^\eta - (k^{ic})^\gamma (l^{ic})^\eta \right\} - w [l^f - l^{ic}] - r [k^f - k^{ic}] = b \theta \psi A
\quad (4.58)
$$

Although, one cannot isolate $z^f$ in an obvious way from this expression, it is easy to see that $z^f$ is a function of parameters only. We have thus characterize our analytical solution for a stationary equilibrium without entry and exit. Given a specific vector of prices $\zeta = (p_F, w, r)$, the equilibrium allocations are given by 4.52, 4.53, 4.54, 4.55, 4.56, 4.57 and 4.58.

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 4:** If the above assumptions (1 to 8) are satisfied then there exists a unique stationary equilibrium associated with each price vector.

**Proof.**

Without entry and exit and when the problem of the firm is entirely static, there is a continuum of equilibria, each indexed by a different vector of prices. To see this, we proceed as follows:

First, given any $\mu$, there exists a unique aggregate input-output vector of prices satisfying $p_F^* = D(Q^*), \ r^* = r(K^*), \ w^* = w(L^*)$.

Second, from the assumptions above, we know that aggregate demand is strictly decreasing in $Q$, and supply functions are strictly increasing in their respective input. Thus, there exists a unique set of allocations $Q^* = Q^*, \ K^* = K^d$ and $L^* = L^d$ associated to any specific vector of prices.

Third, for a specific production function we show below that decision rules $z^{ic}$ and $z^f$ are strictly decreasing in $p_F$ and strictly increasing with $w$ and $r$. Thus there exists a unique set of decision rules $(z^{ic^*}, z^{f^*})$ associated with any vector of prices.

Fourth, as is shown below, the proportion $f$ is strictly decreasing in $z^f$ and so there is a unique proportion $f$ associated to each $z^{f^*}$. Similarly, proportion $iu$ is strictly increasing with $z^{ic}$ and there is also a unique proportion $iu$ associated to each $z^{ic^*}$. However, the proportion $ic$ need not arise uniquely for different price vectors. Nevertheless, since the $iu$ and $f$-proportions are unique, there is a unique
overall size-distribution of firms for this economy associated with the unique price vector we started with and we are back where we began.

In what follows we derive some comparative statics for the parameters of interest. The results of the analysis are summarized in Table 4.1.

**Table 4.1: Comparative Statics**

<table>
<thead>
<tr>
<th></th>
<th>$k^{iu}$</th>
<th>$l^{iu}$</th>
<th>$k^{ic}$</th>
<th>$k^{f}$</th>
<th>$l^{f}$</th>
<th>$z^{ic}$</th>
<th>$z^{f}$</th>
<th>$B$</th>
<th>$iu$</th>
<th>$ic$</th>
<th>$f$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_F$</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>$w$</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+/-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\theta$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+/-</td>
<td>0</td>
<td>+/-</td>
<td>+/</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$l^{ic}$</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>+/-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$b$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>$\psi A$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>+</td>
<td>+</td>
<td>0</td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>-</td>
<td>-</td>
<td>0</td>
<td>+</td>
<td>+/-</td>
<td>0</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

Note: The table should be read as the effect of an increase in the parameters in column 1 on the variables in row 1.

To compute the comparative statics found in Table 4.1 we first need to examine the effect of a change in each of the shock-thresholds $z^{ic}$ and $z^{f}$ on the proportion of the three sectors $s = \{iu, ic, f\}$.

### 4.4.1 Effect of a Change in the Shock-Thresholds

We examine the effect of a change in $z^{f}$ on the proportion of the formal sector, $f$, using Leibniz rule:

\[
\frac{df}{dz^{f}} = \frac{d}{dz^{f}} \int_{z^{f}}^{1} G(z' | z) \mu(z) dz
\]

\[
= (\frac{d(1)}{dz^{f}})G(1 | z) \mu(z) - (\frac{dz^{f}}{dz^{f}})G(z^{f} | z) \mu(z) + \int_{z^{f}}^{1} \frac{d}{dz^{f}} G(z' | z) \mu(z) dz
\]

\[
= -G(z^{f} | z) \mu(z) < 0
\]
The first and third terms after the second equality disappear as the upper bound on shocks and $G(z' | z)$ are both exogenously given.

A change in $z^f$ has the inverse effect on the ic-sector. To see this:

$$
\frac{\partial c}{\partial z^f} = \frac{d}{dz^f} \int_{z^f}^{z_{ic}} G(z' | z) \mu(z) dz
$$

$$
= (\frac{d}{dz^f})G(z^f | z) \mu(z) - (\frac{d}{dz^f})G(z^{ic} | z) \mu(z) + \int_{z^f}^{z_{ic}} \frac{d}{dz^f} G(z' | z) \mu(z) dz
$$

$$
= G(z^f | z) \mu(z) > 0
$$

The second term cancels out because $z^{ic}$ is not affected by $z^f$. The third term vanishes for a similar reason than the one mentionned above. Finally a change in $z^f$ does not affect the proportion of the iu-sector.

The formal sector is not affected by a change in $z^{ic}$. For the informal unconstrained sector, we have that:

$$
\frac{\partial u}{\partial z^{ic}} = \frac{d}{dz^{ic}} \int_0^{z^{ic}} G(z' | z) \mu(z) dz
$$

$$
= (\frac{d}{dz^{ic}})G(z^{ic} | z) - (\frac{d}{dz^{ic}})G(0 | z) + \int_0^{z^{ic}} \frac{d}{dz^{ic}} G(z' | z) \mu(z) dz
$$

$$
= G(z^{ic} | z) \mu(z) dz > 0
$$

The converse is expected for the ic-sector.

### 4.4.2 Effect of a Change in the Size-Threshold

In the context of our example the effect of a change in $l^{ic}$ on $k^{iu}$, $l^{iu}$, $k^f$ and $l^f$ is zero. So a change in the size-threshold does not affect how unconstrained firms choose their inputs. Since the equilibrium bribe is a function of the capital and labor of formal firms it is not affected by a change in $l^{ic}$. The capital of informal-constrained firms changes with $l^{ic}$:
\[
\frac{dk^ic}{dl^ic} = \left( \frac{\gamma p_F(z)^{1-\gamma-\eta}}{r} \right)^{\frac{1}{1-\gamma}} \frac{\eta}{1-\gamma} (l^{ic})^{\frac{\eta}{1-\gamma}-1} > 0 \tag{4.59}
\]

However, without any dynamic behavior associated to investment, capital-labor ratios will stay constant across firms as we have shown in Section 4.3. To examine how an increase in \(l^{ic}\) affects the proportion of firms in the formal sector, we use the product rule and proceed in two steps. First, we examine the effect of an increase in \(l^{ic}\) on \(z^f\) by taking the derivative of Equation 4.58 with respect to \(l^{ic}\):

\[
p_F (1 - \gamma - \eta) (z^f)^{-\gamma-\eta} \frac{dk^f}{dl^f} \left[ \lambda (k^f)^{\gamma} (l^f)^{\eta} - (k^{ic})^{\gamma} l^{ic} \right] +
\]

\[
p_F (z^f)^{1-\gamma-\eta} \gamma \left( \lambda (k^f)^{\gamma-1} (l^f)^{\eta} \frac{dk^f}{dl^f} - (k^{ic})^{\gamma-1} (l^{ic})^{\eta} \frac{dk^{ic}}{dl^{ic}} \right) \frac{dz^f}{dl^f} \]

\[
+ p_F (z^f)^{1-\gamma-\eta} \eta \left( \lambda (k^f)^{\gamma} (l^f)^{\eta-1} \frac{dl^f}{dl^f} \frac{dz^f}{dl^f} - (k^{ic})^{\gamma} (l^{ic})^{\eta-1} \right)
\]

\[
- w \left[ \frac{dl^f}{dl^f} \frac{dz^f}{dl^f} - 1 \right] - r \left[ \frac{dk^f}{dl^f} - \frac{dk^{ic}}{dl^{ic}} \right] \frac{dz^f}{dl^f} = 0
\]

Using the fact that in equilibrium \(w = p_F F_k(z^f, k^f, l^f)\), \(r = p_F F_k(z^f, k^f, l^f)\) and \(r = p_F F_k(z^{ic}, k^{ic}, l^{ic})\) and rearranging we get:

\[
\frac{dz^f}{dl^{ic}} = \frac{p_F (z^f)^{1-\gamma-\eta} \eta (k^{ic}(z^f))^\gamma (l^{ic})^{\eta-1} - w}{p_F (1 - \gamma - \eta)(z^f)^{-\gamma-\eta} \Lambda} > 0 \tag{4.60}
\]

where \(\Lambda = \{ \lambda k^f(z^f)^{\gamma} l^f(z^f)^{\eta} - k^{ic}(z^f)^{\gamma} (l^{ic})^{\eta} \} \)

Hence, collecting terms in the product rule we get:

\[
\frac{df}{dl^{ic}} = \frac{df}{dz^f} \frac{dz^f}{dl^{ic}} < 0 \tag{4.61}
\]

An increase in the size-threshold decreases the proportion of formal firms. If the bounds associated to medium firms fall below the size-threshold then it means

\[\text{Note however that this is not true for } p_F(z^f)^{1-\gamma-\eta} \eta (k^{ic})^{\gamma} (l^{ic})^{\eta-1} - w \text{ as labor is chosen optimally only at } z^{ic} \text{ for an ic-firm. Above } z^{ic}, \text{ a constrained firm has a higher marginal product of labor than its marginal cost due to decreasing returns.} \]
that such a change would alleviate the missing middle but this would be associated to efficiency losses: indeed, those formal firms which were hiring optimally in the formal sector would now be constraining their hiring of inputs. These results should be contrasted with the existing literature. Keen and Mintz (2004) examine the efficiency implications of the choice of an optimal 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 neither does it account for evasion and corruption. In our model, entrepreneurs hide under the threshold where they can substitute capital for labor. This generates an efficiency loss as well as a decrease in tax revenues. Setting $l^{ic}$ to a low value is the best option in our context. Reducing the scope for hiding might generate welfare gains beyond the gains from a simple accounting over enforcement costs and tax revenues collected. Moreover, at high $l^{ic}$ relatively large and productive firms would constrained their labor thereby generating larger efficiency losses than smaller firms would at a lower $l^{ic}$.

We should expect the converse effect of a change in the size-threshold on the proportion of iu-firms. Let us first examine the effect of $l^{ic}$ on $z^{ic}$:

$$
\frac{dz^{ic}}{dl^{ic}} = \left\{ \frac{1}{p_F} \left[ \frac{r}{\gamma} \right] \left[ \frac{w}{\eta} \right]^{1-\gamma} \right\} \frac{1}{(1-\gamma-\eta)} > 0 \quad (4.62)
$$

This and the fact that $\frac{diu}{dz^{ic}} > 0$ yields:

$$
\frac{diu}{dl^{ic}} = \frac{diu}{dz^{ic}} \frac{dz^{ic}}{dl^{ic}} > 0 \quad (4.63)
$$

As expected, an increase in the size-threshold increases the proportion of iu-firms in this economy. Using the previous results in this section, the overall effect of an increase in the size-threshold on the proportion of ic-firms is:

$$
\frac{dic}{dl^{ic}} = \frac{dic}{dz^{ic}} \frac{dz^{ic}}{dl^{ic}} + \frac{dic}{dz^{f}} \frac{dz^{f}}{dl^{ic}} \geq 0 \quad (4.64)
$$
The first term after the equality is negative and the second, positive. Hence, an increase in the size-threshold into formality has ambiguous effect on the proportion of the informal-constrained sector.

