DAVID BARDEY
Universidad de los Andes and Toulouse School of Economics

DANIEL MEJÍA
Universidad de los Andes

Informality and Optimal Public Policy

ABSTRACT This article explores optimal public policy, in terms of the level of tax enforcement and the supply of public goods, in an economy characterized by a huge informal sector. We define informality as the set of productive activities that do not comply (totally or partially) with government regulations. The government intervenes as a Stackelberg leader and has to decide how to allocate public expenditures, using funds collected through the tax system, between the provision of a public good, which can only be used for formal activities, and enforcement effort, aimed at detecting informal firms that evade taxes. Taking the public policy as given, a representative household, owner of a representative firm, decides how to divide a fixed supply of labor between formal and informal activities. Our results show that the greater the distortions in the tax collection process, the larger is the informal sector. Finally, we derive the properties of the optimal public policy. In particular, we show that the shadow cost of public funds represents the rationale of enforcement spending.

JEL Codes: K10, K20, K42, O17

Keywords: Informality, public good, enforcement

The size of the informal sector is close to 40 percent in many developing economies, especially in many Latin American countries.[1] Moreover, informality levels are on the rise in many developing economies. According to Perry and others, informality levels grew in most Latin American countries between the late 1980s and the early 2000s.[2] As pointed out in La Porta and Shleifer, the informal sector has extremely low productivity compared to the formal economy, since informal firms tend to be smaller and inefficient.[3]

ACKNOWLEDGMENTS The authors are extremely grateful to Nicolás Idrobo for his excellent research assistance.

1. Schneider (2005). According to La Porta and Shleifer (2014), it is about 50 percent in the poorest countries.

2. Perry and others (2007). The measures analyzed include the percentage of employed workers, the percentage of salaried workers, the percentage of the labor force that lacks pensions, and the percentage of self-employed workers.

3. La Porta and Shleifer (2014).

This article explores optimal public policy, in terms of the level of tax enforcement and the supply of public goods, in an economy characterized by a huge informal sector, where the size of the informal sector is endogenously determined by the interaction between the government and individual decisions. The model we propose to analyze this issue builds on the work of Mejía and Posada.[4] In particular, we define informality as the group of productive activities that, ex ante, do not comply with government regulations. This noncompliance with government regulations can occur in different dimensions of economic activity, such as tax evasion, social security payments, minimum wages, sanitary and environmental regulations, and so forth. Furthermore, noncompliance with norms or regulations may be partial or total.[5] One of the most salient costs of being informal is the lack of access to some government-provided services, such as recourse to the judicial system to resolve contract-related disputes or the impossibility of participating in public training programs. Therefore, at the individual level, the decision to become informal can be viewed as a rational response to the system of incentives established through government enforcement, on the one hand, and the provision of some public goods (which can only be accessed if the firm or the individual fully complies with government regulations) on the other hand. In particular, firms might decide to become informal if the tax rate is too high for formal activities to be profitable (compared with informal activities, where taxes can be partially or totally evaded). Nevertheless, although such decisions might be rational at an individual level, this collective-action problem may generate aggregate inefficiencies.[6]

In this article, we develop a model in which a representative individual (or household) decides the amount of time he or she allocates to operate in a formal firm (or technology) and in an informal firm. By working with the formal technology, the individual can take advantage of a public good in the production process, but has to pay taxes. When operating with the informal firm, the individual does not pay taxes but cannot take advantage of the public good in the production process.[7] Additionally, the firm can be

4. Mejía and Posada (2011).

5. We focus on the informal sector and do not consider the so-called underground economy. Namely, while the latter involves a violation of the penal code, the former does not. In particular, tax avoidance in many countries or the violation of environmental laws is not necessarily investigated and punished by the penal system.

6. Loayza (2007).

7. We assume that the public good is not a pure public good in the sense that the government can partially or totally exclude informal firms from using it. Thus we are formally referring to a club good.

detected evading taxes with a probability that is increasing in both the level of government enforcement and the size of the informal firm's activities. In case of detection, the informal firm has to pay the evaded taxes plus a fine of a given size.

We consider a nonpaternalistic government that maximizes the representative household's welfare.[8] We assume that the state acts as a Stackelberg leader and decides how to allocate public expenditures, collected through the tax system, between the provision of a public good (which can only be used by firms operating in the formal sector) and enforcement activities, aimed at detecting informal firms that are evading taxes.[9] When deciding the optimal provision of the public good, as well as the enforcement level, the government takes into account how the representative family reacts to these decisions. In this context, and taking into account the shadow cost of public funds, we characterize the optimal carrot-and-stick policy.

