Public Disclosure Authorize

Public Disclosure Authorized

Public Disclosure Authorized

# Does It Pay Firms to Register for Taxes?

## The Impact of Formality on Firm Profitability

*David McKenzie*
*Yaye Seynabou Sakho*

## Abstract

This paper estimates the impact of registering for taxes on firm profits in Bolivia, the country with the highest levels of informality in Latin America. A new survey of micro and small firms enables the authors to control for a rich set of measures of owner ability and business motivations that can affect both profits and the decision to formalize. The paper identifies the impact of tax registration on business profitability using the distance of a firm from the tax office where registration occurs, conditional on the distance to the city center, as an instrument for registration. Proximity to the tax office provides firms with more information about registration, but is argued to not directly affect profits. The findings show that tax

registration leads to significantly higher profits for the firms that the instrument affects. However, there is also evidence of heterogeneous effects of tax formality on profits. Tax registration is found to increase profits for the mid-size firms in the sample, but to lower profits for both the smaller and larger firms, in contrast to the standard view that formality increases profits. The analysis shows that owners of large firms who have managed to stay informal have higher entrepreneurial ability than formal firm owners, in contrast to the standard view (correct among smaller firms) that informal firm owners have low ability.

# Does it pay firms to register for taxes?
# The impact of formality on firm profitability

David McKenzie, *World Bank, BREAD and IZA*[#]
Yaye Seynabou Sakho, *World Bank*

## 1. Introduction

Informality is pervasive amongst firms in developing countries, yet simple comparisons of formal and informal firms usually reveal that formal firms are more productive and profitable. A series of high-profile sector studies by the McKinsey Global Institute around the world comparing the operation of formal and informal firms concluded that informality has a very negative impact on productivity, even going so far as to conclude that "in Portugal and Turkey, for instance, informality accounts for nearly 50 percent of the overall productivity gap with the United States" (Farrell 2004). However, such estimates ignore the fact that formality is a choice of firms – the lower productivity of informal firms may therefore just reflect less productive firms choosing to remain informal rather than be the consequence of informality.

This paper seeks to provide more credible evidence of the impact of being formal on business profits, using new survey data on firms in urban Bolivia, the country with the highest levels of informality in Latin America (World Bank, 2007). We control for a rich set of owner characteristics, which provide a more comprehensive set of measures of owner ability, background, and motivations for entering business than existing studies of formality. We then identify the impact of registering for a tax identification number on firm profits by using the GPS-measured distance of a firm from the tax office where registration for taxes occurs as an instrument for whether or not a firm is registered for taxes. After controlling for the distance of the firm to the city center and other locational characteristics, we argue that this distance affects the information a firm has about registration, but does not independently affect profits. We show our results are robust to several potential threats to this identification assumption.

We find that registering to pay taxes leads to significantly higher profits for the firms that the instrument affects. Nevertheless, we also provide evidence of heterogeneous effects of formality on profitability. In particular, we find that although registering for taxes

increases profits for firms in the middle size group in our sample (2 to 5 workers and the middle tercile of capital stock), registering for taxes is associated with lower profits for firms smaller than this, and for firms larger than this. The main benefit of registering for taxes appears to be an increase in the customer base through the ability to issue tax receipts – we find no evidence of increased access to finance. Very small firms are too small to benefit from issuing tax receipts, while owners of large informal firms have high ability and can achieve a large customer base through their own business skills.

This research adds to a nascent literature on the micro-level impacts of informality on firms. Despite increasing research issue concerning the nature of informality, its determinants, and its macro effects (e.g. Loayza (1996), Schneider and Enste (2000), Maloney (2004, 2006), Perry et al. (2007)), there are currently few studies which attempt to provide rigorous estimates of the impact of formality on firms themselves. In a recent study, Fajnzylber, Maloney, and Rojas (2006a) use propensity-score matching methods and control function approaches to estimate the impact of paying taxes and belonging to business associations on the performance of Mexican micro enterprises. They find relatively large impacts, with paying taxes estimated to increase business profits by at least 20 percent, and belonging to business associations estimated to increase business profits by at least 10 percent. However, their identification relies on assumptions about the formality status of firms being determined either on the basis of a set of observable variables or through a specific functional form in the estimation equation. If firms select into formality on the basis of unobserved owner ability or firm productivity, these may overestimate the impact of becoming formal.

A second set of recent studies have looked at the short-term impacts of increases in formality induced by business simplification procedures. Monteiro and Assunção (2006) and Fajnzylber, Maloney, and Rojas (2006b) use difference-in-differences and regression discontinuity designs respectively to obtain non-experimental estimates of the impact of the SIMPLES program in Brazil. They find higher revenues for firms that operate with a license, but do not examine whether these revenues are offset by higher expenses (such as taxes), or whether these higher revenues are also accompanied by higher profits. Two

recent studies of a business simplification program in Mexico (Kaplan et al, 2006 and Bruhn, 2006) did not find much impact of reforms on registration amongst existing microenterprises. This is consistent with the view that firms weigh the costs and benefits of becoming formal, with small firms seeing few benefits.

Our research builds on these existing studies in several important ways. The instrumental variables approach we use here provides a source of identification driven by differences in information across firms, which together with the rich set of owner characteristics as controls, provides a new, and perhaps more credible, approach to identifying the impact of formality. Secondly, this study is the first we are aware of to empirically examine the heterogeneity of impacts by firm size, with our finding of negative impacts on larger firms in particular a new one. Finally, we also examine some of the channels through which tax registration affects profits, which can aid in designing policies intended to make registration more attractive for firms.

The remainder of the paper is structured as follows. Section 2 describes the new survey data used; Section 3 describes the registration process in Bolivia and the measure of formality we choose to focus on; Section 4 provides a simple model of a firm's formalization decision and uses this to motivate our identification strategy; Section 5 provides the main empirical results; and Section 6 concludes.

## 2. Data

The data come from the Bolivian *Encuesta de Productividad de Empresas*, a survey of micro- and small enterprises designed by the authors and carried out during March 2007.[1] The survey covered the four largest cities in Bolivia - La Paz, El Alto, Santa Cruz and Cochabamba – with additional surveying conducted in several rural areas. We restrict our analysis to urban areas in this paper, and in particular, to firms within a 10 kilometer radius of the city's tax office.[2] The resulting sample consists of 469 firms.

---

[1] The survey was conducted by the Bolivian survey firm *Encuestas y Estudios.*
[2] The total sample including rural areas is 629 firms. We also dropped the few firms surveyed which were not owned by the person interviewed (in order that we can control for owner characteristics), and a few firms in transportation with above 20 workers.

## 2.1 Sampling Design

Six industries were chosen for the survey: grocery stores, restaurants and food sales, manufacturing of clothing from wool and cloth, transportation of passengers and cargo, manufacturing of clothing from camelid wool (from llamas and alpacas), and manufacturing of furniture from wood. The industries were chosen to represent a large portion of the self-employed and small employers, and to encompass a diversity of sectors. According to the 2005 MECOVI (Bolivian Living Standards Measurement Survey), the industries chosen include four of the top five industries[3] for urban small and medium enterprises and cover approximately 40 percent of all self-employed and employers.