4.4.3 Effect of a Change in the Bargaining Power of Entrepreneurs

We now examine the effect of a change in the bargaining power of entrepreneurs over the surplus from evasion. Given that no bribes are exchanged in the informal sector, a change in bargaining power has no effect on \( k_{iu} \), \( l_{iu} \), \( z_{ic} \) and \( k_{ic} \). In the formal sector, a change in bargaining power translates into a change in the same direction for both inputs:

\[
\frac{dk^f}{d\theta} = b(\tau_0 - \tau_p) \left( \frac{r}{1-\gamma-\eta} \right) \lambda \frac{1}{\frac{1}{1-\gamma-\eta}} \left[ p_F \left( \frac{\gamma}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right) \right]^{\frac{1}{1-\gamma-\eta}} > 0 \tag{4.65}
\]

\[
\frac{dl^f}{d\theta} = b(\tau_0 - \tau_p) \left( \frac{r}{1-\gamma-\eta} \right) \lambda \frac{1}{\frac{1}{1-\gamma-\eta}} \left[ p_F \left( \frac{\gamma}{r} \right)^{\frac{\gamma}{1-\eta}} \left( \frac{\eta}{w} \right) \right]^{\frac{1}{1-\gamma-\eta}} > 0 \tag{4.66}
\]

The equilibrium bribe is a function of both inputs \( k^f \) and \( l^f \) which are both functions of \( \theta \):

\[
\frac{dB}{d\theta} = (\tau_0 - \tau_p) p_F \left( z_{1-\gamma-\eta} \right) \left\{ -k^{\gamma-\eta} - (1-\theta) \left[ p_F \left( \frac{\gamma}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right) \right] \right\} + \psi A
\]

Let us multiply the second term on the right hand side (RHS) by \( \frac{\theta}{\eta} \) to get:

\[
\frac{dB}{d\theta} = (\tau_0 - \tau_p) p_F \left( z_{1-\gamma-\eta} k^{\gamma-\eta} \right) \left\{ -1 + \left( \frac{1-\theta}{\theta} \right) \left[ \frac{\gamma}{\eta} \left( k^f / k^f \right) + \eta \left( l^f / l^f \right) \right] \right\} + \psi A \geq 0 \tag{4.67}
\]

where

\[
\varepsilon_{k^f / \theta} = \frac{\theta}{\eta} \frac{dk^f}{d\theta} > 0
\]

\[24\] The second term could be negative for very low \( z^f \) and in such a case no firms are constrained.

\[25\] It is possible to show that with a certain set of parameters the clustering effect in the ic-sector increases with an increase in \( l^f \).
and

\[ \varepsilon_{f/\theta} = \frac{\varepsilon}{\theta} \frac{d\varphi}{d\theta} > 0 \]

A change in the bargaining power as ambiguous effects on the size of the bribe. Indeed, as can be noted from Equation 4.67, the first term on the RHS (call it RHS1) is negative while the second and third terms (call them RHS2 and RHS3) are positive. When enforcement is greater than zero bureaucrats pass on the expected fine, \( \psi A \), to entrepreneurs and this, the higher the bargaining power of an entrepreneur. In the limit, when an entrepreneur has all bargaining power, it must assume the total expected fine from corruption. In general, if the elasticities of the inputs with respect to the bargaining power are large, the bargaining power of a firm has to be high to reduce the bribe. Conversely, when the elasticities are small, entrepreneurs do not change their allocations by much and have more leverage to negotiate a lower bribe. To see this, we explore four cases.

**Case 1:** \( \psi A \rightarrow 0 \), large elasticities

When enforcement is low, we only need to consider RHS1 and RHS2. We have,

\[ \frac{dB}{d\theta} < 0 \text{ if } \theta > \frac{\gamma \varepsilon_{f/\theta} + \eta \varepsilon_{f/\theta}}{1 + \gamma \varepsilon_{f/\theta} + \eta \varepsilon_{f/\theta}} \]  

(4.68)

If the elasticities are large the RHS of Equation 4.68 goes to one and \( \theta \) has to be very large in order for an increase in bargaining power to decrease the bribe. In this case, entrepreneurs could increase their allocation of capital and labor by so much in response to an increase in bargaining power that they need to pay a bigger bribe than if their bargaining power had not changed.

**Case 2:** \( \psi A \rightarrow 0 \), small elasticities

In this case, Equation 4.68 goes to zero. If the elasticities are small, an entrepreneur does not change its allocation of capital and labor by much in response to the change in its bargaining power but it has more power to negotiate a lower bribe.
**Case 3**: $\psi A > 0$, **large elasticities** As before, let us obtain a condition that relates the equilibrium bribe to $\theta$:

\[
\frac{dB}{d\theta} < 0 \text{ if } \theta > \frac{\gamma \varepsilon_{k/\theta} + \eta \varepsilon_{l/\theta}}{1 + \gamma \varepsilon_{k/\theta} + \eta \varepsilon_{l/\theta} - \frac{\psi A}{\tau p z^k g^l f^l}}
\] (4.69)

Again if elasticities are large, the RHS of Equation 4.69 goes to one, and entrepreneurs need a large bargaining power to negotiate lower bribes. Note that the extra term in the denominator makes this requirement more difficult to satisfy than in case 1. However, if $\psi A \geq \tau p z^k g^l f^l$, an entrepreneur prefers to repay its surplus from evasion than a bribe.\[26\]

**Case 4**: $\psi A > 0$, **small elasticities** When elasticities are small, the bribe normally decreases with an increase in bargaining power. This is because an entrepreneur does not change its allocations by much in response to the change in bargaining power.

Finally, we examine the effect of a change in bargaining power on the shock-threshold into formality, $z^f$. Taking the derivative of Equation 4.58 with respect to $\theta$ we have:

\[
p_F (1 - \gamma - \eta) (z^f)^{1 - \gamma - \eta} \frac{dz^f}{d\theta} + \left\{ b (\tilde{\tau}_0 - \tau_p) (k^f)^{\gamma (l^f)^\eta} + \lambda \left[ \gamma (k^f)^{\gamma (l^f)^\eta} \frac{dk^f}{d\theta} + \eta (k^f)^{\gamma (l^f)^\eta - 1} \frac{dl^f}{d\theta} \right] \right\} \\
- \omega \frac{dz^f}{d\sigma} - r \frac{dk^f}{d\sigma} = b \psi A
\]

Using the equilibrium conditions for wages and the rental rate to eliminate some terms in the above expression we get:

\[
\frac{dz^f}{d\theta} = \frac{b \left[ \psi A - (\tilde{\tau}_0 - \tau_p) p_F (z^f)^{1 - \gamma - \eta} (k^f)^{\gamma (l^f)^\eta} \right]}{p_F (1 - \gamma - \eta) (z^f)^{1 - \gamma - \eta} \Lambda} \geq 0 \quad (4.70)
\]

\[26\]In such a situation, an entrepreneur pays a bribe higher than its liabilities and this, for any level of bargaining power.
An increase in bargaining power has ambiguous effects on the shock threshold. Note however that no bribes are exchanged in equilibrium if \( \psi A \geq \tau p_F(z^f)^\lambda (k^f)^\gamma (l^f)^\eta \).
This implies that in a situation where bribes are exchanged, the formal sector increases with an increase in bargaining power, i.e., \( \frac{df}{db} > 0 \).

4.4.4 Effect of a Change in the Audit Probability

The audit probability, \( b \), only affects the choices of formal entrepreneurs in the analytical example. The choice of inputs is inversely related to a change in audit probability as can be inferred from the next two equations:

\[
\frac{dk^f}{db} = -\frac{(\tau_0 - \tau_p)(1 - \theta)}{(1 - \gamma - \eta)} \lambda^{\frac{1}{1 - \gamma - \eta}} z \left[ p_F \left( \frac{\gamma}{\tau} \right)^{1 - \eta} \left( \frac{\eta}{w} \right)^\eta \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \tag{4.71}
\]

\[
\frac{dl^f}{db} = -\frac{(\tau_0 - \tau_p)(1 - \theta)}{(1 - \gamma - \eta)} \lambda^{\frac{1}{1 - \gamma - \eta}} z \left[ p_F \left( \frac{\gamma}{\tau} \right)^\gamma \left( \frac{\eta}{w} \right)^{1 - \gamma} \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \tag{4.72}
\]

Since formal entrepreneurs reduce the amount of labor and capital that they hire when the audit probability is increased, they can only afford to pay lower bribes:

\[
\frac{dB}{db} = \frac{(1 - \theta)(\tau_0 - \tau_p) p_F(z)^{1 - \gamma - \eta} (k^f)^\gamma (l^f)^\eta}{b} \left[ \gamma \varepsilon_{k/f} + \eta \varepsilon_{l/f} \right] < 0 \tag{4.73}
\]

where

\[
\varepsilon_{k/f} = \frac{b}{k^f} \frac{dk^f}{db} < 0
\]

and

\[
\varepsilon_{l/f} = \frac{b}{l^f} \frac{dl^f}{db} < 0
\]

Finally, we examine the effect of a change in the audit probability on the shock-threshold into formality \( z^f \). Taking the derivative of Equation 4.58 with respect to \( b \), we have:

\[
p_F(1 - \gamma - \eta)(z^f)^{-\gamma - \eta} \frac{dz^f}{db} \Lambda
\]
Using the equilibrium conditions for wages and the rental rate to eliminate some terms, we are left with the following expression:

\[
\frac{dz_f}{db} = \frac{(\tau_0 - \tau_p)(1 - \theta)\gamma}{p_F(1 - \gamma - \eta)(z_f^{-\gamma - \eta}A)} > 0 \quad (4.74)
\]

An increase in the audit probability increases the shock-threshold into formality. This implies that the formal sector shrinks and more firms cluster in the ic-sector. There is no effect on the iu-sector.

### 4.4.5 Effect of a Change in the Official Tax Rate

In the example, the variables related to the informal sector are unaffected by a change in the official tax rate, \( \tau_0 \). However, the effect of a change in \( \tau_0 \) on inputs in the formal sector is similar to the effect of a change in the audit probability. An entrepreneur reacts to a change in \( \tau_0 \) with an inversely proportional response in the choice of his inputs as shown in the following two expressions:

\[
\frac{dk_f}{d\tau_0} = -\frac{b}{2} (2 - \theta) \frac{1}{(1 - \gamma - \eta)} \lambda^{\frac{1}{\gamma + \eta} - 1} z \left[ p_F \left( \frac{\gamma}{r} \right)^{1 - \eta} \left( \frac{w}{r} \right)^{\eta} \right]^{\frac{1}{\gamma + \eta}} < 0 \quad (4.75)
\]

\[
\frac{dl_f}{d\tau_0} = -\frac{b}{2} (2 - \theta) \frac{1}{(1 - \gamma - \eta)} \lambda^{\frac{1}{\gamma + \eta} - 1} z \left[ p_F \left( \frac{\gamma}{r} \right)^{1 - \eta} \left( \frac{w}{r} \right)^{1 - \gamma} \right]^{\frac{1}{\gamma + \eta}} < 0 \quad (4.76)
\]

The equilibrium bribe, however, behaves differently under a change in \( \tau_0 \) than under a change in \( p_a \). Taking a derivative of \( B \) with respect to \( \tau_0 \), we have:
\[
\frac{dB}{d\tau_0} = \frac{(1-\theta)}{2} p_F z^{1-\gamma-\eta}(k^f)^\gamma(l^f)^\eta [1 + \gamma \varepsilon_{k/\tau} + \eta \varepsilon_{l/\tau}] \geq 0
\]  
(4.77)

where

\[
\varepsilon_{k/\tau} = \frac{\tau}{k^f} \frac{dk^f}{d\tau_0} < 0
\]

and

\[
\varepsilon_{l/\tau} = \frac{\tau}{l^f} \frac{dl^f}{d\tau_0} < 0
\]

As noted from Equation 4.77, an increase in the official tax rate has ambiguous effect on the level of the bribe. The overall effect depends on the magnitude of the elasticities of the inputs with respect to the tax rate. Suppose these elasticities are small (inelastic), the level of the bribe then increases with the tax rate, that is, 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 big, firms decrease their hiring of inputs in response to a tax increase, and this leaves less profits to be taxed from, and in turn, lower bribes to be exchanged with bureaucrats.

Finally we examine the effect of a change in \(\tau_0\) on the shock-threshold into formality, \(z^f\). Taking the derivative of Equation 4.58 with respect to \(\tau_0\) and using the equilibrium conditions for wages and the rental rate, we obtain:

\[
\frac{dz^f}{d\tau_0} = \left[ \frac{\frac{b}{z}(2-\theta)}{(1-\gamma-\eta)\Lambda} \right] z^f (k^f)^\gamma (l^f)^\eta > 0
\]  
(4.78)

An increase in the official tax rate has the expected effect of shrinking the formal sector. Hence, more firms will cluster in the ic-sector.

### 4.4.6 Effect of a Change in the Expected Fine

An increase in the expected fine does not affect capital and labor in any sector. An increase in the level of enforcement increases the level of the bribe because corrupt bureaucrats ask for a bigger compensations to let firms evade, and this the higher the bargaining power of an entrepreneur. To see this take a derivative of \(B\) with
Examining the effect of the expected fine on the shock-threshold into formality, we take a derivative of $z^f$ with respect to $\psi A$ to get:

$$\frac{dz^f}{d\psi A} = \frac{b\theta}{p_F(1-\gamma-\eta)(z^f)^{-\gamma-\eta} \Lambda} > 0$$

(4.80)

This result implies that the formal sector shrinks when enforcement is increased and firms cluster in the ic-sector.