Our article contributes to the growing literature on informality by providing a new view in which the size of the informal sector is endogenously determined by the interaction between the government and individual decisions. While Cerda and Saravia focus on (heterogeneous) firms' decisions to allocate time and factors between both sectors, we consider that this decision comes from workers and thus is made by a representative consumer.[10] Our theoretical setup allows us to point out the link between the size of the informal sector and the shadow cost of public funds, in line with the empirical evidence presented by Auriol and Warlters.[11] Indeed, we show that the greater the distortions, the larger the informal sector. This result is also consistent with the findings of Adaman and Mumcu, who show, in a global game framework, that inefficiencies and the low level of trustworthiness of the public sector induce an equilibrium with high levels of informality.[12] Despite the fact that the state

8. The debate concerning the state's objective in the presence of an informal sector (where a norm is partially or totally violated) is quite complex. It might seem natural to assume that the state's objective is to maximize formal production and discourage informal production (see, for instance, Mejía and Posada, 2011, for a positive analysis using this assumption). However, a less paternalistic state might simply maximize the representative household's utility. For instance, this assumption may reflect a democratic political system in which the government has been elected in function of the majority's preferences in terms of enforcement.

9. In this sense, this is not a pure public good, as it is possible to exclude users.

10. Cerda and Saravia (2013). Another difference is that their model is intensive and extensive, while we focus on an intensive approach.

11. Auriol and Warlters (2005).

12. Adaman and Mumcu (2010). See also Torgler and Schneider (2009) for empirical evidence.

cannot observe the level of informal production, it is optimal to spend all the budget on public goods and nothing on enforcement activities as long as there is no shadow cost of public funds. In other words, our model reveals that the shadow cost of public funds constitutes the rationale for spending on enforcement activities.

In addition to the works quoted above, our paper is related to and borrows from a number of strands of the literature. Cremer and Gahvari determine the optimal tax design in the presence of tax evasion; they provide sufficient conditions under which tax evasion decreases the optimal tax rate, while showing that an increase in the optimal tax rate is also possible.[13] Boadway and Sato also explore this issue, but they explicitly consider the presence of an informal sector, where the size of the informal sector is endogenous and mainly determined by public policies, as in our setup.[14] Nevertheless, both of these articles focus on optimal taxation issues, whereas our analysis is devoted to the optimal public policy subject to a balanced budget constraint (for example, the choice between the carrot—the provision of a public good—and the stick—enforcement activities).

Besfamille, de Donder, and Lozachmeur analyze the relation between tax enforcement, aggregate output, and government revenue when imperfectly competitive firms evade a specific output tax.[15] They reveal that aggregate output decreases with the level of tax enforcement. Government revenue increases with enforcement when the tax rate is low, but when the tax rate is high, government revenue is either inversely U-shaped or decreasing in the level of enforcement. In line with Besfamille, de Donder, and Lozachmeur, our article analyzes the relationship between the tax level, enforcement, and government revenue. We find that the size of the tax distortion (for example, the shadow cost of public funds) is inversely related to the size of the formal sector, the tax rate, and the optimal provision of the public good and is positively related to the size of the informal sector and optimal enforcement activities.[16]

Finally, our enforcement effort variable can be seen as a stochastic way to tax the informal sector.[17] Even though we do not have tax threshold effects,

13. Cremer and Gahvari (1994).
14. Boadway and Sato (2009).
15. Besfamille, de Donder, and Lozachmeur (2009).
16. There are obviously other channels to explain the size of the informal sector. See, for instance, D'Erasmo (2015) for a channel related to the access to credit.
17. Mubiru (2010).

the fact that the probability of detection increases with the size of the informal sector also constitutes an incentive to limit its size.[18]

The rest of our article is organized as follows. The next section presents the setup. We then characterize the optimal public policy in various contexts. First we solve the central planner's problem when he is able to choose the time devoted by the representative family in both sectors. Next we consider the decentralized model with tax distortions. The last section presents some concluding remarks.

The Model

The production of the final good in the formal sector, y_f, depends on the amount of labor allocated by the representative household to this sector, l_f, and on a public good, b, produced by the state. The production function in the formal sector is described by $y_f = f(l_f, b)$. We assume that the formal sector is characterized by positive and decreasing returns to each input: $f_{l_f}, f_b > 0$ and $f_{l_f l_f}, f_{bb} < 0$. Additionally, we assume that formal labor and the public good are complementary in the production of the final good: $f_{l_f b} > 0$. Alternatively, the final good can be produced in the informal sector, that is, $y_i = g(l_i)$, with $g_{l_i} > 0$ and $g_{l_i l_i} < 0$. We consider that there is no positive externality from the formal to the informal sector.[19] Rather, it is the relative advantage generated by the public good between the two sectors and not the absolute one that is relevant for the representative household's optimal allocation decisions. As the benefit generated by the public good in the formal sector is endogenous, the assumption of no externality to the informal sector is not restrictive.

