

ECONOMIC
RESEARCH
FORUM



منتدى
البحوث
الاقتصادية

2014

working paper series

FIRMS INFORMALITY: A MODEL
AND EMPIRICAL EVIDENCE FOR LEBANON

Rawaa Harati

Working Paper No. 883



FIRMS INFORMALITY: A MODEL AND EMPIRICAL EVIDENCE FOR LEBANON

Rawaa Harati

Working Paper 883

December 2014

We thank the ERF for the use of data, and Konstatin Shamruk for the useful comments. The views expressed here are those of the authors and do not necessarily reflect the views of the Economic Research Forum.

Send correspondence to:

Rawaa Harati

Universite Paris I Pantheon-Sorbonne

rawaa.harati@malix.univ-paris1.fr

First published in 2014 by
The Economic Research Forum (ERF)
21 Al-Sad Al-Aaly Street
Dokki, Giza
Egypt
www.erf.org.eg

Copyright © The Economic Research Forum, 2014

All rights reserved. No part of this publication may be reproduced in any form or by any electronic or mechanical means, including information storage and retrieval systems, without permission in writing from the publisher.

The findings, interpretations and conclusions expressed in this publication are entirely those of the author(s) and should not be attributed to the Economic Research Forum, members of its Board of Trustees, or its donors.

Abstract

This paper tests the implications of a simple equilibrium model that describes the informal economy in Lebanon. The results suggest that informal firms are less productive, managed by less educated entrepreneurs, are smaller and face less inspection from the tax department and the social security audits, which we verify in the data. The paper also investigates how the quality of legal enforcement, captured by the likelihood of detection in the informal sector, affects firms' decisions to be formal or informal. Since enforcement may be endogenous, we instrument this variable with the distance between the city where the firm is located and the closest enforcement offices.

JEL classification: H2, H3, K4

Keywords: Informal market, Development Economics, Tax avoidance, MENA

ملخص

تختبر هذه الورقة الآثار المترتبة على نموذج توازن بسيط وهو أن يصف الاقتصاد غير الرسمي في لبنان. وتشير النتائج إلى أن الشركات غير الرسمية هي أقل إنتاجية، تدار من قبل رجال الأعمال الأقل تعليماً، هي أصغر حجماً وأقل تواجه التفتيش من مصلحة الضرائب ومراجعة الحسابات الضمان الاجتماعي، وهو ما تحقق في البيانات. تحقق الورقة أيضاً في نوعية إنفاذ القوانين، التي استولت عليها من احتمال الكشف في القطاع غير الرسمي، ويؤثر على قرارات الشركات على أن تكون رسمية أو غير رسمية. منذ إنفاذ قد يكون الذاتية، فإننا الصك هذا المتغير مع المسافة بين المدينة التي تقع فيها الشركة وأقرب مكاتب الإنفاذ.

1. Introduction

One feature of developing economies is the importance of untaxed and unregulated activities. Recent estimates of the size of the informal economies as a percentage of the GDP suggested high statistics in developing countries, reaching 60% in some of them (Schneider 2000).

Informality of enterprises represents a lag in the development of these countries, as informal firms are usually smaller, employ less workers, generate less income, and are less capital intensive, i.e. are less involved in production of manufacturing goods, helping to boost the competitiveness of overall economy, let alone the imposed burden on public finances.

Also in Lebanon, research on this topic is scarce and like most countries in the region, Lebanon's economy is dominated by MSE's and the 1996 census of buildings and establishments conducted by the Central Administration of Statistics (CAS) evaluates that cumulatively, enterprises with less than 50 employees generate the majority of employment opportunities in the country, so MSE's are a major partner in the Lebanese development, and a large part of the jobs created in this sector are informal (ERF research report). Hence one important motivation of this study.

We follow the model by De paula and Scheinkman (2011), that links the probability of firm's informality to the ability of the entrepreneur, using a more realistic form of probability function, and we show how formal status affects the size and capital intensity of a firm in Lebanon.

Furthermore we allow for the possibility that enforcement is related to firm's characteristics, which creates a non-trivial relationship between formality, entrepreneur's ability and firm's size, and may induce a bias in the De paula- Scheinkman estimations. We treat this problem by controlling for the probability of inspection and the size of fines in the model. We study the robustness of our results to different proxies of the controls. We also attempt to gauge the effect of degree of enforcement on the size of informal economy: Our result show that in Beirut and MontLiban the two biggest Lebanese departments, enforcement negatively affects the degree of informality. Most of our results are verified in the literature.

An extensive array of literature—both theoretical and empirical—studies the causes of informality. The literature shows evidence that points to some determinants of informal activity. Loayza and Rigolini (2006) argue that regulatory burdens may be a leading cause of informal activity. Chong and Gradstein (2004) focus on inequality and poor institutions. Perry et al. (2007) find a significant relationship between the number of times that a firm is visited by tax inspectors and the fractions of sales and workers that go unreported. Almost all scholars find an important entry cost for firms and workers into the formal sector. This cost is not only financial a financial burden in the form of paying taxes, but also a cost of time because being formal can also be time consuming. Doing Business Reports of the World Bank show that the necessary procedures for registering and getting a license in order to be legally formal can take several weeks in some developing countries while taking only a couple of days or hours in most developed countries. Furthermore, usually, bribing officials is necessary to get things done. To avoid these costs, many firms prefer to stay informal. The typical answers given by informal agents in developing countries to explain the reason why they choose to stay informal are: first, the services offered by the government in exchange of the payment of taxes, as infrastructure and public services, are deemed inadequate to the tax contributions of the private sector; second, the taxpayers are rarely confident in the officials' benevolency, which makes them underreport or not report at all their revenues. Also low monitoring and enforcement may tempt firms to avoid paying taxes either because they do not expect to be audited or because they can always find a way to bribe the tax collector to avoid paying the full amount of tax.