### 4.4.7 Effect of a Change in Aggregate Demand

As the next set of equations shows, inputs in all sector change in the same direction as aggregate demand, $p_F$:

$$\frac{dk_{iu}^i}{dp_F} = z (p_F)^{1-\eta-\gamma} \left[ \left( \frac{\gamma}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^\eta \right]^{\frac{1}{1-\gamma-\eta}} > 0$$

(4.81)

$$\frac{dl^f}{dp_F} = z (p_F)^{1-\gamma-\eta} \left[ \left( \frac{\gamma}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^\eta \right]^{\frac{1}{1-\gamma-\eta}} > 0$$

(4.82)

$$\frac{dk_{ic}^i}{dp_F} = (p_F)^{1-\gamma} \left[ \left( \frac{\gamma}{r} \right) (\frac{\eta}{w})^{1-\gamma} \eta \right]^{\frac{1}{1-\gamma-\eta}} > 0$$

(4.83)

$$\frac{dk_f^i}{dp_F} = z (p_F)^{1-\eta-\gamma} \left[ \lambda \left( \frac{\gamma}{r} \right)^{1-\eta} \left( \frac{\eta}{w} \right)^\eta \right]^{\frac{1}{1-\gamma-\eta}} > 0$$

(4.84)

$$\frac{dl_f^i}{dp_F} = z (p_F)^{1-\gamma-\eta} \left[ \lambda \left( \frac{\gamma}{r} \right)^{1-\gamma} \left( \frac{\eta}{w} \right)^\gamma \right]^{\frac{1}{1-\gamma-\eta}} > 0$$

(4.85)

As expected, the effect on the equilibrium bribe will be similar to the effect on the inputs in the formal sector.
\[
\frac{dB}{dp_F} = (1 - \theta)(\tau_0 - \tau_p)z^{1-\gamma-\eta}(k^f)^\eta(1^f)^\eta \left[ 1 + \gamma \varepsilon_{k/f p_F} + \eta \varepsilon_{l/f p_F} \right] > 0 \quad (4.86)
\]

where \(\varepsilon_{k/f p_F}\) and \(\varepsilon_{l/f p_F}\) are the elasticities of the inputs in the formal sector with respect to the price of output.

We examine also how each sector’s proportion will vary with a change in \(p_F\). In order to accomplish this task, we must first describe how each of the two shock-thresholds \(z^{ic}\) and \(z^f\) are affected by aggregate demand. As before, we use Equation 4.58 and the equilibrium conditions for wages and the rate of rental of capital to solve for the effect of \(p_F\) on \(z^f\):

\[
\frac{dz^f}{dp_F} = \frac{-z^f}{p_F(1 - \gamma - \eta)} < 0 \quad (4.87)
\]

The shock-threshold into formality is thus inversely related to \(p_F\). We observe the same effect on \(z^{ic}\) as can be inferred using Equation 4.57 and taking a derivative with respect to \(p_F\):

\[
\frac{dz^{ic}}{dp_F} = \frac{-1}{1 - \gamma - \eta} \left( p_F \right)^{-1 - \gamma - \eta - 1 + \gamma - \frac{1}{\eta}} \left\{ r \left[ \frac{r}{w} \right]^{1-\gamma} \right\} < 0 \quad (4.88)
\]

Given this set of results on the shock-thresholds, it is easy to see that under an increase in aggregate demand, the formal sector expands:

\[
\frac{df}{dp_F} = \frac{df}{dz^f} \frac{dz^f}{dp_F} > 0
\]

while the iu-sector shrinks:

\[
\frac{diu}{dp_F} = \frac{diu}{dz^{ic}} \frac{dz^{ic}}{dp_F} < 0
\]

However, the effect on the ic-sector is not as obvious as observed from Equation 4.89.
\[ \frac{\text{dic}}{dp_F} = \frac{\text{dic} \ dz^f}{dz^f \ dp_F} + \frac{\text{dic} \ dz^{ic}}{dz^{ic} \ dp_F} \geq 0 \quad (4.89) \]

The first term on the RHS of Equation 4.89 is negative while the second term is positive. The overall change in ic-sector depends on the relative magnitude of entry from the iu-sector and exit towards the formal sector. Basically, we want to check in what direction goes the following relationship:

\[ \left| \frac{\text{dic} \ dz^f}{dz^f \ dp_F} \right| \approx \left| \frac{\text{dic} \ dz^{ic}}{dz^{ic} \ dp_F} \right| \]

\[ \Rightarrow \left| G(z^f) \frac{(-1)z^f}{(1-\gamma - \eta)p_F} \right| \approx -G(z^{ic}) \left( \frac{-1}{1-\gamma - \eta} \right) \left( \frac{1}{p_F} \right)^{\frac{1}{1-\gamma - \eta} + 1} \left| \left( t^{ic} \left( \frac{r}{\gamma} \left( \frac{w}{\eta} \right)^{1-\gamma} \right) \right)^{\frac{1}{1-\gamma - \eta}} \right| \]

Using a uniform distribution and cancelling \( p_F \) and \( 1 - \gamma - \eta \) on each side, we have that:

\[ \Rightarrow \left| (z^f) (1-z^f) \right| \approx \left| -z^{ic} (-1) t^{ic} \left[ \frac{1}{p_F} \left( \frac{r}{\gamma} \left( \frac{w}{\eta} \right)^{1-\gamma} \right) \right]^{\frac{1}{1-\gamma - \eta}} \right| \]

\[ \Rightarrow \left| (z^f)^2 \right| \approx \left| (z^{ic})^2 \right| \]

In our setup, \( z^{ic} \) and \( z^f \) are fixed functions of the parameters and \( z^f > z^{ic} \) otherwise there is no ic-sector anyways. We thus observe a greater influx in the formal sector from the ic-sector than from the iu-sector to the ic-sector, that is:

\[ \left| \frac{df}{dp_F} \right| > \left| \frac{diu}{dp_F} \right| \]

which implies that the clustering effect would be alleviated by an increase in \( p_F \), i.e., \( \frac{\text{dic}}{dp_F} < 0 \).
4.4.8 Effect of a Change in Wages

A change in wage does not affect inputs’ choices in the ic-sector but it is inversely related to the choices of inputs in the iu-sector and in the formal sector as inferred from the next four expressions:

\[
\frac{dl^{iu}}{dw} = z \left( -\frac{1 - \gamma}{1 - \gamma - \eta} \right) \left( \frac{1}{w} \right)^{\frac{\eta}{1 - \gamma - \eta}} \left[ p_F \left( \frac{r}{r} \right)^{\gamma} \eta^{1 - \gamma} \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \quad (4.90)
\]

\[
\frac{dk^{iu}}{dw} = z \left( -\frac{\eta}{1 - \gamma - \eta} \right) \left( \frac{1}{w} \right)^{\frac{\eta}{1 - \gamma - \eta}} \left[ p_F \left( \frac{r}{r} \right)^{1 - \eta} \eta^{\eta} \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \quad (4.91)
\]

\[
\frac{dl^{f}}{dw} = z \left( -\frac{1 - \gamma}{1 - \gamma - \eta} \right) \left( \frac{1}{w} \right)^{\frac{\eta}{1 - \gamma - \eta}} \left[ \Lambda p_F \left( \frac{r}{r} \right)^{\gamma} \eta^{1 - \gamma} \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \quad (4.92)
\]

\[
\frac{dk^{f}}{dw} = z \left( -\frac{\eta}{1 - \gamma - \eta} \right) \left( \frac{1}{w} \right)^{\frac{\eta}{1 - \gamma - \eta}} \left[ \Lambda p_F \left( \frac{r}{r} \right)^{1 - \eta} \eta^{\eta} \right]^{\frac{1}{1 - \gamma - \eta}} < 0 \quad (4.93)
\]

As before, the equilibrium bribe moves in the same direction as the inputs from the formal sector.

\[
\frac{dB}{dw} = \frac{(1 - \theta)}{w} \left( \tau_0 - \tau_p \right) p_F z^{1 - \gamma - \eta} (k^f)^{\gamma} (l^f)^{\eta} [\gamma \varepsilon_{k/w} + \eta \varepsilon_{l/w}] < 0 \quad (4.94)
\]

where \( \varepsilon_{k/w} \) and \( \varepsilon_{l/w} \) are the elasticities of the inputs in the formal sector with respect to wages. Next we examine the effect of a change in wages on sectors’ proportions. We first look at the effect of \( w \) on \( z^{ic} \) and \( z^f \) taking the appropriate derivative in Equation 4.57 and Equation 4.58\(^{27}\).

\(^{27}\)Again, to solve for the derivative of \( z^f \) with respect to \( w \) we use the equilibrium conditions for wages and the rate of rental.
\[
\frac{dz^{ic}}{dw} = \frac{1 - \gamma}{1 - \gamma - \eta} w^{\gamma - 1} \left[ \frac{1}{p_F} \left( \frac{r}{\gamma} \right)^\gamma \left( \frac{1}{\eta} \right)^{1-\gamma} \right]^{\frac{1}{1-\gamma-\eta}} > 0 \quad (4.95)
\]

\[
\frac{dz^f}{dw} = \frac{l^f - l^{ic}}{p_F (1 - \gamma - \eta) (z^f)^{-\gamma - \eta} \Lambda} > 0 \quad (4.96)
\]

With this last set of results in hand we find that under an increase in wages the iu-sector expands:

\[
\frac{diu}{dw} = \frac{diu}{dz^{ic}} \frac{dz^{ic}}{dw} > 0
\]

while the firmal sector shrinks:

\[
\frac{df}{dw} = \frac{df}{dz^f} \frac{dz^f}{dw} < 0
\]

Once again, the effect on the ic-sector is ambiguous as seen from the expression below:

\[
\frac{dic}{dw} = \frac{dic}{dz^{ic}} \frac{dz^{ic}}{dw} + \frac{dic}{dz^f} \frac{dz^f}{dw} \leq 0
\]

The first term on the RHS of Equation 4.4.8 is negative while the second is positive. We thus need to compare:

\[
\left| \frac{dic}{dz^{ic}} \frac{dz^{ic}}{dw} \right| \leq \left| \frac{dic}{dz^f} \frac{dz^f}{dw} \right|
\]

\[
-G(z^{ic}) \frac{1 - \gamma}{1 - \gamma - \eta} w^{\gamma - 1} l^{ic} \left[ \frac{1}{p_F} \left( \frac{r}{\gamma} \right)^\gamma \left( \frac{1}{\eta} \right)^{1-\gamma} \right]^{\frac{1}{1-\gamma-\eta}} \leq \left| G(z^f) \frac{l^f - l^{ic}}{p_F (1 - \gamma - \eta) (z^f)^{-\gamma - \eta} \Lambda} \right|
\]

Using a uniform distribution and cancelling \(1 - \gamma - \eta\) on each side, we have:

\[
-\frac{z^{ic}}{1 - \gamma} w^{-1} l^{ic} \left[ \frac{1}{p_F} \left( \frac{r}{\gamma} \right)^\gamma \left( \frac{w}{\eta} \right)^{1-\gamma} \right]^{\frac{1}{1-\gamma-\eta}} \leq \left| (z^f)^{1 + \gamma + \eta} \frac{l^f - l^{ic}}{p_F} \right|
\]

112
Further simplifying, we get:

\[ \left| -\left(z_{ic} \right)^2 \frac{1 - \gamma}{w} \right| \leq \left| \frac{\left(z_f \right)^{1+\gamma+\eta} \left(I_f - I_{ic} \right)}{p_F \Lambda} \right| \]

Or,

\[ \left| -\left(z_{ic} \right)^2 \frac{1 - \gamma}{w} \right| \leq \left| \frac{\left(z_f \right)^{1+\gamma+\eta} \left(I_f - I_{ic} \right)}{\left(\left(z_f \right)^{\gamma+\eta} \left[p_P \left(\frac{\gamma}{\gamma}\right)^{\gamma \left(\frac{\eta}{w}\right)} \right] - p_F \left(\frac{\gamma}{z}\right)^{\gamma \left(l_{ic} \right)^{\eta}} \right)^{\frac{1}{\gamma}}} \right| \]

It is possible to show that the RHS of the above expression is always greater than LHS except for really high wages\(^{28}\). Hence, the clustering effect is exacerbated: the decrease in proportion of the \(f\)-sector is not compensated by the increase in the proportion of the \(i_u\)-sector.