We denote $p(l_i, e)$ the probability that the state detects the informal firm evading taxes. This probability depends on the size of the informal firm, l_i, and on the state's allocation of resources to enforcement activities, e. We assume $p_e > 0$, $p_{ee} < 0$, $p_{l_i} > 0$, and $p_{l_i l_i} > 0$. In words, the probability that an informal firm is detected evading taxes is increasing in the state's enforcement efforts, but with decreasing returns. The probability of detection is increasing and convex in the size of the informal firm. We also assume that $p(.)$ satisfies the following Inada condition: $(\partial p(0, l_i)/\partial e = +\infty)$. Furthermore, we consider

18. See Kanbur and Keen (2014).

19. The term *public good* is often used in the informality literature. Nevertheless, in our framework, as long as the informal sector cannot benefit from the public good, strictly speaking, we should refer to b as a *club good*.

that $p_{el_i} > 0$, which means that the marginal effect of enforcement on the probability of detection is increasing in the size of the informal firm (the level of enforcement).

We assume that the cost to an informal firm of being detected evading taxes consists in a fine of size $\phi g(l_i)$. This fine may be interpreted as the opportunity cost of not being able to produce if the firm is closed down for a certain amount of time or as a pure fine that the firm has to pay if it is detected.[20] The representative household takes the strategic variables of the state (tax rate, enforcement, and public good provision) as given. We consider that the household's total labor supply, L_s, is exogenously given. Thus the household's decision consists only in allocating the total labor supply between the formal and informal sectors in order to maximize expected income. The problem faced by the representative household can be written as follows:

$$\max_{\{l_f, l_i\}} (1-\tau) f\left(l_f, b\right) + g\left(l_i\right) - p\left(l_i, e\right)\phi g\left(l_i\right), \tag{1}$$

subject to

$$L_s = l_f + l_i, \tag{2}$$

where $(1 - \tau)f(l_f, b)$ represents the net income (after taxes) earned in the formal sector, while $g(l_i)$ denotes the income earned in the informal sector. These income levels must be reduced by $p(l_i, e)\phi g(l_i)$, the expected cost of the fine if $l_i > 0$.

The optimal allocation of labor between the formal and informal sectors is given by[21]

$$g_{l_i}\left(l_i\right)\left[1 - p\left(l_i, e\right)\phi\right] - p_{l_i}\left(l_i, e\right)\phi g\left(l_i\right) = (1-\tau) f_{l_f}\left(l_f, b\right). \tag{3}$$

The household's optimal allocation of labor between the formal and informal sectors is such that the expected net marginal benefit from allocating an extra

20. In this article, since we consider a risk-neutral representative household, we do not focus on the trade-off between enforcement effort and the size of the fine. Finally, we consider that the fine takes a finite value due to a limited liability argument. Otherwise, the solution would be straightforward and would have followed an argument à la Becker: the fine would be equal to infinity. This limited liability argument also explains why, in practice, the amount of resources collected by states through fines is rather limited.

21. We adopt an intensive approach and thus focus our attention on an interior solution.

unit of time to the informal sector is equal to its net marginal benefit in the formal sector. On the one hand, the marginal benefit from allocating an extra unit of time to the informal sector (the left-hand side of equation 3) is given by the marginal productivity in the informal sector, g_{l_i}, times $[1 - p(l_i, e)\,\phi]$, minus the marginal increase in the probability of being detected, p_{l_i}, times the size of the fine that has to be paid if the firm is detected, $\phi g(l_i)$. On the other hand, the net marginal benefit from allocating an extra unit of time to the formal sector (the right-hand side of equation 3) is simply composed of the marginal productivity of labor in the formal sector net of taxes, $(1 - \tau) f_{l_f}, (l_f, b)$.

Remark 1: All other things equal, the level of production in the formal sector increases with the level of provision of public goods and enforcement activities.

Proof: See appendix.

A higher provision of the public good makes the formal production more attractive, ceteris paribus, owing to the fact that labor and the public good are complementary in the production process. As the amount of enforcement efforts increases, the incentive to allocate time to the informal sector decreases because the probability of being caught evading taxes and having to pay the fine increases. Therefore, both instruments allow the state to reduce the size of the informal sector.

The Optimal Public Policy

In this section we distinguish between several cases. We start with a first-best analysis where the state acts as a central planner and can directly choose the representative household's labor supply between the formal and informal sectors, l_f and l_i, respectively. In the second-best allocation, we consider the situation in which the state cannot choose or impose the household labor supply between the two sectors. Nevertheless, it still behaves as a Stackelberg leader in the sense that it chooses the vector (b, e, τ) taking into account the optimal reaction of the representative household to its choices.

First-Best Analysis

In a first-best analysis, the state can directly choose the household's allocation of time between the formal and informal sectors. In this context, enforcement activities, as captured by e, should be interpreted as the cost generated by the

burden of the proof. The state's objective is the maximization of the household's utility, that is,

$$\max_{\{l_f,b,e\}} W^{FB} = (1-\tau) f\left(l_f, b\right) + g\left(L_s - l_f\right)\left[1 - p\left(l_i, e\right)\phi g\left(l_i\right)\right], \tag{4}$$

subject to the following budget constraint:

$$\tau\left(1-\lambda\right) f\left(l_f, b\right) + p\left(l_i, e\right)\phi g\left(l_i\right) \geq b + e. \tag{5}$$

The budget constraint says that the tax collected plus the fine earned through the detection process is equal to the sum of the expenditures on the public good and detection efforts. An easy way to capture these distortions is to consider that for each unit of tax collected, a proportion λ is lost.[22] Because a fraction λ of the taxes collected is lost, the first term on the left-hand side of the state's budget constraint is scaled down by a fraction 1 – λ.