The sample frame consisted of a geographic information database maintained by the survey firm. This database is based on a census of all economic establishments in these cities carried out in August 2005, and includes enterprises operating within households. This was supplemented with data for the transportation sector on all firms that have registered their cars. This provided a reasonably comprehensive sampling frame for urban areas.[4] The sample was stratified across cities and firm size, in order to include a mix of micro firms (with less than five workers), and small firms with 5-20 workers. The sample used here is almost equally divided across the four cities. Groceries, Food, and Transportation each constitute about 20 percent of the sample, Clothing from wool and cloth and manufacturing from wood each constitute 15 percent of the sample, and clothing from camelids the remaining 10 percent.

## 2.2 Sample characteristics

Table 1 provides summary statistics for the 469 firms used in this paper. The median firm has been in business for 9 years, although 27 percent have been in business 3 years or

---

[3] The only one of the top five industries not covered is construction, where it was felt there was little overlap between formal and informal firms, and where the broad industry grouping is very heterogeneous, covering skilled engineers and architects along with less skilled workers.

[4] A sample frame was not available for rural areas, and therefore snowball sampling methods were used to survey camelid and wood firms in rural areas. The lack of a representative sampling frame provides a further reason for excluding rural areas in this study.

less. The median firm in the sample has 2 paid workers, with only 20 percent of the sample having five or more paid workers. Mean monthly profits are 1628 Bolivianos ($US211).[5] Profits were measured through a direct question, following the recommendations made in de Mel, McKenzie and Woodruff (2007). Half the firm owners are female. Average education levels for owners are quite high, but there is substantial variation. The median owner has 12 years of education, but 20 percent of firm owners have less than 6 years education, and 31 percent have more than 12 years.

### 2.3. Measuring Owner Ability and Background

A main concern for our empirical work is that there are characteristics of firm owners which affect both profitability and the decision of whether or not to formally register the business. One such factor could be owner ability, which may determine the size of the gain in profits from becoming formal. Other such factors could be family background and wealth, which might affect the ability of the firm to meet the costs of formalizing, and the utility associated with being formal. Our instrumental variables strategy aims to overcome such concerns, but we also attempt to directly measure and control for these variables much more than is possible with variables such as gender, own education, age, marital status and ethnicity found in standard firm surveys.

We use three different measures of owner ability and motivation. The first is mother's education, consisting of dummy variables for mother having no education (36 percent of firm owners), and for mother having 9 or more years education (16 percent). The second measure is entrepreneurial self-efficacy. Self-efficacy is the first principal component of ten questions intended to measure the self-assessed ability to perform certain tasks. Owners were asked on a four point scale how confident they are that they could do ten business tasks. Examples include their ability to estimate accurately the costs of a new project; to resolve a difficult dispute with a client or supplier in another city; to hire good employees to expand their business; to sell a product to a new client; and to price their business correctly if they wish to sell it. Such a measure is more closely tied to business skills than years of education, and has a correlation of only 0.11 with years of education.

---

[5] During March 2007, 1 USD was approximately 7.7 Bolivianos.

The third measure consists of three dummy variables capturing motivations for going into business. Owners who entered self-employment for the chance of business growth may be more likely to become formal than those who entered self-employment in order to have flexibility to care for family or carry out household tasks.

Additional measures of family background are whether or not their father was a business owner (34 percent say yes), and a childhood poverty index, intended to capture family wealth well before the time of business entry. The childhood poverty index is a principal component based on mother's education, questions about the type of floor their house had as a child (60 percent had a dirt floor, 15 percent cement, while others had mosaic, tile, or wood); and the frequency with which they didn't have enough to eat as a child (31 percent say they never went hungry, 54 percent said sometimes, and 15 percent say almost always or always). This index has a correlation of -0.48 with own education, 0.41 with speaking an indigenous language as a child, and -0.15 with log capital stock, suggesting it is measuring family wealth to some extent.

## 3. The registration process in Bolivia in theory and practice

In order to be fully formal, a firm in Bolivia must register with three different Government agencies, while those with employees are required to register their employees with three additional agencies. The three main steps to formalization are

1) Registering with the municipal government (*Alcaldia*) to obtain a municipal business license,
2) Registering with the tax authorities (*Servicios de Impuestos Nacional (SIN)*) to get a tax identification number (NIT), and
3) Registering in the registry of commerce (*Fundempresa*).

Firms with employees must take the further steps of registering them for health benefits with the National health system (CNS), for social security with the Pension fund (AFP), and registering them with the Labor Ministry.

However, out of our sample of 469 firms, only 13 have completed the three steps listed here, and only 5 also have workers registered with the three additional agencies. In

practice then, almost all micro and small enterprises are informal to some extent. The creation of *Fundempresa* in 2002 does not seem well-known to firms in 2007. Only 10 percent of firms in our sample say they know what *Fundempresa* is and what its purpose is, and only 18 firms are registered with it.

Therefore for most firms in our sample, formality consists of first registering for a municipal license, and then registering for a tax identification number (NIT). 57 percent of firms in our sample have a municipal license, and 29 percent a NIT. Only 14 of the 134 firms with a NIT do not have a municipal license. Recent years have seen improvements in the time taken to obtain municipal licenses in some municipalities. Coupled with greater interaction with municipal officials, our focus group surveys revealed that most firms seemed aware of the process necessary to get a municipal license and felt enforcement was stronger at the municipal level. In contrast, no such efforts have been made to simplify the process of obtaining a tax identification number, and, as we will show, many firms do not have good knowledge of this process. As a result, we believe that most firms failing to have a municipal license do not have one as a result of informed choice, whereas some firms without tax identification numbers do not have one as a result of imperfect information. The result is that we will be able to identify the impact of a tax identification number, but not that of a municipal license.

Our survey firm conducted 12 focus group interviews of firms across different sectors and size groups (Encuestas y Estudios, 2007). Firms in the focus groups were asked what they themselves understood formality to be. For the majority of members of our focus groups, formality means *dar facturas* or providing formal receipts, for which a NIT is required. Others mentioned being registered at the municipal level and paying municipal fees. The measures of formality this paper focuses on are therefore the one firms themselves see as defining what it means to be formal.

We asked firms in our focus groups and in the main survey what they saw as the main benefit of having a NIT. One-quarter of surveyed firms with a NIT said the main benefit was to increase their customer base, through being able to issue tax receipts. Clients can

use these tax receipts for claims or tax refunds. The other main benefits according to firms are avoiding fines (19 percent say this), and to be obeying the law (43 percent say this). The benefit of obeying the law may also reflect less uncertainty about fines or other punishments, or reflect some psychic benefit associated with obeying the rules of society.

