American Economic Journal: Macroeconomics 2 (October 2010): 195–221
http://www.aeaweb.org/articles.php?doi=10.1257/mac.2.4.195

Value-Added Taxes, Chain Effects, and Informality[†]

By Áureo de Paula and José A. Scheinkman*

We present an equilibrium model of tax avoidance and test its implications using a survey of firms in Brazil. In the model, the credit method used to collect value-added tax (VAT) creates informality chains—clients or suppliers of informal firms are more likely to be informal. An increase in enforcement in a production stage increases formality downstream and upstream. Various empirical measures of formality of suppliers and buyers, and of enforcement downstream and upstream, are positively correlated with formality. When the VAT is applied in a single stage of production at a rate estimated by the authorities, these chain effects disappear. (*JEL* H25, H26, L14, L21, O14, O17)

In this paper, we investigate determinants of informality. It is difficult to define informal activities unambiguously, but estimates indicate that in the period 1990–1993 approximately 10 percent of gross domestic product (GDP) in the United States was produced by individuals or firms that evaded taxes or engaged in illegal pursuits. These activities may produce 25–35 percent of aggregate output in Latin America, between 13 and 70 percent in Asian countries, and about 15 percent in Organization for Economic Co-operation and Development (OECD) countries (see Table 2 in Friedrich Schneider and Dominick H. Enste 2000). The underground economy is also estimated to have comprised between 6 and 63.5 percent in countries in the former Soviet Union and other Eastern and Central European nations during the first half of the 1990s (see Table 1 in Simon Johnson, Daniel Kaufmann, and Andrei Shleifer 1997).

Informality creates a fiscal problem, but there is also growing evidence that informal firms are less efficient, perhaps because of their necessarily small scale, perhaps because of their lack of access to credit or access to the infrastructure of legal

* de Paula: University of Pennsylvania, Department of Economics, 3718 Locust Walk, 160 McNeil Building, Philadelphia, PA 19104-6297 (e-mail: aureo@sas.upenn.edu). Scheinkman: Princeton University, Department of Economics, 26 Prospect Avenue, Princeton, NJ 08544-5296 and National Bureau of Economic Research (e-mail: joses@princeton.edu). Most of the analysis in this paper appeared earlier in de Paula and Scheinkman (2008). We thank Rita Almeida and Sandra Brandão for their help with the enforcement data, and Joana Monteiro and Juliano Assuncão for providing us with the ECINF dataset and clarifying its methodology. We also thank Julio Cacho, Paulo Natenzon, Justinas Pelenis, and Glen Weyl for research assistance. We benefited from comments by seminar participants at several institutions and conferences and especially from conversations with William Maloney, Nicola Persico, and Ken Wolpin. The current version particularly benefited from detailed comments by the referees. This material is based upon work supported by the National Science Foundation under Award Numbers SES0350770 and SES0718407. Portions of this work were done while Scheinkman was a Fellow of the John Simon Guggenheim Memorial Foundation.

[†] To comment on this article in the online discussion forum, or to view additional materials, visit the articles page at http://www.aeaweb.org/articles.php?doi=10.1257/mac.2.4.1.

protection provided by the state. In many less developed countries, creating incentives for formalization is viewed as an important step to increase aggregate productivity.

We present an equilibrium model of informality and test its implications using a survey of 48,000+ small firms in Brazil. In our model, informality is defined as tax avoidance. Firms in the informal sector avoid paying taxes, but suffer from other limitations.

The model can be seen as a development of James E. Rauch (1991), who relied on the modeling strategy of Robert E. Lucas Jr. (1978) in which managerial ability differs across agents in the economy, and assumed a limitation on the size of informal firms. As in that paper, agents with low managerial ability become informal managers, and those with the highest abilities become formal managers. Managers with more ability would naturally run larger firms. For this reason they choose to join the formal sector, where they do not face size limitation. The marginal firm trades off the cost of paying taxes versus the scale limitations of informal firms. As a result, the marginal firm is smaller in the informal sector than it would be if it joined the formal sector. Thus, as in Rauch (1991); Bernard Fortin, Nicolas Marceau, and Luc Savard (1997); or Era Dabla-Norris, Mark Gradstein, and Gabriela Inchauste (2008), a size gap develops. Managers that are slightly more efficient than the manager of the marginal informal firm run discretely larger firms. In this class of models, entrepreneurs that operate in the informal sector are too inefficient to benefit from the lower capital costs and scale economies afforded to formal entrepreneurs. This agrees with the results from a survey of informal Mexican firms conducted by David Mckenzie and Christopher Woodruff that is reported in Pablo Fajnzylber, William F. Maloney, and Gabriel V. Montes-Rojas (2009), where 75 percent of the respondents reported that they were too small to make it worth their while to become formal.[1]

The main innovation of our theoretical analysis is a focus on the role of the value-added tax (VAT) in transmitting informality. It exploits the idea that collecting VAT according to a credit scheme sets in motion a mechanism for the transmission of informality. VAT is a prevalent form of indirect taxation. More than 120 nations had adopted it by 2000.[2] In the credit or invoice method that is often used, VAT applies to each sale and each establishment receives a credit for the amount of tax paid in the previous stages of the production chain. This credit is then used by the taxpayer against future tax liabilities. Since purchases from informal suppliers do not generate tax credits and informal buyers cannot use tax payments from formal suppliers, there is an incentive for informal (formal) firms to deal with other informal (formal) firms.[3] Our empirical analysis shows that, in fact, various measures of formality of suppliers and purchasers (and its enforcement) are correlated with the formality of a firm. These findings survive when we use instrumental variables to control for possible simultaneity. Even more interestingly, when we look at sectors where Brazilian firms are not subject to the credit system of VAT, but, instead, the VAT is applied at some stage of production at a rate that is estimated by the tax authorities, this chain effect vanishes.

[1] ...presumably relative to cost.

[2] See Appendix 4 in Alan Schenk and Oliver Oldman (2001).

[3] To our knowledge, the only other study to investigate the informal sector in conjunction with a VAT structure is M. Shahe Emran and Joseph E. Stiglitz (2005). Their focus is on the consequences of informality for a revenue neutral tax reform involving value added and trade taxes.

Using data from the Brazilian Ministry of Labor, we also construct measures of enforcement of formality in the labor market. Our estimates show that enforcement in the clients' or suppliers' sectors has a positive and significant effect on the probability of formalization of an economic unit.

Since the mid 1990s, following the lead of the federal government, several Brazilian states introduced Sistema Ingrado de Pagamento de Impostos e Contribuições das Microempresas e Empresas de Pequeno Porte (SIMPLES) programs that simplified and lowered the VAT rates for small firms. São Paulo, the largest and richest state in Brazil, introduced its SIMPLES in 1998. Rio Grande do Sul, another large and relatively rich southern state, started its program in 2005. We use two rounds of the Brazilian survey of small firms to evaluate the impact of the introduction of these state programs. Our results point to a significantly positive impact of the program introduction, increasing formalization directly and, through chain effects, indirectly.

The model in this paper overlooks several alternative reasons for informality, such as the cost of complying with regulations, labor taxes, or the existence of a minimum wage. It also ignores benefits of formality that have been highlighted in the literature, such as access to participation in the legal system and other civic institutions. Considering these omitted costs and benefits should not change the qualitative implications of our model, which does not aim at providing an exhaustive explanation for informality. In addition, we focus on informality from the viewpoint of firms, not workers. Finally, our model ignores partial compliance: firms either pay their taxes in full or not at all. This is a simplification that is also present in our data, which only provides us with binary information on formalization.

Other papers that investigate causes and determinants of informality include Norman V. Loayza (1996); Johnson, Kaufmann, and Shleifer (1997); and Eric Friedman, Johnson, Kaufmann, and Pablo Zoido-Lobaton (2000), all of which provide evidence of an association between the size of the underground economy and higher taxes, more labor market restrictions, and poorer institutions (bureaucracy, corruption, and legal environment). Juliano Assunção and Joana Monteiro (2005) and Fajnzylber, Maloney, and Montes-Rojas (2009) are recent papers that use an earlier (1997) wave of the survey that we employ. Both papers exploit the introduction of the SIMPLES law, which simplified compliance and reduced rates for small firms in Brazil, to make inferences on the relationship between taxes and informality. Although the empirical results in our paper speak to a somewhat different set of questions (for instance, the multistage transmission of informality captured by our model), use data from a different year (2003 versus 1997), and refer to a different definition for formalization,[4] their empirical results are broadly in line with ours. In particular, both Assuncão and Monteiro (2005) and Fajnzylber, Maloney, and Montes-Rojas (2009) find that the enactment of SIMPLES increased formality.

[4] Assuncão and Monteiro (2005) and Fajnzylber, Maloney, and Montes-Rojas (2009) use municipal licensing as proxy for formalization instead of tax registration, the measure we use. Assuncão and Monteiro (2005) recognize that tax registration would be a more appropriate indication of formalization, but opt for licensing because the question on tax registration was only asked for those who indicated that their firm had been "legally constituted"—that is, a contract had been registered with the proper authorities. We do not view this as a problem since, according to Brazilian law, only legally constituted firms are eligible for tax registration.

The remainder of this paper is organized as follows. In Section I, we develop a model with two stages of production and VAT. Section II contains the empirical results obtained using data on informal firms in Brazil and Section III concludes.

I. A Model for Informality

In this section, we present a very simple model that illustrates the transmission of informality across sectors as a result of taxes on value added. In Section II, we document that this mechanism is relevant for informality in Brazil.