As the tax and detection efforts intervene negatively in the objective function, it is straightforward that the budget constraint is binding. Therefore, we write the tax rate, τ, that ensures that the budget constraint holds with equality as

$$\tau = \frac{b + e - p\left(l_i, e\right)\phi g\left(l_i\right)}{\left(1-\lambda\right) f\left(l_f, b\right)}. \tag{6}$$

Replacing τ with equation 6 in the state's objective function, the state's program becomes

$$\max_{\{l_f,b,e\}} W^{FB} = f\left(l_f, b\right) + g\left(L_s - l_f\right) - \frac{b+e}{1-\lambda} + \frac{\lambda}{1-\lambda} p\left(l_i, e\right)\phi g\left(l_i\right). \tag{7}$$

The objective function is then composed of the sum of the production in both sectors, minus the state's expenditures on enforcement and the public good, $b + e$, scaled down by a fraction 1 – λ. Additionally, the fine collected, $p(l_i, e)\phi g(l_i)$, must be added as it relaxes the budget constraint. The expected

22. Laffont and Tirole (1993). To some extent, 1 – λ captures the quality of the tax system: when $1 is collected through tax, λ is lost.

fine collected, $p(l_i, e)\phi g(l_i)$, is multiplied by λ because for each \$1 coming from the fine (that is, not coming from tax distortions), λ is saved.[23]

In the first-best scenario, in addition to maximizing the welfare function with respect to its policy instruments b and e, the state is able to choose the amount of labor supply in the formal sector, l_f. Moreover, since the tax rate is defined by the budget constraint, it is equivalent to maximizing with respect to τ or one of these instruments, that is, b and e.

Proposition 1: The first-best allocation is characterized by

$$\frac{\partial f\left(l_f^*, b^*\right)}{\partial l_f} - \frac{\partial g\left(L_s - l_f^*\right)}{\partial l_f}$$

$$= \frac{\lambda\phi}{1-\lambda}\left[\frac{\partial p\left(L_s - l_f^*, e^*\right)}{\partial l_f} g\left(L_s - l_f^*\right) + \frac{\partial g\left(L_s - l_f^*\right)}{\partial l_f} p\left(L_s - l_f^*, e^*\right)\right];$$

$$\frac{\partial f\left(l_f^*, b^*\right)}{\partial b} = \frac{1}{1-\lambda};$$

$$\frac{\partial p\left(L_s - l_f^*, e^*\right)}{\partial e} g\left(L_s - l_f^*\right) = \frac{1}{\lambda\phi}.$$

The first condition above is an efficiency condition: if the state is able to choose the household's allocation of time between the formal and informal sectors, and if there are no tax distortions (that is, $\lambda = 0$), it would choose an allocation such that the marginal productivity of labor is the same in both sectors. Moreover, this condition reveals that, for a given level of provision of the public good, the state's optimal allocation of labor to the formal sector decreases with the presence of tax distortions. Because taxes create distortions and the fines collected constitute an alternative source of revenue, the state is more lenient on informality compared with the case of no distortions. The second condition says that the state chooses the optimal provision of the public good in such a way that the marginal productivity of the public good is equal to the marginal cost of providing it, weighted by the size of the tax

23. In this setting, we ignore that some distortions may be associated with the fine system. In the appendix, we show that our results remain valid as long as the tax distortion is higher than the distortion introduced by the fine system.

distortion. All other things equal, for higher values of λ, the spending on the public good becomes lower. The third condition points out that even in the case where the state can impose the household's allocation of time between the formal and informal sectors, it allocates positive levels of resources to enforcement in order to increase the probability of detecting informal firms and thus to increase revenue without tax distortions. Moreover, as the probability of detection is increasing and concave in e, the optimal level of enforcement is increasing in the tax distortion, λ. It also increases with the fine rate, ϕ.

We now define the marginal rate of substitution between enforcement and public good provision:

$$\mathrm{MRS}_{e/l_f} = \frac{\left[\partial p\left(L_s - l_f, e\right)/\partial l_f\right] g\left(L_s - l_f\right) + \left[\partial g\left(L_s - l_f\right)/\partial l_f\right] p\left(L_s - l_f, e\right)}{\left[\partial p\left(L_s - l_f, e\right)/\partial e\right] g\left(L_s - l_f\right)}.$$

This marginal rate of substitution describes the trade-off the state faces when choosing l_f and e. For a given level of fines collected, the state can choose to increase the level of enforcement, $g(L_s - l_f)(\partial p(L_s - l_f, \mathrm{e})/\partial e)$, or the level of informality, l_i. In this case, there are two effects at work. First, for a given level of informal production, $g(L_s - l_f)$, it marginally increases the probability of detection, $(\partial p(L_s - l_f, e)/(\partial l_f)$. Second, for a given level of detection of informal firms, it marginally increases the size of the fine collected, $\partial g(L_s - l_f)/\partial l_f$.