### 4.4.9 Effect of a Change in the Rental Rate of Capital

The effects of a change in the rental rate of capital are similar to those from the previous section. The only exception is that capital’s hiring in the \(i_c\)-sector is affected by a change in rental rate. Inputs in all sectors are inversely related to a change in rental rate:

\[
\frac{dI_{iu}}{dr} = z \left( \frac{-\gamma}{1 - \gamma - \eta} \right) \left( \frac{1}{r} \right)^{\gamma \eta + 1} \left[ p_F \left(\frac{\gamma}{\gamma}\right)^{\gamma \left(\frac{\eta}{w}\right)} \right]^{1-\gamma} < 0 \quad (4.97)
\]

\[
\frac{dI_{k_u}}{dr} = z \left( \frac{-1 - \eta}{1 - \gamma - \eta} \right) \left( \frac{1}{r} \right)^{\gamma \eta + 1} \left[ p_F \left(\frac{\gamma}{\gamma}\right)^{1-\eta} \left(\frac{\eta}{w}\right)^{\eta} \right]^{\gamma} < 0 \quad (4.98)
\]

\[
\frac{dI_{k_c}}{dr} = \left[ \frac{-1}{1 - \gamma} \right] \left( \frac{1}{r} \right)^{\gamma + 1} \left(\gamma p_F z^{1-\gamma-\eta} \left(l_{ic} \right)^{\eta} \right)^{\frac{1}{\gamma}} < 0 \quad (4.99)
\]

\(^{28}\)This would imply that all the economy is totally informal unconstrained anyways.
\[
\frac{dk^f}{dr} = z \left( -\frac{1 - \eta}{1 - \gamma - \eta} \right) \left( \frac{1}{r} \right)^{1-\eta \gamma^{-1} \eta} \left[ \lambda p_F(\gamma)^{1-\eta} \left( \frac{\eta}{w} \right) \right]^{1 \gamma^{-1} \eta} < 0 \quad (4.100)
\]

\[
\frac{dl^f}{dr} = z \left( -\gamma \frac{1 - \gamma - \eta}{1 - \gamma - \eta} \right) \left( \frac{1}{r} \right)^{1-\gamma \gamma^{-1} \eta} \left[ \lambda p_F(\gamma)^{1-\gamma} \left( \frac{\eta}{w} \right) \right]^{1 \gamma^{-1} \eta} < 0 \quad (4.101)
\]

The equilibrium bribe moves in the same direction as the inputs from the formal sector in response to a change in the rental rate:

\[
\frac{dB}{dr} = \frac{(1 - \theta)}{r} (\tau_0 - \tau_p) p_F z^k \rho \eta \left[ \gamma \varepsilon_{k/f} \eta + \eta \varepsilon_{l/f} \right] < 0 \quad (4.102)
\]

where \( \varepsilon_{k/f} \) and \( \varepsilon_{l/f} \) are the elasticities of the inputs in the formal sector with respect to the rental rate.

Next we examine the effect of a change in \( r \) on \( z^{ic} \) and \( z^f \):

\[
\frac{dz^{ic}}{dw} = \frac{\gamma}{1 - \gamma - \eta} \left( \frac{1}{1 - \gamma - \eta} \right)^{\gamma \gamma^{-1} \eta - 1} \left[ \frac{1}{p_F} \left( \frac{1}{\gamma} \right)^{\gamma} \left( \frac{w}{\eta} \right)^{1-\gamma} \right]^{\gamma \gamma^{-1} \eta} > 0
\]

\[
\frac{dz^f}{dr} = \frac{[k^f - k^{ic}]}{p_F (1 - \gamma - \eta) (z^f)^{-\gamma - \eta} \lambda} > 0 \quad (4.103)
\]

This last set of results allows us to establish that, in response to an increase in rental rate, the iu-sector expands:

\[
\frac{diu}{dr} = \frac{diu \ dz^{ic}}{dz^{ic} \ dr} > 0
\]

while the formal sector shrinks:

\[
\frac{df}{dr} = \frac{df \ dz^f}{dz^f \ dr} < 0
\]

As always, the effect on the ic-sector’s proportion is ambiguous:

\[
\frac{dic}{dr} = \frac{dic \ dz^{ic}}{dz^{ic} \ dr} + \frac{dic \ dz^f}{dz^f \ dr} \leq 0
\]
However, it is easy to show that, in response to an increase in $r$, the clustering effect is exacerbated: the decrease in proportion of the formal sector is not compensated by the increase in the proportion of the iu-sector.
Chapter 5

Numerical Analysis

5.1 Numerical Methodology

5.1.1 Indirect Inference

The procedure used to choose some of the structural parameters relies on a simulation-based method called Indirect Inference (Gourieroux et al. [1993]; Smith Jr [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, \( \phi \), 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:

\[
\hat{\phi} = \arg \min_{\phi} (m_s(\phi) - m_d)' W (m_s(\phi) - m_d)
\]

where \( W \) is the inverse of 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 \( \phi \) of parameters to be assigned;
2. Iterate over the conditional value functions until joint convergence of the optimal functions.\[1\]

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 $\phi$ if a minimum has not been reached.

Equation 5.1 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.

5.1.2 Simulated Annealing

Annealing is the process by which metals or other materials are slowly cooled down to let atoms reach their lowest energy state, that is, their most stable configuration.\[2\] Simulated annealing uses this idea for the optimization of multi-dimensional systems using a control parameter that has a similar role than temperature in the annealing process. Starting at a high temperature, the algorithm searches for paths to minimums over a range of the order of the temperature. Gradually, the system is cooled down and the search narrows in on a path that leads to a global minimum.

There are 4 elements needed to implement this procedure:

1) the objective function to be minimized and in our case, Equation 5.1.

2) an N-dimensional vector of states. In our case, $\phi$.

3) A control parameter, that is, the temperature.

4) A procedure for taking random steps from $\phi$ to $\Delta \phi$.

Press et al. [1992] use an extension of the Downhill Simplex method to implement this procedure. The Downhill Simplex method searches for a minimum by

\[1\] See below for a description of the value function iteration algorithm

\[2\] For example, crystalline solids as opposed to glass have reached such a ground state.
expanding, contracting or reflecting points in an N+1 dimensional simplex (where
N is the number of parameters to be estimated). To implement Simulated Anneal-
ing, a random variable is added to the value of the function stored and a similar
random variable is subtracted from every new value of the function at the replace-
ment points. At a finite temperature T, the simplex executes a stochastic motion
within the region defined by the size of T in order to sample random points. That
way a downhill move is always accepted when there exists one and uphill moves
are sometimes accepted. This allows circumventing the problem of narrow valleys
and the inefficiency of most algorithm when converging to a minimum. We use a
modified version of Press et al. [1992]’s AMEBSA subroutine.

5.1.3 Value Function Iteration

Our value function iteration algorithm is based on Judd [1998] and the notes by
Karen Kopecky. Consider our problem as described at the beginning of chapter 4
and write Bellman’s equation for the semi-conditional problem:

\[ V^*(k, z) = \max_{k'} \left\{ c^s + \varphi \beta EV(k', z') \right\} \]  (5.2)

Here are the steps to the value function iteration algorithm:

1. Choose a tolerance level for the approximation error, \( \varepsilon \).

2. Discretize the state space by constructing a matrix for each productivity
   shock/capital pair:

\[
\begin{bmatrix}
z_1, k_1 & z_1, k_2 & \ldots & z_1, k_n \\
z_2, k_1 & \ldots & \ldots & \ldots \\
\vdots & \vdots & \ddots & \vdots \\
z_m, k_1 & \ldots & \ldots & z_m, k_n
\end{bmatrix}
\]

where \( m \) is the number of nodes on the productivity shock grid and \( n \) the
number of nodes on the capital grid.

\(^3\)http://www.karenkopecky.net/Teaching/eco613614/
3. Start with an initial guess of the semi-conditional value function, \( V(0) \), and the fully-conditional value functions, \( V^s(0) \) for each status \( s \). These are \( m \times n \) matrices.

4. Update the value function using Equation 5.2. Specifically for each status \( s \),

   (a) Fix the current shock and capital stock at one of the grid points, \((z_i, k_j)\) with \( i = 1, \ldots, m \) and \( j = 1, \ldots, n \).

   (b) For each possible choice of capital next period, \( k_l \) with \( l = 1, \ldots, n \), calculate

   \[
   T^s(i; j) = U(F(z_i, k_j) + (1 - \delta) k_j - k_l) + \beta V_{i,j}(0) \tag{5.3}
   \]

   If consumption is negative for a particular \( k_l \), assign a large negative number to \( T^s(i; j) \). While iterating over \( l = 1, \ldots, n \), store the value of the maximizer, \( k_l \), in a \( m \times n \) matrix \( F(1) \) and store the value of the associated maximum \( T^s(i, j) \) in a \( m \times n \) matrix \( V^s(1) \).

   (c) Choose a new grid point for the current pair of shock and capital stock in step (a) and repeat steps (b) and (c). Once we have completed steps (a) through (c) for each pair \((z, k)\), we have updated value functions \( V^s(1) \) and policy functions \( F(1) \) and continue to the next step.

5. Compute the distance, \( d \), between \( V^s(0) \) and \( V^s(1) \):

   \[
   d = \max_{i,j \in \{(1,1), \ldots, (m,n)\}} |V_{i,j}^s(1) - V_{i,j}^s(0)|
   \]

6. If \( d \leq \varepsilon \) for all \( s \), the value function has converged. We have thus obtained the numerical estimates of the value and policy functions. If \( d > \varepsilon \), return to Step 4, setting the initial guess to the updated value function, i.e. \( V(0) = V(1) \) where \( V(1) \) is a semi-conditional value function, based on the fully-conditional value functions \( V^s(1) \) for each status \( s \) and the corresponding size-thresholds. Keep iterating until the value function has converged.

---

\( ^4 \)Note that we are using the semi-conditional value function to evaluate \( T^s(i; j) \)
5.2 Parametrization

Obtaining a solution for Equation 5.1 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 calibrated by indirect inference to two parameters. We then estimate eight regulation parameters directly from the Ugandan data while taking three technology parameters from the literature. Table 5.1 presents the value and origin of all parameters used in the benchmark calibration. Note that we set the length of a time period in our model to be one year in the data.

Table 5.1: Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Definition and Origin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preferences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\beta$</td>
<td>.97</td>
<td>Discount factor: Assumption</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>.93</td>
<td>Probability of survival: Indirect Inference</td>
</tr>
<tr>
<td>Regulation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mu^c$</td>
<td>30</td>
<td>Size-threshold into formality: data</td>
</tr>
<tr>
<td>$\tau_0$</td>
<td>.144</td>
<td>Official tax rate: data</td>
</tr>
<tr>
<td>$\tau_p$</td>
<td>$\tau_0/2$</td>
<td>Effective tax rate: data</td>
</tr>
<tr>
<td>$\theta$</td>
<td>.98</td>
<td>Bargaining power: data</td>
</tr>
<tr>
<td>$\psi_A$</td>
<td>71$/yr/emp.</td>
<td>Expected fine: data</td>
</tr>
<tr>
<td>$p_a^s$</td>
<td>.53</td>
<td>Probability of audit, small: data</td>
</tr>
<tr>
<td>$p_a^m$</td>
<td>.71</td>
<td>Probability of audit, medium: data</td>
</tr>
<tr>
<td>$p_a^l$</td>
<td>.77</td>
<td>Probability of audit, large: data</td>
</tr>
<tr>
<td>$q$</td>
<td>1</td>
<td>Probability of being asked for a bribe: assumption</td>
</tr>
<tr>
<td>$w_h$</td>
<td>1800$/yr</td>
<td>Tax Officials’ wage: IMF Staff Country Report 1998</td>
</tr>
<tr>
<td>Technology</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\delta$</td>
<td>.07</td>
<td>Depreciation Rate: Assumption</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>.283</td>
<td>Restuccia and Rogerson [2008]</td>
</tr>
<tr>
<td>$\eta$</td>
<td>.567</td>
<td>Restuccia and Rogerson [2008]</td>
</tr>
<tr>
<td>$c_f$</td>
<td>0</td>
<td>Fixed Cost of Production: Assumption</td>
</tr>
<tr>
<td>$p_{zz}$</td>
<td>.85</td>
<td>Probability to stay in current state: Indirect Inference</td>
</tr>
</tbody>
</table>
**Regulation environment**  The size-threshold into formality, $l^c$, is based on the empirical analysis and set at 30 employees.\(^5\) 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, $\tau_p$, 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 entrepreneurs is obtained by estimating Equation 4.22. As can be inferred from Equation 4.22, the coefficient on the surplus from evasion delivers $\theta$ while the expected fine per employee is given by the value of the constant in the regression. Table 5.2 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.\(^6\) Given the small size of bribes compared to the amount of taxes evaded in the data, it is not surprising to find such high values for $\theta$ in Table 5.2.\(^7\) We note from rows 1 to 4 that the constants (column 3) are almost equal to average bribes (column 4). The last three rows show that the level of the bargaining power is not significantly affected by the introduction of covariates.\(^8\) We use an estimated value of $\theta = 0.98$ and $\Psi A = 71\text{usd/yr}$ for the expected fine. As mentioned previously, 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 assume that the poor performance of the audits is due to the fact

---

5This 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].