There are two stages of production: "upstream" and "downstream." All individuals in the model are entrepreneurs and, for simplicity, we assume that they are specialized in one of the stages. Each entrepreneur in the upstream sector is characterized by his ability $\theta_u > 0$. The density of θ_u is $g_u(\cdot)$. An entrepreneur of ability θ_u can produce θ_u units of the intermediate good in the formal sector. An informal entrepreneur, if detected by the authorities, loses all profit. The probability of being detected increases with the size of the firm. While a more general form for this probability of detection can be adopted, we assume here, for simplicity, that there exists an output level $\overline{y} > 0$, such that entrepreneurs that produce no more than $\overline{y} > 0$ are never detected, but those who produce more than $\overline{y} > 0$ are detected with probability 1. Thus, the production function in the informal sector for an entrepreneur of ability θ_u is $\min(\overline{y}, \theta_u)$.

Downstream entrepreneurs are characterized by an ability parameter θ_d with density $g_d(\cdot)$. An agent with ability θ_d, if in the formal sector, produces $\theta_d x^{\alpha}$ units of the formal good using x units of the intermediate good. In analogy to what we assumed for upstream entrepreneurs, we assume that in the informal sector, only a limited amount of input can be used and the production function becomes $\theta_d \min(\overline{x}, x)^{\alpha}$, where $\overline{x} > 0$.

Both g_u and g_d are continuous and there exists $\theta_u < \overline{y}$ for which $g_u(\theta_u) > 0$, and $g_d(\theta_d) > 0$ for $\theta_d > 0$.

The final good is tradeable and has an exogenous price which we use as numeráire. Firms in the formal sector pay an ad valorem tax rate of τ, and we write $\pi = 1 - \tau$. VAT is levied by the credit method. The tax rate applies to each sale and each establishment receives a credit for the amount of tax paid in the previous stages of production. Because of the tax credit, the prices paid for informal and formal goods may be distinct, and we let p_f be the price of the intermediate good in the formal sector and p_i in the informal sector.

We write

$$\Pi_f^u(\theta_u) = \pi p_f \theta_u \tag{1}$$

$$\Pi_i^u(\theta_u) = p_i \min\{\theta_u, \overline{y}\} \tag{2}$$

for the profit of an upstream firm with manager of quality θ_u if it produces in the formal (informal) sector. Downstream firms face a slightly more complicated problem, since they must also choose which intermediate good (formal or informal) to purchase.

Write

$$\Pi_f^d(\theta_d) = \max\{\max_x[\pi(\theta_d x^\alpha - p_f x)], \max_x[\pi\theta_d x^\alpha - p_i x]\} \tag{3}$$

for the profit of a downstream firm with a manager with ability θ_d that chooses to operate in the formal sector. In an analogous manner, write

$$\Pi_i^d(\theta_d) = \max\{\max_x[\theta_d \min(x, \bar{x})^\alpha - p_f x], \max_x[\theta_d \min(x, \bar{x})^\alpha - p_i x]\} \tag{4}$$

for the profit of a downstream firm with a manager of ability θ_d that chooses to operate in the informal sector.

The demand of an informal entrepreneur of ability θ_d facing an input price p is

$$x_i(\theta_d, p) = \min\left(\bar{x}, \left(\frac{\alpha\theta_d}{p}\right)^{1/(1-\alpha)}\right). \tag{5}$$

In turn, a formal entrepreneur demands, if he buys from the formal sector at a unit price p,

$$x_f(\theta_d, p) = \left(\frac{\alpha\theta_d}{p}\right)^{1/(1-\alpha)}, \tag{6}$$

while if he buys from the informal sector, he demands $x_f(\theta_d, p/\pi)$, since the tax credit does not apply.

We now derive aggregate demand and supply of the intermediate good in the formal and informal sectors as a function of prevailing prices. Since we are interested in equilibrium prices, we may restrict the range of prices to $0 < \pi p_f \leq p_i \leq p_f$. In fact, if $\pi p_f > p_i$, profit maximization and equations (3) and (4) imply that both formal and informal entrepreneurs downstream would buy from informal upstream firms. However, every upstream entrepreneur will prefer to produce in the formal sector. Similarly, if $p_i > p_f$, every downstream entrepreneur would prefer to buy from formal firms. However, small θ_u agents would prefer to produce informally. Furthermore, when $\pi p_f \leq p_i \leq p_f$, downstream informal (formal) entrepreneurs weakly prefer to buy from informal (formal) producers. If these inequalities are strict, preferences are also strict.

The following proposition shows the existence of cutoff points for each stage, $\bar{\theta}_u(p_i, p_f)$ and $\bar{\theta}_d(p_i, p_f)$, such that all managers with ability below the cutoff (weakly) prefer informality, and all those with ability above the cutoff points prefer to join the formal sector.

PROPOSITION 1: *(i) If* $\theta_u < \bar{\theta}_u(p_i, p_f) := p_i\bar{y}/\pi p_f \geq \bar{y}$, *then* $\Pi_i^u(\theta_u) \geq \Pi_f^u(\theta_u)$, *and if* $\theta_u > \bar{\theta}_u(p_i, p_f)$, *then* $\Pi_i^u(\theta_u) < \Pi_f^u(\theta_u)$. *(ii) There exists a* $\bar{\theta}_d(p_i, p_f)$, *such that if* $\theta_d < \bar{\theta}_d(p_i, p_f)$, *then* $\Pi_i^d(\theta_d) \geq \Pi_f^d(\theta_d)$, *and if* $\theta_d > \bar{\theta}_d(p_i, p_f)$, *then* $\Pi_i^d(\theta_d) < \Pi_f^d(\theta_d)$.

At the cutoff points, the size and input demand of firms are discontinuous with respect to the quality of the entrepreneur: the marginal informal managers in both sectors are always constrained.

PROPOSITION 2: *(i) If $p_i > \pi p_f$, the output of the smallest upstream formal firm is $p_i \bar{y}/\pi p_f > \bar{y}$. (ii) If $\Pi_f^d(\theta_d) > \Pi_i^d(\theta_d)$, then the optimal choice of the firm with manager of quality θ_d, $x_f(\theta_d, p)$, where $p = p_f$ if the firm's optimal choice is to buy the formal good and $p = p_i/\pi$ if the firm's optimal choice is to buy the informal good, satisfies*

$$x_f(\theta_d, p) \geq \frac{\bar{x}}{\pi} > \bar{x} \geq x_i(\theta_d, p),$$

for any θ. In particular, the output of the smallest downstream formal firm is strictly bigger than the output of the largest downstream informal firm.

The proofs of these propositions and the proof of existence of an equilibrium price vector (p_i, p_f) are in the Appendix.

Comparative Statics.—In the model, all inefficiencies arise from firms that choose to be informal and too small. As exposited in the previous section, these choices depend on the values of parameters such as the tax rate $(1 - \pi)$ and the level of tolerance in the downstream stage $\bar{x}$ and in the upstream stage $\bar{y}$. Although these parameters can, in principle, be optimally chosen by a government, subject to the government's enforcement technology and budget constraint, we will not consider the optimal tax and enforcement problems.[5] Nonetheless, it is useful to consider the effect of these parameters on equilibrium outcomes.

The effect of tax rate changes on the composition of the economy can be disentangled into a direct effect, holding prices constant, and an indirect effect, which operates through the adjustment of equilibrium prices. The direct effect of taxes on $\bar{\theta}_u$ and $\bar{\theta}_d$ is easily seen to be negative:

$$\frac{\partial \bar{\theta}_u}{\partial \pi} < 0 \qquad \frac{\partial \bar{\theta}_d}{\partial \pi} < 0.$$

In other words, for given prices, an increase in the tax rate (decrease in π) leads to an increase in the informal sector in both stages. The first inequality follows because $\bar{\theta}_u = p_i \bar{y}/\pi p_f$. The effect on $\bar{\theta}_d$ is established by noticing that $\bar{\theta}_d$ is defined at the intersection of $\Pi_i^d(\cdot)$ and $\Pi_f^d(\cdot)$. Since $\Pi_i^d(\cdot)$ does not depend on π directly (equation (4)), and $\Pi_f^d(\cdot)$ increases in π (equation (3)), $\bar{\theta}_d$ increases with the tax rate (decreases with π).

[5] An exception is a numerical example involving tax substitution that is discussed in Section IIF.

Since a change in taxes also affects prices, and these in turn affect the cutoffs, the total effect must account for the equilibrium adjustment of prices. The next proposition establishes that increases in taxes increase informality in both stages. In addition, it states comparative statics results for prices and cutoffs in one stage when tolerance to informality in the other stage changes.

PROPOSITION 3:

(i) $$\frac{dp_i}{d\overline{y}} < 0 \qquad \frac{d\overline{\theta}_d}{d\overline{y}} > 0$$

(ii) $$\frac{dp_i}{d\overline{x}} > 0 \qquad \frac{d\overline{\theta}_u}{d\overline{x}} > 0$$

(iii) $$\frac{d\overline{\theta}_u}{d\pi} < 0 \qquad \frac{d\overline{\theta}_d}{d\pi} < 0.$$

The derivation of these results is presented in the Web Appendix. The total effect of taxes is still in the same direction as the direct effect. Also, as expected, higher tolerance in one stage implies a larger informal sector in the other stage. The impact of an increase in tolerance on the informality at the same stage is less clear. In the case of the upstream stage, for example, the direct effect ($\partial\overline{\theta}_u/\partial\overline{y}$) is positive, but the indirect effect (change in prices times reaction of $\overline{\theta}_u$ to prices) is negative (see the Web Appendix). We nevertheless conjecture that the total effect is positive. Numerical simulations support our conjecture. An increase in tolerance in the upstream sector increases the proportion of informal firms upstream and downstream. For $\alpha = 0.7, \tau = 0.15, \overline{x} = 0.1, \theta_d, \theta_u \sim U[0,1]$, Figure 1 shows that as $\overline{y}$ increases, the proportion of upstream firms that are informal increases. As a result the price of the informal intermediate good p_i decreases and some of the downstream formal firms opt for informality. The fall in demand for the formal intermediate good causes a fall in its price p_f. These qualitative features are unchanged for other parameter choices. A symmetric picture arises when we consider changes in the tolerance for informality in the downstream stage, $\overline{x}$.