Finally, combining the three first-order conditions, we obtain:

Remark 2:

$$\frac{\left[\partial f\left(l_f, b\right)/\partial l_f\right] - \left[\partial g\left(L_s - l_f\right)/\partial l_f\right]}{\partial f\left(l_f, b\right)/\partial b} = \left(1 - \lambda\right)\mathrm{MRS}_{e/l_f}.$$

To understand the previous equality, consider the following function:

$$\Gamma\left(l_f, b\right) = \frac{\partial f\left(l_f, b\right)}{\partial l_f} - \frac{\partial g\left(L_s - l_f\right)}{\partial l_f}.$$

With λ equal to 0, we have $\Gamma(l_f^*, b^*)$, which can be considered a "pure" efficiency condition, in that the time devoted to each sector is chosen to equate their marginal productivities. Technically, it comes from the fact that $e^* = 0$, implying that $\partial p(L_s - l_f^*, e^* = 0)/\partial e \to +\infty$ and therefore $\mathrm{MRS}_{e/l_f}(e, l_f) = 0$.

In words, it means that the presence of tax distortions obligates the state to alter the pure efficiency condition. These distortions depend on the marginal rate of substitution between the size of the informal sector and the level of enforcement efforts. This marginal rate of substitution is informative about how much the state needs to increase its spending on enforcement to outweigh a reduction of the size of the informal sector in order to keep the expected fine constant.

To summarize, in the presence of tax distortions, the state may optimally tolerate a larger level of informality in order to reduce the negative impact of distortions on the state's revenues. On the one hand, more informality is associated with less expenditure on the public good, reducing the effect of tax distortions. On the other hand, because the probability of detection is increasing in the size of the informal firm, more informality implies more fines collected by the state and still fewer tax distortions. These results are summarized in the following corollary.

Corollary 2: The presence of tax distortions in developing countries may provide a rationale for a relatively large informal sector.

This remark is consistent with estimations by Auriol and Warlters for African developing countries.[24] Applying the standard Devarajan and others' 1-2-3 model to a database of thirty-eight countries, these authors find a strong positive relationship between the marginal cost of public funds and the informality levels for the countries in their sample.[25] Our first-best allocation results allow us to provide an explanation for this evidence. Even when the state can choose the household's allocation of time between the formal and informal sectors, it may be optimal for the state to tolerate a larger informal sector in the presence of higher tax distortions. The shadow cost of public funds may constitute a piece of the puzzle that explains the so-called broken contract between the state and the citizens in developing countries that are characterized by high levels of informality.[26]

Second-Best Analysis

In the second-best analysis, we consider a decentralized economy in which the state cannot choose the household's allocation of labor between the formal and informal sectors. However, we assume that the state is a Stackelberg leader

24. Auriol and Warlters (2005).
25. Devarajan and others (1994).
26. See Perry and others (2007).

and chooses the optimal allocation of tax revenues between the provision of the public good and enforcement efforts, taking into account the household's reaction to these choices.

In this context, the state's objective becomes

$$\max_{\{l_f,b,e\}} W^{SB} = f\left(\hat{l}_f, b\right) + g\left(L_s - \hat{l}_f\right) - \frac{b+e}{1-\lambda} + \frac{\lambda}{1-\lambda} p\left(L_s - \hat{l}_f, e\right)\phi g\left(L_s - \hat{l}_f\right),$$

where $\hat{l}_f(e, \tau, b)$ is implicitly determined by the household's reaction function (equation 3).

The first-order conditions are as follows:

$$\frac{\partial f(.)}{\partial b} - \frac{1}{1-\lambda} = \left\{\frac{\partial f(.)}{\partial \hat{l}_f} - \frac{\partial g(.)}{\partial \hat{l}_f} - \frac{\lambda\phi}{1-\lambda}\left[\frac{\partial p(.)}{\partial \hat{l}_f} g\left(l_i\right) + \frac{\partial g(.)}{\partial \hat{l}_f} p\left(l_i, e^*\right)\right]\right\}\frac{\partial \hat{l}_f}{\partial b}$$

and

$$\frac{\lambda\left[\partial p(.)/\partial e\right] g\left(l_i\right)\phi - 1}{1-\lambda}$$

$$= \left\{\frac{\partial f(.)}{\partial \hat{l}_f} - \frac{\partial g(.)}{\partial \hat{l}_f} - \frac{\lambda\phi}{1-\lambda}\left[\frac{\partial p(.)}{\partial \hat{l}_f} g\left(l_i\right) + \frac{\partial g(.)}{\partial \hat{l}_f} p\left(l_i, e^*\right)\right]\right\}\frac{\partial \hat{l}_f}{\partial e}.$$

In order to understand some features of the optimal policy, consider first that the state wants to implement the same level of formality (relative to informality) in the first-best as in the second-best analysis (that is, $l_f^* = l_f^{**}$). In this case, we have

—*Lemma 1*: If $l_f^* = l_f^{**}$, then the optimal public policy consists of $e^{**} > e^*$ and $b^{**} > b^*$.