## 4. Modeling the choice of firms to become formal and our identification strategy
### 4.1. The choice of formality
Profit-maximizing firms will choose whether or not to register for taxes if the expected present discounted value of the net benefits from doing so outweighs the upfront costs. That is, a firm will choose to get a tax identification number if and only if:

$$\sum_{t=1}^{T} \delta^{t} E\left(\pi_{F,t} - \pi_{I,t}\right) + \theta_{law-abiding} > C_{Money} + C_{Time} + C_{Information} \tag{1}$$

where $\pi_{F,t}$ denotes the firm's profits if it is formally registered at time t, and $\pi_{U,t}$ denotes the firm's profits if it is not formally registered at time t. $\theta_{law-abiding}$ denotes the utility benefit to firm owners from obeying the law and feeling they are contributing to national welfare through paying taxes. $C_{Money}$, $C_{Time}$, and $C_{Information}$ denote the monetary, time, and information costs from registering. The monetary costs also implicitly include the shadow value of capital for liquidity-constrained firms. For simplicity of exposition we assume here that the value of being formal at time *t* does not depend on formality status in previous periods. If it does, the problem can be written as a dynamic optimization problem with a value function, and the same intuition as exposited here will still apply.

Equation (1) shows the key challenge of identifying the impact of formality on firm profitability. The choice to register for taxes or become formal will itself depend on the impact of formality on business profits ($\pi_{F,t}$ - $\pi_{U,t}$). As a result, if there are no costs to registering and no non-monetary benefits from doing so, all firms for which it is profitable to be registered will have done so, and only firms for which informality does not pay will remain informal. If there are homogeneous treatment effects, everyone will

either be formal or everyone informal. If there are heterogeneous effects, those who are informal will not be appropriate controls for those who are formal.

However, equation (1) also shows that some firms for which it is profitable to become formal will not do so if the initial costs of registering are too high. In the Bolivian case, the monetary costs of registering for a NIT are very low (zero for the registration itself, with only the costs of photocopying and obtaining accompanying documents needed). However, the time and information costs will depend on how closely located the firm is to the office where they have to register, and on how much general information is available about the registration process. This provides a potential source of identification: comparing two firms with the same potential net change in profits (and same psychic benefits) from obtaining a tax identification number, but with different time and information costs will enable estimation of the increase in profits resulting from formalization.

## 4.2 Identification

Our basic identification strategy is therefore to use the geographic location of the firm relative to that of the tax office as a source of variation on the information and time costs of registering, assuming that after other locational controls are added, that distance to the tax office has no independent impact on profits.

The latitude and longitude of each firm was obtained using GPS receivers. In each city there is a single location of the tax office where firms can register for a NIT. We also obtained the coordinates of these tax offices, and of the city centers. We then calculate the straight-line distance from the firm to the nearest tax office and to the city center. Since the sample is contained in dense urban neighborhoods, straight-line distances will be good approximations for actual travel distances (Gibson and McKenzie, 2007). We have restricted the sample to firms within a 10 kilometer radius of the tax office: the mean firm is 3.2 kilometers away. The assumption is that for two firms of equal distance from the city center, the one that is closer to the tax office has better knowledge of the registration process.

Our survey indeed reveals limited knowledge about the process of getting a tax number. 39 percent of firms without a NIT say they don't know what a tax number is. Only 44 percent of unregistered firms say they know where the nearest tax office is, and only 31 percent are able to give its location to within 3 kilometers. Table 2 uses the sample of unregistered firms to examine whether there is a relationship between distance to the tax office and knowledge of where the tax office is. Column 2 shows a large, but not significant effect on whether the firm owner says they know where the office is. Since some firm owners may claim to know where it is, but not actually know, we also asked them to tell us the distance to the office. Columns 3 and 4 show that there is a large and significant effect of distance to the tax office on this measure of ignorance.

Another reason that distance to the tax office could determine whether or not a firm registers is that the chance of receiving a visit from a tax inspector may be greater the closer a firm is to the tax office. Column 7 of Table 2 shows that this is indeed the case when we don't condition on firm and location characteristics. However, in all of our regressions we will include controls for city, industry, the average tax enforcement rate in each city*industry pair, and the average tax enforcement rate in a 1 kilometer radius around the firm.[6] Column 8 shows that after including these controls, there is no effect of distance to the tax office on enforcement.

We thus believe that after the inclusion of firm, owner, and locational controls, distance to the tax office is determining the information a firm has about registration, but is not having an independent effect on its profits. We examine possible threats to this exclusion restriction in the robustness section below.

### 4.3. First-stage: does distance to the tax office predict formality?
Table 3 then examines whether log distance to the tax office predicts whether or not a firm has a tax identification number, after controlling for firm owner, firm, and location

---

[6] In both cases we do not include the firm's own tax enforcement in calculating the average of firms around it.

characteristics. We show marginal results from estimation of a probit equation where having a NIT is the dependent variable. We present two sets of results. Column 1 does not control for firm size, and shows a strong and significant effect of log distance to the tax office on having a NIT. The marginal effect shows that being 3.2 kilometers away from the tax office (the mean) is associated with a 13 percentage point reduction in the likelihood of having a NIT. Since only 29 percent of firms have a NIT, this is a sizeable effect. Column 2 shows this result also holds after conditioning on firm size, as measured by log of capital stock, and dummy variables for the number of paid workers being in different categories. As one would expect, larger firms are more likely to be formal.

Column 1 also highlights the role of owner ability in the formality decision. Owners with more education, higher entrepreneurial self-efficacy, and who went into self-employment for business growth reasons are more likely to have a NIT. After controlling for owner ability, there is no significant effect of gender on the formality decision. Secondly, column 1 shows that firms are more likely to have a NIT when enforcement is higher: both the inspection rate within their city*industry pair and within a 1 kilometer radius of the firm are positively and significantly associated with having a NIT.

In columns 3 and 4 of Table 3, we then examine whether a similar strategy can be used to estimate the impact of a municipal license. Municipal licenses are obtained at the municipal office (*Alcaldia*). However, while there is a weak negative correlation between log distance to the municipal office and having a municipal license (-0.10), this correlation disappears after controlling for firm, owner, and locational characteristics. This is consistent with the view that firms have reasonable knowledge about how and where to get a municipal license, so that distance does not provide sufficient variation in information access to enable identification in this case.

## 5. The impact of registering for taxes on profits

We now turn to estimation of the impact of a NIT on firm profits. We focus on profits as the dependent variable for two reasons. First, from the point of view of a firm deciding on whether or not to register, the impact on profits is a key determinant. Secondly, whilst from a social welfare point of view one would also like to know whether formality improves productivity, estimation of firm productivity is much more complicated than profitability, and relies heavily on assumptions that are unlikely to be credible with the types of firms surveyed here (see the critique in Katayama, Lu, and Tybout, 2006).

The basic estimation equation of interest is, for firm $i$:

$$\ln(profits)_i = \alpha + \beta NIT_i + \delta'X_i + \gamma'Z_i + \theta'L_i + \varepsilon_i \tag{2}$$

where $NIT_i$ is a dummy variable taking value one if firm $i$ has a tax identification number, $X_i$, $Z_i$, and $L_i$ are owner, firm, and locational characteristics respectively. We instrument the tax identification number using log distance to the tax office. We employ two methods of estimating (2). The first is two-stage least squares, which has the advantage of not imposing distributional assumptions on the error term. The second method is maximum-likelihood estimation (MLE) of a treatment effects model, which takes account of the fact that $NIT$ is a binary variable, and jointly estimates a probit equation for $NIT$ along with equation (2).[7] In both cases we estimate the equation with and without controls for firm size.

