What Determines Firms' Decisions to Formalize?

Evidence from Rural Indonesia

Neil McCulloch, Günther G. Schulze and Janina Voss

What Determines Firms' Decisions to Formalize?

Evidence from Rural Indonesia[∗]

Neil McCulloch,[♣]   Günther G. Schulze[♠]   and Janina Voss[♥]

November 15, 2010

Abstract

In this paper we analyze the decision of small and micro firms to formalize, i.e. to obtain business and other licenses in rural Indonesia. We use the rural investment climate survey (RICS) that consists of non-farm rural enterprises, most of them microenterprises, and analyze the effect of formalization on tax payments, corruption, access to credit and revenue, taking into account the endogeneity of the formalization decision to such benefits and costs. We show, contrary to most of the literature, that formalization reduces tax and corruption payments. The benefits of formalization, and therefore the likelihood of being formal, also depend on characteristics such as firm size, as well as the education and ethnicity of the owner.

JEL classification: O17, O18

Keywords: Formalization, rural development, rural investment climate, informal sector

[♣] Institute of Development Studies at the University of Sussex, Brighton BN1 9RE, U.K.

[♠] Corresponding author: University of Freiburg, Dept. of International Economic Policy, 79085 Freiburg, Germany, guenther.schulze@vwl.uni-freiburg.de .

[♥] formerly University of Freiburg, Dept. of International Economic Policy, 79085 Freiburg, Germany.

1. Introduction

Most developing countries are characterized by large informal sectors.[1] These large informal sectors are potentially detrimental to economic development as they escape government taxation and regulation. The former means that governments have fewer resources for financing development, whilst the latter can lead to welfare losses due to the lack of minimum safety and labor standards in the informal sector. Moreover, having a larger share of informal firms in an economy may reduce growth because of their lack of access to formal credit institutions, government contracts and other institutions. Yet, a large informal economy may also be created by misguided government policies. Auriol and Walter (2004) argue that myopic governments create large informal sectors by erecting high entry barriers, as they find the resulting small numbers of large tax payers easier to tax. Similarly, Loayza (1997) finds that large informal sectors are a consequence of excessive taxation, overregulation and weak government institutions resulting in low growth rates.[2]

Despite extensive research on the informal sector and the substantial policy relevance of the determinants of informality (Perry et al. 2007, Jütting and Laiglesia 2009), relatively little quantifiable evidence exists on the various channels through which formality provides benefits to and causes costs for firms.[3] In order for policymakers to be able to design appropriate policies to reduce the size of the informal sector, it is necessary to understand why some firms decide to go formal, whilst others decide to avoid being registered and licensed. De Soto (1989) has argued that it is complex and costly procedures that prevent many firms from becoming formal. McKenzie and Woodruff (2006), among others, challenge this view. They argue that the costs of registering do not establish a significant barrier; rather many firms are too small to reap the benefits of formalization. These benefits are seen as access to credit and government provided services and contracts as well as subcontracting relations with formal firms (De Paula and Scheinkman 2007). Other authors regard the evasion of taxes and other regulations as the main motivation for staying informal (Gërxhani 2004 and Loayza et al. 2005).

---

[1] See Enste and Schneider (2000) and Schneider (2005) for reviews of the size, causes and consequences of shadow economies around the world.

[2] Djankov et al. (2006) show that countries with low levels of regulation of the business sector (based on the World Bank's "Doing Business" indicators) grow significantly faster than countries with excessive regulation. Loyaza et al. (2005) provide similar evidence.

[3] Another strand of the literature studies the motivation of workers to work in the formal or informal sector (rather than the decision parameters of firms to operate in the formal or informal sector), e.g. Pradhan and van Soest (1995), Jütting and Laiglesia (2009).

Firm characteristics may also influence the decision to formalize – for example, it is frequently found that formal firms tend to be larger and older. However, it is not clear which way the causality runs between firm characteristics and formality. Levenson and Maloney (1998) illustrate this by treating formality as an input into a firm's production function. In exchange for the benefits of formalization – enforceability of contracts, credibility, access to capital and access to public risk-pooling mechanisms – a firm has to pay initial entry costs and periodic 'taxes', such as reporting requirements or insurance payments. They assume that emerging firms face uncertainty over their managerial ability and production costs, which resolves over time. Given that entry costs are assumed to be identical, the firms that become successful over time, i.e. those with more able entrepreneurs, will find it profitable to become formal and thus the older and larger firms will have a higher probability of being formal. They find evidence supporting this hypothesis from the Mexican National Micro Enterprise Survey 1992. Jäckle and Lee (2006) partly corroborate Levenson and Maloney's hypothesis in a panel approach that captures firm dynamics using the Peruvian Living Standard Measurement Survey for 1994, 1997, and 2000. They show that larger and fast growing firms are more likely to become formal, but they do not find firm age to be a significant determinant of formality. These studies focus on two firm characteristics, yet manager characteristics such as gender, ethnicity and religion, may also influence formality. For example, if certain groups are more subject to scrutiny than others, then the costs of informality for these groups will be higher and so they will be more likely to formalize.

In general, if formality is considered an input in the firm's production function, it clearly affects firm revenue, costs and longevity (considering the high failure rates of newly emerging firms in developing countries). Thus there is an endogeneity problem with respect to formal status: older and more successful firms may be more likely to go formal, but formal firms may also become more successful and long-lived.

Fajnzylber et al. (2010) solve the endogeneity problem in a quasi-experimental regression discontinuity approach. The introduction of the Brazilian SIMPLES program, which aimed at reducing registration costs and tax rates for small firms, led to higher formalization rates. Comparing the firms established shortly before and after the introduction of the program they show that formal firms do better in terms of revenue, investment and employment. They do not find formal firms to have better access to credit; instead they argue that formally established firms choose a different production scale (and possibly technology) at the time of start up.[4] McKenzie and Sakho (2010), by contrast, use an instrumental variable approach to

---

[4] They also find that the implementation of SIMPLES has only a local effect and that the increase of registration decayed relatively quickly. Fajnzylber et al. (2010) analyze urban microenterprises while our focus is on rural non-farm microenterprises.

address the endogeneity problem. They analyze Bolivian micro- and small enterprises in urban areas that are within 10 kilometers of the city's tax office and use the distance to the tax office as an instrument for formalization. They show that formalization increases profits, but that this effect is limited to medium sized firms (2-5 workers) while very small and larger firms lose. McKenzie and Sakho argue that the tax identification number, their measure of formality, allows firms to issue tax receipts; this increases the customer base for middle sized firms while very small firms are too small to profit from this tool and firms with six or more employees have a large customer base already. This 'marketing tool' comes at the cost of a higher probability of paying taxes for firms of all sizes thus leading to the observed nonlinearity of profit changes with respect to firm size.[5] They find no evidence that formality improves access to credit.[6]

This paper examines the determinants of formalization in Indonesia. We contribute to this emerging literature and complement it in a number of ways. First, our analysis focuses on the channels through which formalization provides net benefits to firms, i.e. its effects on tax payments, corruption, and access to credit and to government contracts as well as on revenues. We also study whether the costs of registration deter potential candidates for formal status from acquiring it. Second, our empirical approach addresses heterogeneity in net benefits, arguing that the costs and benefits may be systematically related to firm characteristics. Extending the approach of McKenzie and Sakho (2010), we estimate heterogeneous impacts of formalization using a wide array of firm, manager and community characteristics. Indeed we identify firm characteristics that significantly influence the net gains from formalization. Finally, while the existing literature on the determinants of formalization of firms has mostly studied urban microenterprises in Latin America, we analyze rural nonfarm microenterprises in an Asian country.

We use the Indonesian Rural Investment Climate Survey (RICS),[7] which covers 2500 small and micro non-farm enterprises and their communities in rural Indonesia. Indonesia is particularly suited for such analysis because its diversity allows studying systematically the

---

[5] Although, using propensity score matching they do not find that taxes as a share of profits increases significantly with a tax identification number.

[6] In a related vein, Bruhn (2010) and Kaplan et al. (2007) study the effect of the Mexican licensing reform 2002-2006 on the number of newly registered firms. The Rapid Business Opening System (SARE) reduced the time to open a business from an average of 67 days to 2 days.. Both studies show that the number of registered businesses increased by approximately 5 percent in the eligible industries. Bruhn finds that this increase is not due to informal firms becoming formal but rather wage earners becoming self-employed, while Kaplan et al. show that this increase was largely confined to the initial period after the introduction of SARE. Both papers study the macroeconomic effects of this relatively moderate increase of the formal sector (on profits, prices and competition), but they do not investigate the channels through which formality provides benefits to the firms, the issue we are concerned with.

[7] The survey was conducted by the World Bank and one of a series of RICS in other countries. See www.worldbank.org/id/rica for details.

role of ethnicity, religion and community characteristics. We employ an IV approach to account for the endogeneity of firm formality and we interact the formalization variable with a number of firm and owner characteristics to explore the different effects of formalization on heterogeneous firms. We find that formalization reduces tax payments and the incidence of corruption as well as its magnitude, but that this effect is non-uniform and depends, inter alia, on firm size and the gender of the manager. We find that larger firms are more likely to be formal, as are firms with non-local managers or those that are Chinese or of a non-dominant ethnic group which are easier target for officials' scrutiny. Thus formality serves as a protective device to avoid arousing government's attention.

The remainder of the paper is structured as follows. Section 2 presents the RICS dataset, explains licensing practices in Indonesia which affect our definition of formality. Section 3 introduces the costs and benefits of formalization. Section 4 lays out the estimation strategy, section 5 presents our results after which the final section concludes.

2. Data and Definitions

The Indonesian Rural Investment Climate Survey

The Indonesian Rural Investment Climate Survey (RICS) is a quantitative survey of households, non-farm enterprises (NFE) and communities that attempts to measure the business environment faced by rural firms. It was conducted in 2006 in six districts (*Kabupaten*) purposively chosen to reflect the distinct geographical environments faced by local enterprises in different parts of Indonesia (World Bank, 2006).[8, 9]

The sampling within these districts was done in three stages.[10] Based on a 2006 village census, 30 villages were selected with probabilities of selection proportional to their number of businesses. In a second step census blocks were constructed in each village from which in the third stage a sample of 'household enterprises' (located in the place of residence),

---

[8] The survey locations were: Labuhan Batu, North Sumatra – a plantation area; Kutai, East Kalimantan – an area rich in mineral resources; Barru, South Sulawesi – a forest fringe area; Malang, East Java – a rich agricultural area; Badung, Bali – a semi-urban agglomeration area and; Sumbawa, NTB – a dryland area.

[9] There are over 400 districts in Indonesia. Because the focus of the survey was the exploration of the investment climate at the district level, the RICS takes a large representative sample in each district to facilitate cross-district comparisons. The survey is not, therefore, nationally representative (which would have required extensive but sparse sampling), but the choice of regions attempts to encompass the main types of economic geography in the country.

[10] A more detailed review of the sampling strategy, the data collection and data processing is provided by LPEM-FEUI in the Data User Guide (2006a) and the Field Survey Report (2006b). A summary of the sampling procedure and the main results of the RICS can be found in Schulze and Quadros (2007).

'standalone enterprises' (not located in the place of residence) and 'pure households' (i.e. with no household enterprise) was drawn. The stratified sampling design requires the use of sampling weights in the empirical analysis in order to obtain unbiased estimation results that are representative of the population of NFEs at the district level (World Bank, 2008).

The survey consisted of three linked questionnaires: (1) the household questionnaire, collecting information on household economic activity and consumption patterns, (2) the enterprise questionnaire, that aims at investigating the rural investment climate for NFEs by looking at their costs, revenues and constraints, and (3) the community questionnaire, that provides information on local infrastructure and governance. From these we constructed an enterprise dataset which contains 2137 micro enterprises (1-4 employees), 263 small enterprises (5-19 employees), 51 medium-sized enterprises (20-99 employees), and nine large enterprises (> 100 employees).[11] Since our analysis focuses on formalization, we exclude the 60 medium and large firms from our sample.