II. Empirical Application

In this section, we explore implications of our theoretical framework using a dataset of informal firms in Brazil. Tax noncompliance is an important phenomenon in Brazil. Schneider and Enste (2000) estimate that informality represents more than one-quarter of the Brazilian economy. Its VAT system was established in the 1960s, and VAT represents approximately 35 percent of total tax collection.

A. *Data*

Our principal data source is the ECINF survey (Pesquisa de Economia Informal Urbana) on informal firms realized by the Brazilian Statistics Bureau (IBGE). We used the 2003 edition of that survey, collected in October 2003, which contains

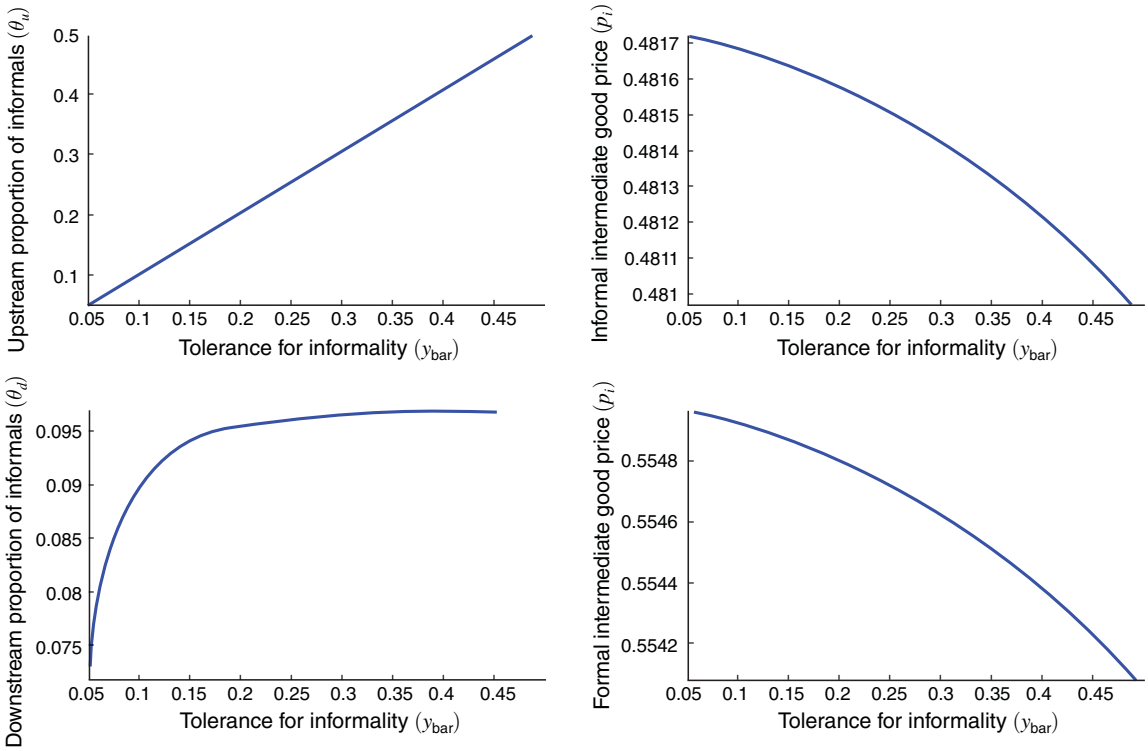

FIGURE 1. $\alpha = 0.7, \tau = 0.15, \bar{x} = 0.05, \theta_d, \theta_u \sim U[0,1]$

information on 48,701 entrepreneurs in urban regions from all states in the Brazilian federation. We also used the 1997 edition for the analysis present in Section IIE. The survey focused on units with five or less employees.[6] The sampling strategy uses the demographic census as a frame. First, preliminary interviews screened households for the presence of at least one entrepreneur with a business employing five or less people, for possible inclusion in the survey. The sampling was done in two stages: in each state (of a total of 27), the primary sampling units (census tracts) were stratified geographically in three strata (state capital, other census tracts in the capital's metropolitan area, and remaining census tracts). In a second step, the primary sampling units were stratified according to levels of income within the geographical stratum. Census tracts were randomly selected with a probability proportional to the number of households in the tract. For each selected census tract, 16 households were randomly chosen for interviews.[7] Interviewees were told that the information collected for the survey was confidential and would only be utilized for statistical purposes and, in fact, a vast majority declared that their firm was informal.

[6] The Brazilian Small Business Administration (SEBRAE) defines small businesses as those with less than 10 workers in commerce or services or less than 20 workers in all other sectors. According to SEBRAE's Boletim Estatístico de Micros e Pequenas Empresas: Primeiro Semestre de 2005 (http://www.sebrae.com.br), in 2002 small businesses accounted for 93.6 percent of the total number of firms, employed 36.2 percent of the workers, and responded for 10.3 percent of wages in 2002.

[7] The census tract is a sub-municipal geographical partition. For the 1997 edition of the survey, there were, on average, 3.1 such census tracts per municipality covered by the survey, with as many as 62 tracts in the municipality of Goiânia and as few as 1.7 tracts per municipality in the State of Maranhão. For more information on the sampling strategies employed, see Rosângela Almeida and Zélia Bianchini (1998).

Table 1—Variable Description

Variable	Description	Obs.	Mean	SD
Tax reg	1 = Tax registration	47,920	0.131	0.338
Tax sub	1 = Tax substitution	48,316	0.121	0.326
Large cl	1 = Large client	47,918	0.041	0.199
Small cl	1 = Small client	47,918	0.070	0.256
Outside hh	1 = Outside household	47,923	0.637	0.481
# employees	Number of employees	47,926	1.477	1.047
Revenue	Revenue in Oct/2003 (R$1,000)	47,188	2.091	6.299
Other job	1 = Owner has other job	48,290	0.125	0.330
Bank loan	1 = Bank loan	47,905	0.062	0.242
Education	Education level (owner)	48,255	4.367	1.884
Age	Age (owner)	48,316	41.026	12.312
Gender	Gender (owner)	48,314	0.644	0.479
Homeowner × # rooms	Homeowner × number of rooms	48,042	4.889	3.316
Sup enf	Supplier enforcement	47,846	0.012	0.010
Cl enf	Client enforcement	47,846	0.010	0.010
Log wage	Log of mean wage (R$1,000)	6,485	−1.830	0.855
Supplier formal	Formalization among suppliers	47,846	0.159	0.035
Client formal	Formalization among clients	47,846	0.127	0.039

An ideal dataset for testing our model would contain information on the production chain associated with each firm. Although the ECINF contains certain characteristics of a firm's clientele (whether they were predominantly large or small companies, persons, or governmental institutions), this information is quite limited. To complement these data, we used the input-output matrix information available from IBGE. We computed intersectoral technical input coefficients and measures of output sectoral destination using the 2003 Brazilian national accounts.[8]

B. *Description of Variables*

We eliminated firms with owners who were less than 15 years old and the observations lacking education or gender information. Entrepreneurs who claimed that their main client was a governmental institution, which comprised less than 1 percent of the original data, were also discarded. This restricted our sample to around 48,000 observations.

Table 1 summarizes the principle variables used in this paper. The first variable indicates formalization. It is a dummy variable that equals one if the firm is registered with the Brazilian tax authorities.[9] For firms in economic sectors that qualify for tax substitution (see Section IIF for an explanation), *tax sub* takes the value one.

[8] Tables 1 and 2 under "Tabelas de Recursos e Usos" available under National Accounts are available on http://www.ibge.gov.br for 2003. The construction of technical coefficients follows the European System of Integrated Economic Accounts (ESA) specifications (see Thijs ten Raa 2005).

[9] The tax registry is the Cadastro Nacional de Pessoas Juridícas, which replaced the previous system, the Cadastro Geral de Contribuintes (CGC), used in the 1997 survey. This variable is the most representative of formalization for our purposes, but we have nonetheless experimented with using "legally constituted firms" and obtained virtually identical results. This is not surprising, since, as we already mentioned, the latter is a prerequisite for tax registration, and the correlation between the two measures of informality is 0.98.

Table 2—Education

1 =	No education
2 =	Reads and writes
3 =	Some primary education
4 =	Graduated primary school
5 =	Some secondary education
6 =	Graduated secondary school
7 =	Some college education
8 =	Graduated college

The next two variables are dummies for firms that sell their products mainly to large firms (*large cl*) or small firms (*small cl*) (where large firms are those with more than five employees). Other alternatives are persons or ignored. *Outsidehouse* is a dummy that equals one when the activity is performed outside the home. The number of employees (# *employees*) includes the owner. Even though the survey focused on firms with five or less employees, a few units (less than 0.1 percent) employ more than five people due to the lag between screening and interviewing. The variables *revenue*, *other job*, and *bank loan* are self-explanatory. *Education* is a categorical variable with values depicted in Table 2. *Age* of the owner is in years, and *gender* equals one for male. The variable *homeowner* $\times$ # *rooms* is a measure of wealth and is zero for nonhomeowners, and otherwise displays the number of rooms in the house. The variables log *inv* and log *inst* measure the logarithm of investments and capital installations in October/2003 (R$1,000).[10] *Profit* equals revenue minus expenses in October/2003 (also in R$1,000). Log*wage* denotes the logarithm of the total expenditures in wages (in R$1,000) divided by the number of employees in the firm.[11] The variables *cl form* and *sup form* measure formalization among customers and suppliers of a firm (see Section IIC for the construction of these variables).