This relation between the size of the underground economy, the higher tax burden, poor institutions and regulatory requirements was studied by Loayza (1996), Johnson et al. (1997), Friedman (1995) and many others. The literature also emphasized the importance of financial development and the credit market on the size of the informal market and the decision to evade taxes (see Straub 1991). Studying the causality between informality and performance of firms has been especially difficult with regards to the direction of the causality. Loayza and Rigolini (2006) take the average growth of countries over 20 years as a dependent variable, and estimate the impact on growth of informality for over 100 countries. The effect is found to be negative and significant. However, the authors acknowledge this relationship may not necessarily be a causal one; it could be that informality is a symptom of other causes which are themselves at the heart of a weak economic growth rate. La Porta and Shleifer (2011) in their study on 24 African countries, conclude that high productivity comes from formal firms and in particular large formal firms, and productivity jumps sharply when comparing small formal firms to informal firms. So even if it is unclear how the direction of the causality goes, it is well known that informal firms are less productive than formal firms.

The questions we try to answer in this paper are the following: What are the determinants of informality for Lebanese MSE's and how does enforcement influence the enterprise's decision to operate formally or informally? The latter was studied in the literature for crimes and punishment (Becker 1968; and Friedman 1995). The same argument is still valid: if we design institutions that can eliminate all informal firms and workers and turn them into formal, this would no doubt be a good thing, but this decision depends on how much this costs. If reducing the creation of informal enterprises from 10 to 0 requires turning half the population into investigators, it is probably not worth it. Also, compliance can most cheaply be assured with a small probability of detection and large penalties, but law or social convention may prohibit very large penalties (see Scotchmer 1987), also as Feld and Schneider (2010) argue in survey studies on undeclared work, the perceived probability of being detected has a consistent robust significantly negative effect, but for perceived fines and punishment the results are less robust. So it is still a challenge for the enforcement agency to decide on the audit strategy and the level of penalties that minimize its cost and maximize the number of formal firms.

In this paper we try to construct a simple model on enterprise informality that includes the probability of detection and the fines informal firms face after inspection, and we try to generate some empirically testable predictions.

Some attempts were made (De paula and Scheinkman 2011; and Dabla-Norris et al. (2008)) to construct models of informality. These papers were a variant of the model of Dabla-Norris et al. (2008), and to some extent Allingham and Sandmo (1972). But for simplification, these papers assume that the probability of detection is exogenous, and when an informal firm is detected, it loses all its profit. This means that small firms (with small capital) are never detected, but big firms are always detected and once they are, all their profit is taken by the government. This leaves us with an economy where all small firms are informal and all big firms are formal, which can diverge from reality especially in developing countries where we can always find some small firms which are formal for a reason, and also big firms who are completely or partially informal.

We argue that this paper has a more realistic form of the probability of detection. Our model draws from De paula and Scheinkman (2011). Enterprises are considered informal if they avoid paying taxes. We assume that firms use labor and capital as inputs and each firm has its own productivity which can reflect the abilities of the manager (education, experience) and other factors like the accumulation of skills, experience and know-how within the firm (De paula and Scheinkman 2011). Probability of detection is increasing with capital that

represents the visibility of the firm. We do not consider a cut-off in a certain level of capital where all firms are formal or informal. It is true that we assume this probability function increasing in capital but we can observe small formal firms or big informal firms, but the smaller the firm is, the lower the probability of detection it faces and the biggest the firm is, the biggest is the probability of detection. Our penalty function depends on the net revenue of the firm after wage payments. This form is for simplification. We assume that informal entrepreneurs own the capital, so in case of detection the penalty will be taken from their profit net of wages. Several implications of this model are tested empirically with the Lebanese Micro and Small enterprise survey. Finally this paper tests the effect of enforcement on informality, using the geographical distance between the firm and the inspection office as an instrument, in an attempt to correct for endogeneity.

The rest of the paper is organized as follows. In section 2 we develop the model, in section 3 we present the empirical results, and section 4 concludes.

2. A Model of Firms' Informality

We try to construct a simple model so we can generate empirically testable predictions. Following De paula and Scheinkman (2011), we suppose that an entrepreneur chooses his formal status by comparing his profits as a formal or informal business. In both cases his firm uses capital and labor as inputs. We assume that entrepreneur is a proprietary of his capital and when formal, pays a share of his sales revenue to the government

$$\Pi_f = (1 - \tau)\theta k^\alpha l^\beta - wl - rk$$

Informal firms do not pay taxes but run a risk of being inspected and forced to pay a penalty, which can be thought of as a real fine defined by the law, or a bribe that the investigator takes to turn a blind eye to the firm's informality.

$$\Pi_i = [1 - p(k)](\theta k^\alpha l^\beta - wl) - rk$$

where l is the amount of labor employed, k is the capital, w denotes the wage rate and r the capital cost. We assume that factor costs are the same for formal and informal firms. θ is the firm's productivity, which includes the entrepreneur's experience, education, etc.