6Bribes 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.

7Note that the average bribe per worker is approximately half a month of a tax official’s wage. This entails a bribe per firm (for the average firm) which is about 5 times the yearly wage for tax officials.

8There is a potential endogeneity issue in that the bribe and the evasion surplus might be decided at the same time. We use capital investment in 1995 and 1996 as an instrument for the evasion surplus: to the extent that capital markets are imperfect, past investment decisions are correlated to current evasion surplus while being independent of current productivity shocks. The level and significance of the coefficient on the evasion surplus (and thus the value of the bargaining power) are not sensitive to the use of the instrument. An alternative approach is to calibrate the bargaining power parameter through Indirect Inference. However we leave this avenue for future work.
that most tax officials are potentially corruptible. We thus set in the benchmark $q = 1^{[9]}$ Finally we use the IMF Staff Country Report 1998 to assess the average public wage in Uganda in 1997 which is approximately $w_b = \text{US$1800}^{[10]}$

Table 5.2: Bargaining Power and Enforcement Costs

<table>
<thead>
<tr>
<th></th>
<th>Bargaining Power ($\theta$)$^b$</th>
<th>Expected Fine/emp ($\psi A$)$^b$</th>
<th>R_squared</th>
<th>Average Bribe/emp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Firms</td>
<td>.995 [1.95]</td>
<td>32.12 [1.93]</td>
<td>.14</td>
<td>47.16</td>
</tr>
<tr>
<td>Medium Firms</td>
<td>.986** [2.40]</td>
<td>210.39* [1.97]</td>
<td>.50</td>
<td>120.81</td>
</tr>
<tr>
<td>Large Firms</td>
<td>.988** [2.87]</td>
<td>72.87 [1.86]</td>
<td>.20</td>
<td>72.66</td>
</tr>
<tr>
<td>All Firms</td>
<td>.988** [2.18]</td>
<td>99.75** [2.49]</td>
<td>.26</td>
<td>71.35</td>
</tr>
<tr>
<td>All Firms$^d$</td>
<td>.983*** [3.41]</td>
<td>380.98*** [3.27]</td>
<td>.38</td>
<td>71.35</td>
</tr>
<tr>
<td>All Firms$^e$</td>
<td>.978*** [5.73]</td>
<td>476.61*** [4.53]</td>
<td>.58</td>
<td>71.35</td>
</tr>
</tbody>
</table>

$^a$ * significant at 10%, ** significant at 5%, *** significant at 1%; nb. of firms in parenthesis, t-stat in brackets; $^b$ First two columns are imputed from a regression based on Equation 4.22; $^c$ All regressions include dummies for industry-location; $^d$ Covariates include Age, $Age^2$, price of capital, average wage per employee; $^e$ Covariates include Age, $Age^2$, price of capital, average wage per employee and audit intensity; $^f$ Results are not significantly affected if we use total amounts instead of amounts per employee.

Technology Parameterizing technology presents some challenges. We estimate shares into production for capital and labor using the Ugandan data in Chapter 3. 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 technol-

$^9$The report also states that compliance to fill tax reports was low. We thus set the fine paid by entrepreneurs, $FA$, to zero in the regulation equation.

$^{10}$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.
ogy 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 thus 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 depreciation of $\delta = 0.07$. We set the fixed cost of production to zero.\(^{11}\)

As in Hopenhayn and Rogerson [1993], 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. Second, we use the relative demand for labor, equation 4.49, 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.\(^{12}\) This gives us a support for shocks that ranges from $z_{\text{low}} = 1$ to $z_{\text{up}} = 3.8$. 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, $\nu$, is a uniform on $[z_{\text{low}}, z_{\text{up}}]$. Finally, we calibrate the wage using the labor demand for the largest firm in the Ugandan data. We assume that this wage satisfies the workers’ problem for an arbitrary given $D$. Log preferences for workers generate

\(^{11}\)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 [1992a]). 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 the model to data. We prefer to impose greater discipline on the model and use a smaller set of structural parameters for the calibration.

\(^{12}\)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]
a labor supply that is inelastic. We use this fact to examine the pure reallocative effects on output of a change in the environment, i.e., in all experiments below we adjust the wage to keep aggregate labor as in the benchmark. As such, the effect of a change in the environment on output can be considered as a lower bound in the experiments below as part of the productivity change is embedded in the change in the equilibrium wage.

Preferences Finally, Arellano et al. [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 (e.g. 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 $\beta$.

Parameters calibrated by Indirect Inference
This leaves us with the calibration of the survival rate $\varphi$ 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 $\varphi$ 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.\textsuperscript{13} Let the probability that next period productivity shock, $z'$, equals the current period productivity shock, $z$, be given by:

$$p_{zz} = \sum_{j=1}^{n-1} p_{zy} = 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

\textsuperscript{13}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.
where \( j = 1, \ldots, n_z - 1 \), we have:

\[
p_{zy} = \frac{1 - p_{zz}}{n_z - 1}, \quad \forall y_j
\]  

(5.5)

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 Equation 5.5 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 \( j = 1, \ldots, n_z - 1 \).

Notice that, if \( u < v \), the model generates a set of over-identifying restrictions which can be used to test it. We thus supplement the auxiliary model with one other moment that describes aggregate data: average capital-output ratio.\(^{14}\)

Summary for Indirect Inference

In conclusion, we calibrate two parameters by Indirect Inference: the exogenous survival rate of firms, \( \varphi \), and the transition probability \( p_{zz} \) for the ten lowest shocks on the productivity grid. There are four moments being targeted: average number of employees, average number of firms, average growth rate of output and average capital-output ratio. There are two parameters estimated and four moments so this should leave us with 2 degrees of freedom in our test of over-identifying restrictions. However, the choice of estimating the transition probability \( p_{zz} \) for the ten lowest shocks on the productivity grid is rather arbitrary. We thus make the assumption that this arbitrary choice diminishes the degrees of freedom to one.

\(^{14}\)The auxiliary model allows to choose a convenient set of auxiliary parameters which could be anything from moments, impulse responses to regression coefficients. This is the main advantage over Simulated Method of Moments.
5.3 Model Validation

5.3.1 A Test of the Model

We provide a formal test of the model using the over-identifying restrictions from the Indirect Inference procedure. Table 5.3 contrasts the simulated moments obtained from the calibration against the targeted moments in the data. We note first that the average values of the simulated moments are within a close range of the actual moments. Similarly, the standard errors are also within the range of what is observed in the data. The statistic generated by the quantitative model is $\chi^2 = 5.14$ and lies below a critical value of 5.412 for a $\chi^2$-distribution with one degree of freedom and a 2% type I error probability. At such an unconventionally low level of significance, the test rejects the model.

It is important to note that the endeavour of finding a model that would match the data in all dimensions is not only ambitious but also unrealistic. This model fares relatively well compared to other models that have been tested using similar $\chi^2$ tests. For example, Low et al. [2010] develop a model of the effects of disability risk and insurance on consumption and labor decisions in the U.S. Their model is marginally accepted for a probability level of 2.25%, using a similar formal test of overidentifying restrictions with 23 degrees of freedom. Is this solid grounds to reject their work? Probably not, given that their model replicates a wide and interesting array of features that are observed in the U.S. data.\(^{15}\)

It is important to note that there are no “true” models that perfectly describe a DGP. Researchers can however rank models based on how close they fit the data. Hnatkovska et al. [2008] have developed a procedure to conduct such comparisons. Nonetheless, a formal test of the ranking of different channels in terms of their fit to the data is beyond the scope of this thesis and left for future work. We conduct in what follows an informal comparative analysis of the model with and without the distortions associated to regulation costs and/or missing capital market.

Finally, we don’t what is the power of this test. However, our choice of using only one degree of freedom is conservative and should work against us. As such,

\(^{15}\)Other works that have used such a formal test of overidentifying restrictions include Meenagh et al. [2009]
parameters for the auxiliary model in Indirect Inference can be chosen with much leeway. Introducing more moments in our computation would thus certainly work in our favor as the number of degrees of freedom to conduct a Wald test would increase accordingly.

**Table 5.3: Targets and Wald Test**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Benchmark</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>132</td>
<td>124</td>
</tr>
<tr>
<td>Nb. Employees</td>
<td>(213)</td>
<td>(259)</td>
</tr>
<tr>
<td>Average</td>
<td>13.9</td>
<td>13.9</td>
</tr>
<tr>
<td>Age</td>
<td>(13)</td>
<td>(12)</td>
</tr>
<tr>
<td>Capital-Output</td>
<td>1.46</td>
<td>1.34</td>
</tr>
<tr>
<td>Ratio</td>
<td>(3.77)</td>
<td>(3.05)</td>
</tr>
<tr>
<td>Average</td>
<td>.124</td>
<td>.08</td>
</tr>
<tr>
<td>Growth</td>
<td>(.40)</td>
<td>(.54)</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>5.144</td>
<td>—</td>
</tr>
</tbody>
</table>

Critical Value based on a $\chi^2$ distribution with one degree of freedom and a right-tail area of 0.02 is 5.412.

5.3.2 Non-Targeted Statistics

In this section we examine some statistics generated by the model that were not targeted in the calibration. Table 5.4 summarizes these 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 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. Figures 5.1 and 5.2 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)\(^{16}\) This is

\(^{16}\)Note however that we also observe a second mode in the data around 400 employees.
Table 5.4: Non-Targeted Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Size bins</th>
<th>Benchmark Average</th>
<th>Benchmark Std. Err.</th>
<th>Data Average</th>
<th>Std. Err.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Share of firms</td>
<td>Small</td>
<td>.47</td>
<td>.48</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>.12</td>
<td>.20</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.41</td>
<td>.32</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Average Size</td>
<td>Small</td>
<td>17</td>
<td>21</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>59</td>
<td>158</td>
<td>50</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>285</td>
<td>433</td>
<td>334</td>
<td>379</td>
</tr>
<tr>
<td>Average Age</td>
<td>Small</td>
<td>9</td>
<td>15</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>14</td>
<td>38</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>20</td>
<td>28</td>
<td>16</td>
<td>15</td>
</tr>
<tr>
<td>Maximum Age</td>
<td>All</td>
<td>97</td>
<td></td>
<td>74</td>
<td></td>
</tr>
<tr>
<td>Output Growth</td>
<td>Small</td>
<td>.04</td>
<td>.51</td>
<td>.02</td>
<td>.43</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>.19</td>
<td>.74</td>
<td>.13</td>
<td>.59</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.20</td>
<td>.70</td>
<td>.17</td>
<td>.67</td>
</tr>
<tr>
<td>Share of Output</td>
<td>Small</td>
<td>.06</td>
<td></td>
<td>.03</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>.06</td>
<td></td>
<td>.14</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>.88</td>
<td></td>
<td>.84</td>
<td></td>
</tr>
<tr>
<td>Bribes per Sales</td>
<td>All</td>
<td>.02</td>
<td>.02</td>
<td>.02</td>
<td>.10</td>
</tr>
<tr>
<td>Capital/labor</td>
<td>Small</td>
<td>65</td>
<td>89</td>
<td>78</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>Medium</td>
<td>61</td>
<td>79</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td></td>
<td>Large</td>
<td>52</td>
<td>59</td>
<td>50</td>
<td>97</td>
</tr>
</tbody>
</table>

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.