Each firm in the sample is classified into economic activities following the CNAE (Classificação Nacional de Atividades Econômicas) classification.[12] We were able to match these CNAE activities to the sectors in the Brazilian National Accounting System (NAS) using equivalence tables that are available from IBGE. Typically several CNAE activities correspond to a single NAS sector, but there are a few exceptions in which a single activity corresponds to multiple NAS sectors. The use and make tables of the NAS provide vectors of input and output coefficients. Using these NAS vectors, the value of output in each NAS sector and the correspondence tables, we constructed coefficients indicating how much input each CNAE activity uses from each of the NAS sectors and how much output of each CNAE activity is allocated to each of the NAS sectors. We also computed a measure of the proportion of firms in our sample with tax registration in each NAS sector.[13] The ECINF survey

[10] The value of installations refers to owned installations. Rented equipment is not included. Only 7 percent of formal firms and 7 percent of informal firms reported any rented equipment

[11] For comparison, annual GDP per capita in Brazil in 2003 was R$8,694.47.($\log(8.69447/12) = \log(0.72454) = -0.13$).

[12] IBGE's Web site (http://www.ibge.gov.br) provides a description of this classification.

[13] For the less 4 percent of the firms in our sample with a CNAE activity that corresponds to multiple NAS sectors, we counted the firms in each of the multiple NAS sectors.

TABLE 3—ECONOMIC SECTOR (ECINF)

	Frequency	Percent	Description
1	5,130	10.65	Transformation and mineral extraction industry
2	7,000	14.43	Construction
3	14,675	30.47	Retail and repair services
4	4,104	8.54	Lodging and food services
5	4,451	9.24	Transportation and communications
6	3,125	6.48	Real estate and services
7	2,937	6.11	Education, health, and social services
8	4,693	9.67	Other collective, social, and personal services
9	2,199	4.41	Other activities

also has its own aggregate sectoral characterization, and these ECINF sectors are listed in Table 3.

The last two variables in Table 1 are measures of formalization enforcement for suppliers and customers and were constructed as follows. The Brazilian Ministry of Labor, using its own sector classification, reports on the number of firms in each sector and state visited during 2002 by its inspectors.[14] Using data from the IBGE's "Cadastro Central de Empresas," we normalized the number of visits in each state and sector by the number of persons employed in that state and sector.[15] The Ministry of Labor classification has only 11 sectors, and each CNAE activity or NAS sector corresponds to a single Ministry of Labor sector. Hence, we may construct unambiguously output coefficients indicating how much of a given CNAE activity's production is sold to each of the Ministry of Labor sectors. Assuming that a firm's clients are in the same state,[16] we generated an index of client formalization enforcement as a weighted average of the enforcement variables (visits per persons employed), using as weights the output allocation coefficients corresponding to the firm's activity. We used activities' input demand coefficients to obtain an analogous measure of supplier formalization enforcement.

Table 4 contains probit estimates for the formalization variable *tax reg*. The signs obtained for each one of the regressors are as expected. The coefficient of the variable "working outside the home" is positive. The coefficients are also positive for variables related to the size of the firm (number of employees and revenue), credit (bank loan), or the quality of the entrepreneurial input (education, age, or having no additional job). Since women in Brazil are likely to have substantial household duties, the sign on the gender variable is probably related to entrepreneurial input. These variables may also partially control for other determinants of informality, such as opportunities in the labor market.[17] The coefficients on all these variables are statistically significant.

[14] Enforcement data was obtained from the "Relatório de Gestão da Fiscalizacão do FGTS" and was employed by Rita Almeida and Pedro Carneiro (2005).

[15] We also used the number of firms in the state-sector to normalize the number of visits, and obtained roughly similar results.

[16] Only 5.23 percent of the respondents in our sample claim to sell in other states.

[17] With the available data, it is not possible to control for business cycle variations, another potentially relevant determinant of informality that is not captured in our model.

Table 4—Probit Estimates

Dependent variable = tax reg	Coefficient (SE)	Marginal effect
Outside household	0.174***	0.021
	(0.024)	
# Employees	0.407***	0.052
	(0.012)	
Revenue	0.051***	0.006
	(0.005)	
Bank loan	0.379***	0.062
	(0.033)	
Other job	−0.242***	−0.027
	(0.033)	
Education	0.192***	0.025
	(0.006)	
Age	0.036***	0.005
	(0.004)	
Age^2	0.000***	0.000
	(0.000)	
Gender	0.148***	0.018
	(0.020)	
homeowner ×# rooms	0.030***	0.004
	(0.003)	
ECINF sector dummies	Yes	
State dummies	Yes	
N	47,201	
Pseudo-R^2	0.3634	
$\chi^2_{(44)}$	5,435.96	

Note: Standard errors clustered by census tract.
***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

In the estimates we used dummies to control for state and ECINF sector (according to the specification in Table 3).

C. *Chain Effects on Formalization*

One initial approach to investigate the existence of cross-firms effects of formalization was to employ a characterization of a firm's clientele in the ECINF survey. Interviewees were asked to declare whether sales were principally to large firms (more than five employees), small firms, or persons. Sales to firms tend to increase the probability of formalization with a bigger effect being associated with large firms as depicted in Table 5. These correlations are supportive of the existence of a chain effect in formalization.

We also used a composite measure of formalization among a firm's suppliers to examine this chain effect. For a firm in CNAE activity i, we define its suppliers's formality as

TABLE 5—PROBIT ESTIMATES (*Chain effects*)

Dependent variable = Tax reg	Coefficient (SE) [Marginal efficiency]	Coefficient (SE) [Marginal efficiency]	Coefficient (SE) [Marginal efficiency]	Coefficient (SE) [Marginal efficiency]
Large cl	0.373***			0.331***
	(0.049)			(0.049)
	[0.061]			[0.051]
Small cl	0.168***			0.107***
	(0.035)			(0.036)
	[0.024]			[0.014]
Supplierformal		2.803***		0.115
		(0.294)		(0.329)
		[0.358]		[0.014]
Clientformal			4.976***	4.745***
			(0.296)	(0.330)
			[0.618]	[0.587]
N	47,196	46,654	46,749	46,744
Pseudo-R^2	0.3664	0.3657	0.3722	0.3743
χ^2	5,491.36	5,469.05	5,597.23	5,622.48

Notes: Standard errors clustered by census tract. Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, *age*^{2}, *gender*, *homeown_numroom*, *state,* and ECINF sector dummies.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

$$(7) \qquad supplier\,formal_i = \frac{\sum_j tc_{ij} \times formality_j}{\sum_j tc_{ij}},$$

where $formality_j$ is the percentage of firms in NAS sector j that display tax registration[18], and tc_{ij} is the required amount of input from NAS sector j per monetary unit of output produced by a firm in activity i. Some caveats apply. This measure of supplier's formality only accounts for potential suppliers that are present in the survey and, in particular, ignores all suppliers that are large firms. On the other hand, the technical coefficients are obtained from a sample of formal firms. The net impact on the measure of formalization is unclear. Nevertheless, the results of our analysis favor the model; the coefficients attached to this variable are positive and statistically significant. The estimation results are in Table 5. The marginal effect of supplier formalization on the probability of being formal is 0.358.[19]

An analogous strategy measures formalization across a firm's clients. Formalization is now weighted according to the output break up by NAS sector,

$$(8) \qquad client\,formal_i = \frac{\sum_j oa_{ij} \times formality_j}{\sum_j oa_{ij}},$$

[18] NAS sectors that are not sampled in the ECINF survey were excluded in this calculation.

[19] The marginal effect of covariate x_k is defined as $E_x[\partial P(formalization\,|\,x)/\partial x_k]$ for a continuous covariate, and $E_x[P(formalization\,|\,x_{\hat{k}}, x_k = 1) - P(formalization\,|\,x_{\hat{k}}, x_k = 0)]$ for a binary covariate. Here, $x_{\hat{k}}$ denotes the vector of all covariates except the kth one.

where oa_{ij} is the fraction of output of CNAE activity *i* that goes to NAS sector *j*. The results are depicted in Table 5. The coefficient on this composite measure of client formalization is positive and statistically significant, with a marginal impact of 0.618.

We also ran regressions including *large cl*, *small cl*, and the client and supplier formality proxies. The coefficients are all positive and only the supplier-formality coefficient is no longer significant. To address the likely correlation of *large cl* with observable determinants of formality for a given firm, the regressions in Table 5 include ECINF sector and state dummies and other covariates which we viewed as the most natural confounding variables. We have nonetheless tried additional specifications. First, we experimented with a narrower sector classification (the CNAE activity designation) for controls instead of the economic sectors listed in Table 3. The results are unchanged. We also estimated the regressions on subsamples with different sectors (manufacturing and services), and the conclusions are basically unchanged. To handle potential correlation of capital intensity of production across various stages of a given production process, we included a quadratic polynomial on capital (installations) in our regressions, and the marginal effects are essentially unchanged for the main variables (*large cl* and *small cl*).

While the degree of tax compliance among a firm's suppliers and customers seems to affect formalization, an endogeneity problem may arise since suppliers and customers of a firm respond to the degree of tax compliance of that firm. This would tend to bias the estimator upward. Since the variable we use as a proxy for formalization among clients is an imperfect measure of tax compliance, one cannot rule out the possibility of attenuation bias in the opposite direction of the simultaneity bias, as is the case in some models with misclassified categorical regressors (see John Bound, Charles Brown, and Nancy Mathiowetz 2005). To address this potential endogeneity, we ran instrumental variable versions of the results displayed in Table 5, using the average education level in an entrepreneur's census tract as an instrument for the formalization of his clients. For the average education level in the census tract to be a valid instrument, it must only affect one's propensity to be formal through formalization of his or her clients, and not respond to the formalization of a single individual.[20] As an additional instrumental variable, we use the time (in minutes) it takes for the entrepreneur to arrive at a bank branch (*nearest bank*).[21] Since our dependent variable (*tax reg*) and the endogenous variables (*large cl* and *small cl*) are binary, we run our instrumental variable (IV) regressions using a binary probit dropping *small cl* from the regression of interest.[22] The results are presented in Table 6. The coefficient on *large cl* is positive and remains so in the IV version. In fact, the IV version displays an even larger coefficient, which we ascribe

[20] A similar strategy was used in Denise DiPasquale and Edward L. Glaeser (1999).