Table 1: Firm and manager characteristics about here

Table 1 shows the characteristics of the firms and their managers in our sample. The number of firms is distributed fairly equally over the six *Kabupatens* with 400 to 420 firms in each district. Although the sample focuses on rural areas and does not include any major metropolitan centers, the statistical bureau still classifies villages as either 'urban' or 'rural' based on a set of village characteristics (share of agriculture, population density, and several services and facilities associated with urban living – see World Bank (2006)). According to this definition about half of the firms in the sample are located in rural areas, although this share is much higher in Labuhan Batu and much lower in the peri-urban area of Badung. Typically more than half of the sampled firms in each area operate in the trading sector, around 35% in the service sector, with the remainder in the manufacturing sector.[12] Of the 2460 firms in the sample, almost 90 percent are 'micro' firms; as a result the mean number of employees is only 2.6. The mean enterprise age is nine years. However, more than 50 percent of firms in the sample have been operating for six years or less and 12 percent have existed for only one year.

---

[11] These employment groups follow the BPS size classification of enterprises. The number of employees also includes the owner or manager of the enterprise as well as unpaid laborers.

[12] The three sectors comprise the following activities: trade, including wholesale and retail trade; the service sector, comprising of repair shops, hotels, food and beverages, transportation, finance, real estate, health and public services; manufacturing activities, defined as mining and excavation, manufacturing including the processing of agricultural goods, electricity, gas and water provision and construction.

For about 90 percent of firms the owner is also the primary manager of the business.[13] The majority of firms surveyed are managed by men (63%) with a mean age of 41 years. Education is at a very low level for almost half of firm managers: 13 percent state that they did not complete primary school, and a further 30 percent have primary school as their highest level of schooling. Only eight percent of firm managers report that they have a university level education.

Ethnicity is highly diversified in the sample as is true for Indonesia as a whole. However, in each Kabupaten one local ethnicity dominates among firm managers. Only in Kutai and Labuan Batu are there more managers from non-indigenous ethnicities than indigenous managers. Chinese Indonesians play a particularly important role in the Indonesian economy and so are specified separately in Table 2. Although they only account for 1.7% of enterprise managers throughout the six Kabupaten, they represent a much higher share of small firm managers than micro firm managers. Compared to ethnicity, the religion of the manager is quite homogeneous. The share of Islamic managers is more than 90 percent in all Kabupaten, except Badung in Bali, where the majority of managers are Hindus (64%), with only 28 percent being Moslems.

Business Licensing in Indonesia: Defining Formality

Business licensing in Indonesia is characterized by complex, time-consuming, and costly procedures. According to the World Bank's 'Doing Business Report', Indonesia ranks among the worst-performing countries in the world in terms of 'starting a business' and 'dealing with licenses' (World Bank, 2008).[14] The far-reaching decentralization undertaken in 2001 made processes even less transparent. Indonesia has rapidly transformed from decades of authoritarian rule to one of the most decentralized countries in the world with 33 provinces, subdivided into 459 districts (*Kabupaten* and *Kota*, as of 2006). These districts have become responsible for a large share of economic policies, among them business licensing (von Luebke 2006: 2). Hence licensing practices differ throughout Indonesia. Additionally there are still licenses being issued at the provincial or central government level, which further complicates the process by requiring cooperation between the different administrative levels (KPPOD 2008: 31). Local governments tend to see licensing services more as a means of generating local revenues than for regulating markets or collecting information: since

---

[13] In the subsequent analysis, all variables referring to manager characteristics are taking the value of the owner if he is managing the business himself and the primary manager otherwise.

[14] In 2006, when RICS was conducted, it took 12 procedures, 151 days at a cost of 101.7% of GNI per capita to officially start a new business in Indonesia. For dealing with all required licenses to build a warehouse in Jakarta, a SME company with 20 employees has to take into account 10 procedures, 224 days and 370.5% of income per capita (World Bank, 2008).

decentralization about 1600 new regulations and local government licenses have been introduced, at least 30% of which are considered to distort economic activity (ibid: 14).

This complexity makes the identification of formal and informal firms more difficult. In theory business licensing comes as a 'package' in Indonesia. To be fully registered an enterprise must complete a number of administrative processes at the national as well as the local level (TAF 2007: 8f). First, to make sure a new enterprise fulfils all requirements for a formal business or company, it needs a deed of establishment from a notary and a tax identification number (NPWP) from the central government. Second, the firm needs physical permits such as a building permit (IMB) and a nuisance permit (HO). Third, a sectoral license has to be obtained to allow operation in one of the major sectors. The main sectoral licenses are the trade permit (SIUP) and the industrial registration/permit (TDI). Only after those requirements have been fulfilled can a firm process the business registration (TDP) at the local level. Depending on which products or activities a firm deals with, it may additionally need to obtain product- and activity-specific licenses.[15]

The RICS data contains information on the main components of this 'package': building permit (IMB), industrial permit (TDI), trade permit (SIUP), enterprise registration (TDP), and tax identification number (NPWP). In terms of the formalization procedure explained above, a firm would be 'fully registered' if it has a tax number, a building permit (if it operates in a building separate from a household residence), one of the sectoral licenses (trade or industrial permit), and the enterprise registration. According to this definition, only 2% or 48 firms in the sample are fully formalized. Leaving out the tax registration number from the definition – as it is the only license not issued at the district level – results in 2.9% of firms being 'fully registered with the local government'.

In practice, there is considerable confusion about which licenses a business really needs. The trade permit for example, although officially designed for companies engaged in trading activities, is the most common license held even among industrial firms (KPPOD 2008: 32). In the RICS, the trade permit is almost as common among manufacturing firms (13 percent hold it) as the industrial permit (14 percent). Similarly, the building license is legally required only for firms operating in a separate building, but is often issued to firms that operate their business from their dwelling (11 percent) and even firms that do not operate from a fixed location (4 percent). Indeed the belief that certain licenses are not needed may be one reason why such a small number of firms are 'fully formalized'.[16]

---

[15] For a limited liability company these steps are even more numerous. For a detailed description of all processes involved see World Bank (2008: 53ff).

[16] When asked about the reason for not holding a certain permit, only three to five percent of firms consider the respective license to be too expensive and merely one percent criticizes complicated procedures. The main

This paper will therefore rely on a more practical definition of formality: a firm is classified as being formal if it has at least one of the local licenses (IMB, SIUP, TDI or TDP), which is the case for 23.4% or 541 enterprises. Since the policies on local licenses are being determined at the Kabupaten level, there are significant differences across the six districts in the share of formal firms. The lowest share can be observed in Malang (7.4%), followed by Labuhan Batu (17.9%); Badung, Sumbawa and Kutai have shares between 23 and 29 percent, whilst in Barru 42 percent of firms hold at least one of the four main local licenses.

3. Costs and Benefits of Formality

A firm's decision to formalize will depend on "firms perceiving that it is in their self-interest" (Kenyon 2007: 5). We therefore anticipate that firms will choose to formalize when the expected profit of operating formally is larger than the expected profit of operating informally.[17] The formality status of the firm will influence the revenue and cost functions that it faces.

Each firm that wants to 'legally' open a business or register an already existing one faces an initial entry cost in terms of time and fees (Djankov, 2002). We define entry costs by the mean cost and time it takes to register, where registration is defined as holding at least one of the four local licenses (SIUP, TDI, TDP or IMB). After registering, a formal firm has to abide by government regulations that are often complex. Besides the direct monetary costs in terms of taxes (or bribes) to government officials, the firm also has to bear the indirect costs of time spent submitting government documents or fulfilling product or labor standards (cf. Djankov et al. 2003: 66f; Ishengoma and Kappel 2006: 16f).

An informal firm, by contrast, saves on these costs. At the same time, it is more dependent on the 'goodwill' of government officials and the police. This may make bribe payments more unpredictable and potentially higher for informal firms (cf. ibid: 18, Djankov et al. 2003: 71). Anecdotal evidence from a qualitative evaluation of One Stop Shops for Business Licensing in Indonesia suggests that an important reason firms get licenses is to reduce unofficial payments to business inspectors (LabSosio 2008: 41). Informal firms may respond to such costs by staying small or changing location in order to escape detection and harassment by

---

stated reason for firms not obtaining a license is because they think it is not required for them. This is the case for 52% of firms that operate from a fixed location but do not have an IMB, 49 % of trading firms without SIUP, and 48% of manufacturing firms without TDI. Even the enterprise registration certificate, TDP, which is the only license that is needed by all firms no matter where they operate or in which sector, reveals the same picture. Only 7.5% of firms hold an enterprise registration at all – of the remaining 2011 firms, more than 51 % state that TDP it is not required.

[17] If firms are not risk neutral, their decision will also take into account the relative uncertainty of the profit streams under formality and informality, but we do not explore this here.

the police or government officials (cf. ibid; Djankov et al. 2003: 70). For the same reason such firms may avoid investment in conspicuous fixed assets or technology. In this paper, we calculate the costs of operating formally as the (differential) value of levies a firm has to pay, subdivided into total taxes and 'other levies'. Total taxes include the taxes a firm pays to the central, provincial and local governments; 'other levies' comprises payments to security officials, thugs, and sub-district or village officials which we use as a proxy for unofficial payments and bribes.

There are also several potential benefits of operating formally. These can include: the ability to access formal credit markets; access to contract governments; access to legal dispute resolution mechanisms, such as courts; and the ability to enter into collaborative agreements with large firms. None of these benefits are typically available to informal firms. We calculate the benefits of formality using total firm revenue, access to government contracts and access to credit. We did not examine the benefits associated with using the courts, because, of the 2.8% of the firms in our sample that had any kind of payment dispute, none had used the courts to resolve the problem.

Table 2: Mean costs and benefits about here

Table 2 shows the average value of these costs and benefits for formal and informal firms separately. At face value it would appear that the costs, in terms of taxes and levies, are much higher for formal rather than informal firms, but that the benefits – sales to government and better access to credit – are also commensurately larger. However, these differences could be due to other firm or manager characteristics rather than formality itself. In particular, larger average costs and benefits may simply be because formal firms tend to be larger and older (Levenson and Maloney, 1998). Table 3 shows that this is indeed the case.

Table 3: Mean enterprise size and age about here

In addition reverse causality is an issue here: For instance firms could become formal because they are larger tax payers; they could also be older because they have opted to be formal. It is therefore not possible to draw conclusions about the impact of formality on firm revenues and costs from the stylized facts above. Rather we need to estimate the causal effect of licensing on firm revenues and costs taking into account the endogeneity of the formality decision to the net benefits of formalization.

4. Estimation Approach

We seek to estimate the impact of formality on potential costs and benefits of the firm. Our basic equation is the following:

$$Y_i = \beta_0 + \beta_1 L_i + \beta_2 M_i + \beta_3 F_i + \beta_4 C_i + \varepsilon_i \tag{1}$$

where Y_i denotes the respective cost or benefit of firm i, i.e. taxes, 'other levies', firm revenue, sales made to the government, and access to credit. L is the formality dummy, and M, F, and C are controls for manager, firm and location (community) characteristics.[18]

Because formality is the consequence of a deliberate decision, and thus endogenous to costs and benefits that we seek to estimate, we instrument it using 2SLS. In addition to estimating the overall impact of formality on costs and benefits, we are interested in how the impact of formality differs across firms with different characteristics. We therefore enter our first-stage estimate of licensing, $\hat{L}_i$, both directly, in order to estimate the overall effect of licensing, as well as in the form of interaction terms with the firm, manager and location characteristics i.e.

$$Y_i = \beta_0 + \beta_1 \hat{L}_i + \beta_2 \hat{L}_i * M_i + \beta_3 \hat{L}_i * F_i + \beta_4 \hat{L}_i * C_i + \beta_5 M_i + \beta_6 F_i + \beta_7 C_i + \varepsilon_i \tag{2}$$

Our approach is therefore similar to McKenzie and Sakho (2010) who show that the effect of formality on firm profits varies with the number of employees. We use a more comprehensive list of owner, firm and location characteristics that might cause heterogeneous treatment effects on a broader range of potential costs and benefits.