The probability of being detected depends on the size of the firm and I measure this size by the amount of capital the firm has. Even if other ways of measuring the size are possible, we argue that capital stock (installations...) make the firm more visible, and that a firm cannot hide a big stock of capital before an inspection (De paula and Scheinkman, 2011). In case of fraud detection, which happens with probability $p(k)$, the authorities confiscate sales revenue net of wages, assuming that the fine's payment comes after the salaries are paid to the employees. $p(k)$ is an increasing function in k and $1 - p(k)$ is therefore the probability of not being detected.

For simplicity we assume that $p(k) = 1 - k^{-\gamma}$, where γ captures the quality of the country's legal framework, the higher it is the more likely that each informal institution is identified by the authorities (see figure 1). Figure 1 also shows De paula et al.'s probability of detection where all small firms are informal with probability of detection = 0, while all big firms are formal with a probability of detection=1.

Under this configuration, a profit that a rational entrepreneur can extract from a formal and informal enterprise is a function of factor input prices, tax rate, legal framework quality and entrepreneurial ability:

$$\log \Pi_f = \alpha_{0f}(r, w, \tau, \alpha, \beta) + \alpha_{1f} \log \theta$$

$$\log \Pi_i = \alpha_{oi}(r, w, \alpha, \beta, \gamma) + \alpha_{iv} \log \theta$$

The relation between these profits is a direct determinant of the formal status of entrepreneur:

$$\Pi_f = \text{Prob}(\log \frac{\Pi_f}{\Pi_i} > 0)$$

$$\log \frac{\Pi_f}{\Pi_i} = \xi_0(r, w, \alpha, \beta, \gamma, \tau) + \xi_1(r, w, \alpha, \beta, \gamma, \tau) \log \theta$$

The effect of law enforcement on the effect of the entrepreneur's decision to be formal or

informal: $\frac{\partial^2 y}{\partial E \partial \theta} = \frac{1 - \beta}{(\zeta + (1 - \beta)\gamma)^2}$

This coefficient is positive which means that law enforcement increases the effect of entrepreneur's ability on the formality (without law enforcement, we have more educated and capable entrepreneur tend to run formal firms but with law enforcement, even less capable entrepreneur would run formal firms because of the effect of law enforcement)

The coefficient of $\log \theta$ is positive, which means that more productive firms are working in the formal sector.

The capital-labor ratio in this configuration is clearly negatively related to the level of managerial ability:

$$\log \frac{k}{l} = \phi(r, w, \alpha, \beta, \gamma)$$

Capital and labor grow at a slower rate in the informal sector than in the formal sector when managerial ability increases.

The number of employees is a function of factor costs, legal framework and ability:

$$\log l = \psi(r, w, \alpha, \beta, \gamma)$$

The algebraic derivations are given in the appendix.

3. Empirical Application

3.1 Data

Our database is the Micro and Small Enterprises (MSE) which was collected by the Economic Research Forum (ERF) for the Project on Promoting Competitiveness on Micro and Small Enterprises in the MENA Region where MSEs constitute on average 90% of the number of enterprises. The database is the result of a unique field survey performed on the micro and small enterprises (MSEs) in four selected countries of the MENA region: Egypt, Lebanon, Morocco and Turkey. In this paper we only use the Lebanese data. It should be noted that this survey was performed under difficult conditions. The non-availability of (or access to) the basic national database constituted one of several challenges that the Lebanese team managed to overcome. If anything, there is no doubt that the database gathered on the MSE sector in Lebanon will make a substantial contribution and would fill a gap at the national and regional level. Two methodological tools were applied in this study, including a sampling approach in order to obtain a representative sample of enterprises taking into account size of enterprises, geographical distribution and other variables. Also, the sample correction was used in order to generate results at the national level (Hamdan 2003). The survey contains information on households and enterprises gathered from approximately 3,000 questionnaires that were filled by MSEs in Lebanon. The surveys were performed between 2002 and 2004 and provide estimates for the key indicators related to activities,

characteristics of entrepreneurs, legal status of the firm (registration to tax authorities, license acquisition, social insurance), performance of the firm and information on the workers inside these enterprises. The collected data excludes agricultural activities, illegal activities, non-marker activities, domestic services, production for own use, enterprises employing more than 50 workers.