[21] We thank a referee for this suggestion.

[22] Since both our outcome of interest (*tax reg*) and endogenous variables are dichotomous, standard procedures such as two stage least squares or David Rivers and Quang H. Vuong (1988) are inadequate, and we used a bivariate probit to generate our IV estimates as suggested in James J. Heckman (1978). To achieve numerical convergence to a maximum, we had to drop revenue. We repeated the estimation using linearized TSLS and the Rivers-Vuong approach including revenue and the results are qualitatively unchanged.

TABLE 6—IV PROBIT ESTIMATES (*Chain effects*)

Dependent variable = Tax reg	Non-IV coefficient (SE)	IV coefficient (SE)	First stage (IV) dependent variable = large cl	Coefficient (SE)
Large cl	0.343***	4.220***	educurbsec	0.098***
	(0.048)	(0.371)		(0.017)
			nearest bank	-4×10^{-5}
				(4×10^{-5})
N	46,822	33,740		
Pseudo-R^2				0.14

Notes: Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, age^2, *gender*, *homeown_numroom*, *state,* and ECINF sector dummies. The second regression uses the average level of education in the census tract as an instrument. IV results obtained as bivariate probit. Standard errors clustered by census tract.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

to the attenuation effect of imperfect measurement of supplier formalization in the noninstrumented estimation.[23]

We also ran instrumented and noninstrumented probit regressions using a subsample of firms having only large and small firm clients, and using the latter as baseline. The coefficient for the large client dummy is positive in the noninstrumented version of this estimation, and it again increases when we use the instrumental variable. To sum up, instrumental variables are not easy to obtain in the present context. Nevertheless, the results are robust given the available variables in our sample.

D. *The Effect of Enforcement*

The previous results show evidence of correlation in the degree of informality across stages of production. Our model also suggests that increased tolerance towards informality in the upstream stage leads to a reduction in formalization in the downstream stage. Similarly, higher tolerance for informality among downstream firms should be accompanied by a higher degree of tax avoidance in the upstream stage. We use the measures of formalization enforcement in the labor market described in Section IIB as an indicator for monitoring within each state and Ministry of Labor sectors of a firm's suppliers and clients. Our estimates in Table 7 show that enforcement in upstream or downstream stages has a positive and significant effect on the probability of formalization.[24]

[23] A similar phenomenon is observed in David Card (1995), where IV estimates for return to schooling are about twice as large as OLS estimates.

[24] Since the Ministry of Labor uses a different definition of sectors than ECINF, we may, in addition, control for (Ministry of Labor defined) own sector enforcement. The inclusion of this variable leaves the coefficients of the interaction variables roughly unchanged, but it implies a negative association between enforcement in one's own sector and formalization. This is probably a result of simultaneity between own-stage-enforcement and own-stage-informality. If monitoring efforts are concentrated on sectors where formality is low, the coefficient on own-stage-enforcement would be negative. This simultaneity would justify caution with respect to the inclusion of own-stage-enforcement beyond the state and ECINF-sector controls already present in the estimates. While

TABLE 7—PROBIT ESTIMATES (*Enforcement*)

Dependent variable = Tax reg	Coefficient (SE)	Marginal effect	Coefficient (SE)	Marginal effect
Sup_enf	5.607*** (1.463)	0.724		
Cl_enf			11.817*** (1.294)	1.51
N	46,749		46,749	
Pseudo-R^2	0.3628		0.3649	
$\chi^2_{(45)}$	5,410.44		5,482.02	

Notes: Standard errors clustered by census tract. Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, *age*^{2}, *gender*, *homeown_numroom*, *state,* and ECINF sector dummies.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

As a further check on the robustness of the estimated effects of downstream or upstream enforcement on formalization, we have also obtained results for the regressions presented on Table 7 only for those individuals who have become owners within the last year (between October/2002 and October/2003). The results remain highly significant. The marginal effect of *sup enf* in the first regression is 0.589 (standard error = 0.283, *p*-value = 0.038), and the marginal effect of *cl enf* is still positive at 1.12 (standard error = 0.301, *p*-value = 0.000).

E. *The SIMPLES Tax Program*

In 1996, the Brazilian federal government established the SIMPLES tax program, targeted at small firms—those with roughly less than R$1,000,000 in annual revenues. It consolidated taxes and social security contributions in a single payment, and simplified verification and remittance procedures for tax collection. Although states and municipalities were allowed to join the system for the collection of VAT, very few did. More than 20 states eventually established their own state-level versions of the SIMPLES system for the collection of VAT and other state taxes. In 1998, for example, the state of São Paulo established a local version of the SIMPLES program. The system exempted firms with less than R$120,000 annual revenues from the collection of state VAT, and offered reduced rates to firms with at most R$1.2 million in annual revenues. The program provided firms with a significant reduction in VAT. For example, a firm with monthly sales of R$60,000 with input costs of R$20,000 would owe R$7,200 in VAT before the SIMPLES. Under the new program, the VAT would total less than R$1,300.

enforcement policy for a given stage may be responding to informality in that stage, it is nonetheless *implausible* that enforcement in one stage would respond to informality in preceding or subsequent stages in the production chain.

TABLE 8—STATE SIMPLES PROGRAMS (*Before 2003*)

State	State law #	Date	State	State law #	Date
Mato Grosso do Sul[1]	1866	7/8/1998	Goias[1]	13270	6/4/1998
Sao Paulo[2]	10086	11/19/1998	Paraiba	7332	4/28/2003
Rio de Janeiro[2]	3342	12/29/1999	Espirito Santo[1]	5389	4/24/1997
Bahia[2]	7357	11/4/1998	Alagoas[2]	6271	10/3/2001
Parana[1]	246	1/29/2003	Rio Gde do Norte[3]	8296	1/28/2003
Pernambuco[2]	11157	12/29/1997	Distrito Federal[4]	2510	12/29/1999
Ceara[2]	13298	4/2/2003	Sergipe[1]	4185	12/22/1999
Para[3]	6616	1/7/2003	Rondonia	8945	12/30/1999
Maranhao[1]	6904	3/24/1997	Acre[1]	1340	7/19/2000
Santa Catarina[2]	11398	5/8/2000	Amapa	1933	6/17/1998

[1] http://www.telecentros.desenvolvimento.gov.br
[2] Cartilha da Lei Geral (SEBRAE)
[3] Secretaria de Fazenda Estadual

We used the first round of the ECINF survey, collected in 1997, and its 2003 edition to measure the effect of this reduction in VAT on formalization. For comparison, we obtained data for states that established their state SIMPLES before and after 2003. Table 8 displays the state laws and their publication dates for those states that established SIMPLES programs before 2003. Two states, Rio Grande do Sul (RS) and Minas Gerais (MG), implemented SIMPLES after 2003, and others (Amazonas (AM), Piauí (PI), Mato Grosso (MT), Tocantins (TO), and Roraima (RR)) have not established a program.

Table 9 displays results from a probit model where dummy variables for the state, and pre- and post-introduction of the state SIMPLES, are used to assess the variation in the formalization. We applied the same controls we used in our previous formalization regressions.[25] The results point to a positive impact of the program's introduction with a marginal effect of 1.4 percentage points on formalization, an increase in the probability of formalization of approximately one-tenth.

We could not obtain data on eligibility in each states' SIMPLES program, because each state has a different tax code with a myriad of case-specific rules and contingencies. Instead, we took eligibility to participate in the federal SIMPLES as a proxy for the eligibility to participate in the state SIMPLES programs in our sample.[26] Table 10 reports the impact of eligibility for the state SIMPLES. For the eligible firms in states where SIMPLES prevailed by 2003, the marginal effect of selling mostly to large or small clients on the likelihood of having a tax registration, a measure of the chain effect on formalization, is positive for both and significant for the former. The marginal effect associated with having a large client increased by 5.6 percent for firms in states that instituted the SIMPLES between 1997 and 2003. The marginal effect associated with having a small client increased by 1.4 percent. The estimated marginal effect for firms that were not eligible to participate in state SIMPLES is not significantly different from zero. The null hypothesis that H_0 : SIMPLES $\times$ 2003 $\times$ Large Client + Ineligible $\times$ SIMPLES $\times$ 2003 $\times$ Large Client $= 0$ has a

[25] Standard errors are not clustered by census tract since their definition varied between 1997 and 2003.
[26] See the Appendix in Assuncão and Monteiro (2005) for a list of eligible activities for the federal SIMPLES.

TABLE 9—PROBIT ESTIMATES (*SIMPLES*)

Dependent variable = tax reg	Coefficient (SE)	Marginal effect
SIMPLES × 2003	0.101*** (0.030)	0.014
SIMPLES	−0.241*** (0.021)	−0.036
2003	−0.754*** (0.027)	−0.11
N	9,0224	
Pseudo-R^2	0.3215	
$\chi^2_{(21)}$	2,1951.11	

Notes: SIMPLES stands for states that implemented a version of SIMPLES for the VAT. Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, *age*^{2}, *gender*, *homeown_numroom*, and ECINF sector dummies.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

p-value of 0.67. The *p*-value for the null hypothesis that SIMPLES × 2003 × Smal l Client + Ineligible × SIMPLES × 2003 × Small Client = 0 is 0.4577.