—*Proof*: See appendix.

Lemma 1 says that in order to implement the same level of formality as in the first-best analysis, the state has to spend more resources on public good provision as well as on enforcement. this is because, as pointed out in Remark 3, the household's chosen level of formal activity increases in b and e.

Let us consider the case where $\lambda = 0$.

Remark 3: For $\lambda = 0$, we have $e^{**} = 0$.

This remark points out that if taxes do not generate distortions, the state still prefers to spend all its budget on the public good and nothing on enforcement. This result comes from the same mechanism as in Proposition 1 and Remark 1. It says that spending on the public good is a sufficient instrument, in the absence of tax distortions, to maximize the representative household's welfare. Conversely, the shadow cost of public funds constitutes the rationale behind the state's investment in enforcement activities. From the state's perspective, all other things being equal, it becomes more profitable to spend on enforcement activities when λ increases. Indeed, in the second-best analysis, the first-order condition in e implies that the term $1 - (\partial p(.)/\partial e)\lambda\phi$ is negative, which means that the marginal cost of enforcement activities is lower than the marginal benefit, inducing the state to spend a positive amount of resources on enforcement activities.

Rearranging the first-order conditions gives

$$\frac{\partial \hat{l}_f}{\partial b} = \frac{\left[1/(1-\lambda)\right] - \left[\partial f(.)/\partial b\right]}{\left[\partial f(.)/\partial \hat{l}_f\right] + \left[\partial g(.)/\partial \hat{l}_f\right] + \left[\lambda\phi\,\partial p(.)/1 - \lambda\,\partial \hat{l}_f\right]}$$

and

$$\frac{\partial \hat{l}_f}{\partial e} = \frac{\left[1/(1-\lambda)\right] - \left[\lambda\phi/(1-\lambda)\right]\left[\partial p(.)/\partial e\right]}{\left[\partial f(.)/\partial \hat{l}_f\right] + \left[\partial g(.)/\partial \hat{l}_f\right] + \left[\lambda\phi/(1-\lambda)\right]\left[\partial p(.)/\partial \hat{l}_f\right]}.$$

Combining the last two conditions leads to the following proposition:

Proposition 4: In the presence of distortionary taxes and incentive problems, the optimal public policy follows:

$$\frac{\partial \hat{l}_f/\partial b}{\partial \hat{l}_f/\partial e} = \frac{\left[1/(1-\lambda)\right] - \left[\partial f(.)/\partial b\right]}{\left[1/(1-\lambda)\right] - \left[\lambda\phi/(1-\lambda)\right]\left[\partial p(.)/\partial e\right]}.$$

To explain Proposition 4, we first focus on the case when $\lambda = 0$. The previous condition becomes

$$\frac{\partial f(.)}{\partial b} = 1 - \mathrm{MRS}_{e/b}.$$

We can easily recognize the equality obtained in the first-best case between the marginal productivity of the public good in the formal sector and its marginal cost. Nevertheless, this equality is distorted by the marginal rate of substitution between e and b. As long as the state cannot directly choose the level of informality, it has to set the amounts e^{**} and b^{**} taking into account their relative impact on the relative size of the formal and informal sectors.

Concluding Remarks

This article develops a model in which the size of the informal sector is endogenously determined by the interaction of a representative individual and the government. On the one hand, the representative individual has to decide the allocation of time between the formal and informal sectors. In the formal sector, the individual can make use of a public good provided by the government, but has to pay taxes with probability one; in the informal sector, the individual only pays taxes with an endogenously determined probability (which is lower than one), but cannot benefit from the use of the government-provided good. The government, on the other hand, has to decide the allocation of resources (collected through the tax system and fines imposed on those informal firms that are detected) between enforcement activities to detect and penalize informal activities (the stick) and the provision of a public good that can only be used in the formal sector (the carrot). We emphasize the role of tax distortions (the shadow cost of public funds) in the determination of the size of the informal sector, the optimal tax rate, and total production, among other endogenous variables.

Our model could be extended in several ways. First, we consider a representative household. The introduction of some heterogeneity would allow for different levels of consumption of the public good. In practice, the consumption of public goods may follow a U-shaped curve that we cannot take into account in our setup. Second, in this article, we intentionally adopt a black-box approach regarding the shadow cost of public funds. This approach has two advantages. First, we focus on the normative feature of public policy, and we show that this shadow cost of public funds is one of the rationales of the informal sector and the state's enforcement policy. Second, it is easier to connect with the empirical literature quoted in this article. However, it would be interesting to shed light on public policies that would simultaneously reduce the size of the informal sector and weaken one of its rationales. It is on our research agenda.