3.2 Description of variables

Firms with owners who were less than 15 years old were eliminated and also those observations where information on gender or education were lacking. We ended up with a sample of 2638 enterprises. Table 1 contains sample statistics of the variables we consider. Our variable of formalization is a dummy variable that equals one if the firm is registered with the tax authorities and zero otherwise. We created another variable using the social security instead of the tax registration to define informality. Finding a universal definition of informality is hard task: Some studies use the social security approach when considering the workers' side and other studies rely on tax registration. We have both variables in our data so we decide to use both separately to compare and prove one more time that each definition can lead to different results, which is a big complication in the informality literature. This is for robustness checks and also because we are convinced that informality concerns the regulatory status of the employees as well as firms' registration with the tax department. For instance, a firm that is registered with the tax department but does not allow its employees to have social security is still informal. Table 2 shows the percent of firms in each situation. 40 % of these firms pay taxes and have their employees registered with social security. But around 31% of the firms either pay tax and do not declare the employees or the opposite. This is a very common problem in developing countries and it makes defining and measuring informality extremely difficult. This can also give us an idea how efficient enforcement and regulation are in these countries. The reason why such a table exists is because Lebanon is a country where the tax department and the social security do not work together and do not share data. So a firm can be registered with the tax department and hide its employees to avoid registration with social security, and there is no way for the social security to know about it. Outside house is a dummy that equals one when the activity is performed outside the home. The firm size is defined as the number of employees and this number includes the owner. The survey focused on firms with 50 or less employees. The variables revenue and other job are self-explanatory. Education is a variable indicating the number of grades completed. In some tables, this variable is substituted by dummies for primary, secondary and high education. The variable $\ln(\text{capital})$ measures the logarithm of capital installations in the firm. $\ln(\text{wage})$ denotes the logarithm of the total expenditures in salaries divided by the number of employees in the firm. Age of the owner is in years and gender equals 1 for male. Enforcement is a variable reflecting how constraining labor and tax inspections are for firms. Thus we assume that when entrepreneurs perceive these inspections as hard, enforcement is high and these firms are getting being inspected frequently. We are aware that this is a strong assumption, but this is the only enforcement variable that is available and this is how we justify it: Every firm will try to cheat and avoid paying taxes if there was no inspection because by minimizing its cost, the firm will be maximizing its profit which is every manager's goal, so the way a firm perceive inspection can tell us how severe and regular is the enforcement faced by it, even if our proxy is not directly the number of inspections in the last year or in the last 3 months. Most of the surveys do not ask questions concerning inspections, penalties and bribes because it's very sensitive and people can lie or choose not to answer. WBES ask these questions but it is hard to tell how reliable the answers are, and also the answer rate is very low.

The sectoral characterization of the survey can be found in Table 3. The major economic activity of MSE's in Lebanon is trade (72%) followed by other (12.8%) and to a lesser extent

industry (8.8%). We put construction and industry together, but construction do not include more than 0.6% of the surveyed firms, and firms in the category that encompasses hotels and restaurants do not exceed 5.1 % of the firms.

3.3 Empirical results

3.3.1 Informality and firm/entrepreneur characteristics

Table 4 contains probit estimates using the two definitions of informality discussed above: inftax and infss. The majority of signs obtained for the regressors are as suggested by the model, but there is a variation between the two definitions: Firm age is negative for both definitions but only significant for inftax. With time, firms grow and cannot stay in the shadow so some of them register with the tax authorities but once again this does not mean they declare their entire revenues. The registration with the National Social Security Fund (NSSF) is not as significant because employees are probably easier to hide than the firm itself. Firm size is negative for both definitions but only slightly significant for the SS definition. It should be noted that one-person enterprises in Lebanon are exempted from NSSF registration but we have many one person firms that are registered, so either they are not telling the truth, or they register by choice because it benefits them somehow. Maybe that is the reason why this variable is not very significant in this table. And in this survey, for questions concerning the registration, there's a possibility to answer "not required" so to avoid confusion, we assume that firms which can be exempted but did not choose the category "not required" are either formal or informal. The coefficients are significantly negative as the model predicts for education. Age is surprisingly non-significant, sex too because more than 90% of the sample is constituted by male entrepreneurs. Wealthier owners tend to be in a more formal category. Marital status, outsidehouse and otherjob are found to be non-significant, so they are not included in this estimation.

3.3.2 Informality and productivity

Table 5 presents the effect of formalization on installations per worker and the use of technology. In the first two columns, the tax registration definition is used, and the last two columns present results with the social security definition. The model predicts formal firms to be more productive. Since an entrepreneur's true ability is not observable, it makes sense to measure the effect of formalization after controlling for characteristics of the manager and the firm. The coefficients have the expected sign and are statistically significant. In other words, formalization is associated with an increase in installations per worker and investment even after using the control variables and the dummies for sector and regions. Unfortunately, we were constrained with the choice of proxies for firms' performance because very few answered these parameters.

3.3.3 Informality and size of firm

Our model predicts that formalization is associated with bigger firms. Table 6 uses non-linear least squares and finds supportive evidence for this proposition. We suppose as in De paula and Scheinkman (2011) that the observable measure of entrepreneurial quality is: $x = \exp(z'\beta)$ where z contains education, age, age squared, whether the entrepreneur has another job, and gender.

We estimate the regression: $\log(\text{firm size}) = z'\beta + \xi_2(\text{taxreg} \times z'\beta) + \text{controls} + \varepsilon$. The controls include sector and state dummies. ξ_2 is an interaction term between firm registration and the entrepreneur's quality. Table 6 shows that ξ_2 is positive and significant as expected.

3.4 Adding enforcement

Until now, all the results presented do not include any enforcement considerations. Our model assumes that enforcement constitute a cost for informal firms, thus high enforcement is positively related to formality and quality of the entrepreneurs. We estimate $y_i = z'\beta + \phi(E_c) + \xi_3(E_c \times Z'\beta) + controls + \varepsilon$. The dependent variable y_i is the decision to act formally by the enterprise i . Table 7 shows that ξ_3 is positive and significant when using the tax definition. Law enforcement, and entrepreneur's quality are positively correlated to formality. This effect disappears when using the social security definition of formality.