F. *Robustness: Tax Substitution*

Brazilian tax law imposes *tax substitution* ("substituição tributária") in certain activities.[27] Under this tax collection system, VAT is charged at some stage in the production chain at a rate estimated by the tax authorities. This method tends to be adopted for activities with a reduced set of initial producers and many smaller units at the subsequent stages of production, but, in principle, the tax may be applied at any single stage of production. In Brazil, VAT on tires are paid by the tire manufacturer, based on an *estimated* sale price to final consumers. No VAT is due by any downstream seller of tires. The tire may be sold to a distributor which then sells it to an auto parts store that, in turn, sells it to a garage, which finally sells it to a consumer. None of these transactions would generate any VAT credit or obligation. Hence, the chain effect of client formalization should be much less pronounced for firms in which a substantial part of activities are subject to tax substitution.

We ran probit estimates for firms in activities where tax substitution is imposed. These activities (and their CNAE numerical designation) are automobile and auto parts manufacturing (34001, 34002, 35010, 35020, 35030, 35090), production of tires (25010), production and distribution of liquor (15050 and 53030), cigarettes (16000), commercialization of automobiles and tires (50010, 50020, 50030, and 54040), distribution of fuel (50050 and 53065), and oil refining (23010 and 23020).

[27] Tax substitution is not peculiar to Brazil. See Antônio de Pádua Ribeiro (1997).

TABLE 10—PROBIT ESTIMATES (*SIMPLES, ineligible*)

Dependent variable = tax reg	Coefficient	(SE)	Marginal effect
Ineligible × SIMPLES × 2003 × Large client	−0.253	(0.256)	−0.028
SIMPLES × 2003 × Large client	0.332 *	(0.178)	0.056
Ineligible × SIMPLES × 2003 × Small client	−0.198	(0.195)	−0.023
SIMPLES × 2003 × Small client	0.089	(0.129)	0.013
Ineligible × 2003 × Large client	−0.069 **	(0.221)	−0.008
2003 × Large client	0.023	(0.153)	0.003
Ineligible × 2003 × Small client	0.051	(0.171)	0.007
2003 × Small Client	0.005	(0.110)	0.001
Ineligible × SIMPLES × Large client	0.237	(0.202)	0.037
SIMPLES × Large client	−0.214	(0.143)	−0.025
Ineligible × SIMPLES × Small client	0.327 **	(0.081)	0.055
SIMPLES × Small client	−0.222 **	(0.094)	−0.026
Ineligible × Large client	0.378 **	(0.171)	0.065
Large client	0.254 **	(0.122)	0.040
Ineligible × Small client	0.050	(0.127)	0.007
Small client	0.244 ***	(0.080)	0.038
Ineligible × SIMPLES	0.046	(0.029)	0.006
SIMPLES	−0.193 **	(0.018)	−0.028
N		22,312	
Pseudo-R^2		0.3268	
$\chi^2_{(28)}$		22,312	

Notes: SIMPLES stands for states that implemented a version of SIMPLES for the VAT. Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, *age*^{2}, *gender*, *homeown_numroom*, and ECINF sector dummies.

*** Significant at the 1 percent level.

** Significant at the 5 percent level.

* Significant at the 10 percent level.

The results concerning investment and installations, number of employees, and the entrepreneur's education level remain qualitatively as before. In Table 11 we interact tax substitution with our measure of formality of the clients. To facilitate comparisons with the results in Table 6, we, again, consolidate the dummy variables indicating large firm and small firm clients as a single variable. The coefficient of the interaction term is negative and significant. This is implied by the regression run on the subsample of firms involved in the production of goods and services eligible for tax substitution (Table 11). The coefficient on client formalization ceases to be positive. If anything, there is evidence for a *negative* rather than positive coefficient.

Tax substitution raises a natural question. When is it advantageous to replace the VAT by a tax on a single stage? We now discuss a numerical example that illustrates the economic forces that determine the optimal choice. Fix the level $\bar{x}$ of tolerance in the downstream stage and vary the level of tolerance $\bar{y}$ in the upstream stage. We will consider two scenarios. In the first scenario, a sales tax is applied only in the downstream sector. Producers of the intermediary good are tax exempt. In this case, all upstream firms choose to be formal and a single price prevails in the market for the intermediate good. The upstream tolerance level $\bar{y}$ has no effect on equilibrium quantities and, in particular, on the total tax collection T, or on the output net of taxes of the final good. Now, consider a VAT that affects both sectors. To insure that the government's budget constraint is satisfied, for each level of $\bar{y}$ find a VAT rate $\tau(\bar{y})$ such that the total tax collection equals T. It is intuitive that as $\bar{y}$ becomes larger,

TABLE 11—PROBIT ESTIMATES (*Tax substitution*)

Variable	Full sample coefficient (SE)	Tax sub = 1 coefficient (SE)
Large cl	0.428 ***	0.059
	(0.049)	(0.208)
Small cl	0.241 ***	−0.384 ***
	(0.036)	(0.128)
Tax sub_large cl	−0.406 *	
	(0.213)	
Tax sub_small cl	−0.577 ***	
	(0.128)	
Tax sub	0.348 ***	
	(0.030)	
N	46,822	5,732
Pseudo-R^2	0.3697	0.3285
$\chi^2_{(47)}$	5,684.12	959.64

Notes: Standard errors clustered by census tract. Controls include *outside hh*, *n_employee*, *revenue*, *bank loan*, *other job*, *education*, *age*, *age*^{2}, *gender*, *homeown_numroom*, *state*, and ECINF sector dummies.

***Significant at the 1 percent level.
**Significant at the 5 percent level.
*Significant at the 10 percent level.

making informality easier in the upstream sector, the tax rate must be increased to maintain total tax collection. This intuition is verified in this example as illustrated in Figure 2. The tax rate is relatively low when $\overline{y}$ is small, but increases as $\overline{y}$ grows. Furthermore, as shown in Figure 3, for small values of $\overline{y}$, the net output of the final good is larger than it would be if we only taxed the final good. At the other extreme, when $\overline{y}$ is very large, the output of the final good is smaller under a VAT system. Again, this result is quite intuitive. If it is very difficult to impose formality in one stage relative to imposing on the other stage, it may be preferable to only tax the latter. When the stage where enforcement is easiest is downstream, the optimal choice is a sales tax, otherwise forward tax substitution may be desirable.

III. Conclusion

An implication of our model is that informal firms are smaller, less productive, and use less inputs. In addition, informality is transmitted through vertical relationships when VAT is levied through the credit method. Using microdata from surveys conducted in Brazil, we confirmed implications of the model.

In the model, informal firms are less efficient than formal ones, but our analysis is silent concerning the optimal amount of informality. To discuss this question one must also model the cost of enforcement, presumably as a function of a firm's size and the value of tax revenues.

Tax authorities in Brazil impose tax substitution hoping to increase compliance. Our model predicts a decrease in the interaction effect when tax substitution is imposed, but does not make any prediction concerning the effect of tax substitution

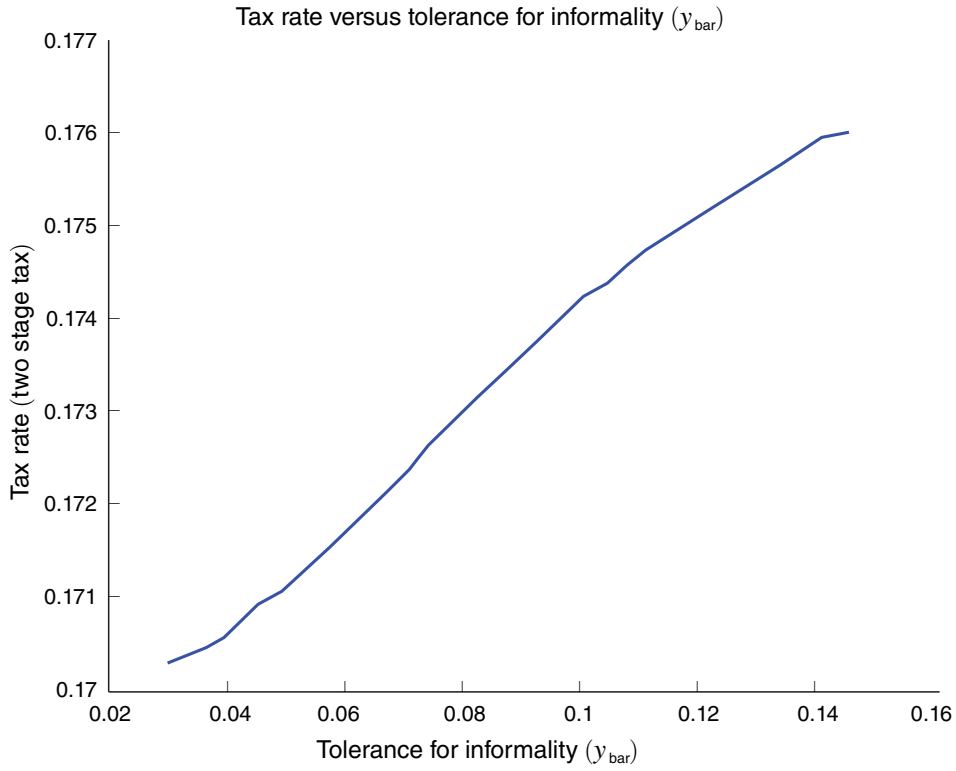

FIGURE 2. $\alpha = 0.4, \tau = 0.17$ (sales tax), $\bar{x} = 0.05, \theta_d, \theta_u \sim U[0,1]$

on the level of informality. The firms in our sample that produce goods or services subject to tax substitution more frequently have individuals as their main clients, and are owned by less-educated entrepreneurs; factors associated with less formality. Nonetheless, the difference in the rate of formalization between these firms and other firms is 7.8 percentage points (with a standard error of 0.4), a very large difference when compared with the average level of 13.2 percent in our sample. This increased formalization reflects the criterion used by Brazilian tax authorities. Tax substitution is imposed when, at some level in the chain, the typical producer is a large firm which is then charged the estimated VAT. If these large firms cannot afford to become informal and pay the estimated VAT, smaller firms in the same production chain will face lower costs to formality than the typical small firm in Brazil and, for this reason, may opt to become formal more often.