Columns 1 and 2 of Table 4 provide the MLE estimates and columns 3 and 4 provide the 2SLS estimates. The estimates show very large, and highly significant, increases in profits from registering for a tax identification number. The dependent variable is log profits, so the MLE estimate in column 2 suggests that firms which obtain a tax identification number have 86 percent higher profits. The 2SLS estimates are larger, but not significantly so. The MLE estimates are significant at the 1 percent level, and the 2SLS estimates at the 5 percent level. Thus according to these estimates, firms earn much larger profits after registering.

---

[7] This was carried out using the treatreg command in STATA.

Several other variables are also notable in Table 4. First, after conditioning on size and other characteristics, male and female owners don't have significantly different profitability. Secondly, there is a strong and significant effect of entrepreneurial self-efficacy on profits – more able owners earn higher profits. Third, profits are lower when the tax inspection rate is higher, which is consistent with higher taxes reducing profits. Fourth, profits are higher for larger firms. Finally, note the point estimate on log capital stock (which excludes land and buildings), suggests a 10 to 12 percent per month return on capital for the mean firm in the sample. This return to capital is comparable to that obtained by McKenzie and Woodruff (2006) for microenterprises in Mexico.

### 5.1. Robustness

In Table 5 we examine the robustness of our results to potential challenges to the identification strategy. Column 1 of Table 5 repeats the MLE estimate from column 2 of Table 4: an 86 percent increase in profits from having a NIT. The first potential threat to our exclusion restriction is that more able entrepreneurs may choose to locate their businesses in locations which are closer to the tax office. The fact that so few unregistered firms know where the tax office is also leads us to believe it extremely unlikely that firms are choosing their business location to be closer to or further from the tax office. If they are, unobserved owner ability will not be orthogonal to distance.

Assuming that observed and unobserved ability have similar correlations with distance, in column 2, we examine the sensitivity of our results to dropping our measures of ability and owner background. The point estimate drops from 0.86 to 0.81, with the difference not significant. Therefore the inclusion of ability measures does not significantly affect the estimated treatment effect, which provides support for the view that more able entrepreneurs are not choosing locations closer to the tax office. Instead, micro and small entrepreneurs typically operate at, or near, their homes, making the location decision of the firm a function of where the owner happens to live. In our sample, 33 percent of firms

are operating out of the dwelling of the owner.[8] The point estimate of the treatment effect is still 0.78 for this sub-sample, although with the smaller sample size, it is not statistically significant.

A second potential threat to the exclusion restriction is that despite controlling for the distance to the city center, distance to the tax office may be correlated with other interactions with the government, which could also affect profitability. To examine the robustness of our results to this possibility, in column 3 we add the distance to the municipal office when municipal licenses are issued as a further control. The mean distance from the municipal office to the tax office is only 1.4 kilometers. We see this variable has no effect on firm profitability, and leads to no significant change in the estimated effect of a NIT. This adds further support to our claim that the distance to the tax office is picking up the specific information and time costs of obtaining a NIT, and not something else to do with other interactions with the government.

Finally, while we have controlled for the general tendency of firms in a particular location to receive visits from a tax inspector, in column 4 of Table 5 we also include whether a firm itself has received a visit from a tax inspector. This is likely to be endogenous to profits, but conditioning on this provides a further check of the robustness of our results to controls for enforcement. Again there is no significant change in the coefficient after adding this control, and if anything, the point estimates show even larger impact of a tax identification number. Thus our results appear robust to several potential threats to our identifying assumption.

## 5.2. Comparison with OLS

Columns 5 through 8 of Table 4 provide OLS estimates of the impact of having a NIT on log profits, for comparison with the treatment effect estimates provided by MLE and 2SLS. We begin by including the same control variables as in columns 1 through 4. Columns 5 and 6 show that OLS provides much lower estimates of the impact of a NIT.

---

[8] Firm owners were surveyed at their business locations, and therefore we do not have the GPS coordinates of the home locations which would allow us to say precisely how many firms are operating within a kilometer or two of their homes.

Without controlling for firm size, having a NIT is associated with 43 percent higher profits. However, after controlling for size, there is no significant effect of a NIT, with the point estimate only being 1.2 percent higher profits.

However, the typical concern in estimating equation (2) by OLS is that there are unobservable firm or owner characteristics which are correlated with the decision to become formal and which also affect profits. Two of the most common such factors would be owner ability and anticipated positive productivity shocks. In both cases, this would lead to a positive bias in the OLS estimate of the impact of having a NIT. That is, after using instrumental variables, we would expect to find *lower*, not higher, impacts of NIT.

To examine whether the direction of bias in the OLS estimates is working the way we would expect, columns 7 and 8 of Table 4 remove our measures of owner ability, background, and reasons for going into business. If wealthier, more able owners, who expect their businesses to grow are the ones more likely to formalize, then removing these variables should lead to larger coefficients on NIT. This is indeed what we see. When we don't control for firm size and we fail to control for these owner characteristics, the impact of a NIT is to increase profits by 53 percent, compared to 43 percent when these controls are included. While our proxies for ability are more comprehensive than most studies, there may still be unobserved ability components that influence both profits and formality status. However, based on this analysis, we would expect such terms to operate in the same way, if anything lowering the OLS estimates further with their inclusion. Thus it does not appear that this simple story of ability bias can explain the difference between the OLS and MLE/2SLS estimates.

**5.3 Treatment effect heterogeneity and the difference from OLS**

OLS will also provide a different estimate than the MLE/2SLS estimates if there are heterogeneous treatment effects. That is, if the impact of obtaining a tax identification number differs across firms. If selection into formality status occurs only in terms of the observable variables included in equation (2), and columns 5 and 6 of Table 4, OLS will

estimate the average treatment effect. In contrast, 2SLS and MLE will estimate local average treatment effects (Imbens and Angrist, 1994). Specifically, they will estimate the impact of registering for taxes on firms for which distance to the tax office influences their formality status. These are firms for which distance makes a difference to whether they know how to register, or which are enough at the margin of registering that a small change in the travel costs of registering will influence their registration decision.

There are good reasons to believe that the reasons for being informal vary with firm size, since many of the costs and benefits of becoming formal will vary with firm size. To explore the possibility of different effects of a NIT, we use propensity-score matching to estimate the impact of having a NIT for different firm size groupings.[9] The same set of variables as were used in the OLS regressions are used for matching purposes. Nearest neighbor matching was used to estimate average treatment effects matching firms to their three nearest neighbors[10].

Table 6 reports the results of propensity score matching. The first two columns report the results of matching, with and without controls for firm size. After controlling for firm size, the average treatment effect of a NIT is a 20.9 percent increase in profits – larger than the OLS coefficient, but smaller than the MLE/2SLS coefficients. Again, with heterogeneous treatment effects the propensity-score matching treatment effect will differ from the average treatment effect estimated by OLS and the local average treatment effect estimated by MLE and 2SLS. We then split firms into three size groupings, based on terciles of capital stock and groupings of 0 to 1, 2 to 5, and 6 and above workers. Although the smaller sample sizes when we split the sample this way lead to some statistically insignificant results, the point estimates are suggestive of heterogeneity of treatment effects. Having a NIT is estimated to *lower* profits for very small firms by 61.6

---

[9] One could also estimate the treatment effects regressions for subsamples. However, with our rich set of controls and smaller samples, convergence of the MLE was a problem for these restricted samples. The propensity-score matching estimators enable the large number of covariates to be employed more efficiently, and thus we use this method to investigate heterogeneity of effects.