The instrument we use is the community average level of licensing. It is quite common to use village or city averages as an instrument (see for instance Dollar et al. 2005); in our case the instrument works very well and better than any other possible instrument.[19] As an aggregate measure, village licensing averages are correlated with the formality status of individual firms, but have little influence on the costs and benefits obtained by individual firms, other than through its influence upon licensing.[20] In constructing the instruments for individual firms we use community averages for licensing excluding the individual firm from the village sample in order to avoid that the community average is influenced by the status of the firm it is intended to instrument.

---

[18] Appendix Table A1 provides the list of variables used in the analysis; Table A2 provides descriptive statistics.

[19] We also explored a range of other community characteristics as instruments (frequency of village meetings, different measures for conflict, local business organizations and education of the village head, registration time and costs) but these instruments did not pass our tests.

[20] Our instrument is in the spirit of McKenzie and Sakho (2010) who use distance to the tax office to instrument for their measure of formality, which is having a tax identification number. If such a distance (in our context to the district office) is the main 'driver' for formalization, it should be closely correlated with village averages. If there are other determinants as well, e.g. because geographical distance does not adequately represent travel time, village averages may actually outperform distance as instrument.

However, it could be argued that villages with high shares of licensed firms may perform better in other fields of economic policy. If village averages of licensing are correlated with a more favorable business climate and this reduces firm costs (or increases their benefits), our results on licensing may be confounded biasing our second stage estimates. To address this problem, each second stage regression controls for village averages of the respective cost or benefit analyzed (again excluding the firm under consideration) that should capture any village-level confounding factor on the respective endogenous variable.[21]

The results of the first stage estimates are summarized in Table A3 in the appendix. The firm, manager and location variables used to predict the fitted value of licensing will also be used to estimate the costs and benefits in the second stage.[22] The Sargan test of overidentification requires the application of at least one more instrument than endogenous variables. To perform this test, some of the weaker instruments that we had used (frequency of village meetings, registration time and cost) were additionally included in the regression. The null hypothesis of instrument validity – denoting that all instruments are uncorrelated with the error term and thus have correctly been excluded – could not be rejected for either combination. However, a redundancy test shows that the weaker instruments do not improve the efficiency of estimates. Village averages of licensing are therefore applied as the only instrument. The high F-statistic of the excluded instrument (between 71 and 126, depending on the dependent variable in the second stage) indicates that there is no problem of weak identification.[23]

5. Results

5.1. The impact of formality on costs and benefits

We now summarize the results from estimating the impact of formality on official and unofficial levies as well as on revenues, access to government contracts, and access to credit. In each case we compare our 2SLS results with OLS.[24] For testing heterogeneity of

---

[21] As a robustness check we also ran all regressions without village averages of the endogenous variable which showed that estimates of the other covariates were robust to the inclusion of village averages.

[22] In order to obtain consistent estimates the same control variables are included in both stages (Wooldridge 2009). To reduce the endogeneity with licensing, actual values of size and input quantities were replaced by dummies. For the number of employees we included three dummies that indicate two/ three or four/ four more than four employees (including the owner; omitted category is the single person microfirm). We constructed quintile dummies for the log of sales and a dummy for total fixed assets above the median. We have also investigated groupings, e.g. quintile dummies for fixed assets. The results were largely unaffected by this.

[23] We use the predicted values from a probit model with the endogenous variable being licensed; we have in addition used a linear prediction, which yielded similar, if not more significant results. We take a conservative approach reporting the less significant results.

[24] Since costs and benefits are being determined simultaneously at the firm-level, one could also estimate a system of equations rather than each equation individually. A seemingly unrelated regression model (SUR) that

treatment effects, interaction terms of the fitted values for licensing and the various control variables were employed.

Each cost and benefit regression will control for the same firm, manager and location characteristics. The manager characteristics are highest level of education, gender, age, ethnicity, a dummy for the manager/ owner being Muslim in an area inhabited predominantly by Muslims and a dummy for a Hindu manager/owner in Bali, a Hindu-dominated area, and a dummy that denotes whether the manager lives in the same village the enterprise is in. The ethnicity dummy equals one if the manager belongs to the local indigenous group. A dummy for managers of Chinese descent is included separately in the regression.

The regressions control for firm size in terms of the number of employees, sales and total value of fixed assets.[25] As these magnitudes might be endogenous to the licensing status we have used dummy variables. Our choice of dummies was guided by aim to create a parsimonious, yet rich enough model to capture the differential effects of licensing.[26] Further firm-level variables included are the number of years the firm has been in operation, and the sector of operation (manufacturing, service or trade). [27] As noted above, we control for how the general investment climate might influence the respective cost or benefit, by including the village averages of the relevant costs and benefits. We also include a dummy for whether villages are classified as rural or urban as well as district dummies. The estimates on access to credit also include a dummy for whether or not there is a bank in the village.

Taxes

The results of estimating the impact of formality on the log of total tax payments are shown in Table 5. OLS estimation that does not control for the endogeneity of licensing (Column 1) reveals a positive and significant association of licensing with tax payments. This correlation, however, may be the result of high taxpayers self-selecting themselves into formality. The IV

takes into account possible correlations of the error term might produce more efficient results. However, a SUR setup does not allow for the use of sampling weights that are applied here to balance the stratified sampling design. Even though they might be less efficient, individual OLS estimates of the same cost and benefit equations outside SUR will still be unbiased and consistent (Greene, 2007). Sampling weights on the other hand will correct for sample stratification bias. Applying sampling weights will thus be preferred over the SUR setup in the following analysis.

[25] Total fixed asset value includes buildings, land, equipment and machinery, furniture, storage facilities, and vehicles. Although number of employees, log total sales, and log of total fixed asset value are different measures for firm size, no additional multicollinearity could be detected when entering all three variables at the same time. Since the information given by each of them is slightly different, including them all was preferred to using a compound index of firm size.

[26] For instance we analyzed fixed assets in quartiles and quintiles, but since the quartiles above the mean behaved very similar and so did the quartiles below the mean, we opted for a single dummy indicating fixed assets below or above the mean.

[27] Trade is the excluded category.

estimates show that formality is on average associated with a significant *decrease* in firms' total tax payments! The bias in the OLS result has exactly the sign we expect. If the underlying effect of licensing is to reduce tax payments, then firms with characteristics that make them likely to pay high taxes will want to get a license. OLS does not take into account this endogenous positive selection effect and therefore significantly underestimates the tax reducing effect of obtaining a license.

Table 4: Taxes about here

The effect of licensing on taxes varies with the level of firm sales. The results in Column 3 show that licensing reduces tax payments for firms in the lowest sales quintile the most, while the effect is significantly smaller for the fourth and fifth sales quintiles.[28] In other words, small enterprises benefit from a substantial reduction in tax payments if they are licensed. Column 4 also shows that the licensing effect is significantly smaller for firms with female managers. Yet women managers/owners pay significantly less taxes to begin with. Trading and manufacturing firms benefit from licensing through tax reduction more than double compared to trading firms, even though for manufacturing the difference is significant only at the 15% level (Column 5).

Aside from licensing, several other characteristics influence the amount of taxes paid. Better educated and older owners tend to pay more tax (probably because these characteristics are correlated with firm performance). Female owners, by contrast, pay less. Unsurprisingly, "large" firms pay significantly more tax – i.e. firms with five employees or more, with fixed assets values above the median and firms in the two highest sales quintiles. Firms in the manufacturing sector pay more. The general investment climate also appears to matter, with firms that are situated in villages with high average tax payments paying more.[29]

On average, formalization reduces tax payments rather than increasing them. Our results thus contradict conventional wisdom that informality keeps firms under the tax administration's radar screen (Gërxhani 2004, Loayza et al. 2005). A possible explanation for this result may be found in the tax collection process in Indonesia. In his case study on local governance in six Kabupatens, von Luebke (2006) found that tax collection practices are inefficient and to a large extent based on rough estimates or personal negotiations. The results above suggest that licensing may increase the 'bargaining power' of very small

---

[28] Interaction with sales quintiles rather than simply with the value of firm sales was used to alleviate multicollinearity.

[29] Since only roughly half of our firms in the sample paid taxes, we ran a probit model on who paid taxes. Results are reported in Table A4 in the appendix (Columns 1 and 2). We do not find a significant effect of licensing on the probability to pay taxes.

enterprises in such negotiations and thus help to reduce the amount of taxes paid by these firms.

Corruption

Our questionnaire also asked firms about the payment of 'other levies' aside from taxes. These include payments to security officials (i.e. the police), *preman* (organized thugs that extort money), as well as other 'informal' payments requested by sub-district or village officials. In short, 'other levies' are bribes and extortion payments.

Table 5: Unofficial levies about here

In estimating the impact of formality on 'other levies', we need to take account of the relatively low incidence of such payments in our dataset. Only 27 percent of firms report positive values for 'other levies'. In order to address this we run regressions on the full sample (columns 1-3 of Table 5) and on the reduced sample with only positive corruption payments (column 4). We also run a probit regression on who pays unofficial levies at all (see below). Column 1 displays the simple OLS results, with no instrumentation of licensing. As before, we see a positive but insignificant association between licensing and other levies. Column 2 shows the 2SLS results in which licensing has been instrumented: The sign of the licensing variable is now negative and strongly significant, suggesting that, other things equal, having a license reduces the amount of other levies paid. Again, informality does not prevent microenterprises from being targeted by officials, thugs etc.; rather formality reduces both formal and informal payments that they have to make. We could not detect significant interaction effects of formality on corruption implying that licensing has a similar effect on corruption across all types of firms.[30] Other levies are different therefore from taxes, where the benefits of licensing accrued more strongly to smaller firms. The corruption reducing effect of licensing is even stronger in the subsample of firms that pay positive amounts of bribes (column 4).

Several other interesting results emerge. Female owners/managers appear to be less subject to extortion — predominantly male officers or thugs might be more reserved in approaching female managers for other levies — as are owners that live in the same village as the enterprise (i.e. locals). Also owners that belong to the majority religion have to pay

---

[30] There is some evidence that rural firms and firms with Muslim managers in Muslim dominated areas profit more from being licensed; however the effects are significant only at the 16 and 17 percent level, respectively.

less (Hindus in Bali, Muslims in all other districts). That shows that corruption is not a uniform phenomenon, but targets minority groups more heavily than the dominating group.[31]

As in the case of taxes, larger firms pay more unofficial levies. The number of employees is the most 'visible' of all firm size measures. An enterprise with a larger number of employees will be more 'exposed' to security officers, thugs or other officials asking for unofficial payments and less able to change location to escape such harassment. Firms with three or more employees pay significantly more. Likewise firms in the highest two sales quintiles and with fixed assets above the median pay significantly more. We again find a strongly significant environmental effect as expressed by the variable that measures the average corruption level in a village. This variable captures a number of effects such as attitudes of officials, the quality of village governance and the extent of accountability of officials. Comparing the results of 2SLS regressions on the full and the censored sample show largely the same results; yet the extortion-reducing effect of having a license becomes even bigger.