APPENDIX

PROOF OF PROPOSITION 1:

(i) is immediate from equations (1) and (2). To show that (ii) holds note that θ_d enters the definition of the profit function of formal firms exactly as an output price and hence, from the properties of profit functions with respect to output prices, we know that its derivative with respect to θ_d is proportional to $x_f(\theta_d, p)$ which goes to infinity as $\theta_d \to \infty$. Furthermore, the function $\Pi_i^d(\theta_d)$ is convex and, since supply functions of firms must slope up, if, conditional on informality, the choice of a firm

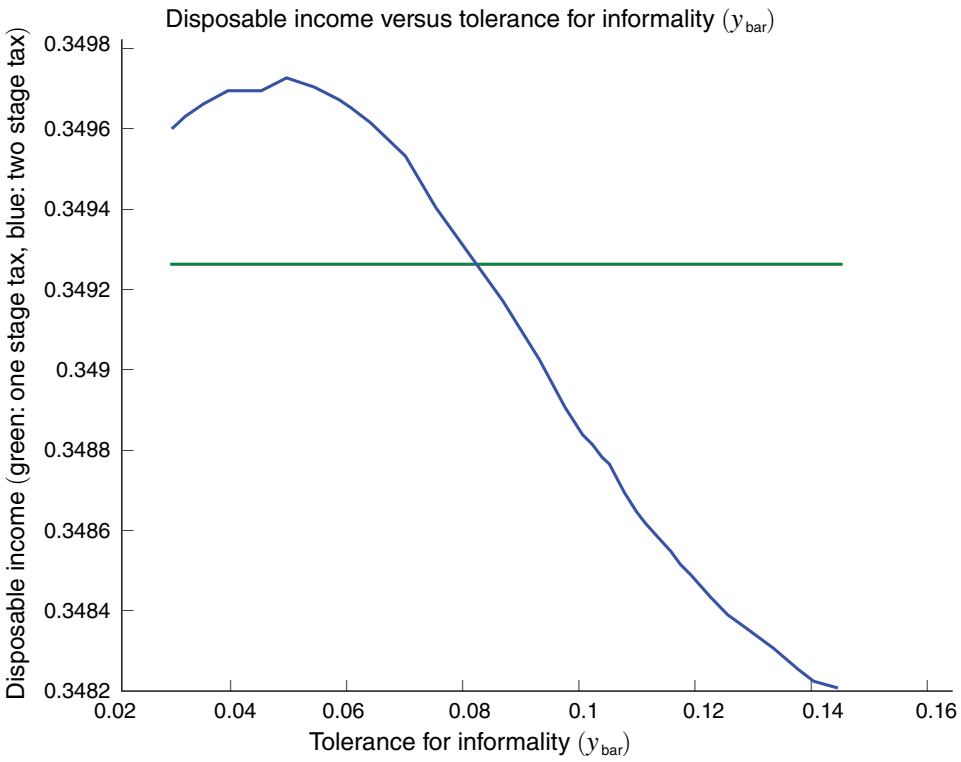

FIGURE 3. $\alpha = 0.4, \tau = 0.17$ (sales tax), $\bar{x} = 0.05, \theta_d, \theta_u \sim U[0,1]$

of ability θ satisfies $x_i(\theta) = \bar{x}$, then the optimal choice conditional on informality, $x_i(\theta_d) = \bar{x}$ for $\theta_d \geq \theta$, and as a consequence, $\Pi_i^d(\theta_d)$ is linear for $\theta_d \geq \theta$. In addition, whenever $x_i(\theta_d) < \bar{x}$, the informal firm's constraint is not binding. In this case, since $p_f \geq p_i$,

$$\Pi_i^d \;=\; \varphi(p_i) \;>\; \varphi(p_f),$$

where $\varphi(p) \;=\; [\alpha^{\alpha/(1-\alpha)} - \alpha^{1/(1-\alpha)}](q\theta_d/p^{\alpha})^{1/(1-\alpha)}$. Since

$$\Pi_f^d \;=\; \max\{\pi\varphi(p_f), \pi^{1/(1-\alpha)}\varphi(p_i)\},$$

then $\Pi_i^d(\theta_d) > \Pi_f^d(\theta_d)$, provided $\theta_d > 0$.

PROOF OF PROPOSITION 2:
(i) is obvious. For (ii), suppose first that it is optimal for the firm with manager of quality θ_d to buy the formal good. If $\pi x_f(\theta_d, p_f) < \bar{x}$, since

$$\theta_d(\pi x_f(\theta_d,p_f))^{\alpha} \;-\; \pi p_f x_f(\theta_d,p_f) \;\geq\; \pi(\theta_d x_f^{\alpha}(\theta_d,p_f) \;-\; p_f x_f(\theta_d,p_f)),$$

the firm would prefer to be in the informal sector and buy $\pi x_f(\theta_d, p_f)$ of formal inputs. If the firm bought the informal good and $\pi x_f(\theta_d, p_i/\pi) < \bar{x}$, since

$$q\theta_d(\pi x_f(\theta_d,p_i/\pi))^{\alpha} \;-\; \pi p_i x_f(\theta_d,p_i/\pi) \;\geq\; \pi q\,\theta_d x_f^{\alpha}(\theta_d,p_i/\pi) \;-\; \pi p_i x_f(\theta_d,p_i/\pi),$$

the firm would prefer to be in the informal sector and buy $\pi x_f(\theta_d, p_i/\pi)$ of informal inputs. Consequently, $\pi x_f(\theta_d, p_f) \geq \overline{x}$ and $\overline{x} \geq x_i(\theta_d, p)$ for any θ_d. Furthermore, the entrepreneur $\overline{\theta}_d(p_i, p_f)$ must be indifferent between being formal or informal. Since informal (formal) entrepreneurs weakly prefer to buy from informal (formal) suppliers, we must have:

$$\overline{\theta}_d(p_i,p_f)\overline{x}^{\alpha} - p_i\overline{x} = \pi[\overline{\theta}_d(p_i,p_f)x_f^{\alpha}(\overline{\theta}_d(p_i,p_f)) - p_f x_f(\overline{\theta}_d(p_i,p_f))]. \tag{9}$$

In particular, choosing $\theta_d = \overline{\theta}_d(p_i, p_f)$ and $p = p_f$ shows the discontinuity at the marginal firms.

Existence of Equilibrium.—Because of the possibility of indifference, we have supply and demand correspondences instead of functions. We will write $\mathcal{S}(p_i, p_f)$ for the set of possible aggregate supply vectors $(s_i(p_i, p_f), s_f(p_i, p_f))$ obtained from the choices of profit maximizing entrepreneurs in the upstream stage. If $p_i \neq \pi p_f$, the set $\mathcal{S}(p_i, p_f)$ contains a single vector (s_i, s_f) given by

$$s_i = \int_0^{\frac{p_i\overline{y}}{\pi p_f}} \min\{\theta, \overline{y}\} g_u(\theta)\, d\theta \tag{10}$$

$$s_f = \int_{\frac{p_i\overline{y}}{\pi p_f}}^{\infty} \theta g_u(\theta)\, d\theta. \tag{11}$$

If $\pi p_f = p_i = 0$, then $\mathcal{S}(p_i, \underline{p}_f) = \{0\}$. Finally, when $\pi p_f = p_i \neq 0$, a point $(s_i, s_f) \in \mathcal{S}(p_i, p_f)$, if there exists a $\theta_u \leq \overline{y}$, such that[28]

$$s_i = \int_0^{\theta_u} \theta g_u(\theta)\, d\theta. \tag{12}$$

$$s_f = \int_{\overline{\theta}_u}^{\infty} \theta g_u(\theta)\, d\theta. \tag{13}$$

We write $\mathcal{X}(p_i, p_f)$ for the set of possible aggregate demand vectors $(x_i(p_i, p_f), x_f(p_i, p_f))$ obtained from the choices of profit-maximizing entrepreneurs in the downstream stage.

When $\pi p_f = p_i$, formal firms are indifferent between buying the formal or informal input, but informal firms prefer buying from informal firms. Hence, we can allocate all formal firms with managers below a certain threshold to buying in the

[28] In principle, we could assign any subset of entrepreneurs with productivity below $\overline{y}$ to the informal sector, but there is always an interval containing the origin that would produce exactly the same aggregate output.

informal sector with the complement interval assigned to purchase in the formal sector.[29] In this case, a point $(x_i, x_f) \in \mathcal{X}(p_i, p_f)$, if there exists a $\gamma \geq \overline{\theta}_d(p_i, p_f)$, such that

$$x_i = \int_0^{\overline{\theta}_d(p_i,p_f)} x_i(\theta, p_i) g_d(\theta)\, d\theta + \int_{\overline{\theta}_d(p_i,p_f)}^{\gamma} x_f\left(\theta, \frac{p_i}{\pi}\right) g_d(\theta)\, d\theta \tag{14}$$

$$x_f = \int_{\gamma}^{\infty} x_f(\theta, p_f) g_d(\theta)\, d\theta. \tag{15}$$

If $\pi p_f < p_i < p_f$, formal (informal) firms prefer to buy from formal (informal) firms. In this case, a point $(x_i, x_f) \in \mathcal{X}(p_i, p_f)$ if

$$x_i = \int_0^{\overline{\theta}_d(p_i,p_f)} x_i(\theta, p_i) g_d(\theta)\, d\theta \tag{16}$$

$$x_f = \int_{\overline{\theta}_d(p_i,p_f)}^{\infty} x_f(\theta, p_f) g_d(\theta)\, d\theta. \tag{17}$$

If $p_f = p_i$, informal firms are indifferent, but formal firms prefer buying from formal firms. Hence, we may assign informal firms arbitrarily to buying formal or informal inputs. In this case, a point $(x_i, x_f) \in \mathcal{X}(p_i, p_f)$, if there exists $\gamma \leq \overline{\theta}_d(p_i, p_f)$, such that

$$x_i = \int_0^{\gamma} x_i(\theta_d, p_i) g_d(\theta)\, d\theta \tag{18}$$

$$x_f = \int_0^{\overline{\theta}_d(p_i,p_f)} x_i(\theta_d, p_i) g_d(\theta)\, d\theta + \int_{\overline{\theta}_d(p_i,p_f)}^{\infty} x_f(\theta_d, p_i) g_d(\theta)\, d\theta. \tag{19}$$

An equilibrium is a vector (p_i, p_f), such that $\exists z \in \mathcal{X}(p_i, p_f) \cap S(p_i, p_f)$. We will present the proof of existence of an equilibrium price in two steps. First we will set $p_i = \mu p_f$ with $\pi \leq \mu \leq 1$. For each μ we will show that there exists a unique $p_i(\mu)$, such that if $(p_i, p_f) = (p_i(\mu), (p_i(\mu)/\mu))$, then the sum of aggregate supply of the formal and informal intermediate goods equals the sum of aggregate demands. We then show that there exits a unique μ^*, such that $(p_i(\mu^*), (p_i(\mu^*)/\mu^*))$ is an equilibrium. We will use the following preliminary result which is proven in the Web Appendix.