Since enforcement may be endogenous, we instrument this variable with the distance between the city where the firm is located and surrounding enforcement offices (Almeida and Carneiro 2009). We estimate $Y_{jc} = \beta E_c + X_{jy} + \varepsilon_{jc}$, where Y is the outcome of interest (informality in our case), for firm j in city c , E_c is enforcement, X_j is a vector of firm characteristics. We include in z_c the distance (in minutes) between the firm and its corresponding audit office since E_c is potentially correlated with ε_{jc} (more violation of law attracts more enforcement or better developed areas have better institutions and higher enforcement).

Table 8 reports the OLS estimates. Enforcement is negative and significant. In theory, the sign of the OLS bias is not clear. As Almeida and Carneiro (2009) argue, on the one hand, there can be more inspections in cities where informal employment is more prevalent. On the other hand, it could also happen that inspections are more frequent where institutions are more developed, and this happens in richer cities with low levels of informality.

In table 9, we adopt an instrumental variable strategy. We assume that the instrument (the average distance from the audit office to the city where the firm is located) measures the cost of enforcement in each city and that it is not correlated with the dependent variable of interest (except through enforcement). The IV estimates are significant and have the expected sign.

This strategy worked only using tax definition and only in Beirut and MontLiban. This can be explained by the fact that Lebanon is a small country. Thus in smaller departments, no big variation exists for the distance between the audit offices and the firms. Also in Lebanon, there are more labor audit offices than tax offices which is why we present only the results for Beirut and MontLiban considering the tax definition of informality. For both departments, the elasticity of informality with respect to enforcement is -0.37 and -0.49 which is quite reasonable.

4. Conclusion

This paper contributes to the growing body of evidence backing policies that aim at increasing incentives for formalization, which in turn is an important step in boosting firms' productivity. Our model shows that informal firms are smaller, less productive, managed by less educated managers and also less inspected by the authorities. These results are supported by enterprise data from Lebanon. An attempt was made to instrument enforcement in Lebanon with the distance between the firm and the audit offices but the instrument was only found valid in two of the biggest departments in Lebanon: Beirut and MontLiban. The reason why it does not work in the other departments is because they are too small that there is no enough variation in distances between the audit offices and the firms. It would make more sense to apply this instrument on data collected for larger countries. Enforcement is found important but it is not the only way to treat informality. As the World Bank reports suggest, a package of carrots and sticks is necessary everywhere to push people to be more formal, and especially when we deal with small firms. Encouraging them to be formal with less barriers to register can be more effective than inspecting them every day and making them pay high

penalties they cannot even afford. Doing Business Reports for Lebanon show that even in 2014, Lebanon is ranked very low in the ease of doing business and complying with the regulation. Another problem that Salti and Chaaban (2010) and Salti and Chaaban (2012) highlight is the role of sectarianism in the allocation of public expenditure in postwar Lebanon. Instead of allocating public expenditures and distribution of funds according to socioeconomic needs, it is exclusively determined by a rule of sectarian balance. Because of this inequality, we can argue that in regions that are in need and which are left without help, there is a low incentive for firm development and even starting a business, and if this happens, these firms will most likely be informal.

This paper uses two definitions of informality that can be found in the literature: tax definition and social security definition. Some differences in the results were found when using the tax and social security definition, which can be worrisome. Conclusions cannot be made before considering both definitions. Also finding a universal definition of informality is necessary. Our data does not include information about fines and bribes paid to inspectors. We tried to impute this information from the WBES data, but unfortunately very few firms accept to answer this question. Further research and data should be able to improve this lack of information.

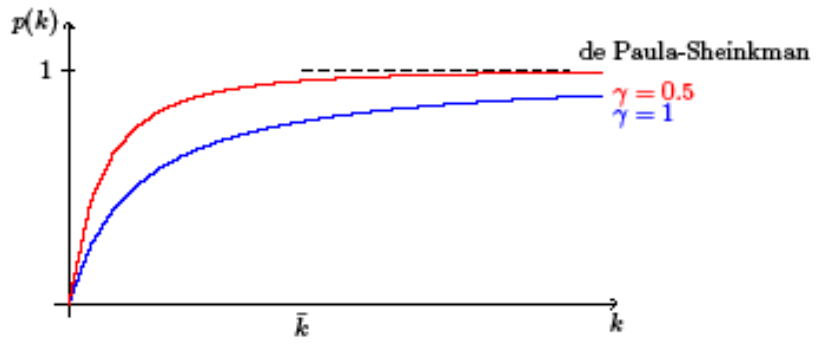
This paper is limited to MSEs and it does not include firms which have more than 50 employees. Even if only few firms with this size exist in Lebanon, it is important to know their behavior overlooking informality. This paper is also limited with the definition of informality as firms lacking registration. Other types of tax evasion exist and cannot be taken into account with this data.