[10] Matching was carried out using the nnmatch command in STATA. See Abadie, Drukker, Herr and Imbens (2004).

percent, increase profits for slightly larger firms by 40.8 percent, but then *lower* profits for firms with 6 or more workers in the top tercile of capital stock, by 54.6 percent.

What explains this possible nonlinear relationship between formality and profitability? The explanation for very small firms is clear. These firms are too small to immediately benefit from formalization. Registering for taxes immediately involves more costs in the form of tax payments, but these firms are too small to benefit from increased customer base or better access to credit. However, if these firms plan on growing over time, the current cost of formalizing may be justified in terms of the anticipated future benefits. Consistent with this, we find the firm owners who are formal at this small size to be of much higher ability. Firms who are formal in the middle size group (2 to 5 workers and the middle tercile of capital stock) are big enough to enjoy some of the benefits of formalizing, and tax registration increases profits for this group. The mid-sized formal firm owners are still of higher ability than those who remain informal.

A new insight from this work concerns the somewhat larger firms – those with 6 or more workers and in the top tercile of capital stock. The propensity score matching estimate shows a large and statistically significant *negative* effect of tax formality on profits for this group of firms. The firm owners who have managed to get to this size and remain informal have higher entrepreneurial self-efficacy than firm owners of this size with registered firms. These firms have likely figured out ways to avoid inspections and access many of the potential benefits of formality, without having to pay taxes. Only 25 percent of these larger informal firms received a visit by tax inspectors in 2006, compared to 77 percent of formal firms of the same size.

Our finding that the owners of larger informal firms have higher entrepreneurial ability than owners of larger formal firms is in contrast to the prediction of Rauch (1991), that smaller and informal businesses are more likely to be run by less-talented entrepreneurs. In his model, enforcement only occurs for larger firms, and more talented entrepreneurs are endogenously allocated to bigger firms, where they are compensated for the costs of regulation. However, in practice, the talent of an entrepreneur may also affect his or her

ability to evade enforcement. In this case, more talented entrepreneurs may have more incentive to stay informal – consistent with what we see here.

This heterogeneity in the effect of tax registration seems a likely explanation for the differences in the effect of a NIT over the different methods. OLS averages the effect of a NIT over all firm sizes: the smaller and larger firms who have negative effects, and the middle-sized group of firms who have positive effects. As a result, the average over all these groups is close to zero. Propensity-score matching places more weight on the firms in the middle size group, since these firms are more likely to have similar propensity scores. The MLE and 2SLS estimates will estimate the effect for firms at the margin of becoming formal, for whom information and travel costs make a difference. These are unlikely to be the very small or largest firms, but rather a subset of the middle-sized group, for whom the gains to formality are very high indeed.

**5.4. How does having a NIT change profits?**
Our treatment effects regression has found effects of a NIT on profits for some firms. In Table 7 we explore several avenues through which a NIT may change profits. Maximum-likelihood treatment effects regressions using log distance to the tax office as an instrument for a NIT are used, with our standard set of control variables, including firm size.

First and foremost, the main cost of registering for taxes is that we would expect firms with a NIT to be more likely to be paying taxes. The first two columns of Table 7 show this to be the case. Firms with a NIT are much more likely to pay taxes, and pay a larger share of their profits as taxes.

The main benefit of a NIT according to the firms themselves is the ability to attract more customers by issuing tax receipts. Columns 3 and 4 of Table 7 provide evidence for this positive effect. Having a NIT results in higher sales, and a greater likelihood of issuing tax receipts. This increase in customers is not coming through more contracting with the government, multinationals, or large firms: 97 percent of our firms make no sales to any

of these entities. Further suggestive evidence of more customers is seen through industry-specific questions asked to firms in grocery sales and transportation. Grocery firms with a tax number are 19 percentage points less likely to have many periods during the day without customers than firms without a tax number, controlling for firm size and firm and owner characteristics. Transport firms with a tax number spend a larger share of each day with customers and less time idle.

Another often-discussed potential benefit of formality is better access to credit (e.g. Straub (2005)). However, Table 7 shows no significant effect of a NIT on the use of trade credit, the provision of working capital from suppliers or customers, and on the likelihood of having a bank loan. There are two likely reasons for this lack of effect. The first is that only a minority of firms use the financial system. Access to credit for many firms is a significant issue, with 56 percent of firms saying the procedures for accessing credit are a medium or very severe obstacle to their growth, and 67 percent saying the cost of credit is a medium or very severe obstacle to their growth. Secondly, our focus groups and country discussions suggest that when Banks do decide to grant credit, they are not concerned with the tax status of a firm, and if anything, are just concerned with whether or not they have a municipal license.

A final way in which being registered may increase profits is by lowering the costs of corruption. Only 2 percent of firms give this as the major benefit of having a NIT. Nevertheless, the last column of Table 7 shows a sizeable, although not quite significant, effect of having a NIT on whether a firm believes corruption is a constraint to their business growth. There is therefore some suggestive evidence for this channel operating.

## 6. Conclusions

Profit-maximizing firms will choose whether or not to formalize according to whether or not doing so increases their profitability. This provides a challenge for identifying the impact of formality on profits, since firms self-select into formality status. However, similar firms which have different access to information about formalizing may end up with different formality status, even though they have the same net benefits from

formalizing. As a result, we can estimate the impact of registering for taxes on firms for whom the choice to formalize is affected by their distance to the tax office. We find large effects of registering for taxes on the profits of these firms, with the effects appearing to come mainly through increases in their customer base.

However, we also find evidence that the effect of formalizing differs across firms. While there is a group of firms which can benefit from formalizing, formality lowers profits for very small firms – who are too small to benefit -  and for the larger firms in our sample. Firm owners with enough ability to grow their firm to a size of six workers or more and achieve a reasonably large capital stock would have lower profits from formalizing – they would pay more taxes, but are already able to reach a large customer base without formalizing.

The irony then is that registering for taxes would seem to benefit most the informal firms who don't know how to formalize. Conceptually, those who are informal by choice will expect lower profits from formalizing otherwise they would be formal, whereas those who are informal due to ignorance of the procedures needed to formalize stand to benefit from learning how to. Consequently, the results of this study highlight the need for better information provision by Governments to induce firms at the margin to formalize. They also suggest the need for policies which target the high ability owners of larger informal firms, through providing more benefits of formality and tighter enforcement.

**References**

Abadie, Alberto, David Drukker, Jane Herr and Guido Imbens (2004) "Implementing Matching Estimators for Average Treatment Effects in Stata", *Stata Journal* 4(3): 290-311.

Bruhn, Miriam. 2006. "License to sell: the effect of business registration reform on entrepreneurial activity in Mexico." Mimeo. Cambridge, Mass.: MIT.