Appendix

Proof of Remark 1

Consider the following implicit function:

$$\phi\left(l_f, b, e\right) \overset{def}{=} (1-\tau) f_{l_f}\left(l_f, b\right) - g_{l_i}\left(l_i\right)\left[1 - p\left(l_i, e\right)\phi\right] + p_{l_i}\left(l_i, e\right)\phi g\left(l_i\right).$$

The implicit function theorem yields

$$\frac{dl_f}{\partial b} = -\frac{\left(\partial\phi/\partial b\right)}{\left(\partial\phi/\partial l_f\right)}.$$

Since $(\partial\phi/\partial l_f)$ is negative owing to the second-order conditions, we have

$$\text{sign}\left(\frac{dl_f}{\partial b}\right) = \text{sign}\left(\frac{\partial\phi}{\partial b}\right),$$

with

$$\frac{\partial\phi}{\partial b} = (1-\tau)\frac{\partial^2 f\left(l_f, b\right)}{\partial l_f \partial b} > 0.$$

Similarly, we have

$$\frac{dl_f}{\partial e} = -\frac{\left(\partial\phi/\partial e\right)}{\left(\partial\phi/\partial l_f\right)},$$

with

$$\frac{\partial\phi}{\partial e} = \left[\frac{\partial p(.)}{\partial e} g_{l_i}\left(l_i\right) + \frac{\partial^2 p(.)}{\partial l_f \partial b} g\left(l_i\right)\right]\phi > 0.$$

Q.E.D.

Proof of Lemma 1

Consider the following function:

$$\psi\left(e^{**}, b^{**}, l_f^{**}, \lambda\right) \overset{\text{def}}{=} \frac{\partial f(.)}{\partial \hat{l}_f} - \frac{\partial g(.)}{\partial \hat{l}_f} - \frac{\lambda\phi}{1-\lambda}\left[\frac{\partial p(.)}{\partial \hat{l}_f} g\left(l_i\right) + \frac{\partial g(.)}{\partial \hat{l}_f} p\left(l_i, e^*\right)\right].$$

In the first-best, we have $\Psi(e^{**}, b^{**}, l_f^{**}, \lambda) = 0$. Moreover, we have

$$\begin{cases} \dfrac{\partial f\left(l_f^*, b^*\right)}{\partial b} = \dfrac{1}{1-\lambda}; \\ g\left(L_s - l_f^*\right) = \dfrac{\lambda\phi}{1-\lambda} \dfrac{1}{\partial p\left(L_s - l_f^*, e^*\right)/\partial e}. \end{cases}$$

The second-best is characterized by

$$\begin{cases} \dfrac{\partial f\left(\hat{l}_f, b^{**}\right)}{\partial b} = \dfrac{1}{1-\lambda} - \psi\left(e^{**}, b^{**}, \hat{l}_f, \lambda\right)\dfrac{\partial \hat{l}_f}{\partial b}; \\ g\left(L_s - \hat{l}_f\right) = \left[\dfrac{\lambda\phi}{1-\lambda} - \psi\left(e^{**}, b^{**}, \hat{l}_f, \lambda\right)\dfrac{\partial \hat{l}_f}{\partial e}\right]\dfrac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e}. \end{cases}$$

Therefore, we obtain that

$$\begin{aligned} & g\left(L_s - \hat{l}_f\right) - g\left(L_s - l_f^*\right) \\ &= \left[\frac{\lambda\phi}{1-\lambda} - \psi\left(e^{**}, b^{**}, \hat{l}_f, \lambda\right)\frac{\partial \hat{l}_f}{\partial e}\right]\frac{1}{\partial p\left(L_s - l_f^*, e^{**}\right)/\partial e} \\ &\quad - \frac{\lambda\phi}{1-\lambda}\frac{1}{\partial p\left(L_s - l_f^*, e^*\right)/\partial e} \\ &= \frac{\lambda\phi}{1-\lambda}\left[\frac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e} - \frac{1}{\partial p\left(L_s - l_f^*, e^{**}\right)/\partial e}\right] \\ &\quad - \psi\left(e^{**}, b^{**}, \hat{l}_f, \lambda\right)\frac{\partial \hat{l}_f}{\partial e}\frac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e}. \end{aligned}$$

We have

$$g\left(L_s - \hat{l}_f\right) - g\left(L_s - l_f^*\right) \geq 0 \Leftrightarrow$$

$$\frac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e} - \frac{1}{\partial p\left(L_s - l_f^*, e^*\right)/\partial e}$$

$$\leq \psi\left(e^{**}, b^{**}, \hat{l}_f, \lambda\right)\frac{\partial \hat{l}_f}{\partial e}\frac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e}.$$

This implies that

$$\frac{1}{\partial p\left(L_s - l_f, e^{**}\right)/\partial e} \leq \frac{1}{\partial p\left(L_s - l_f^*, e^*\right)/\partial e}.$$

Because $p_{el_i} > 0$, we have $e^{**} > e^*$. The same reasoning applies for the public spending. Q.E.D.