References

- Allingham, M., and A. Sandmo. 1972. Income tax evasion: A theoretical analysis. *Journal of Public Economics* 1, 323–38.
- Almeida, R., and P. Carneiro. 2009. Enforcement of labor regulation and firm size. *Journal of Comparative Economics* 76(2):169–217.
- Becker, G. 1968. Crime and punishment: An economic approach. *The Journal of Political Economy* 76, 169–217.
- Chong, A., and M. Gradstein. 2004. Inequality, institutions, and informality. Inter-American Development Bank Working Paper no. 516.
- Dabla-Norris, E., M. Gradstein, and G. Inchauste. 2008. What causes firms to hide output? The determinants of informality. *Journal of Development Economics* 85, 1–27.
- De paula, A., and J. Scheinkman. 2011. The informal sector: An equilibrium and some empirical evidence from Brazil. *Review of Income and Wealth* 57, special issue.
- Feld, L., and F. Schneider. 2010. Survey on the shadow economy and undeclared earnings in OECD countries. Invited Paper written for publication in the *German Economic Review*, Department of Economics, University of Linz, Linz, Austria.
- Friedman, D. 1995. Rational criminals and profit-maximizing police: The economic analysis of law and law enforcement. *The Economics of Human Behavior*, Cambridge University Press.
- Hamdan, K. 2003. Micro and small enterprises in Lebanon. ERF Research Report 0417.
- Johnson, S., D. Kaufmann, and A. Shleifer. 1997. The unofficial economy in transition. Brookings Papers on Economic Activity 2.
- La Porta, R., and A. Shleifer. 2011. The unofficial economy in Africa. NBER Working Paper 16821.
- Loayza, N. 1996. The economics of the informal sector: A simple model and some empirical evidences from Latin America. Carnegie Rochester Conference Series on Public Policy 45, 129–62.
- Loayza, V., and J. Rigolini. 2006. Informality trends and cycles. The World Bank Policy Research Working Paper Series 4078.
- Lucas, R. 1978. On the size distribution of business firms. *The Bell Journal of Economics* 9(2):508–23.
- Perry, G., O. Arias, P. Faynzylber, A. Mason, and J. Saavedra. 2007. Informality exit and exclusion. Washington DC, World Bank, Latin American and Caribbean Studies.
- Rauch, J. 1991. Modeling the informal sector formally. *Journal of Development Economics* 35, 33–47.
- Salti, N., and J. Chaaban. 2010. The role of sectarianism in the allocation of public expenditure in postwar Lebanon. *International Journal of Middle East Studies* 42, 637–55.
- Salti, N., and J. Chaaban. 2012. The political economy of attracting public funds: The case of Lebanon. *International Journal of Development and Conict* .
- Schneider, F. 2000. Shadow economies: Sizes, causes and consequences. *Journal of Economic Literature* 38, 77–114.

- Scotchmer, S. 1987. Audit classes and tax enforcement policy. *American Economic Review Papers and Proceedings* 77(2):229–33.
- Straub, S. 1991. Informal sector: The credit market channel. *Journal of Development Economics* 78, 299–321.

Figure 1: Probability of Detection Function



Appendix

Table 1: Variable Description Lebanon

Variable	Description	Obs.	Mean	Std. Dev.
Taxreg	1= tax registration	2638	0,606	0,489
Ssreg	1= social security registration	2638	0,494	0,500
Outsidehouse	1= outside household	2638	0,971	0,168
Firm Size	number of employees	2638	2,465	3,579
Firm Age	age of the firm	2668	12,13	11,68
Revenue	Revenue per month at interview time	2238	2400,314	8789,918
Otherjob	1= owner has another job	2638	0,076	0,266
Education	nb. Of years Education (owner)	2638	10,028	3,855
Age	age (owner)	2638	40,924	12,693
Gender	1= male	2638	0,920	0,271
Loginst	log of installations per worker	1949	8,220	1,834
Loginv	log of investment per worker	2654	4,17	3,59

Table 2: Registration Status

SSreg but no TAXreg	9.85% (272)
TAXreg but no SSreg	21.09% (583)
SSreg and TAXreg	40.19% (1111)
NO reg	28.87% (798)

Notes: ssreg presents social security registration and taxreg is registration with the tax department.

Table 3: Economic Sector Lebanon

Description	Freq.	%
1 Trade	2,036	72,43
2 Industry and Construction	271	9,64
3 Hotel and Restoration	143	5,09
4 Other activities	361	12,84

Table 4: Probit Estimates

	Dep. Var= infntax coeff.	Std.Err	Dep. Var= infss coeff.	Std.Err
Firm Age	-0.024***	(0.003)	-0.009	(0.005)
Firm Size	-0.0124**	(0,018)	-0.052*	(0,022)
Secondary Educ.	-0,431***	(0,098)	-0,255*	(0,111)
High Educ.	-0,820***	(0,124)	-0,503***	(0,142)
Age	-0,0246	(0,015)	0,031	(0,019)
Age ²	0	(0)	0	(0)
Male	0,041	(0,124)	-0,332*	(0,136)
Wealth	-0.184**	(0.061)	-0.169**	(0.062)
Sector_Dummies		YES		YES
State_Dummies		YES		YES
N		2628		1481
Pseudo-R ²		0.16		0.13
Chi2 (18)		291		135.84

Notes: standard errors in parentheses.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Standard errors clustered by city.