De Mel, Suresh, David McKenzie and Christopher Woodruff (2007) "Measuring Microenterprise Profits: Don't ask how the sausage is made", *World Bank Policy Research Working Paper Series* No. 4229.

Encuestas y Estudios (2007) "Informe Estudio Cualitativo: Informalidad y Productividad" (Qualitative Study Report: Informality and Productivity), Report submitted to the World Bank, February 2007.

Farrell, Diana. -2004. "The hidden dangers of the informal economy." *The McKinsey Quarterly* (Number 3). http://www.mckinseyquarterly.com/article_page.aspx?ar=1448&L2=7&L3=10&srid=6&gp=1

Fajnzylber, Pablo, William Maloney, and Gabriel Montes Rojas. 2006a. "Releasing Constraints to Growth or Pushing on a String? The Impact of Credit, Training, Business Associations and Taxes on the Performance of Mexican Micro-Firms." World Bank Policy Research Working Paper No. 3807. Washington, D.C.

_____. 2006b. "Does Formality Improve Micro-Firm Performance? Quasi-Experimental Evidence from the Brazilian SIMPLES program."

Gibson, John and David McKenzie (2007) "Using the Global Positioning System (GPS) for Better Economics and Better Policy", forthcoming, *World Bank Research Observer.*

Imbens, Guido and Joshua Angrist (1994) "Identification and Estimation of Local Average Treatment Effects", *Econometrica* 62(2): 467-75.

Kaplan, David, Eduardo Piedra, and Enrique Seira. 2006. "Are Burdensome Registration Procedures an Important Barrier on Firm Creation? Evidence from Mexico." Mimeo. Washington, D.C.: World Bank.

Katayama, Hajime, Shihua Lu and James Tybout (2006) "Firm-level productivity studies: Illusions and a Solution", Mimeo. Pennsylvania State University.

Loayza, Norman (1996) "The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America", *Carnegie-Rochester Conference Series on Public Policy* 45: 129-62.

Maloney, William (2004) "Informality Revisited", *World Development* 32(7): 1159-78.

Maloney, William (2006) "Informality Deconstructed", Mimeo. World Bank.

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(1): 3-42.

Monteiro, Joana and Juliano Assunção. 2006. "Outgoing the shadows:

estimating the impact of bureaucracy simplification and tax cut on formality and investment." Mimeo. PUC Rio

Perry, Guillermo, William Maloney, Omar Arias, Pablo Fajnzylber, Andrew Mason and Jaime Saavedra (2007) *Informality: Exit and Exclusion*. World Bank Latin America and Caribbean Studies: World Bank, Washington D.C.

Rauch, James (1991) "Modeling the informal sector formally", *Journal of Development Economics* 35(1): 33-47.

Schneider, Friedrich. and Dominik Enste (2000) "Shadow Economies: Size, Causes and Consequences", *Journal of Economic Literature* 38(1): 77-114.

Straub, Stéphane (2005) "Informal sector: the credit market channel", *Journal of Development Economics* 78(2): 299-321.

World Bank (2007) *Setting Bolivia's Course on the Road to Formality*, Mimeo. The World Bank.

**TABLE 1: SUMMARY STATISTICS**

| | # obs | Mean | S.D. | 25th | 50th | 75th |
|---|---|---|---|---|---|---|
| Female | 469 | 0.50 | | | | |
| Age of Owner | 465 | 42.0 | 12.1 | 32 | 42 | 50 |
| Married | 469 | 0.68 | | | | |
| Spoke indigenous language as child | 469 | 0.30 | | | | |
| Years of Education | 469 | 10.5 | 4.5 | 6.5 | 12 | 14 |
| Mother had no education | 469 | 0.36 | | | | |
| Mother had 9 or more years education | 469 | 0.16 | | | | |
| Entrepreneurial Self-efficacy | 469 | 0.00 | 1.94 | -1.55 | 0.02 | 1.59 |
| Went into self-employment to care for family | 469 | 0.62 | | | | |
| Went into self-employment for flexible hours | 469 | 0.57 | | | | |
| Went into self-employment for business growth | 469 | 0.68 | | | | |
| Childhood poverty index | 469 | 0.00 | 1.44 | -1.42 | 0.34 | 1.43 |
| Father was a business owner | 469 | 0.34 | | | | |
| Distance to the Tax Office (km) | 469 | 3.22 | 1.90 | 1.77 | 2.97 | 4.29 |
| Distance to the city center (km) | 469 | 3.25 | 1.98 | 1.61 | 3.00 | 4.44 |
| Distance to the municipal office (km) | 469 | 3.30 | 1.75 | 1.95 | 3.38 | 4.28 |
| Average tax inspection rate in city*industry | 469 | 0.34 | 0.17 | 0.16 | 0.39 | 0.43 |
| Tax inspection rate in 1km radius of firm | 459 | 0.36 | 0.25 | 0.17 | 0.33 | 0.50 |
| Firm has a municipal licence | 469 | 0.57 | | | | |
| Firm has a tax identification number (NIT) | 469 | 0.29 | | | | |
| Age of the firm (years) | 466 | 12.0 | 13.0 | 3 | 9 | 17 |
| Firm is 3 years or less in business | 469 | 0.27 | | | | |
| Number of Workers | 469 | 3.20 | 3.11 | 1 | 2 | 5 |
| Number of Paid Workers | 466 | 2.30 | 3.00 | 0 | 1 | 4 |
| Zero paid workers | 469 | 0.35 | | | | |
| One to Four Paid workers | 469 | 0.45 | | | | |
| Five to Ten Paid Workers | 469 | 0.17 | | | | |
| Eleven or More Paid Workers | 469 | 0.03 | | | | |
| Log Capital Stock (excluding land and buildings) | 418 | 8.90 | 2.15 | 7.65 | 9.24 | 10.50 |
| Monthly Profits (Bolivianos) | 416 | 1628 | 2104 | 500 | 843 | 2000 |
| Log Monthly Profits | 416 | 6.81 | 1.09 | 6.21 | 6.74 | 7.60 |

Source: World Bank Bolivian *Encuesta de Productividad de Empresas 2007.*

**TABLE 2: DO UNREGISTERED FIRMS CLOSER TO THE TAX OFFICE HAVE BETTER KNOWLEDGE?**

Marginal Coefficient from Probit estimation

| | Doesn't know where tax office is | | Doesn't know where tax office is or is more than 3km off in guess | | Firm was been visited by a tax inspector in 2006 | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (7) | (8) |
| Log distance to tax office | 0.0268 | 0.0783 | 0.0763** | 0.0939* | -0.0732** | 0.0122 |
| | (0.041) | (0.066) | (0.038) | (0.055) | (0.030) | (0.038) |
| Firm owner controls | no | yes | no | yes | no | yes |
| Firm characteristics | no | yes | no | yes | no | yes |
| City dummies | no | yes | no | yes | no | yes |
| Industry dummies | no | yes | no | yes | no | yes |
| Observations | 335 | 325 | 335 | 325 | 335 | 325 |

*Notes:* Robust standard errors in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% levels.
Firm owner controls are age, gender, marital status, indigenous language use, education, mother's education,
Childhood poverty, reasons for going into business, and entrepreneurial self-efficacy.
Firm characteristics are a dummy for firm age of 3 years or less, log distance to city center, and average
tax enforcement rates for the city*industry and for a 1km radius around the firm.