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. In Table 5.4 we note that 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.\[17\]

\[17\] The median size in the model (51 employees) is higher than in the data (35 employees).
Figure 5.1: Size Distribution of Firms: Benchmark v.s. Data

Figure 5.2: Size Distribution of Firms ( Logs): Benchmark v.s. Data
**Growth and Shares of Output** Table 5.5 exhibits the OLS estimates of a regression of the growth rate of output on labor using the simulated data from the benchmark. The table also reproduces the results of Table 3.12 for the sake of comparison. In chapter 3 we have shown that a 1% increase in the number of employees generates a 0.7% increase in firms’ growth rate of sales. It is easy to see that a similar effect but of smaller magnitude is also present in the simulated data.

**Table 5.5: Growth Rate of Output and Labor: OLS Estimates**

<table>
<thead>
<tr>
<th>Dep. Var.: Growth rate of Output</th>
<th>Benchmark</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.14***</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Labor</td>
<td>2.5e-4*</td>
<td>3.4e-4**</td>
</tr>
<tr>
<td></td>
<td>(1.4e-4)</td>
<td>(1.4e-4)</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Observations</td>
<td>208</td>
<td>208</td>
</tr>
<tr>
<td>F-test</td>
<td>3.22*</td>
<td>5.82**</td>
</tr>
<tr>
<td>p-value</td>
<td>[0.07]</td>
<td>[0.02]</td>
</tr>
</tbody>
</table>

*significant at 10 % level; **significant at 5 % level; ***significant at 1 % level; Std.dev. in parenthesis;

Small Ugandan firms exhibit lower growth rates than their larger counterparts. This is a peculiar aspect of the Ugandan data: patterns in firms’ growth rates actually go against some findings in the literature on firms dynamics (Evans [1987]; Hall [1987]). Interestingly, the quantitative model delivers patterns in growth rates that are close to the data. Small firms refrain from growing because of the rise in audit probability and regulation costs, and larger firms can experience higher growth rates because they take advantage of the accumulation of capital during the earlier phases of their lives.

**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 5.4 that the average capital-labor ratio of small 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 4.39: for any given shock $z$, an entrepreneur faces higher regulation costs the higher his 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. 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 such a situation larger firms may levy more debt if they have more capital.

Figure 5.3 shows the correlation of capital-labor ratios for all firms in the left panel and for small, medium and large firms with less than 150 employees in the right panel, using the simulated data from the benchmark. This figure should be compared to Figure 3.10. Figure 5.3 documents a clustering of firms at the size threshold of 30 employees where small firms substitute capital for labor. We also note the presence of a gap between 30 and 42 employees.

![Figure 5.3: Capital-Labor Ratios v.s. Firm Size](image)

---

18 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.

19 We show in the empirical analysis that small firms are heavily credit constrained and do not levy any debt while large firms have some access to credit in Uganda.

20 The results presented in this section are based on a simulation where there are no medium constrained firms. The reason for this is that, according to Table 3.5 the jump in audit probability between medium and large firms is not significant. Results based on a simulation containing medium constrained firms do not vary significantly from those presented here (available upon request). The only differences are 1) an upward slope on Figure 5.3 for medium firms because of a clustering of firms at 75 employees and 2) another gap in the size distribution of firms between 75 and 101 employees.
Column 1 and 2 in Table 5.6 present the results of OLS 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 follow the pattern observed in the data but are of smaller magnitude. The Wald statistic F1 rejects the null that the coefficients are jointly zero and the Wald statistic F2 rejects the null that the coefficients are equal both for the Ugandan data and the simulated data.

Table 5.6: Capital-Labor Ratio and Labor: OLS Results (Benchmark)

<table>
<thead>
<tr>
<th></th>
<th>Dependent Variable: ln(K/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data^c</td>
</tr>
<tr>
<td>Smaller firms</td>
<td>.80***</td>
</tr>
<tr>
<td></td>
<td>(.23)</td>
</tr>
<tr>
<td>Larger firms</td>
<td>.56***</td>
</tr>
<tr>
<td></td>
<td>(.13)</td>
</tr>
<tr>
<td>R_squared</td>
<td>.09</td>
</tr>
<tr>
<td>F1^b</td>
<td>10.54***</td>
</tr>
<tr>
<td></td>
<td>[0.0]</td>
</tr>
<tr>
<td>F2^c</td>
<td>3.90**</td>
</tr>
<tr>
<td></td>
<td>[.05]</td>
</tr>
<tr>
<td>Nb. Obs.</td>
<td>206</td>
</tr>
</tbody>
</table>

^a* significant at 10%; ** significant at 5%; ***significant at 1%; Std. Dev. in parenthesis and p-value in brackets;
^b F1: Wald test stat with \( H_0 \): coefficients are jointly zero;
^c F2: Wald test stat with \( H_0 \): the coefficients on size and capital are equal;
^d Smaller firms have 30 employees or less and larger firms have more than 30 employees;
^e 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. Equation 4.22 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.

**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.

### 5.4 Positive Analysis

Next, we examine the implications of the distortions we have added to the model of Hopenhayn [1992a]. 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 a jump in audit probability. Finally, case 4 is what we have described as the First-Best. In cases 3 and 4, we use a rental rate for capital \( r = 0.21 \) which is based on the average commercial banks lending rate in Uganda in 1997 as reported by the IMF in the 1998 Staff Country Report for Uganda. We discuss below some issues related to the sensitivity of the choice of \( r \).

Table 5.7 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 the appendix, Table A.1 gives the average and standard errors for the variables of Table 5.7.

**Pure Reallocative Effect on Labor** There are two assumptions in our model that deserve further comments. First, we have chosen a hand-to-mouth budget constraint and a log-utility function for workers. Second, we have assumed that any exiting firm is automatically replaced by a new entrant. These assumptions are made for the following reason. For a given disutility parameter, \( D \), aggregate labor supply must remain at the level of the benchmark in all subsequent experiments. To ensure that the labor market clears, we must thus adjust the wage in all experiments. This allows us to examine the reallocative effect of a distortion on labor across firms. More particularly, we focus solely on the effect of a distortion within a fixed
number of firms and a fixed number of employees, without having to worry about the effect of a distortion on the aggregate level of employees or firms. One caveat is that part of the effect of a distortion on productivity is captured by the adjusted wage. The reader should keep this in mind when interpreting the figures obtained for the change in value of output per worker in all experiments as these are lower bounds on the potential impact of a policy change.

Table 5.7: Model Validation

<table>
<thead>
<tr>
<th>Experiment Variable</th>
<th>Case 2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Case 3&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Case 4&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>wage</td>
<td>14.3</td>
<td>50.2</td>
<td>67.2</td>
</tr>
<tr>
<td>Share small</td>
<td>-18.7</td>
<td>49.7</td>
<td>28.2</td>
</tr>
<tr>
<td>Share medium</td>
<td>72.5</td>
<td>-22.9</td>
<td>59.7</td>
</tr>
<tr>
<td>Share Large</td>
<td>-.07</td>
<td>-50.1</td>
<td>-50.0</td>
</tr>
<tr>
<td>Agg. Capital</td>
<td>22.3</td>
<td>60.9</td>
<td>78.9</td>
</tr>
<tr>
<td>Agg. Output</td>
<td>6.1</td>
<td>50.5</td>
<td>55.1</td>
</tr>
<tr>
<td>Agg. Taxes</td>
<td>-100</td>
<td>50.5</td>
<td>-100</td>
</tr>
<tr>
<td>Agg. Bribe</td>
<td>-100</td>
<td>50.5</td>
<td>-100</td>
</tr>
<tr>
<td>Exit</td>
<td>0</td>
<td>50.1</td>
<td>50.1</td>
</tr>
<tr>
<td>K/L small&lt;sup&gt;d&lt;/sup&gt;</td>
<td>60.0</td>
<td>59.4</td>
<td>59.3</td>
</tr>
<tr>
<td>K/L medium&lt;sup&gt;d&lt;/sup&gt;</td>
<td>68.0</td>
<td>59.0</td>
<td>59.3</td>
</tr>
<tr>
<td>K/L large&lt;sup&gt;d&lt;/sup&gt;</td>
<td>53.9</td>
<td>59.0</td>
<td>59.3</td>
</tr>
</tbody>
</table>

<sup>a</sup> Case 2: No market for capital and no taxes;  
<sup>b</sup> Case 3: Perfect market for capital and audit distortion;  
<sup>c</sup> Case 4: Perfect market for capital and no taxes;  
<sup>d</sup> K/L are in millions of consumption unit per unit of labor;  
<sup>e</sup> See Table A.1 for the average and standard errors of each variable presented here.

Case 2

From Table 5.7 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 5.7 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 5.4 both clustering and gap in the capital-labor distribution have vanished in the graph for Case 2. This should again be compared to Figure 3.10. Examining the regression results with the simulated data from case 2 in column 3 of Table 5.6 we note that the coefficient on small firms is not significantly different than for medium firms (F2=1.88). 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 5.4:** Capital-Labor Ratios v.s. Firm Size (Case 2)

![In(K/L) v.s. Firm Size Case 2](image)

**Case 3**

Introducing perfect capital market entails large changes in aggregate variables. Table 5.7 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. Due to the regulation distortion, the average capital-labor ratio is higher for small firms than for medium firms. 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 5.5. Other implications from the theory can be noted from Figure 5.5. First, Figure 5.5 documents no dispersion within each productivity class. Observations for firms with a similar productivity shock lie on top of each other in Figure 5.5. Finally, let us compare the dispersion of firms’ size between Case 1, 2 and 3 to assess which of the distortion generates a higher dispersion of resources. The standard deviation of firms size equals 206 employees in the benchmark, 199 employees in Case 2 and 293 employees in Case 3. Based on these results it seems that the regulation distortion [credit constraints] increases [restrains] the dispersion of resources across productivity class. It is interesting to note that the standard deviation of firm size in the data is 259 employees, about halfway between the benchmark and Case 3.

**Figure 5.5:** Capital-Labor Ratios v.s. Firm Size (Case 3)
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 5.7). 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 identical 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 5.7. 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 proxy for the output that could be generated by a relatively undistorted economy. 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 benchmark and an undistorted economy by 92%.

Sensitivity to the Interest Rate

The results of Case 3 and 4 are obviously sensitive to the choice of the interest rate. We have thus run simulations for various interest rates where $r \in [0.04, 1.0]$ in the context of the First-Best (Case 4). Figure 5.6 plots productivity as measured as the level of output in real terms against interest rate. Productivity drops drastically over the range 0.04 to 0.2. The largest gains in productivity from the introduction of capital market are obviously happening at low interest rates in our model. Using an interest rate of 4%, the output is about twice the amount of aggregate sales found in Case 4. However, if we use an interest rate twice the size of Uganda’s commercial lending rate, say $r=40\%$, we observe a decrease in productivity of about 21%. It is interesting to note that, for interest rates beyond a value of $r=0.65$, gains in productivity are compromised and fall under a situation where entrepreneurs can finance investment with their own
savings, i.e. the benchmark economy. This is because we are depicting an extreme case where entrepreneurs must rent capital at the given rental rate.\footnote{One interesting problem on our research agenda is to examine the implications of a market for capital where entrepreneurs can either use their capital to finance their investment or as collateral to borrow from lending institutions.}

**Figure 5.6: Productivity versus Interest Rate**

![Productivity versus Interest Rate](image)

### 5.5 Normative Analysis: Policy Experiments

#### 5.5.1 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 5.8 where we tabulate percentage changes in aggregate variables relative to the benchmark. Capital-labor ratios are presented in levels.