Table 5: Investment, Installations Using Tax and SS

	(1)	(2)	(3)	(4)
	loginstp	loginvpw	loginstp	loginvpw
inf_tax	-0.368*** (0.114)	-0.129*** (0.030)		
ssreg			-0.042 (0.000)	-0.091** (0.034)
firm_age	0.0171 *** (0.000)	0.000 (0.001)	0.019*** (0.000)	0.000 (0.001)
firm_size	0.165 *** (0.020)	0.017*** (0.003)	0.169*** (0.021)	0.011** (0.003)
secondary_educ	0.336** (0.116)	0.047* (0.02)	0.336** (0.116)	0.033 (0.022)
high_educ	0.738** (0.158)	0.136*** (0.037)	0.738*** (0.158)	0.108** (0.037)
manager_age	0.041** (0.015)	-0.006 (0.004)	0.041** (0.015)	-0.004 (0.004)
manager_ageSQ	-0.000* (0.000)	0.000 (0.000)	-0.000* (0.000)	0.000 (0.000)
manager_sex	0.163* (0.163)	0.027 (0.026)	0.163 (0.160)	0.023 (0.026)
wealth	0.366 (0.072)	0.152*** (0.019)	0.366 (0.072)	0.155*** (0.019)
sector_dummies		YES YES		YES YES
state_dummies		YES YES		YES YES
N	2483	2475	2483	2484
pseudo-R2	0.25	0.2	0.25	0.20

Notes: standard errors in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Log of Number of Workers (=Dep. Var)

Coefficient	
(Std.Err.)	
ξ^2	0,212*** (0,05)
Education	0,026*** (0,003)
Otherjob	0,056 (1,41)
Age	0,012*** (0,002)
Agesq	-0,000*** (0)
Gender	0,092* (0,037)
Sector_Dummies	yes
State_Dummies	yes
N	2774
R ²	0.52

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Formality (=Dep. Var)

	tax def Coefficient (Std. Err.)	ss def Coefficient (Std. Err.)
ξ_3	0.201*** (0.032)	-0.032 (0,03)
Education	0.020*** (0.001)	0.019*** (0,002)
Otherjob	-0.021 (0.024)	-0.054 (0,035)
Age	0.011*** (0.001)	0.009*** (0,002)
Agesq	-0*** (0)	-0** (0)
Gender	0.032 (0.022)	-0.048 (0,033)
Sector_Dummies	yes	yes
State_Dummies	yes	yes
N	2507	2681
R ²	0,68	0,58

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: OLS

Dep. Var. = informality	taxreg		ssreg	
	Coeff.	St. Err.	Coeff.	St. Err.
Enf	-0,127***	0,011	-0,055**	0,016
Firm Size	-0,009**	0,003	-0,011***	0,003
Manager Educ	-0,026***	0,002	-0,012***	0,002
Manager Age	-0,001**	0,000	0,000	0,000
Manager Sex	0,005	0,003	0,111**	0,035
Firm Age	-0,006***	0,000	-0,001	0,001
Sector_Dummies	YES		YES	
Obs	2045		2168	
R ²	0,15		0,03	

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: IV with First Stage as Probit

	MontLiban		Beirut	
	Coeff.	Std. Err.	Coeff.	Std. Err.
Dep. Var. = informality				
Enf	-0,370***	(0,102)	-0,494***	(0,089)
Firm Size	-0,003	(0,004)	-0,004**	(0,005)
Sales	-0,008	(0,014)	0,014	(0,016)
Firm Age	-0,006***	(0,001)	-0,003*	(0,001)
Manager Educ	-0,023***	(0,005)	-0,016*	0,006
Manager Sex	0,036***	(0,050)	-0,062	(0,075)
Manager Age	-0,000	(0,001)	0,000	(0,001)
Sector Dummies		YES		YES
Dep. Var. = Enf				
Distance	-0,047***	(0,010)	-0,004***	0,001
DistanceSQ	0,000***	(0,000)	0,004***	0,001
Firm Size	0,009***	(0,011)	0,015**	0,007
Sales	-0,000	(0,000)		
Firm Age	-0,011*	(0,005)	0,005**	0,002
Manager Educ	0,012	(0,014)	0,027***	0,006
Manager Sex	-0,332*	(0,160)	-0,191**	0,091
Manager Age	-0,004	(0,004)	-0,002	0,002
Sector_Dummies		YES		YES
Obs		695		260
Prob > chi2		0,000		0,000

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: MSE Distribution in Mohafazat

Mohafazat	number of MSEs	percentage of total
Beirut	357	12,74
MontLiban	1014	36,07
North	601	21,38
Bekaa	413	14,69
South	255	9,07
Nabatieh	170	6,05
Total	2811	100

Table 11: MSE Distribution per Mohafaza and Firm Size

Mohafazat	size				Total
	1	[2-4]	[5-9]	[10-49]	
Beirut	115	183	41	18	357
MontLiban	418	486	70	40	1014
North	321	256	17	7	601
Bekaa	197	193	14	9	413
South	103	137	12	3	255
Nabatieh	89	81	3	1	170
Total	1239	1336	157	78	2810

Table 12: Registration

Social Insurance	Tax Department			Total
	No	Yes	Not Required	
No	798	537	46	1381
Yes	24	296	10	330
Not Required	248	402	403	1053
Total	1070	1235	459	2774

if answer is yes for tax registration, the card member is acquired.