**TABLE 3: FIRST-STAGE - DOES DISTANCE PREDICT FORMALITY?**
Marginal effects from probit regression for having a Tax identification number (NIT) (columns 1 and 2) and for having a municipal licence (columns 3 and 4).

| | NIT | | Municipal Licence | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Female | -0.0949 | 0.0298 | -0.0769 | 0.00275 |
| | (0.059) | (0.063) | (0.071) | (0.082) |
| Age of Owner | 0.00312 | 0.00453** | 0.00130 | 0.00128 |
| | (0.0021) | (0.0022) | (0.0026) | (0.0029) |
| Owner is Married | 0.0344 | 0.00435 | 0.0608 | 0.0786 |
| | (0.048) | (0.051) | (0.059) | (0.064) |
| Owner speaks Indigenous language | 0.0170 | -0.0251 | 0.00726 | -0.0404 |
| | (0.058) | (0.055) | (0.067) | (0.073) |
| Years of Education | 0.0180*** | 0.0135** | -0.00255 | -0.00589 |
| | (0.0066) | (0.0067) | (0.0078) | (0.0085) |
| Mother had no education | 0.0119 | 0.0821 | 0.00486 | 0.0529 |
| | (0.066) | (0.069) | (0.073) | (0.080) |
| Mother had 9 or more years education | 0.0266 | 0.0496 | -0.0331 | -0.00365 |
| | (0.069) | (0.076) | (0.082) | (0.090) |
| Entrepreneurial Self-efficacy | 0.0258** | 0.0200 | 0.0171 | 0.0147 |
| | (0.012) | (0.013) | (0.015) | (0.016) |
| Went into self-employment to care for family | -0.0711 | -0.0188 | -0.0698 | -0.0905 |
| | (0.051) | (0.054) | (0.059) | (0.066) |
| Went into self-employment for flexible hours | -0.00630 | -0.0331 | 0.0336 | 0.0598 |
| | (0.048) | (0.052) | (0.057) | (0.063) |
| Went into self-employment for business growth | 0.0859* | 0.0211 | -0.0283 | -0.0696 |
| | (0.047) | (0.050) | (0.058) | (0.063) |
| Childhood poverty index | 0.00429 | -0.00216 | -0.0186 | -0.0182 |
| | (0.022) | (0.022) | (0.026) | (0.029) |
| Father owned a business | 0.0330 | -0.0164 | 0.0730 | 0.0643 |
| | (0.049) | (0.050) | (0.057) | (0.062) |
| **Log distance to the Tax office** | **-0.137*** | **-0.122*** | | |
| | **(0.036)** | **(0.038)** | | |
| Log distance to the city center | 0.0551 | 0.0682 | -0.0629 | -0.0543 |
| | (0.040) | (0.045) | (0.046) | (0.052) |
| Average tax inspection rate in city*industry | 1.059*** | 0.986*** | 0.245 | -0.0700 |
| | (0.36) | (0.35) | (0.40) | (0.43) |
| Tax inspection rate in 1km radius of firm | 0.254** | 0.216** | 0.399*** | 0.457*** |
| | (0.10) | (0.11) | (0.13) | (0.14) |
| Firm is 3 years or less in age | -0.0926* | -0.0413 | -0.171*** | -0.160** |
| | (0.049) | (0.050) | (0.060) | (0.065) |
| Firm has one to four paid workers | | 0.244*** | | 0.170** |
| | | (0.063) | | (0.067) |
| Firm has five to ten paid workers | | 0.521*** | | 0.235*** |
| | | (0.10) | | (0.083) |
| Firm has eleven or more paid workers | | 0.696*** | | 0.326*** |
| | | (0.14) | | (0.11) |
| Log Capital Stock (excluding land and buildings) | | 0.0595*** | | 0.0525*** |
| | | (0.017) | | (0.019) |
| **Log distance to municipal authorities** | | | **-0.00250** | **0.0635** |
| | | | **(0.045)** | **(0.053)** |
| Number of Firms | 455 | 406 | 455 | 406 |

*Notes:* Robust standard errors in parentheses.
 *, **, and *** denote significance at the 10%, 5% and 1% levels.
Probit also contains industry and city dummies.