**Table 5.8: Policy Experiments**

<table>
<thead>
<tr>
<th>Experiment Variable</th>
<th>Smooth Auditing (% from the benchmark)</th>
<th>Honest Bureaucracy (% from the benchmark)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau_c )</td>
<td>-</td>
<td>-.0463</td>
</tr>
<tr>
<td>wage</td>
<td>.2</td>
<td>.546</td>
</tr>
<tr>
<td>Share Small</td>
<td>-10.0</td>
<td>-3.3</td>
</tr>
<tr>
<td>Share Medium</td>
<td>41.1</td>
<td>24.7</td>
</tr>
<tr>
<td>Share Large</td>
<td>-.8</td>
<td>-3.5</td>
</tr>
<tr>
<td>Agg. Capital</td>
<td>1.0</td>
<td>45.6</td>
</tr>
<tr>
<td>Agg. Output</td>
<td>.3</td>
<td>15</td>
</tr>
<tr>
<td>Agg. Taxes</td>
<td>.3</td>
<td>0</td>
</tr>
<tr>
<td>Agg. Bribe</td>
<td>-7.4</td>
<td>-100</td>
</tr>
<tr>
<td>K/L small</td>
<td>61.5</td>
<td>71.2</td>
</tr>
<tr>
<td>K/L medium</td>
<td>68.2</td>
<td>97.3</td>
</tr>
<tr>
<td>K/L large</td>
<td>53.3</td>
<td>64.9</td>
</tr>
</tbody>
</table>

This table presents % 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.
5.5.2 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 or non-corrupt equilibrium. The equivalent of the aggregate bribe from the benchmark is now paid to tax officials out of tax revenues.\(^{22}\)

The effects of a change from a corrupt to a honest tax administration are summarized in column 2 of Table 5.8. 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 \( (\tau_c = -0.0463) \) and this in turn stimulates aggregate demand and entices entrepreneurs to scale up their operations. 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 output of 45.6% and 15% respectively. It is interesting to compare costs of inducing honest behavior to the gains of such a policy. In chapter 3, we have noted that the aggregate bribe amounts to 2.4% of aggregate sales. Thus the gains are about 6 times the costs.

Examining the capital-labor ratios across firm size we note that the increase in

\(^{22}\)One way to see this is that the equivalent of the aggregate bribe is being paid as a commission when bureaucrats bring back taxes to the tax agency. We assume that such a commission is sufficient to lead tax officials to collect tax revenues that were previously evaded and not accept bribes. However, we are aware that there is an issue of asymmetry of information. Ideally, the agency would see two reports: one from the firm and one from the tax official who audited that firm. Even in this case there is possible capture by one of the bargainers since they both have an incentive to pay the other to send a false report to the agency and pocket some of the liabilities.
aggregate capital is mainly due to an increase in capital accumulated by medium firms. This implies that the increase in production is also due in part 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%. However, the remaining large firms are larger than before and still produce the bulk of the output.

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 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 \( \theta \) on the burden from regulation \( \lambda^s \) is the same.\(^{23}\) If \( \theta = 0 \) then entrepreneurs who are audited pay double the amount of regulation costs they were paying in the benchmark.\(^{24}\) 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 in an aggregate efficiency loss.

This means that the degree of honesty of the bureaucracy and firms’ bargaining power must be targeted through a different set of policies. On the one hand, there are institutional changes that affect both the degree of honesty and the bargaining power of firms at the same time. For example, the appointment of a lax/rigorous judiciary is a signal to the rest of the governmental administration of a weaker/stronger enforcement of the rule of law. The degree of freedom of speech of journalists or the degree of independence of watchdogs such as the ombudsman or audit agencies would also affect how easily bureaucrats can conceal their deed, leaving entrepreneurs stronger or at a disadvantage in the bargaining process over bribe and evasion.

On the other hand, the degree of honesty of bureaucrats and firms bargaining power can also be targeted through different channels. Firms bargaining power in our model does not affect the level of evasion but only determines who keeps the surplus from evasion in the absence of other corruption deterrence mechanism.

---

\(^{23}\)Indeed, the burden from regulation is: \( \lambda^s = 1 - \tau_p - p_s^\prime (1 - q\theta)(\tau_0 - \tau_p) \)

\(^{24}\)Note that when \( \theta = 0 \), the second term in Equation 4.22 drops out. From the point of view of the entrepreneurs the effect of a change in \( q \) or \( \theta \) is the same.
As such, bargaining power does not stems from institutional factors only. There are other forces at play. Svensson (2003) suggests that bargaining power is linked to firm’s mobility. Firms with higher sunk costs have a harder time dealing with bureaucrats because these entrepreneurs cannot close their plant and move to another location. [Azam et al. 2004] argue that more productive entrepreneurs are also more efficient at bargaining. In this paper, we argue that another determinant of bargaining power is the degree of connectedness of an entrepreneur with governmental officials. As we have documented in the empirical section, larger firms pay less bribes per employee than medium firms even though they meet more often with bureaucrats. This probably results from the fact that owners of larger firms have been in the market for a longer time and have better relationships with the administration than smaller and younger firms. These channels are rather indirect and it might prove difficult to affect firms’ bargaining power. Contrarily, increasing the honesty of the bureaucracy through the wage/supervision structure is a more direct method to reduce evasion and corruption and it could entail clear efficiency gains as we have shown in this section.
Chapter 6

Conclusion

Cross-country differences in levels and growth rates of output are large and persistent (Hall and Jones 1999, Klenow and Rodriguez-Clare 1997). As argued by Prescott [1998] and Easterly and Levine [2001], much of this variation is attributable to heterogeneity in total factor productivity (TFP). In order to develop a theory of TFP, increasing attention has been devoted to the effect of institutional distortions on the reallocation process of productive resources across productive units and efficiency (Restuccia and Rogerson 2008, Alfaro et al. 2007). Restuccia and Rogerson [2008] emphasize the necessity to restrict the magnitude and types of distortions studied, to obtain better measures of specific distortions and to evaluate their aggregate consequences. This was the purpose of this thesis. We have used micro-level data on Ugandan firms to identify two typical distortions in LDCS and investigated the extent to which these distortions result in a mis-allocation of resources and a loss in measured productivity.

In Chapter 3, we have identified, among competing explanations, the driving force behind the allocation of resources in Uganda: a sharp increase in the probability of being audited for taxes around a size threshold of 30 employees. We have also shown evidence that regulation costs such as tax and bribe payments increase significantly for medium firms (between 31 and 75 employees). Not only this translates in a significant break in the density distribution of firm sizes around 30 employees but the data also suggests that entrepreneurs, who would otherwise grow their firms beyond that size threshold, choose ex-ante suboptimal capital-
labor allocations and that smaller firms exhibit lower growth rates of output than their larger counterparts. We have argued that only relatively large productivity shocks outweigh the costs of increasing employment beyond 30 employees. This results in reduced output growth at lower size levels, a clustering of firms at or below the 30-employees threshold and to a significant discontinuity in the level of capital-labor ratios of firms above or below the threshold. The data has allowed examining two other potential explanations: 1) a price differential across factors and across firm sizes and 2) a more labor-intensive technology for smaller firms. However, these evidence go against the pattern in capital-labor ratios that we observed for small and medium firms.

In Chapter 4 we have presented an adaptation of Hopenhayn [1992a]'s growth model of heterogeneous firms with entry and exit to include the salient features of the Ugandan environment described in Chapter 3. Consistent with the empirical evidence we have assumed that entrepreneurs evade part of their tax liabilities, that tax officials are corruptible and that, with some probability, they accept bribes when auditing entrepreneurs. We have introduced a frequency of tax audits that is increasing with size. Finally, we have assumed that entrepreneurs must finance their investment in capital with their own savings. Once we have defined what is an equilibrium for our benchmark economy, we have noted an implication of the Euler equation in this case: entrepreneurs invest at a decreasing rate due to the fact that the audit probability increases with size. We have then compared the benchmark economy to an economy without distortions in order to examine each distortion separately. There were 3 implications stemming from this First-Best case. First, the first-best capital-labor ratio is constant across firms. Second, the size-dependent auditing scheme acts as a progressive tax: it decreases the amount of labor employed at any given firms but relatively more in larger firms. Third, when firms must use their own cash-flow to finance their investment, there is a non-degenerate distribution of resources across firms within a productivity class (i.e., a given level of productivity shock).

Finally, we have introduced some simplifying assumptions to solve the model analytically and derived comparative statics for the parameters of interest. We have shown that an increase in the output price [input prices] increases [decreases] allocations of capital and labor in all sectors as well as the level of bribes paid, and
it alleviates [exacerbates] the clustering of firms below the size-threshold where the audit probability increases. A change in firms’ bargaining power over the surplus from tax evasion affects proportionally the allocations of capital and labor at firms dealing with tax officials. However, such a change has an ambiguous effect on the value of bribes paid as it depends on the magnitude of the elasticities of the inputs with respect to the bargaining power. We have documented similar patterns for the effect of a change in the tax rate. A change in the size-threshold where the audit probability increases does not affect how firms choose their inputs except for those firms right at the size-threshold. Since the bribe is a function of capital and labor it is thus not affected by such a change. An increase in the audit probability reduces the allocations of capital and labor as well as the size of the bribes paid at firms involved with tax officials and it increases the clustering of firms at the size-threshold. Finally, the expected fine faced by tax officials who get caught receiving a bribe does not affect the allocations of capital and labor. However, an increase in the expected fine increases the level of bribes paid which in turn entices firms to cluster at the size-threshold.

In Chapter 5 we have estimated the full-blown version of the model. The parameters for the regulation environment were calibrated using firm-level information from Ugandan data. We have calibrated the technology parameters based on the existing literature. Finally, we have assessed the values of the parameters for the processes of firms survival and transition of shocks using an Indirect Inference approach. This simulation-based method allowed generating a set of over-identifying restrictions to test the fit of the model. The test of over-identifying restrictions rejects the null that the moments from the simulated data are equal to the moments from the observed data. There are a few potential reasons for the rejection of the model that should be noted. First, the model overestimates the size of the missing middle on Figure 5.1 and 5.2. This is due to the stark depiction of the audit intensity (a step function) imposed in our model. The use of a polynomial to describe the audit intensity will probably reduce the gap observed between the simulated and the actual missing middle. Second, the pattern in simulated capital-labor ratio for small and medium firms mimics the data but the pattern for medium and large firms is quite different in the simulated than in the actual data. A better description of the interaction between inefficient credit markets and the audit distortion is in
order for future work.

We noted however that the model does a reasonable job at reproducing some features of the data such as patterns by firm size in capital-labor ratios and output growth. We also observe in the simulated data a clustering of small firms just before the exogenous size threshold. According to our theory these small firms substitute capital for labor while waiting for a productivity shock that will offset the cost of growing. These choices generate a gap in the size distribution of firms and lead to the so-called missing middle.

We have examined the theoretical implications of the model by sequentially removing the regulation distortion and the missing credit market assumption. In this way we have shown that the increase in audit probability is key in generating the clustering of firms at the size-threshold. We found that removing all regulation distortions would close the gap in output per worker between the Ugandan benchmark and the First-Best economy by 11%. However introducing perfect capital markets generates much larger gains and explains roughly 92% of the output difference between the benchmark and the undistorted economy. One important conclusion from this work is that firms adjust their input mix optimally to avoid productivity losses in the presence of a local institutional distortion such as a rise in the audit probability around a size-threshold and the resulting firms size distribution is a by-product of their optimal responses. In comparison, it seems that inefficiencies associated to credit markets generate a much greater efficiency loss and attention should somewhat be redirected toward financial frictions in trying to increase TFP.

Having established the ability of the model to rationalize some relevant features of the data, we have designed and implemented two counterfactual experiments. First, we have shown that only small efficiency gains are associated to 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 have changed wage compensation of auditors so as to make them indifferent between being honest or corrupt. We have argued that bribing may serve as a useful second-best mechanism in that it buys entrepreneurs a tax rebate. However, a wage that makes auditors in-
different between being honest or corrupt brings about larger efficiency gains than corruption. Finally, we have examined how these policy changes may affect the gap in output per worker between Uganda and an undistorted economy. Flattening the distribution of audit probabilities would only reduce this gap by a mere 0.5% while the no-corruption wage would reduce this gap by 27%. In order to study the pure reallocation effect of the policies aggregate labor supply was held constant throughout all our simulations. Therefore part of the productivity variation from each policy change is captured by the variation in the adjusted workers’ wage. The aggregate output gains must thus be interpreted as lower bounds on the potential effects of these policies.

Before closing this thesis we remind the reader and ourselves that this work presents some limitations. These limitations should not be regarded as failures but as opportunities to further our knowledge about this topic. First, the modelisation of the size threshold exactly at 30 employees is a rather stark depiction of the reality. From a close examination of the data, it is more appropriate to argue that the jump in the audit probability and its consequences lie somewhere in the vicinity of 30 employees. The use of a polynomial for $p_a$ which would depend on labor could be an alternative approach to develop the framework of Chapter 4 and Chapter 5. However, we show in Chapter 3 that labor is not a statistically significant determinant of the audit probability on the entire size domain. Moreover, it is not clear beforehand whether a polynomial would allow generating the capital-labor substitution that we observe in the data. Our discrete depiction of the distortion associated to the rise in audit probability thus seems a reasonable approximation.