The first order conditions for the informal firm are as follows:

$$\begin{cases} (1-p)\alpha\theta k^{\alpha-1}l^\beta - r - p'(\theta k^\alpha l^\beta - wl) = 0 \\ (1-p)(\beta\theta k^\alpha l^{\beta-1} - w) = 0 \end{cases}$$

$$\begin{cases} k^{\alpha-1}l^\beta = \frac{r}{\alpha\theta} \cdot \frac{\alpha}{(1-p)\alpha - p'(1-\beta)k} \\ k^{\alpha-1}l^\beta = \frac{w}{\beta\theta} \end{cases}$$

Denote $\xi = 1 - \alpha - \beta$. Bringing the system above to the log-linear form and solving for $\log k$ and $\log l$ we get:

$$\log k = \frac{1-\beta}{\xi} \log \left[\frac{\alpha}{r} \frac{\alpha(1-p) - (1-\beta)p'k}{\alpha} \right] + \frac{\beta}{\xi} \log \frac{\beta}{w} + \frac{1}{\xi} \log \theta$$

$$\log l = \frac{\alpha}{\xi} \log \left[\frac{\alpha}{r} \frac{\alpha(1-p) - (1-\beta)p'k}{\alpha} \right] + \frac{1-\alpha}{\xi} \log \frac{\beta}{w} + \frac{1}{\xi} \log \theta$$

$$\log \frac{k}{l} = \log \frac{\alpha}{r} - \log \frac{\beta}{w} + \log \frac{\alpha(1-p) - (1-\beta)p'k}{\alpha}$$

Under this particular form of probability, we have $p'k = \gamma k^{-\gamma}$ and $1-p = k^{-\gamma}$. Plugging these values into equation for capital demand and solving for $\log k$ yields:

$$\log k = \frac{1-\beta}{\xi + \gamma(1-\beta)} \log \frac{\alpha}{r} + \frac{\beta}{\xi + \gamma(1-\beta)} \log \frac{\beta}{w} + \frac{1}{\xi + \gamma(1-\beta)} \log \theta + \frac{1-\beta}{\xi + \gamma(1-\beta)} \log \frac{\alpha - (1-\beta)\gamma}{\alpha}$$

$$\log l = \frac{\alpha}{\xi + \gamma(1-\beta)} \log \frac{\alpha}{r} + \frac{\gamma + (1-\alpha)}{\xi + \gamma(1-\beta)} \log \frac{\beta}{w} + \frac{1+\gamma}{\xi + \gamma(1-\beta)} \log \theta + \frac{\alpha}{\xi + \gamma(1-\beta)} \log \frac{\alpha - (1-\beta)\gamma}{\alpha}$$

$$\log \frac{k}{l} = \frac{\xi}{\xi + \gamma(1-\beta)} \log \frac{\alpha}{r} - \frac{\xi + \gamma}{\xi + \gamma(1-\beta)} \log \frac{\beta}{w} - \frac{\gamma}{\xi + \gamma(1-\beta)} \log \theta + \frac{\xi}{\xi + \gamma(1-\beta)} \log \frac{\alpha - (1-\beta)\gamma}{\alpha}$$

Under the same assumption on the form of probability of detection function, and the first-order-condition on capital, maximum profit as function of optimal labor and capital choices is:

$$\Pi_i = (\xi + \gamma(1-\beta))\theta k^{\alpha-\gamma} l^\beta$$

$$\log \Pi_i = \log(\xi + (1-\beta)\gamma) + \log \theta + (\alpha - \gamma) \log k + \beta \log l$$

$$\log \Pi_i = \log(\xi + (1-\beta)\gamma) + \frac{\alpha - \gamma(1-\beta)}{\xi + (1-\beta)\gamma} \log \frac{\alpha - (1-\beta)\gamma}{\alpha} + \frac{\alpha - \gamma(1-\beta)}{\xi + (1-\beta)\gamma} \log \frac{\alpha}{r} + \frac{\beta}{\xi + (1-\beta)\gamma} \log \frac{\beta}{w} + \frac{1}{\xi + (1-\beta)\gamma} \log \theta$$

As in De paula and Scheinkman (2011) expression for the maximum attainable profit of the formal enterprise is given by:

$$\log \Pi_f = \frac{1}{\xi} \log(1-\tau) + \frac{1}{\xi} \log \theta + \frac{\alpha}{\xi} \log \frac{\alpha}{r} + \frac{\beta}{\xi} \log \frac{\beta}{w} + \log \xi$$

$$\log \frac{\Pi_f}{\Pi_i} = \frac{(1-\beta)\gamma}{\xi(\xi+(1-\beta)\gamma)} \log \theta + \frac{1}{\xi} \log(1-\tau) + \frac{(1-\beta)^2\gamma}{\xi(\xi+(1-\beta)\gamma)} \log \frac{\alpha}{r} + \frac{\beta(1-\beta)\gamma}{\xi(\xi+(1-\beta)\gamma)} \log \frac{\beta}{w} + \log \xi - \log(\xi+(1-\beta)\gamma) - \frac{\alpha-\gamma(1-\beta)}{\xi+(1-\beta)\gamma} \log \frac{\alpha-(1-\beta)\gamma}{\alpha}$$

Now we try to find level of θ , at which the entrepreneur is indifferent between acting formally or informally. As $\log \frac{\Pi_f}{\Pi_i}$ is an increasing linear function of $\log \theta$, there should exist

the value of the argument $\log \theta$ at which $\log \frac{\Pi_f}{\Pi_i} = 0$. Denoting this value $\log \bar{\theta}$ we get the

following expression:

$$\log \bar{\theta} = \frac{\xi}{(1-\beta)\gamma} \left((\alpha-(1-\beta)\gamma) \log \frac{\alpha-(1-\beta)\gamma}{\alpha} + (\xi+(1-\beta)\gamma) \log \frac{\xi+(1-\beta)\gamma}{\xi} \right) - (1-\beta) \log \frac{\alpha}{r} - \beta \log \frac{\beta}{w} - \frac{\xi+(1-\beta)\gamma}{(1-\beta)\gamma} \log(1-\tau)$$