**TABLE 4: TREATMENT EFFECT OF A TAX IDENTIFICATION NUMBER ON PROFITS**
Dependent variable: Log monthly profit

| | (1) MLE | (2) MLE | (3) 2SLS | (4) 2SLS | (5) OLS | (6) OLS | (7) OLS | (8) OLS |
|---|---|---|---|---|---|---|---|---|
| Tax identification number (NIT) | 1.121*** | 0.864*** | 1.616** | 1.447** | 0.431*** | 0.0119 | 0.533*** | 0.0835 |
| | (0.40) | (0.28) | (0.68) | (0.68) | (0.12) | (0.12) | (0.12) | (0.13) |
| Female | -0.286** | -0.100 | -0.243 | -0.102 | -0.346** | -0.0974 | -0.449*** | -0.154 |
| | (0.14) | (0.14) | (0.18) | (0.16) | (0.15) | (0.14) | (0.15) | (0.14) |
| Age of Owner | -0.0135*** | -0.0103** | -0.0151*** | -0.0122** | -0.0113** | -0.00757 | -0.0127** | -0.00787 |
| | (0.0051) | (0.0048) | (0.0058) | (0.0058) | (0.0049) | (0.0047) | (0.0050) | (0.0050) |
| Owner is Married | 0.0855 | 0.102 | 0.0490 | 0.0634 | 0.137 | 0.159 | 0.0750 | 0.111 |
| | (0.12) | (0.11) | (0.13) | (0.13) | (0.11) | (0.10) | (0.11) | (0.10) |
| Owner speaks Indigenous language | 0.0629 | 0.0807 | 0.0413 | 0.0869 | 0.0931 | 0.0716 | -0.0142 | -0.0368 |
| | (0.13) | (0.12) | (0.15) | (0.15) | (0.13) | (0.13) | (0.14) | (0.14) |
| Years of Education | -0.0154 | -0.0192 | -0.0250 | -0.0270 | -0.00201 | -0.00778 | 0.000771 | -0.0101 |
| | (0.017) | (0.015) | (0.021) | (0.019) | (0.016) | (0.014) | (0.015) | (0.014) |
| Mother had no education | 0.0930 | 0.0711 | 0.0922 | 0.0514 | 0.0940 | 0.0999 | | |
| | (0.14) | (0.13) | (0.15) | (0.15) | (0.13) | (0.13) | | |
| Mother had 9 or more years education | 0.302* | 0.207 | 0.302 | 0.199 | 0.301* | 0.218 | | |
| | (0.16) | (0.15) | (0.19) | (0.17) | (0.18) | (0.15) | | |
| Entrepreneurial Self-efficacy | 0.0932*** | 0.0916*** | 0.0779** | 0.0812** | 0.115*** | 0.107*** | | |
| | (0.031) | (0.027) | (0.039) | (0.032) | (0.029) | (0.026) | | |
| Went into self-employment to care for family | -0.174 | -0.144 | -0.151 | -0.139 | -0.205* | -0.150 | | |
| | (0.11) | (0.11) | (0.12) | (0.12) | (0.11) | (0.11) | | |
| Went into self-employment for flexible hours | -0.0486 | -0.0314 | -0.0583 | -0.0439 | -0.0352 | -0.0133 | | |
| | (0.11) | (0.11) | (0.11) | (0.11) | (0.11) | (0.10) | | |
| Went into self-employment for business growth | 0.0352 | -0.00708 | -0.00304 | -0.0312 | 0.0887 | 0.0282 | | |
| | (0.12) | (0.11) | (0.13) | (0.12) | (0.11) | (0.10) | | |
| Childhood poverty index | -0.0790 | -0.0696 | -0.0788 | -0.0689 | -0.0794* | -0.0708 | | |
| | (0.050) | (0.048) | (0.054) | (0.055) | (0.048) | (0.045) | | |
| Father owned a business | -0.200* | -0.265** | -0.198* | -0.244** | -0.203* | -0.297*** | | |
| | (0.11) | (0.11) | (0.12) | (0.12) | (0.11) | (0.10) | | |
| Log distance to the city center | 0.0222 | 0.0309 | 0.0220 | 0.0345 | 0.0226 | 0.0255 | 0.0457 | 0.0578 |
| | (0.089) | (0.086) | (0.092) | (0.095) | (0.086) | (0.089) | (0.093) | (0.094) |
| Average tax inspection rate in city*industry | -1.905** | -2.077*** | -2.385** | -2.515*** | -1.235* | -1.438** | -1.710** | -1.978*** |
| | (0.85) | (0.74) | (0.99) | (0.87) | (0.72) | (0.66) | (0.77) | (0.71) |
| Tax inspection rate in 1km radius of firm | 0.0225 | -0.138 | -0.104 | -0.270 | 0.199 | 0.0553 | 0.360 | 0.199 |
| | (0.26) | (0.24) | (0.32) | (0.29) | (0.25) | (0.24) | (0.25) | (0.23) |
| Firm is 3 years or less in age | -0.194 | -0.0704 | -0.155 | -0.0442 | -0.248** | -0.109 | -0.263** | -0.104 |
| | (0.12) | (0.11) | (0.14) | (0.13) | (0.12) | (0.11) | (0.13) | (0.12) |
| Firm has one to four paid workers | | 0.0596 | | -0.0431 | | 0.209 | | 0.160 |
| | | (0.12) | | (0.18) | | (0.13) | | (0.13) |
| Firm has five to ten paid workers | | 0.387** | | 0.197 | | 0.664*** | | 0.683*** |
| | | (0.19) | | (0.30) | | (0.17) | | (0.18) |
| Firm has eleven or more paid workers | | 0.596* | | 0.282 | | 1.054*** | | 1.104*** |
| | | (0.35) | | (0.51) | | (0.25) | | (0.26) |
| Log Capital Stock (excluding land and buildings) | | 0.128*** | | 0.102** | | 0.166*** | | 0.178*** |
| | | (0.031) | | (0.043) | | (0.033) | | (0.034) |
| Constant | 8.189*** | 6.875*** | 8.471*** | 7.435*** | 7.794*** | 6.060*** | 8.020*** | 6.122*** |
| | (0.60) | (0.66) | (0.71) | (0.87) | (0.55) | (0.62) | (0.58) | (0.65) |
| Observations | 404 | 369 | 404 | 369 | 404 | 369 | 404 | 369 |
| First-stage F-statistic: | | | 12.63 | 12.56 | | | | |
| R-squared | | | | | 0.29 | 0.42 | 0.21 | 0.35 |

*Notes:* Robust standard errors in parentheses.
*, **, and *** denote significance at the 10%, 5% and 1% levels.
Log distance to the tax office is used as the instrument in columns 1-4.

**TABLE 5: ROBUSTNESS**

Dependent variable: Log monthly profits

|  | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Tax identification number (NIT) | 0.864*** | 0.807*** | 0.870*** | 0.937*** |
|  | (0.28) | (0.28) | (0.27) | (0.24) |
| Log distance to Municipal Office |  |  | 0.0316 | 0.0133 |
|  |  |  | (0.090) | (0.088) |
| Own firm visited by tax inspector in last year |  |  |  | -0.267* |
|  |  |  |  | (0.14) |
| Owner ability and background controls | Yes | No | Yes | Yes |
|  |  |  |  |  |
| Observations | 369 | 369 | 369 | 369 |

Notes: Column (1) replicates column (2) of Table 4.

Column (2) drops mother's education, self-efficacy, reasons for going into self-employment, childhood poverty, and father's business background.

Apart from dropping these ability and background variables in column (2), all columns include the same controls as column (2) of Table 4.

 *, **, and *** denote significance at the 10%, 5% and 1% levels.

Estimates are MLE using log distance to the tax office as an instrument for NIT.

**TABLE 6: HETEROGENEITY OF TREATMENT EFFECTS**
**Propensity Score Matching Estimates of Impact of Having an NIT**

| | All firms | | By Firm Size | | |
| | without size controls | with size controls | 0-1 workers Bottom tercile K | 2-5 workers Middle tercile K | 6+ workers Top tercile K |
|---|---|---|---|---|---|
| NIT | 0.484 | 0.209 | -0.616 | 0.408 | -0.546 |
| s.e. | 0.124 | 0.130 | 0.597 | 0.274 | 0.257 |
| p-value | 0.000 | 0.110 | 0.302 | 0.136 | 0.033 |
| | | | | | |
| Proportion of firms with NIT | 0.270 | 0.266 | 0.066 | 0.246 | 0.721 |
| Mean self-efficacy of firms with NIT | 0.55 | -0.14 | 0.92 | 0.64 | 0.74 |
| Mean self-efficacy of firms without NIT | -0.15 | 0.52 | -0.94 | -0.06 | 1.72 |

**TABLE 7: HOW DOES FORMALITY AFFECT PROFITS?**
**Treatment effects regressions.**

| | Pay Taxes | Taxes as share of profits | Log sales in February | Issue tax receipts | Use of trade credit | Working capital from customers or suppliers | Bank loan in 2005 or 2006 | Corruption is an obstacle to growth |
|---|---|---|---|---|---|---|---|---|
| Having a NIT | 0.831*** | 0.482*** | 0.911* | 0.439*** | 0.122 | -0.148 | -0.004 | -0.332 |
| | (0.188) | (0.063) | (0.471) | (0.055) | (0.37) | (1.14) | (0.02) | -1.61 |
| | | | | | | | | |
| Mean of dependent variable | 0.323 | 0.141 | 8.03 | 0.116 | 0.163 | 0.195 | 0.303 | 0.571 |

*Notes:* Robust standard errors in parentheses. *, **, and *** denote significance at the 10%, 5% and 1% levels.
The same control variables as Table 4 are used, and log distance to the tax office is used as an instrument.