Another strong assumption in the model is the imposition of the same rental market for capital across firm size. We have shown in Chapter 3 that small firms are credit constrained while larger firms have access to some credit in Uganda. One reason why we have imposed the same constraints on all firms is due to our focus being on small and medium firms. Another reason is to simplify the code for the numerical exercise. Letting the market for capital function adequately beyond a certain firm size entails huge efficiency gains. Starker institutional constraints would thus be needed to generate a missing middle of the order that we observed in the Ugandan data. Nonetheless, it is reasonable to assume that the missing middle in Uganda is generated by a mixture of both institutional constraints and differen-
tial in factor prices. The first force refrains firms from growing while the second one entices entrepreneurs to pick up size. Another extension of the model could consider the explicit modelling of financial markets. This would allow examining intermediary cases of financial development. Based on Buera et al. [2009], we will develop a setup where entrepreneurs can either use their capital or rent it from the capital market. However, in this version of the model, the level of financial development of a country will dictate how much capital entrepreneurs can borrow based on an incentive compatibility constraint. One interesting outcome in this line of research would be the observation of an endogenous size threshold beyond which entrepreneurs have just enough collateral to start borrowing, as observed in the data.

We note that even if the test of overidentifying restrictions fails to reject the null that the moments from the simulated data are equal to the moments of the actual data the model cannot be interpreted as being the true DGP. However one step in the right direction would be to compare and rank different mispecified models, nested or non-nested, in terms of their fit to the data. Hnatkovska et al. [2008] have developed a procedure to conduct such comparisons. In Chapter 5 we have compared, albeit informally, the model with and without the institutional constraint and with and without a market for capital that functions adequately. A formal test of the ranking of these different channels in terms of their fit to the data is beyond the scope of this thesis and left for future work.

One outcome of this thesis is that it sets the basis for an ambitious and diversified research agenda. 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 where the audit probability increases. Such a phenomenon is not specific to the environment at hand. One should suspect similar clusters to arise wherever a sharp change in policy conflicts with agents’ incentives. For example, one should observe a clustering in the distribution of agents when a steep change in unemployment benefits policy or in a progressive tax scheme on individual income conflicts with agents incentives to work. Lemieux and Milligan [2008] report such

---

1 We discuss in Chapter 3, based on a setup where collateral is needed to borrow, why a rental rate differential in favor of larger firms would induce patterns in capital-labor ratios that go against what we observe in the data for small and medium firms.
evidence for a sharp discontinuity based on age in the assignment of social assistance benefits in Quebec and find that more generous benefits reduce employment. A model of optimal behaviour could shed light on the mechanism that leads the recipients of assistance benefits to opt out from work. Our model can be adapted to other contexts. Another extension to our model is to endogenize the size-threshold around which the clustering takes place. This would allow examining the optimal decision of the government in setting its taxation objectives. Our model could be extended to include an early stage in the timing schedule where individuals would decide whether to become entrepreneurs or rent-seeking bureaucrats based on the expected payoffs of each career. Such a framework would allow examining the impact of remuneration incentives on the relationship between the allocation of resources and the allocation of talent.
Bibliography


153


154


Appendix A

A.1 Empirics

![Graph showing Local polynomial smooth of Audit Intensity v.s. Firm Size.](image)

**Figure A.1:** Audit Intensity v.s. Firm Size in Uganda in 1997
Figure A.2: Density of Firm Size with a Bandwidth of 5 Employees

Figure A.3: Capital-Labor Ratio v.s. Firm Size in Uganda in 1997
A.2 Theory

A.2.1 Effect of $l^c$ on $i^c$

We need to compare:

$$
\left| \frac{d i^c}{d l^c} \right| \gg \left| \frac{d z^f}{d l^c} \right|
$$

$$
\left| -G(z^c | z) \left\{ \frac{1}{p_F} \left[ \frac{r}{\gamma} \right] \left[ \frac{w}{\eta} \right]^{1-\gamma} \right\}^{\frac{1}{1-\gamma-\eta}} \right|

\gg \left| G(z^f | z) \frac{p_F (z^f)^{1-\gamma-\eta} \eta (\lambda^c (z^f))^{\gamma} (\mu^c)^{\eta-1} - w}{p_F (1 - \gamma - \eta) (z^f)^{-\gamma-\eta} GG} \right|
$$

Using a uniform distribution, we have:

$$
\left| -z^c \left\{ \frac{1}{p_F} \left[ \frac{r}{\gamma} \right] \left[ \frac{w}{\eta} \right]^{1-\gamma} \right\}^{\frac{1}{1-\gamma-\eta}} \right|

\gg \left| z^f \frac{p_F (z^f)^{1-\gamma-\eta} \eta (\lambda^c (z^f))^{\gamma} (\mu^c)^{\eta-1} - w}{p_F (1 - \gamma - \eta) (z^f)^{-\gamma-\eta} GG} \right|
$$

Replacing for $z^c$ on the LHS and simplifying on the RHS:

$$
\left| - \frac{(z^c)^2}{l^c} \right| \gg \left| \frac{p_F z^f \eta (\lambda^c (z^f))^{\gamma} (\mu^c)^{\eta-1} - w (z^f)^{1+\gamma+\eta}}{p_F (1 - \gamma - \eta) GG} \right|
$$

Rewriting the RHS only in terms of shocks:

$$
\left| - \frac{(z^c)^2}{l^c} \right| \gg \left| \frac{\eta p_F^{\frac{1}{\gamma}} (z^f)^{\frac{1-\gamma+\eta}{\gamma}} \left( \frac{r}{\gamma} \right)^{\gamma} (\mu^c)^{-\lambda} - w (z^f)^{1+\gamma+\eta}}{p_F (1 - \gamma - \eta) GG} \right|
$$

or,

$$
\left| - \frac{(z^c)^2}{l^c} \right| \gg \left| \frac{\eta p_F^{\frac{1}{\gamma}} (z^f)^{\frac{1-\gamma+\eta}{\gamma}} \left( \frac{r}{\gamma} \right)^{\gamma} (\mu^c)^{-\lambda} - w (z^f)^{1+\gamma+\eta}}{p_F (1 - \gamma - \eta) GG} \right|
$$

which we rewrite as:

\[161\]
\[ \left| \frac{-\left(\frac{z_{ic}}{l_{ic}}\right)^2}{l_{ic}} \right| \leq \left| \frac{\eta p_F \gamma \left(\frac{z^f}{r} \right)^{\gamma + \frac{1}{1-\gamma}} \left(\frac{z_{ic}}{l_{ic}}\right)^{\frac{1}{1-\gamma}} \left( l_{ic}\right)^{-\lambda} - w \left(z^f\right)^{1+\gamma + \eta} }{(1 - \gamma - \eta) \left( \left(\frac{z^f}{r}\right)^{\gamma + \eta} [HH p_F \left(\frac{z^f}{r}\right)^{\gamma} \left(\frac{\eta}{w}\right)^{\eta}] \right)^{\frac{1}{1-\gamma}}} \right| \]

The LHS and RHS of the above expression are functions of parameters only.\(^1\)

### A.2.2 Effect of \(r\) on \(ic\)

To sign the last expression, we need to compare:

\[ \frac{\frac{d}{dz} \frac{dz_{ic}}{dz} \frac{dz^f}{dr}}{\frac{dz_{ic}}{dz} \frac{dz^f}{dr}} \leq \frac{\frac{G(z_{ic})}{dz_{ic}} \frac{dz^f}{dr}}{\frac{G(z^f)}{dz} \frac{dz^f}{dr}} \]

Using a uniform distribution and plugging the expression for each terms:

\[ \left| -\frac{z_{ic}}{1 - \gamma - \eta} r^{\frac{1}{1-\gamma}} \left[ \frac{1}{p_F} \left(\frac{1}{\gamma} \right)^{\gamma} \left(\frac{w}{\eta}\right)^{1-\gamma} \right]^{\frac{1}{1-\gamma}} \leq \left| \frac{z^f}{p_F \left(1 - \gamma - \eta\right)} \frac{\left[k^f - k_{ic}\right]}{GG} \right| \]

Further simplifying:

\[ \left| -\frac{\left(\frac{z_{ic}}{l_{ic}}\right)^2}{\frac{r}{l_{ic}}} \right| \leq \left| \left(\frac{z^f}{r}\right)^{1+\gamma + \eta} \left[ \frac{k^f - k_{ic}}{p_F GG} \right] \right| \]

Replacing term GG, we get\(^2\)

---

\(^1\)We show in an earlier version of the paper that the clustering effect is exacerbated by an increase in the size-threshold. Results available upon request.

\(^2\)We show in an earlier version of the paper that above a certain value for \(r\), there is only an informal unconstrained sector. Thus the clustering effect is exacerbated by an increase in the rate of rental of capital: the decrease in proportion of the formal sector is not compensated by the increase of the proportion of the iu-sector.
\[-\left(\frac{z_f^c}{r} \right)^2 \frac{\gamma_f}{r} \leq \left(\frac{z_f^c}{r} \right)^{1+\gamma_f + \eta} \frac{[k_f^c - k_i^c]}{\left(\left(\frac{z_f^c}{r} \right)^{\gamma_f + \eta} \left[H H p_F \left(\frac{\lambda^c}{r} \right)^{\gamma_f + \eta} \right)^{\frac{1}{r}} - \left[p_F \left(\frac{r}{z_f^c} \right)^{\gamma_f + \eta} \right)^{\frac{1}{r}} \right]}{1 - \gamma_f + \eta}\right]
### A.3 Numerical Results

#### Table A.1: Model Validation, Coefficient and Standard Errors

<table>
<thead>
<tr>
<th>Experiment Variable</th>
<th>Benchmark</th>
<th>Case 2&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Case 3&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Case 4&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>wage</td>
<td>0.526</td>
<td>0.601</td>
<td>0.789</td>
<td>0.879</td>
</tr>
<tr>
<td>Share small</td>
<td>0.469</td>
<td>0.3812</td>
<td>0.7019</td>
<td>0.6014</td>
</tr>
<tr>
<td>Share medium</td>
<td>0.1216</td>
<td>0.2097</td>
<td>0.937</td>
<td>0.1942</td>
</tr>
<tr>
<td>Share Large</td>
<td>0.4094</td>
<td>0.4091</td>
<td>0.2044</td>
<td>0.2044</td>
</tr>
<tr>
<td>Average Capital</td>
<td>151</td>
<td>185</td>
<td>244</td>
<td>271</td>
</tr>
<tr>
<td></td>
<td>(205)</td>
<td>(239)</td>
<td>(572)</td>
<td>(632)</td>
</tr>
<tr>
<td>Average Output</td>
<td>131</td>
<td>139</td>
<td>198</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>(218)</td>
<td>(220)</td>
<td>(465)</td>
<td>(476)</td>
</tr>
<tr>
<td>Average Taxes</td>
<td>9.4</td>
<td>0.0</td>
<td>14</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>(15.7)</td>
<td>(0.0)</td>
<td>(33)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>Average Bribe</td>
<td>3.2</td>
<td>0.0</td>
<td>4.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td>(5.4)</td>
<td>(0.0)</td>
<td>(11.5)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>Exit</td>
<td>724</td>
<td>724</td>
<td>1087</td>
<td>1087</td>
</tr>
<tr>
<td>K/L small</td>
<td>83.5</td>
<td>75.5</td>
<td>59.4</td>
<td>59.3</td>
</tr>
<tr>
<td></td>
<td>(89.2)</td>
<td>(103.6)</td>
<td>(2.02)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>K/L medium</td>
<td>77.5</td>
<td>85.6</td>
<td>59.0</td>
<td>59.3</td>
</tr>
<tr>
<td></td>
<td>(78.9)</td>
<td>(95.9)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
<tr>
<td>K/L large</td>
<td>66.7</td>
<td>67.8</td>
<td>59.0</td>
<td>59.3</td>
</tr>
<tr>
<td></td>
<td>(59.3)</td>
<td>(71.8)</td>
<td>(0.0)</td>
<td>(0.0)</td>
</tr>
</tbody>
</table>

Case 2: No market for capital and no taxes; [b] Case 3: Perfect market for capital and audit distortion; [c] Case 4: Perfect market for capital and no taxes; [d] Capital, Output, Taxes, Bribe and K/L are in consumption units; [e] Standard Error in parenthesis.