LEMMA 4: *If $\pi p_f < p_i < p_f$, then $\overline{\theta}_d(p_i, p_f)$ decreases with p_i, and it increases with p_f. Further, if $\pi \leq \mu \leq 1$, then, $\overline{\theta}_d(p_i, (p_i/\mu))$ increases with p_i.*

[29] As before, these assignments can reproduce the demands realized by any arbitrary assignment of firms to each sector.

We now return to the equilibrium analysis. For $\mu = \pi\,(p_i = \pi p_f)$, the sum of the aggregate supply always equals

$$\int_0^{\infty} \theta g_u(\theta)\, d\theta. \tag{20}$$

On the other hand, the sum of aggregate demands always equals

$$\int_0^{\overline{\theta}_d\left(p_i, \frac{p_i}{\pi}\right)} x_i(\theta, p_i) g_d(\theta)\, d\theta \;+\; \int_{\overline{\theta}_d\left(p_i, \frac{p_i}{\pi}\right)}^{\infty} x_f\left(\theta, \frac{p_i}{\pi}\right) g_d(\theta)\, d\theta. \tag{21}$$

It is easy to check that this last expression goes to zero as $p_i \to \infty$ and to ∞ as $p_i \to 0$. Furthermore, since demand of any type decreases with the price of the input, and, from Proposition 2 $x_f(\overline{\theta}_d, p_i/\pi)\overline{x}$, using the Lemma above, it is immediate that aggregate demand is monotonically decreasing with p_i. Hence, there exists a unique $p_i(\pi)$ for which the sum of supplies equal the sum of demands.

For $\pi < \mu \leq 1$, using expressions (10) and (11), we obtain that the sum of the aggregate supplies is

$$\int_0^{\frac{\mu\overline{y}}{\pi}} \max\{\theta, \overline{y}\} g_u(\theta)\, d\theta \;+\; \int_{\frac{\mu\overline{y}}{\pi}}^{\infty} \theta g_u(\theta)\, d\theta. \tag{22}$$

On the other hand, using equations (16) and (17), the sum of the aggregate demands equals

$$\int_0^{\overline{\theta}_d\left(p_i, \frac{p_i}{\mu}\right)} x_i(\theta_d, p_i) g_d(\theta)\, d\theta \;+\; \int_{\overline{\theta}_d\left(p_i, \frac{p_i}{\mu}\right)}^{\infty} x_f\left(\theta_d, \frac{p_i}{\mu}\right) g_d(\theta)\, d\theta. \tag{23}$$

Just as before, the result in the Lemma insures the monotonicity properties that yield the existence of a unique $p_i(\mu)$ that equates the sum of aggregate demands with that of aggregate supplies.

An increase in μ always decreases aggregate supply since it causes some firms in the upstream sector to switch from formal to informal. In addition, an increase in μ increases the demand by formal firms at each p_i and causes some firms to switch from informal to formal in the downstream sector. Thus, at each p_i, aggregate demand goes up. Hence, $p_i(\mu)$ increases with μ.

The supply of the informal sector, when $p_i = \pi p_f$, is some amount in the interval $[0, \int_0^{\overline{y}} \theta g_u(\theta)\, d\theta]$. The demand is some number in the interval $[\int_0^{\overline{\theta}_d(p_i, p_i/\pi)} x_i(\theta_d, p_i) g_d(\theta)\, d\theta, \int_0^{\overline{\theta}_d(p_i, p_i/\pi)} x_i(\theta_d, p_i) g_d(\theta)\, d\theta + \int_{\overline{\theta}_d(p_i, p_i/\pi)}^{\infty} x_f(\theta_d, p_i/\pi) g_d(\theta)\, d\theta]$.

If these intervals overlap at $p_i = p_i(\pi)/\pi$, then $(p_i(\pi), p_i(\pi)/\pi)$ is an equilibrium. This will happen whenever the tolerance for informality in the upstream sector $(\bar{y})$ is high enough.

If these intervals do not overlap the informal supply of the intermediate good must necessarily go up with μ. On the other hand, the informal demand at $p_i(\mu), (p_i(\mu)/\mu)$ will go down since $p_i(\mu)$ goes up and the relative price of the formal good goes down. At $\mu = 1$, the supply of the informal good is $\int_0^{\bar{y}/\pi} \max\{\theta, \bar{y}\} g_u(\theta)\, d\theta$ whereas the demand is any number in the interval $[0, \int_0^{\bar{\theta}_d(p_i, p_f)} x_i(\theta_d, p_f) g_d(\theta)\, d\theta]$. Hence, there always exists a unique μ^* such that $(p_i(\mu^*), p_i(\mu^*)/\mu^*)$ is an equilibrium.

REFERENCES

Almeida, Rita, and Pedro Carneiro. 2005. "Enforcement of Regulation, Informal Labor and Firm Performance." Institute for the Study of Labor (IZA) Discussion Paper 1759.

Almeida, Rosângela, and Zélia Bianchini. 1998. "Aspectos de Amostragem da Pesquisa de Economia Informal Urbana." Instituto Brasileiro de Geografia e Estatística (IBGE) Texto para Discusso 89.

Assunção, Juliano, and Joana Monteiro. 2005. "Outgoing the Shadows: Estimating the Impact of Bureaucracy and Tax Reduction on Formality." Unpublished.

Bound, John, Charles Brown, and Nancy Mathiowetz. 2005. "Measurement Error in Survey Data." In **Handbook of Econometrics**. Vol. 5, ed. James J. Heckman and Edward Leamer, 3705–3843. Amsterdam: North-Holland.

Card, David. 1995. "Using Geographic Variation in College Proximity to Estimate the Return to Schooling." In *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*, ed. Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky, 201–22. Toronto: University of Toronto Press.

Dabla-Norris, Era, Mark Gradstein, and Gabriela Inchauste. 2008. "What Causes Firms to Hide Output? The Determinants of Informality." *Journal of Development Economics*, 85(1–2): 1–27.

de Paula, Áureo, and José A. Scheinkman. 2007. "The Informal Sector." National Bureau of Economic Research Working Paper 13486.

DiPasquale, Denise, and Edward L. Glaeser. 1999. "Incentives and Social Capital: Are Homeowners Better Citizens?" *Journal of Urban Economics*, 45(2): 354–84.

Emran, M. Shahe, and Joseph E. Stiglitz. 2005. "On Selective Indirect Tax Reform in Developing Countries." *Journal of Public Economics*, 89(4): 599–623.

Fajnzylber, Pablo, William F. Maloney, and Gabriel V. Montes-Rojas. 2009. "Does Formality Improve Micro-Firm Performance? Quasi-Experimental Evidence from the Brazilian SIMPLES Program." Institute for the Study of Labor (IZA) Working Paper 4531.

Fortin, Bernard, Nicolas Marceau, and Luc Savard. 1997. "Taxation, Wage Controls and the Informal Sector." *Journal of Public Economics*, 66(2): 293–312.

Friedman, Eric, Simon Johnson, Daniel Kaufmann, and Pablo Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics*, 76(3): 459–93.

Heckman, James J. 1978. "Dummy Endogenous Variables in a Simultaneous Equation System." *Econometrica*, 46(4): 931–59.

Instituto Brasileiro de Geografia e Estatística (IBGE). 2003. *Sistema de Contas Nacionais Brasil*. Séries Relatórios Metodológicos número 24.

Johnson, Simon, Daniel Kaufmann, and Andrei Shleifer. 1997. "The Unofficial Economy in Transition." In *Brookings Papers on Economic Activity 1997*, Vol. 2, ed. William C. Brainard and George L. Perry, 159–221. Washington, DC: Brookings Institution Press.

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

Lucas, Robert E., Jr. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9(2): 508–23.

Rauch, James E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics*, 35(1): 33–47.

Ribeiro, Antônio de Pádua. 1997. "Substituição tributária para frente." *Revista CEJ, Composição do Conselho da Justiça Federal*, 1(3): 105–10.

Rivers, Douglas, and Quang H. Vuong. 1988. "Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models." *Journal of Econometrics*, 39(3): 347–66.

Schenk, Alan, and Oliver Oldman. 2001. *Value Added Taxes: A Comparative Approach in Theory and Practice.* Ardsley, NY: Transnational Publishers.

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

Tait, Alan A., ed. 1991. "Value-Added Tax: Administrative and Policy Issues." International Monetary Fund Occasional Paper 88.

ten Raa, Thijs. 2005. The Economics of Input-Output Analysis. Cambridge, UK: Cambridge University Press.