The analysis of who pays bribes at all tells the same story (cf. Table A4 in the appendix, columns 3 and 4). While in a simple probit regression having a license is insignificantly positively correlated with the probability of paying 'unofficial levies'; it does exert a strongly negative influence on the probability to have to pay bribes if appropriately instrumented: The marginal effect is 0.73 ! Thus, licenses are held predominantly by firms that are more likely to be targeted for extortion. The other variables have very similar influences on the probability of having to pay bribes as on the quantity of corruption payments (conditional of being subject to extortion at all): Firms owned by women or locals (i.e. people residing in the village of the enterprise) are less likely to be targeted as are members of the majority religion. Larger firms are more likely to be targeted and again there is a strong environmental effect at the village level on the probability of being subject to corruption payments.[32]

Firm Performance and Business Expansion

As noted above, being licensed may also provide firms with access to new markets and customers, as well as potentially better access to finance. We therefore explore the impact of licensing on three variables, total sales, the share of sales to government, and access to

---

[31] In other specifications we found Chinese managers to pay significantly more and the members of the majority indigenous ethnicity to pay significantly less, thereby providing additional evidence of a 'positive' discrimination of the majority groups. The results reported in Table 5 are very stable with respect to inclusion of other variables.

[32] We also ran a Heckman selection model which resulted in very similar results. Results are available upon request. Significant collinearity existed between the inverse Mills ratio and the explanatory variables of the outcome stage regressions. Evidence from Monte Carlo simulations suggests that, where such collinearity exists, it may be more robust to simply estimate using OLS on the censored dataset (i.e. in our case the sample of firms that pay positive levies) (Puhani, 2000).

credit. OLS estimates can be expected to overestimate the impact of formality, because formalization may contribute to firm growth and larger firms are more likely to formalize.

Total sales

It turns out that overall sales are not affected by the formality status of a firm. Results are reported in Table 6, columns 1-3. Indeed, the licensing variable changes its sign when instrumented but it never reaches conventional significance levels. This result masks a size pattern: larger firms, as measured by employment categories, profit from formalization in terms of sales whereas smaller firms lose; yet the firms in the largest sales quintile do not gain (column 3). This finding corresponds with the evidence from a qualitative survey of firms using licensing services in Indonesia that found that smaller enterprises feel that they are not able to take advantage of the opportunities opened up by formality to increase revenue, such as the facilitation of trading across a wider region or the use of licenses as a sign of legality and trustworthiness to establish new business relations (LabSosio 2008: 43ff). The largest firms may enjoy these benefits even without a license, so that formality does not increase their revenue (but provides other benefits, see above). This finding is in line with McKenzie and Sakho (2010) who find nonlinearity of profit changes from formalization in firm size.

Table 6: Sales and sales to the government about here

The control variables have the expected signs: Firms with a larger work force have higher sales as do older firms and firms with a better educated manager/owner. We observe a sector pattern with trade firms having larger sales as in particular trading firms and again there is a strongly significant effect of the business environment as measured by the average sales level.

Access to government contracts

Government institutions in Indonesia usually require licenses for businesses who want to interact with them (LabSosio 2008: 43). If this is the case, formal firms should have a higher probability to sell to the government and should sell more to it. However, on average, we find no impact of licensing on the share of sales to governments either from our OLS or 2SLS estimates (**Table 6**, Columns 4 and 5). Only three percent of all firms in our sample are dealing with the government, of which manufacturing firms are highly overrepresented: 10 percent of the 250 manufacturing firms in the sample have contracts with the government[33] Such a concentration makes the overall effect insignificant. Yet we find a highly significant

---

[33] We also ran a probit regression on the probability to have government contracts (not reported) which showed clearly that manufacturing firms are much more likely to deal with the government. The formal status had no overall impact on that probability.

and very strong effect of licensing for government contracts in the manufacturing sector (column 6). We also find a clear size effect (not reported): Firms with fixed assets above the mean profit significantly from having a license while the others do not. Lastly the firms with a Muslim manager/owner profit from licensing more than other firms (even though they are not overrepresented in the set of suppliers to the government).

We also ran a probit model on the probability to have a government contract. Results are reported in Table A4 in the appendix. It turns out that overall a formal status does not improve access to government contracts. Interaction effects (for simplicity displayed in a linear probability model) show that the effect of licenses increases the probability by 88 percent for manufacturing firms; it increases the likelihood also for owned by Muslims and firms with fixed assets above the median. These findings corroborate earlier findings on the influence of licenses on the size of government contracts. However, we treat these results with some caution since only 74 firms (3%) in the RICS sample report positive sales to governments – small and micro firms are typically not suppliers to local government.

Credit

Access to credit has widely been considered as a central argument in favor of licensing. The regulations of the Bank of Indonesia require a firm to have a legal business status in the form of a TDP or a SIUP in order to apply for credit from a commercial bank (WB 2006: 51). In practice, banks in Indonesia do not seem to consider licensing as a central determinant for the approval of a loan request. Decisions are rather based on a survey of business feasibility that banks conduct themselves and the collateral or other securities a firm has to offer (ibid: 50, LabSosio 2008: 44ff). Alternative loan sources such as cooperatives or private lending are widely used by small firms, since they are more flexible and have lower requirements.

We measure access to credit in two ways. The first dummy variable (credit) equals one for all firms that have had a loan approved in the last twelve months. The counterfactual group only contains those firms who state that they currently need additional funding for their enterprise. According to this definition, 82 percent of firms have constrained access to credit in the sense that they need additional financial support, but did not receive credit. The second dummy variable (hascredit) equals one if the firm has any loan from a bank in its current asset and loan structure.

Table 7: Access to Credit about here

The results of probit regressions on the access to credit are presented in Table 7. In accordance with the literature, licenses turn out not to be decisive for access to credit (cf.

Fajnzylber et al. 2010 and McKenzie and Sakho 2010). Larger firms, both in terms of the number of employees and in the value of fixed assets are more likely to have access to credit. The latter variable may capture the availability of collateral. Environmental factors play a role as expected: Those firms residing in villages with good access to credit are more likely to have access to credit as do firms located in a village with a bank. Older firms are more likely to have bank loans in their portfolio and firms with owners residing in the same village are more likely to have received a loan in the last 12 months. Interaction effects of licensing with gender, sales quintile dummies or input variables turn out not to be statistically significant at usual levels at the sample means, but have been partly significant for other values of the exogenous variables (cf. Ai and Norton 2003).

An important issue that is not captured sufficiently in the presented regression is the problem of information on access to finance. From the 57 percent of firms in the sample who state that they need additional funding, only 15 percent have applied for a loan in the last twelve months, but of these, only eight percent report that their loan requests have been rejected. One reason for the fact that most firms do not even apply for a loan, despite needing additional funding, may be a lack of information. A third of firms in the RICS stated that they would not apply for a loan from a formal financial institution, because they do not know where or how to apply.

5.2. The Determinants of Formality

The empirical analysis presented in the previous section has shown that the impact of licensing on both costs and benefits depends on a range of firm and manager characteristics. We therefore conclude by estimating the probability of being formal in reduced form, in order to identify whether the characteristics which reduce costs (and enhance benefits) are indeed associated with being formal.

$$\Pr(L_i = 1) = \delta_0 + \delta_1 M_i^S + \delta_2 F_i^S + \delta_3 C_i^S + \delta_4 EC_i + \mu_i \tag{3}$$

Note that equation (3) is quite different from the first stage of our 2SLS procedure, since it contains neither the village level licensing instrument nor the general investment climate variables. Moreover, it does not contain the costs and benefits directly, but rather those manager, firm and location characteristics (M^S, F^S and C^S) that, in interaction with formality, were shown to influence the costs and benefits. *EC* denotes the entry costs into the formal sector. Again, we face a potential endogeneity problem since some firm characteristics that are dependent on firm performance (e.g. firm size and enterprise age)

may be endogenous as they are also affected by the firm's formality status. To mitigate this, we enter potentially endogenous variables as quintile dummies.

Table 8: Firms' Decision to Formalize about here

The results are presented in Table 8. The estimates in Column (1) only use the firm characteristics as independent variables; in column (2), manager characteristics have been added as a proxy for the information available to the firm about licensing procedures. Column (3) adds the existence of banks, as this turned out as significant determinant for the access to credit. The coefficient on education shows the expected positive impact on a firm's probability of being formal. Although slightly different in point estimates, the results of the other coefficients are robust to the inclusion of education.

We have included one-time costs of entering the formal sector (*EC*), proxied by registration costs and time.[34] Both indicators have insignificant coefficients close to zero (output omitted). This outcome could be explained by the fact that uncertainty over costs and licensing procedures, as well as incomplete information, are more important in determining firms' decisions to formalize than the actual amount paid or time waited (cf. Section 3.2; LabSosio 2008: 36).[35] It could also imply that the cost of business registration is on average not the relevant constraint, but that the differences in benefits derived from formality (in terms of lower taxes and informal payments) drive the pattern of formalization.[36]

The probability of being formal increases significantly with all measures of firm size. Firms with three or more employees are between 6 and 9 percent more likely to have a license (after controlling for owner characteristics), firms in the highest sales quintile are around 16 percent more likely to be formal and firms with fixed assets above the mean are around eight percent more likely to have gone formal. Enterprise age seems not to be correlated significantly with formality.

We find clear evidence that better educated owners/ managers are significantly and sizably more likely to formalize their firms. For instance graduates of the junior high school are 4.5 percent more likely to register their firm than owners with primary education only. If formal educational attainment is a predictor for unobservable ability our estimates provide

---

[34] The expected average registration time and cost for each firm was calculated as a weighted mean over employment quartiles, sector and sub-districts, based on the existing data on registration processes provided by the licensed firms in RICS. For those sub-districts that had no data entries on time and costs, district averages have been used.

[35] A more meaningful indicator would be a variable that captures uncertainty over licensing processes by calculating the deviation of the observed time and cost in RICS from officially announced values. However, no data was available on these official figures.

[36] This finding is in line with Bruhn (2010) who finds that simplification of licensing procedures in Mexico did not lead to informal firms becoming formal (but rather provided incentives for wage earners to open small scale businesses). Cf. McKenzie and Woodruff (2006) for similar evidence.

compelling evidence for more productive entrepreneurs to become formal. This supports Levinson and Maloney's notion that more productive managers choose to become formal. Female managers are no less likely than male manager to be licensed, despite gaining less in reduced taxes. For managers who belong to the local indigenous ethnic group, the probability of being formal decreases by about 5 percent (at 13 and 14 percent significance level). By contrast, Chinese managers are more than 35 % more likely to register their firm. This is surprising since Chinese managers of formal firms were found to make a lower share of their sales to governments. However if managers of a different ethnicity, and especially Chinese managers, are under closer scrutiny of local authorities (or are less trusted by other businesspeople) they may have a stronger incentive to get licensed. A similar case could be made for managers that do not live in the place of business, who are significantly more likely to be formal than those living in the village (by around 15 percent).

For rural firms, licensing reduces the amount of taxes paid and the likelihood of paying other levies at all. One might therefore expect rural firms to be more likely to be licensed. In fact being rural reduces the probability of being formal by between 9 and 14 percent. This may reflect higher information, time and monetary costs of obtaining licenses for those in rural areas (Mackenzie and Sakho 2010). Thus, the benefits from licensing will have to be high enough for rural firms in order to outweigh the additional transaction costs of entering the formal sector.

6. Conclusion

In this paper we have analyzed the determinants of going formal for small and micro non-farm enterprises in rural Indonesia. After taking into account the endogeneity of licensing we find that licensing does provide advantages in terms of reduced tax and corruption payments. This result is in stark contrast to the conventional wisdom that formalization will increase tax payments as firms appear on the tax authorities' 'radar screen'.

We have also shown that the benefits of formalization depend on the size of firm and other firm characteristics such as the ethnicity, religion and gender of the manager, the sector the firm operates in, and whether they are located in rural or peri-urban areas. Access to credit and to government contracts seem, on average, to be unaffected by the formality status of a firm. This may be due to the fact that most firms in our sample are discouraged borrowers who do not apply for credit and do not sell to the government, quite independent of their formal status; this result is in line with the findings in the still nascent literature.

We have also analyzed the probability of being formal. Larger firms are more likely to get licensed, as are firms that are owned or managed by more educated individuals. This

suggests that the more visible and the more productive firms will tend to become formal. Firms in rural areas, those run by managers who belong to the local indigenous group, and those who live in the village where their companies are located, are less likely to become formal, while firms that are owned or managed by Chinese Indonesians are more likely to be formal. This suggests that local authorities may favor locals and members of their own ethnicity and put Chinese managers under closer scrutiny.

In short, we find that formality serves as a protective device against excessive rent extraction; it lowers rather than increases payments to the government and bribes. As a consequence those groups that are preferred targets of such extraction – large firms, minority firms and firms without a strong local support system – are more likely to be formal.

References

Ai, Chunrong and Edward Norton (2003) Interaction terms in logit and probit models, *Economics Letters* 80: 123–129.

Angelini, John and Kenichi Hirose (2004): Extension if Social Security Coverage for the Informal Economy in Indonesia. ILO Working Paper 11, December 2004, Geneva.

Auriol, Emmanuelle and Michael Walters (2004): Taxation base in developing countries, *Journal of Public Economics* 89: 625-646.

Bruhn, Miriam (2010) License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico, *Review of Economics and Statistics*, (doi:10.1162/REST_a_00059), forthcoming.

De Paula, Aureo and Jose Scheinkman (2007) The Informal Sector, NBER Working Paper # 13486.

De Soto, Hernando (1989) *The Other Path: Invisible Revolution in the Third World*, New York: Harper and Row.

Djankov, Simeon, Ira Lieberman, Joyita Mukherjee and Tatiana Nenova (2003): Going Informal: Benefits and Costs. In: Belev, Boyan (ed.) *The Informal Economy in the EU Accession Countries*, Sofia: Center for the Studies of Democracy, 63-80.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de-Silanes and Andrei Shleifer (2002): The Regulation of Entry. *Quarterly Journal of Economics*, Vol. CXVII (1): 1-37.

Djankov, Simeon, Caralee McLiesh and Rita Ramalho (2006) Regulation and Growth, *Economic Letters*, 92: 395–401.

Dollar, David, Mary Hallward-Driemeier and Taye Mengistae (2005). 'Investment Climate and Firm Performance in Developing Economies', *Economic Development and Cultural Change*, 54(1):1-31.

Enste, Dominik and Friedrich Schneider (2000) Shadow Economies: Sizes, Causes, and Consequences, *Journal of Economic Literature*, 38: 77-114.

Fajnzylber, Pablo, William Maloney and Gabriel Montes Rojas (2010): Does Formality Improve Micro-Firm Performance? Evidence from the Brazilian SIMPLES Program. *Journal of Development Economics*, doi:10.1016/j.jdeveco.2010.01.009.

Gërxhani, Klarita (2004) The Informal Sector in Developed and Less Developed Countries: a Literature Survey, *Public Choice*, 120, 267–300.

Greene, William . (2007): *Econometric Analysis*, 6th ed. Pearson Education, New Jersey.

Hussmanns, Ralf and Farhad Mehran (1999) Statistical definition of the informal sector – International standards and national practices. http://www.gdrc.org/informal/huss0772.pdf (12/05/2008)

Ishengoma, Ester and Robert Kappel (2006): Economic Growth and Poverty: Does Formalization of Informal Enterprises Matter? GIGA Working Paper No. 20, Hamburg.

Jäckle, Annette . and Carmen Lee (2006): Firm Dynamics and Institutional Participation: A Case Study on Informality of Micro-Enterprises in Peru, *Economic Development and Cultural Change*, 54: 557–578.

Jütting, Johannes and Juan Laiglesia [eds.] (2009) *Is Informal Normal?* Towards more and better jobs in developing countries, OECD: Paris.

Kaplan, David, Eduardo Piedra and Enrique Seira (2007) Are Burdensome Registration Procedures an Important Barrier on Firm Creation? Evidence from Mexico, Centro de Investigation Economica, Mexico DF 10700, discussion paper 07-01.

Kenyon, Thomas (2007): A Framework for Thinking about Enterprise Formalization Policies in Developing Countries. World Bank Policy Research Working Paper 4235, Washington: World Bank.

KPPOD (Regional Autonomy Watch) and The Asia Foundation (2008): Local Economic Governance in Indonesia. http://k: od.org/ind/datapdf/rating2007/rating07.pdf (23/07/2008)

LabSosio and The Asia Foundation (2008): A Qualitative Evaluation of the Impact of One Stop Shop Business Licensing Service on Local Economies and Governance, mimeo. Jakarta.

Levenson, Alec and William Maloney (1998): The informal sector, firm dynamics and institutional participation. World Bank Working Paper 1988: Washington: World Bank.

Loayza, Norman (1997): The economics of the informal sector. A simple model and some empirical evidence from Latin America. World Bank Policy Research Working Paper 1727, Washington: World Bank.

Loayza, Norman, Ana Maria Oviedo and Luis Serven (2005) The Impact of Regulation on Growth and Informality: Cross-country Evidence, World Bank Development Policy Working Paper WPS 1941, Washington: World Bank.

LPEM-FEUI (2006a): RICS Data User Guide and Documentation. Institute for Economic and Social Research, Faculty of Economics, University of Indonesia, Jakarta, mimeo.

LPEM-FEUI (2006b): Field Survey Report. Indonesian Rural Investment Climate Survey January 22-February 6, 2006. Institute for Economic and Social Research, Faculty of Economics, University of Indonesia, Jakarta, mimeo

Luebke, Christian von (2006): Leadership and Voice in Local Governance: Political Economy of District Business Regulations in Central Java, West Sumatrea, Bali and NTB. Paper presented for the Rural Investment Climate Assessment, World Bank, Jakarta, processed.

McCulloch, Neil (ed.) (2009) *Rural Investment Climate in Indonesia*, Institute of Southeast Asian Studies: Singapore.

McKenzie, David and Christopher Woodruff (2006) Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises, *Economic Development and Cultural Change*, 55: 3-42.

McKenzie, David and Seynabou Yaye Sakho (2010): Does it pay firms to register for taxes? The impact of formality on firm profitability, *Journal of Development Economics*, 91(1): 15-24.

Perry, Guillermo, William Maloney and Omar Arias (eds.) (2007): *Informality: Exit and Exclusion*. World Bank Publications, Washington: World Bank,1-20.

Pradhan, Menno and Arthur van Soest (1995) Formal and informal sector employment in urban areas of Bolivia, *Labour Economics*, 2: 275-97.

Puhani, Patrick . (2000): The Heckman Correction for Sample Selection and its Critique. In: *Journal of Economic Surveys*,14(1): 53-68.

Schneider, Friedrich (2005) Shadow Economies around the World: What do we really know?, *European Journal of Political Economy* 21(3): 598-642.

Schulze, Günther G. and Susanne Quadros (2007): Report on the Results of the Rural Investment Climate Assessment (RICA) in Indonesia 2006. University of Freiburg, mimeo.

The Asia Foundation (2007): Making Sense of Business Licensing in Indonesia: A Review of Business Licensing Policy and Survey of One Stop Shop Service Centers, Jakarta.

World Bank (2006) Revitalizing the rural economy: An assessment of the investment climate faced by non-farm enterprises at the district level. Consultative Draft. World Bank Office Jakarta. http://siteresources.worldbank.org/INTINDONESIA/Resources/ Publication/280016-1152870963030/RICAFINALe.pdf (09/04/2008)

World Bank (2007) Doing Business 2008: Country Report Indonesia. International Financial Corporation, Jakarta.

World Bank (2008) *The Rural Investment Climate*. Analysis and Findings. Report No.44009-GLB, Washington.

Wooldridge, Jeffrey (2009*): Econometric Analysis of Cross Section and Panel Data*. 2nd ed., MIT Press, London and Cambridge/Mass.

Tables

Table 1: Firm and manager characteristics

| | District | | | | | |
	Labuan B.	Malang	Badung	Sumbawa	Kutai K.	Barru
Firm characteristics						
# of firms	418	410	399	414	409	411
% rural	80	45	16	47	61	45
% manufacturing	9	18	9	11	6	8
% trade	55	43	62	58	54	54
% services	36	39	29	31	40	38
Mean no. of employees	2.3	2.6	3.3	2.5	2.2	2.6
Mean age of firm (years)	7.2	11.8	6.8	8.9	7.6	12
Median age of firm (years)	4	8	5	6	5	10
Manager characteristics						
Age	41	44	39	42	41	43
% female	32	34	42	51	37	29
% indigenous ethnicity	42	93	65	75	30	92
% Chinese-Indonesian	4	1	2	3	0	0
Highest level of education						
% did not complete primary	5	26	7	16	13	13
% primary	34	39	19	35	32	23
% junior secondary	26	15	15	16	20	18
% senior secondary	27	12	35	27	24	29
% vocational/university	8	8	24	6	11	16

Table 2: Mean costs and benefits

	Formal	Informal
Mean of log total taxes	4.47	1.90
Mean of log total 'other levies'	2.06	1.15
Mean % of sales to government	3.65	0.49
Access to credit	0.24	0.15

Table 3: Mean enterprise size and age

	Formal	Informal
Mean number of employees	3.93	2.23
Mean of log total sales	11.03	9.92
Mean enterprise age	10.77	8.54

Table 4: Taxes

VARIABLES	(1) ln taxes	(2) ln taxes	(3) ln taxes	(4) ln taxes	(5) ln taxes
Licensed	**0.892****				
	(0.412)				
Licensed (instrumented)		**-1.647***	**-3.858*****	**-2.224****	**-1.029**
		(0.984)	**(1.395)**	**(0.947)**	**(1.109)**
Licensed(iv)*sales2			**1.411**		
			(1.219)		
Licensed(iv)*sales3			**2.092**		
			(1.385)		
Licensed(iv)*sales4			**2.015***		
			(1.198)		
Licensed(iv)*sales5			**3.227***		
			(1.673)		
Licensed(iv)*Female				**2.152*****	
				(0.579)	
Licensed(iv)* manufacturing					**-1.577**
					(1.097)
Licensed(iv)*services					**-1.357***
					(0.723)
Education	0.346***	0.486***	0.491***	0.491***	0.491***
	(0.0593)	(0.105)	(0.104)	(0.101)	(0.109)
Age	0.0154	0.0242*	0.0261*	0.0235	0.0265*
	(0.0102)	(0.0136)	(0.0133)	(0.0143)	(0.0140)
Female	-0.354	-0.373*	-0.425*	-0.766***	-0.365*
	(0.235)	(0.219)	(0.231)	(0.244)	(0.219)
Resides in village	0.163	-0.375	-0.287	-0.349	-0.357
	(0.512)	(0.466)	(0.496)	(0.423)	(0.469)
Indigenous	-0.179	-0.284	-0.313	-0.303	-0.285
	(0.277)	(0.270)	(0.276)	(0.273)	(0.273)
Chinese	0.0792	0.732	0.397	0.0187	0.385
	(0.751)	(0.778)	(0.851)	(0.720)	(0.766)
Hindu	0.201	0.0487	0.0426	0.0565	0.0559
	(0.492)	(0.478)	(0.501)	(0.451)	(0.506)
Islam	-0.500	-0.706	-0.855	-0.751	-0.731
	(0.787)	(0.712)	(0.710)	(0.680)	(0.688)
Employment dummy2	-0.337	-0.230	-0.205	-0.224	-0.220
	(0.299)	(0.296)	(0.292)	(0.288)	(0.292)
Employment dummy3	-0.170	-0.00250	0.0317	0.00398	-0.0110
	(0.260)	(0.281)	(0.275)	(0.274)	(0.283)
Employment dummy4	0.563**	0.962***	0.866**	1.043***	1.097***
	(0.284)	(0.345)	(0.333)	(0.348)	(0.341)
sales quintile2	-0.375	-0.419	-0.579*	-0.445	-0.421
	(0.318)	(0.300)	(0.343)	(0.293)	(0.305)
sales quintile3	0.134	0.0942	-0.130	0.110	0.131
	(0.326)	(0.319)	(0.374)	(0.316)	(0.324)
sales quintile4	0.578*	0.655*	0.451	0.612*	0.697**
	(0.327)	(0.337)	(0.377)	(0.336)	(0.338)
sales quintile5	1.012**	1.276***	0.600	1.176***	1.216**
	(0.449)	(0.462)	(0.627)	(0.437)	(0.467)
Fixed assets above median	0.548*	0.735***	0.801***	0.688**	0.732***

	(0.284)	(0.277)	(0.270)	(0.285)	(0.269)
Age of firm	-0.0159	-0.0150	-0.0152	-0.0139	-0.0165
	(0.0149)	(0.0150)	(0.0151)	(0.0159)	(0.0153)
Rural	0.136	-0.131	-0.225	-0.0931	-0.109
	(0.215)	(0.223)	(0.219)	(0.216)	(0.221)
Manufacturing	0.668	0.768**	0.781**	0.736*	1.048**
	(0.410)	(0.386)	(0.374)	(0.401)	(0.500)
Service	0.252	0.212	0.183	0.190	0.476*
	(0.239)	(0.240)	(0.227)	(0.242)	(0.279)
Village Avg Taxes	0.527***	0.593***	0.608***	0.593***	0.593***
	(0.0903)	(0.0769)	(0.0771)	(0.0784)	(0.0758)
Constant	-0.836	-0.972	-0.929	-0.786	-1.227
	(1.217)	(1.278)	(1.323)	(1.230)	(1.294)
District controls	yes	yes	yes	yes	yes
Observations	1,668	1,668	1,668	1,668	1,668
R-squared	0.402	0.396	0.402	0.403	0.399

Robust standard errors in parentheses, clustering at the village level, *** p<0.01, ** p<0.05, * p<0.1

Table 5: Unofficial levies

	(1)	(2)	(3)	(4)
VARIABLES	Ln other levies	Ln other levies	Ln other levies	Ln other levies
Licensed	**-0.00112**			
	(0.243)			
Licensed (instrumented)		**-2.482***	**-1.976**	**-3.547**
		(0.716)	**(0.979)**	**(1.543)**
Licensed(iv)*sales2			**-0.241**	
			(0.935)	
Licensed(iv)*sales3			**-0.431**	
			(1.102)	
Licensed(iv)*sales4			**-0.376**	
			(0.786)	
Licensed(iv)*sales5			**-0.877**	
			(1.258)	
sales quintile2	0.0144	-0.0183	0.00919	-0.267
	(0.295)	(0.283)	(0.316)	(0.350)
sales quintile3	-0.0777	-0.117	-0.0758	0.397
	(0.224)	(0.226)	(0.244)	(0.394)
sales quintile4	0.500	0.576*	0.598	1.014**
	(0.313)	(0.319)	(0.367)	(0.442)
sales quintile5	0.482	0.723**	0.930**	0.797*
	(0.293)	(0.312)	(0.416)	(0.427)
Education	-0.0971	0.0486	0.0475	0.217**
	(0.0616)	(0.0824)	(0.0836)	(0.106)
Age	-0.0178**	-0.00936	-0.00993	0.00895
	(0.00832)	(0.00801)	(0.00809)	(0.00913)
Female	-0.441*	-0.465**	-0.449*	-0.651***
	(0.242)	(0.230)	(0.229)	(0.222)

Resides in village	-0.837***	-1.356***	-1.393***	-1.598***
	(0.305)	(0.365)	(0.383)	(0.378)
Indigenous	0.194	0.0892	0.0941	0.205
	(0.233)	(0.224)	(0.231)	(0.342)
Chinese	0.0666	0.791	0.914	1.560
	(0.760)	(0.633)	(0.670)	(1.238)
Hindu	-1.325***	-1.437***	-1.429***	-0.686*
	(0.349)	(0.358)	(0.371)	(0.356)
Islam	-1.377	-1.559*	-1.510*	-1.524**
	(0.832)	(0.812)	(0.829)	(0.646)
Employment dummy2	0.173	0.258	0.250	0.547*
	(0.182)	(0.182)	(0.185)	(0.308)
Employment dummy3	0.447**	0.610***	0.601***	0.759**
	(0.201)	(0.214)	(0.214)	(0.356)
Employment dummy4	0.574*	0.948***	0.979***	1.065***
	(0.316)	(0.324)	(0.328)	(0.382)
Fixed assets above median	0.339*	0.536***	0.517***	0.518
	(0.194)	(0.195)	(0.191)	(0.352)
Age of firm	-0.00249	-0.00214	-0.00210	-0.00118
	(0.0123)	(0.0120)	(0.0121)	(0.00973)
Rural	0.278*	-0.0299	-0.00293	-0.786*
	(0.164)	(0.163)	(0.161)	(0.422)
Manufacturing	-0.278	-0.193	-0.193	0.335
	(0.292)	(0.272)	(0.275)	(0.619)
Service	0.0297	-0.0170	-0.0136	0.491*
	(0.186)	(0.175)	(0.175)	(0.261)
Village Avg Other levies	0.594***	0.630***	0.633***	0.178
	(0.158)	(0.145)	(0.146)	(0.156)
Constant	3.359***	3.294***	3.279***	4.311***
	(0.785)	(0.713)	(0.679)	(0.797)
Distric controls	yes	yes	yes	yes
Observations	1,668	1,668	1,668	458
R-squared	0.476	0.488	0.489	0.475

Robust standard errors in parentheses, clustering at the village level, *** p<0.01, ** p<0.05, * p<0.1

	sales			Sales to government			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	OLS	2SLS	2SLS	OLS	2SLS	2SLS	2SLS
Licensed	**0.237**			**0.953**			
	(0.223)			**(1.733)**			
Licensed (instrumented)		**-0.453**	**-1.257****		**0.0123**	**-0.881**	**-3.129**
		(0.494)	**(0.502)**		**(2.493)**	**(2.072)**	**(3.404)**
Licensed(iv)* employmentd2			**0.766**				
			(0.786)				
Licensed(iv)* employmentd3			**1.210****				
			(0.483)				
Licensed(iv)* employmentd4			**0.814**				
			(0.612)				
Licensed(iv)* manufacturing						**40.70*****	
						(7.141)	
Licensed(iv)* services						**0.527**	
						(3.257)	
Licensed(iv)* Islam							**4.480***
							(2.680)
Education	0.103***	0.146***	0.148***	0.711**	0.766*	0.414	0.810**
	(0.0267)	(0.0351)	(0.0361)	(0.313)	(0.414)	(0.409)	(0.407)
Age	-0.00788*	-0.00561	-0.00561	0.0315	0.0347	-0.00327	0.0377
	(0.00427)	(0.00528)	(0.00536)	(0.0422)	(0.0442)	(0.0336)	(0.0427)
Female	-0.172	-0.184	-0.185	0.400	0.389	0.284	0.362
	(0.140)	(0.144)	(0.146)	(0.748)	(0.741)	(0.700)	(0.745)
Indigenous	-0.0669	-0.0892	-0.0857	-1.463	-1.505	-1.352*	-1.428
	(0.154)	(0.156)	(0.167)	(0.907)	(0.960)	(0.797)	(0.927)
Chinese	-0.179	0.0323	-0.0193	-3.258	-2.987	-1.619	-1.685
	(0.434)	(0.469)	(0.521)	(2.065)	(1.994)	(2.031)	(2.019)
Hindu	-0.307	-0.343*	-0.335*	-1.389	-1.450	-3.208	-1.957
	(0.197)	(0.186)	(0.197)	(2.657)	(2.601)	(2.743)	(2.816)
Islam	-0.225	-0.265	-0.308	1.042	0.984	0.675	-0.114
	(0.251)	(0.235)	(0.257)	(1.824)	(1.726)	(1.671)	(2.163)
Resides in village	-0.177	-0.322	-0.312	2.341	2.145	2.785	1.605
	(0.231)	(0.278)	(0.275)	(1.523)	(1.537)	(1.753)	(1.319)
Employment dummy2	0.148	0.177	0.0818	-0.262	-0.227	-0.266	-0.233
	(0.179)	(0.171)	(0.174)	(0.556)	(0.520)	(0.500)	(0.528)
Employment dummy3	0.580***	0.640***	0.443**	-0.176	-0.117	-0.000867	-0.186
	(0.177)	(0.174)	(0.197)	(0.835)	(0.757)	(0.752)	(0.723)
Employment dummy4	1.690***	1.826***	1.714***	2.951	3.088	1.563	3.535
	(0.232)	(0.213)	(0.358)	(2.711)	(2.643)	(2.492)	(2.795)
Fixed assets above median	0.0542	0.113	0.132	0.926**	1.002*	0.769	0.913*
	(0.0923)	(0.0964)	(0.0951)	(0.456)	(0.566)	(0.467)	(0.548)
Age of firm	0.0216***	0.0219***	0.0227***	-0.0195	-0.0193	0.00881	-0.0218
	(0.00685)	(0.00680)	(0.00658)	(0.0398)	(0.0403)	(0.0345)	(0.0405)
Rural	-0.0496	-0.108	-0.124	-0.0428	-0.165	-0.113	-0.0461
	(0.142)	(0.158)	(0.157)	(1.024)	(0.974)	(0.611)	(0.902)
Manufacturing	-0.322	-0.311	-0.309	9.729**	**9.763****	2.603	9.783**
	(0.209)	(0.213)	(0.213)	(4.758)	(4.767)	(2.460)	(4.772)
Service	-0.803***	-0.827***	-0.832***	0.735	0.721	0.522	0.761

	(0.111)	(0.114)	(0.112)	(0.847)	(0.823)	(0.452)	(0.799)
sales quintile2				0.384	0.369	0.803	0.277
				(0.719)	(0.716)	(0.542)	(0.723)
sales quintile3				2.147	2.131	2.108	2.005
				(1.351)	(1.362)	(1.337)	(1.293)
sales quintile4				-1.089	-1.055	-0.589	-1.143
				(1.346)	(1.303)	(0.924)	(1.314)
sales quintile5				-1.106	-1.012	-0.852	-0.973
				(1.862)	(1.873)	(1.932)	(1.852)
Village Avg Sales	0.415***	0.460***	0.462***				
	(0.116)	(0.122)	(0.121)				
Village Avg Sales to gov				-0.0265	-0.0219	0.00674	-0.0247
				(0.0818)	(0.0756)	(0.0665)	(0.0727)
Constant	6.004***	5.546***	5.586***	-3.764	-3.732	-0.634	-2.388
	(1.287)	(1.338)	(1.322)	(5.042)	(4.940)	(4.053)	(5.278)
District controls	yes	yes	yes	yes	yes	yes	yes
Observations	1,668	1,668	1,668	1,668	1,668	1,668	1,668
R-squared	0.409	0.408	0.411	0.165	0.163	0.270	0.166

Robust standard errors in parentheses, clustering at the village level, *** p<0.01, ** p<0.05, * p<0.1

Table 7: Access to Credit

VARIABLES	(1) credit	(2) credit	(3) hascredit	(4) hascredit
Licensed	**0.0919**		**0.107**	
	(0.0636)		**(0.0662)**	
Licensed (instrumented)		**-0.0934**		**-0.148**
		(0.143)		**(0.0979)**
Education	-0.0182	-0.00725	0.0123	0.0266**
	(0.0118)	(0.0149)	(0.00955)	(0.0107)
Age	-0.00227	-0.00199	0.000709	0.000917
	(0.00142)	(0.00141)	(0.00107)	(0.00116)
Female	0.0855*	0.0809	0.000307	-0.00771
	(0.0482)	(0.0503)	(0.0254)	(0.0273)
Resides in village	0.111***	0.0911***	0.0437	0.00905
	(0.0209)	(0.0268)	(0.0301)	(0.0396)
Indigenous	0.0241	0.00536	0.0359	0.0146
	(0.0385)	(0.0418)	(0.0310)	(0.0347)
Chinese	-0.00672	0.0428	-0.0587	0.000716
	(0.134)	(0.175)	(0.0548)	(0.104)
Hindu	0.213**	0.209**	0.275	0.254
	(0.0927)	(0.0899)	(0.190)	(0.183)
Islam	-0.0519	-0.0484	-0.232	-0.190
	(0.0911)	(0.0879)	(0.188)	(0.161)
Employment dummy2	0.0161	0.0218	0.0310	0.0397
	(0.0324)	(0.0330)	(0.0275)	(0.0290)
Employment dummy3	-0.00628	-0.00325	0.104**	0.111**

	(0.0452)	(0.0448)	(0.0478)	(0.0486)
Employment dummy4	0.266*	0.317**	0.135	0.216*
	(0.152)	(0.160)	(0.106)	(0.122)
sales quintile2	-0.0608	-0.0730**	-0.00285	-0.0225
	(0.0370)	(0.0363)	(0.0410)	(0.0404)
sales quintile3	0.0265	0.0137	0.0659	0.0504
	(0.0454)	(0.0414)	(0.0453)	(0.0463)
sales quintile4	0.00180	-0.000420	0.0256	0.0259
	(0.0432)	(0.0447)	(0.0429)	(0.0473)
sales quintile5	-0.0303	-0.0315	0.116*	0.116
	(0.0493)	(0.0507)	(0.0701)	(0.0748)
Fixed assets above median	0.0454*	0.0617**	0.0358	0.0566**
	(0.0265)	(0.0257)	(0.0231)	(0.0230)
Age of firm	0.00289	0.00330	0.00351**	0.00415***
	(0.00263)	(0.00266)	(0.00157)	(0.00154)
Rural	-0.0161	-0.0425	0.0137	-0.0194
	(0.0313)	(0.0350)	(0.0235)	(0.0279)
Bank	0.0389	0.0632*	0.00830	0.0426
	(0.0344)	(0.0367)	(0.0240)	(0.0260)
Manufacturing	-0.0161	-0.0217	-0.0936***	-0.0919***
	(0.0693)	(0.0678)	(0.0206)	(0.0233)
Service	0.0292	0.0307	7.99e-05	0.00561
	(0.0473)	(0.0482)	(0.0327)	(0.0322)
Village Avg Credit	0.320**	0.299**	0.239***	0.214***
	(0.133)	(0.137)	(0.0750)	(0.0716)
District controls	yes	yes	yes	yes
Observations	1,015	1,015	1,015	1,015
Pseudo R squared	0.2081	0.2036	0.2286	0.2204

Marginal effects for probit regressions, endogenous variables credit and hascredit (at sample means). For dummy variables we report the effect of changing the value from zero to one. Robust standard errors in parentheses, clustering at the village level *** p<0.01, ** p<0.05, * p<0.1

Table 8: Firms' Decision to Formalize

VARIABLES	(1) Licensed	(2) Licensed	(3) Licensed
Education		0.0449***	0.0408***
		(0.00728)	(0.00780)
Age		0.00350***	0.00330***
		(0.00110)	(0.00106)
Female		-0.0106	-0.00779
		(0.0325)	(0.0317)
Resides in village		-0.142***	-0.155***
		(0.0679)	(0.0636)
Indigenous		-0.0490†	-0.0447†
		(0.0324)	(0.0306)
Chinese		0.384**	0.354**
		(0.154)	(0.144)
Hindu		-0.00405	-0.00484
		(0.0505)	(0.0476)
Islam		-0.00246	-0.0165
		(0.0964)	(0.0946)
Employment dummy2	0.0163	0.0305	0.0329
	(0.0409)	(0.0365)	(0.0327)
Employment dummy3	0.0757*	0.0620*	0.0670**
	(0.0449)	(0.0394)	(0.0366)
Employment dummy4	0.204***	0.0767*	0.0891**
	(0.0714)	(0.0506)	(0.0474)
sales quintile2	0.00465	-0.00514	-0.0174
	(0.0477)	(0.0356)	(0.0330)
sales quintile3	-0.0276	-0.00446	-0.0116
	(0.0472)	(0.0488)	(0.0462)
sales quintile4	0.0871	0.0931†	0.0913†
	(0.0694)	(0.0689)	(0.0687)
sales quintile5	0.210**	0.172**	0.161*
	(0.110)	(0.0993)	(0.0964)
Fixed assets above median	0.0789***	0.0877***	0.0820***
	(0.0227)	(0.0222)	(0.0220)
Age of firm	-0.000177	0.00114	0.00114
	(0.00195)	(0.00168)	(0.00173)
Rural	-0.185***	-0.137***	-0.0898**
	(0.0461)	(0.0467)	(0.0402)
Bank			0.107***
			(0.0339)
Manufacturing	0.0172	0.00919	0.00545
	(0.0455)	(0.0386)	(0.0355)
Service	-0.00671	-0.00415	-0.00364
	(0.0300)	(0.0228)	(0.0218)
District controls	yes	yes	yes
Observations	2,164	1,901	1,873
Pseudo R 2	0.2474	0.3300	0.3475
Wald chi 2	333.37***	755.86***	959.67***

Marginal effects of a probit regression with robust standard errors in parentheses, clustering at the village level, endogenous variable LICENSED, *** p<0.01, ** p<0.05, * p<0.1, † p<0.14

Appendix

Table A1. List of Variables

Variable Name	Description
Age	Owner/ manager age in years
Bank	dummy for whether a financial institution is available in the village area
Licensed	dummy =1 for firms that hold at least one of the four most common local licenses (TDI, SIUP, IMB or TDP)
Licensed (instrumented)	Instrumented value for licensed, different values for each regression (cf. Table A3)
Licensed(iv)*sales3	Interaction of the instrumented license variable with the dummy for the third sales quintile (other interactions analogous)
Pay other levies	dummy for firms with positive payments of 'other levies'
Credit	access to credit (dummy =1 for firms that have received a loan or credit in the last 12 months; dummy=0 for all firms that need additional funding)
Education	manager education (ordinal scale)
# Employees	total number of employees
Age of firm	Age of the enterprise in years
Fixed assets above median	dummy = 1 for firms with fixed assets above median
Female	dummy =1 for female manager
Fully formalized	dummy =1 for firms being fully formalized
Sales to gov	share of sales made to government
Indigenous	dummy =1 for manager who belongs to a local indigenous group
Islam	dummy = 1 for Muslim owner/ manager in Muslim dominated areas
Hindu	dummy = 1 for Hindu owner/ manager in Hindu dominated area (Badung)
Chinese	Dummy ==1 for manager/ owner of Indonesian-Chinese descent
Ln other levies	log of total 'other levies' at central, provincial and district level (+1)
Ln fixed assets	log of total fixed assets 2005 (includes land, buildings, equipment, furniture, vehicles) (+1)
Ln sales	log of total sales 2005
Sales quintile2-5	Dummies for quintiles 2-5 of Ln sales
Ln taxes	log of total taxes at central, provincial and district level (+1)
Manufacturing	dummy for firm in manufacturing sector
Mean cost	average expected registration costs averaged over employment, location and sector
Mean time	expected registration time, averaged over employment, location and sector
Employment dummy 2-4	Dummies for 2/ 3-4/ 5 or more employees (including owner)
Qlnsales1-5	dummies for ln sales quintiles
Resides in village	dummy =1 for owner/ manager who lives in the same village as the business operate
Rural	dummy for rural =1 and urban =0
Service	dummy for firm in service sector
Pays tax	dummy for firms with positive tax payments
Village Avg Licensed	village average of Licensed (excluding the firm considered)
Village Avg Other levies	village average of Ln other levies (excluding the firm considered)

Village Avg Credit	village average of credit (exclluding the firm considered)
Village Avg Sales to gov	village average of Sales to gov (exclluding the firm considered)
Village Avg Sales	village average of Ln sales (exclluding the firm considered)
Village Avg Taxes	village average of Ln taxes (exclluding the firm considered)

Table A2 Descriptive Statistics

Variable Name	Obs.	Mean	St.Dev.	Min	Max
Fully formalized	1668	.0173861	.1307443	0	1
Licensed	1668	.220024	.4143867	0	1
Village Licensed	1668	.2247662	.2307698	0	1
Pays tax	1668	.4784173	.4996838	0	1
Ln taxes	1668	2.565023	2.991886	0	12.11424
Village Ln taxes	1668	2.44259	1.572285	0	9.98461
Pay other levies	1668	.2745803	.4464364	0	1
Ln other levies	1668	1.364606	2.390503	0	12.26435
Village Other levies	1668	1.312532	1.430844	0	5.759327
Age	1668	41.91187	11.75985	16	85
Female	1668	.381295	.4858504	0	1
Indigenous	1668	.6792566	.4669023	0	1
Islam	1668	.8135492	.3895869	0	1
Hindu	1668	.1007194	.3010471	0	1
Chinese	1668	.0203837	.1413513	0	1
Resides in village	1668	.8944844	.3073087	0	1
# Employees	1668	2.59952	2.174797	1	19
Age of firm	1668	8.905875	8.412085	1	56
Ln sales	1668	10.16535	1.583465	4.49981	16.99356
Village ln sales	1668	10.17525	.7514047	8.330511	13.80638
Sales to gov	1668	1.118705	8.19632	0	100
Village Sales to gov	1668	.9720125	2.44819	0	33.33333
Ln fixed assets	1668	8.028777	4.095533	0	18.1355
Rural	1668	.4898082	.500046	0	1
Bank	1650	.6030303	.4894179	0	1
Credit	1032	.1579457	.3648669	0	1
Hascredit	1668	.103717	.3049845	0	1
Village Credit	1653	.1905161	.1872667	0	1
Mean Time	1525	531.1666	490.6214	13.5	2150
Mean Cost	1588	10.35739	8.958577	1	60

For consistency descriptive statistics refer t0 the data set used in the regression tables 4 to 7.

Table A3 First stage regression

VARIABLES	(1) Licensed	(2) Licensed	(3) Licensed	(4) Licensed	(5) Licensed
Education	0.0403***	0.0409***	0.0431***	0.0403***	0.0370***
	(0.00740)	(0.00739)	(0.00826)	(0.00723)	(0.00928)
Age	0.00373***	0.00375***	0.00354***	0.00373***	0.00142
	(0.00105)	(0.00104)	(0.00107)	(0.00106)	(0.00111)
Female	-0.0110	-0.0110	-0.0197	-0.00975	-0.0317
	(0.0263)	(0.0265)	(0.0254)	(0.0261)	(0.0272)
Resides in village	-0.206***	-0.204***	-0.210***	-0.202***	-0.192*
	(0.0620)	(0.0631)	(0.0661)	(0.0614)	(0.107)
Indigenous	-0.0231	-0.0214	-0.0229	-0.0237	-0.0370
	(0.0267)	(0.0267)	(0.0272)	(0.0264)	(0.0274)
Chinese	0.104	0.105	0.139	0.108	0.0780
	(0.137)	(0.138)	(0.146)	(0.140)	(0.117)
Hindu	-0.0597	-0.0563	-0.0584	-0.0538	-0.0156
	(0.0446)	(0.0438)	(0.0438)	(0.0446)	(0.0411)
Islam	-0.0676	-0.0675	-0.0620	-0.0700	0.0515
	(0.120)	(0.119)	(0.115)	(0.121)	(0.0482)
Employment dummy2	0.0498	0.0490	0.0519	0.0569*	0.0307
	(0.0313)	(0.0312)	(0.0349)	(0.0310)	(0.0289)
Employment dummy3	0.0929**	0.0941**	0.116**	0.0992**	0.0352
	(0.0459)	(0.0460)	(0.0499)	(0.0466)	(0.0320)
Employment dummy4	0.135***	0.134***	0.196***	0.144***	0.174**
	(0.0477)	(0.0477)	(0.0488)	(0.0481)	(0.0867)
sales quintile2	-0.0246	-0.0229		-0.0228	-0.0565**
	(0.0323)	(0.0327)		(0.0320)	(0.0245)
sales quintile3	-0.0220	-0.0233		-0.0250	-0.0646**
	(0.0398)	(0.0396)		(0.0394)	(0.0309)
sales quintile4	0.0306	0.0278		0.0267	-0.00483
	(0.0531)	(0.0534)		(0.0534)	(0.0466)
sales quintile5	0.0854	0.0815		0.0758	-0.00563
	(0.0745)	(0.0747)		(0.0748)	(0.0543)
Fixed assets above median	0.0837***	0.0860***	0.0855***	0.0847***	0.0570**
	(0.0202)	(0.0200)	(0.0185)	(0.0197)	(0.0223)
Age of firm	-3.13e-05	-8.56e-05	0.000457	-0.000140	0.00340**
	(0.00141)	(0.00138)	(0.00136)	(0.00138)	(0.00169)
Rural	-0.0367	-0.0380	-0.0329	-0.0290	-0.0202
	(0.0354)	(0.0338)	(0.0351)	(0.0326)	(0.0310)
Village Avg Licensed	**0.424*****	**0.424*****	**0.391*****	**0.465*****	**0.390*****
	(0.0697)	(0.0728)	(0.0687)	(0.0786)	(0.0717)
Manufacturing	-0.0115	-0.0123	-0.0255	-0.0174	0.00133
	(0.0353)	(0.0353)	(0.0369)	(0.0331)	(0.0495)
Service	-0.0194	-0.0217	-0.0384*	-0.0174	0.0243
	(0.0276)	(0.0277)	(0.0226)	(0.0276)	(0.0271)
Village Avg Taxes	0.00275				
	(0.00897)				
Village Avg Other levies		0.00892			
		(0.00978)			
Village Avg Sales			0.0248		
			(0.0175)		
Village Avg Sales to gov				-0.00783*	
				(0.00446)	
Bank					0.0522**
					(0.0240)

Village Avg Credit					-0.0968*
					(0.0556)
District controls	yes	yes	yes	yes	yes
Observations	1,668	1,668	1,668	1,668	1,015
Pseudo R2	0.4022	0.4027	0.3941	0.4045	0.4223
Wald chi2	969.31***	935.25***	924.58***	973.76***	472.67***

Marginal effects of a probit regression with robust standard errors in parentheses, clustering at the village level, endogenous variable LICENSED, *** p<0.01, ** p<0.05, * p<0.1

Table A4: Probability to pay taxes and corruption payments

VARIABLES	(1) taxdum	(2) taxdum	(3) corrptdum	(4) corrptdum
Licensed	**0.121**		**-0.0191**	
	(0.0802)		**(0.0662)**	
Licensed (instrumented)		**-0.352**		**-0.725****
		(0.216)		**(0.287)**
Education	0.0593***	0.0844***	-0.0230	0.0195
	(0.0150)	(0.0207)	(0.0176)	(0.0246)
Age	0.00551**	0.00695**	-0.00548***	-0.00309
	(0.00258)	(0.00288)	(0.00211)	(0.00235)
Female	-0.0342	-0.0390	-0.0839	-0.0909*
	(0.0560)	(0.0533)	(0.0582)	(0.0539)
Resides in village	-0.00108	-0.0972	-0.131	-0.293***
	(0.123)	(0.123)	(0.0831)	(0.0935)
Indigenous	-0.0841	-0.104	0.0491	0.0247
	(0.0748)	(0.0730)	(0.0643)	(0.0641)
Chinese	0.0749	0.200	-0.0709	0.180
	(0.171)	(0.165)	(0.105)	(0.200)
Hindu	0.0870	0.0624	-0.203***	-0.220***
	(0.135)	(0.132)	(0.0774)	(0.0742)
Islam	-0.0754	-0.111	-0.333**	-0.373***
	(0.110)	(0.0981)	(0.146)	(0.142)
Employment dummy2	-0.107	-0.0914	0.0184	0.0316
	(0.0700)	(0.0700)	(0.0594)	(0.0600)
Employment dummy3	-0.0918	-0.0573	0.0836	0.127*
	(0.0595)	(0.0603)	(0.0641)	(0.0681)
Employment dummy4	0.00418	0.0758	0.0664	0.163
	(0.0795)	(0.0915)	(0.0995)	(0.108)
sales quintile2	-0.127	-0.132	0.0322	0.0207
	(0.0900)	(0.0877)	(0.0858)	(0.0826)
sales quintile3	-0.0156	-0.0240	0.00208	-0.000831
	(0.0787)	(0.0786)	(0.0703)	(0.0726)
sales quintile4	0.118*	0.128*	0.111	0.141*
	(0.0671)	(0.0658)	(0.0696)	(0.0768)
sales quintile5	0.0891	0.133	0.178**	0.265***
	(0.0986)	(0.0934)	(0.0752)	(0.0891)
Fixed assets above median	0.180***	0.213***	0.0629	0.118*

	(0.0620)	(0.0679)	(0.0776)	(0.0634)
Age of firm	-0.00540	-0.00512	-0.00210	-0.00216
	(0.00338)	(0.00333)	(0.00327)	(0.00321)
Rural	0.0459	0.00297	0.146**	0.0432
	(0.0459)	(0.0439)	(0.0654)	(0.0617)
Manufacturing	0.134	0.143	-0.126**	-0.108*
	(0.101)	(0.0966)	(0.0602)	(0.0566)
Service	0.0588	0.0485	-0.0615	-0.0752
	(0.0565)	(0.0560)	(0.0542)	(0.0520)
Village Avg Taxes	0.130***	0.145***		
	(0.0194)	(0.0195)		
Village Avg Other levies			0.145***	0.157***
			(0.0388)	(0.0360)
District controls	yes	yes	yes	yes
Observations	1,668	1,668	1,668	1,668
Pseudo R sqrd	0.2570	0.2566	0.3674	0.3828
Wald statistic	754.14***	847.28***	531.95***	548.56***

Probit estimation, marginal effects at sample mean, for dummy variables for discrete change from 0 to 1.
Robust standard errors in parentheses, clustering at the village level, *** p<0.01, ** p<0.05, * p<0.1 Tax
dummy and corruption dummy have sample means of 0.469 and 0.273 respectively.

Table A5: Probability of having a government contract

VARIABLES	(1) probit	(2) probit	(3) 2SLS	(4) 2SLS	(5) 2SLS	(6) 2SLS
Licensed	**0.00208**					
	(0.00248)					
Licensed (instrumented)		**-0.000823**	**0.00325**	**-0.0366**	**-0.0496**	**-0.110**
		(0.00103)	**(0.0517)**	**(0.0408)**	**(0.0539)**	**(0.0829)**
Licensed(iv)* manufacturing				**0.880***		
				(0.182)		
Licensed(iv)* service				**0.0713**		
				(0.0558)		
Licensed(iv)*Islam					**0.0777***	
					(0.0458)	
Licensed(iv)* fixed assets above median						**0.129***
						(0.0654)
Education	0.000366	0.000504	0.0191**	0.0114	0.0197**	0.0210**
	(0.000293)	(0.000403)	(0.00802)	(0.00707)	(0.00792)	(0.00826)
Age	-1.18e-05	-4.58e-06	0.000798	-1.44e-05	0.000861	0.000899
	(1.36e-05)	(1.47e-05)	(0.000698)	(0.000445)	(0.000681)	(0.000726)
Female	0.000199	0.000223	0.00963	0.00699	0.00912	0.00731
	(0.000406)	(0.000467)	(0.0128)	(0.0122)	(0.0126)	(0.0126)

Resides in village	0.000235	-2.78e-05	0.000693	0.0106	-0.00918	-0.0117
	(0.000249)	(0.000534)	(0.0272)	(0.0214)	(0.0248)	(0.0295)
Indigenous	-0.00133	-0.00168	-0.0216	-0.0305**	-0.0232	-0.0268*
	(0.00131)	(0.00164)	(0.0176)	(0.0154)	(0.0179)	(0.0160)
Chinese	0.00120	0.00250	-0.0170	0.0265	0.00568	-0.0308
	(0.00208)	(0.00374)	(0.0524)	(0.0368)	(0.0579)	(0.0563)
Islam	0.000740	0.000653	0.0327	0.0298	0.0142	0.0293
	(0.000829)	(0.000793)	(0.0269)	(0.0221)	(0.0298)	(0.0288)
Employment dummy2	-0.000603	-0.000700	-0.00830	-0.00859	-0.00817	-0.00782
	(0.000530)	(0.000603)	(0.00925)	(0.00860)	(0.00936)	(0.00914)
Employment dummy3	-7.91e-05	-8.17e-05	-0.00519	-0.00329	-0.00658	-0.00194
	(0.000368)	(0.000363)	(0.0132)	(0.0130)	(0.0129)	(0.0134)
Employment dummy4	0.00107	0.00121	0.0417	0.00656	0.0497	0.0400
	(0.00161)	(0.00180)	(0.0326)	(0.0267)	(0.0326)	(0.0312)
sales quintile2	0.00308	0.00343	0.0101	0.0177	0.00826	0.00964
	(0.00311)	(0.00329)	(0.0139)	(0.0112)	(0.0136)	(0.0138)
sales quintile3	0.00410	0.00442	0.0245	0.0224	0.0223	0.0249
	(0.00408)	(0.00414)	(0.0154)	(0.0146)	(0.0147)	(0.0155)
sales quintile4	0.000378	0.000927	-0.00901	-0.000777	-0.0103	-0.00835
	(0.00101)	(0.00169)	(0.0180)	(0.0118)	(0.0175)	(0.0177)
sales quintile5	0.00102	0.00199	-0.00130	0.00491	-0.000540	0.00432
	(0.00192)	(0.00303)	(0.0188)	(0.0196)	(0.0187)	(0.0200)
Fixed assets above median	0.000741	0.00115	0.0198*	0.0133	0.0178	-0.00137
	(0.000671)	(0.000962)	(0.0110)	(0.00882)	(0.0108)	(0.0121)
Age of firm	2.35e-05	3.04e-05	0.000111	0.000610	3.36e-05	2.73e-05
	(2.28e-05)	(2.70e-05)	(0.000811)	(0.000698)	(0.000809)	(0.000814)
Rural	8.46e-05	-0.000208	-0.00145	-0.00294	0.000245	-0.00365
	(0.000409)	(0.000539)	(0.0189)	(0.0117)	(0.0180)	(0.0185)
Manufacturing	0.198***	0.199***	0.170**	0.0153	0.171**	0.167**
	(0.0764)	(0.0766)	(0.0747)	(0.0232)	(0.0746)	(0.0740)
Service	0.00445	0.00433	0.0226*	0.00835	0.0237*	0.0217
	(0.00320)	(0.00307)	(0.0131)	(0.00870)	(0.0128)	(0.0132)
Village Avg Sales to gov	7.74e-05	9.09e-05	0.00131	0.00146	0.00117	0.00123
	(0.000106)	(0.000124)	(0.00168)	(0.00163)	(0.00161)	(0.00166)
Constant			-0.106*	-0.0453	-0.0859	-0.0832
			(0.0640)	(0.0487)	(0.0630)	(0.0599)
District controls	yes	yes	yes	yes	yes	yes
Observations	1,668	1,668	1,668	1,668	1,668	1,668
Pseudo/ R-squared	0.5351	0.5253	0.166	0.318	0.168	0.174

Columns 1-2: Probit estimation, marginal effects at sample mean, for dummy variables for discrete change from 0 to 1.
Columns 3-6: Linear probability model with instrumented licensed variable and selected interactions. Robust standard errors in parentheses, clustering at the village level, *** p<0.01, ** p<0.05, * p<0.1