Robustness Check

In this section, we explore the possibility of introducing a distortion, λ_f, at the fine level. The state's objective is the maximization of the household's utility, that is,

$$\max_{\{l_f, b, e\}} W^{FB} = (1-\tau) f\left(l_f, b\right) + g\left(L_s - l_f\right)\left[1 - p\left(l_i, e\right)\phi g\left(l_i\right)\right],$$

subject to the following budget constraint:

$$\tau\left(1-\lambda\right) f\left(l_f, b\right) + p\left(l_i, e\right)\phi g\left(l_i\right)\left(1-\lambda_f\right) \geq b + e,$$

which can be rewritten as

$$\tau = \frac{b + e - p\left(l_i, e\right)\phi g\left(l_i\right)\left(1-\lambda_f\right)}{\left(1-\lambda\right) f\left(l_f, b\right)}.$$

Replacing the previous expression in the state's objective function, we obtain

$$\max_{\{l_f, b, e\}} W^{FB} = f\left(l_f, b\right) + g\left(L_s - l_f\right) - \frac{b+e}{1-\lambda} + \frac{\lambda - \lambda_f}{1-\lambda} p\left(l_i, e\right)\phi g\left(l_i\right).$$

The first-order conditions yield

$$\frac{\partial f\left(l_f^*, b^*\right)}{\partial l_f} - \frac{\partial g\left(L_s - l_f^*\right)}{\partial l_f}$$

$$= \frac{\left(\lambda - \lambda_f\right)\phi}{1-\lambda}\left[\begin{array}{l} \dfrac{\partial p\left(L_s - l_f^*, e^*\right)}{\partial l_f} g\left(L_s - l_f^*\right) \\ \\ + \dfrac{\partial g\left(L_s - l_f^*\right)}{\partial l_f} p\left(l_f^*, e^*\right) \end{array}\right];$$

$$\frac{\partial f\left(l_f^*, b^*\right)}{\partial b} = \frac{1}{1-\lambda};$$

$$\frac{\partial p\left(L_s - l_f^*, e^*\right)}{\partial e} g\left(L_s - l_f^*\right) = \frac{1}{\left(\lambda - \lambda_f\right)\phi}.$$

The amount of the public good supplied remains unchanged. Moreover, our results remain qualitatively unchanged as long as $\lambda - \lambda_f > 0$. Otherwise, the optimal policy involves zero enforcement effort, and consequently the optimal informal labor supply would be higher.

References

Adaman, Fikret, and Ayse Mumcu. 2010. "Perceptions on Governance Effectiveness and Informality: A Self-Fulfilling Equilibrium." Working Paper 2010/05. Bogazici University.

Auriol, Emmanuelle, and Michael Warlters. 2005. "Taxation Base in Developing Countries." *Journal of Public Economics* 89(4): 625–46.

Besfamille, Martin, Philippe de Donder, and Jean-Marie Lozachmeur. 2009. "Tax Enforcement May Decrease Government Revenue." *Economics Bulletin* 29(4): 2665–72.

Boadway, Robin, and Motohiro Sato. 2009. "Optimal Tax Design and Enforcement with an Informal Sector." *American Economic Journal: Economic Policy* 1(1): 1–27.

Cerda, Rodrigo, and Diego Saravia. 2013. "Optimal Taxation with Heterogeneous Firms and Informal Sector." *Journal of Macroeconomics* 35(C): 39–61.

Cremer, Helmuth, and Firouz Gahvari. 1994. "Tax Evasion, Concealment, and the Optimal Linear Income Tax." *Scandinavian Journal of Economics* 96(2): 219–39.

D'Erasmo, Pablo. 2015. "Access to Credit and the Size of the Formal Sector in Brazil." *Economía* 16(2): 143–99.

Devarajan, Shantayanan, Delfin S. Go, Jeffrey D. Lewis, and others. 1994. "Policy Lessons from a Simple Open Economy Model." Policy Research Working Paper 1375. Washington: World Bank.

Kanbur, Ravi, and Michael Keen. 2014. "Thresholds, Informality, and Partitions of Compliance." *International Tax and Public Finance* 21(4): 536–59.

Laffont, Jean-Jacques, and Jean Tirole. 1993. *A Theory of Incentives in Procurement and Regulation.* MIT Press.

La Porta, Rafael, and Andrei Shleifer. 2014. "Informality and Development." Working Paper 20205. Cambridge, Mass.: National Bureau of Economic Research.

Loayza, Norman. 2007. "The Causes and Consequences of Informality in Peru." Working Paper 18. Lima: Reserve Bank of Peru.

Mejía, Daniel, and Carlos E. Posada. 2011. "Informalidad: Teoría e implicaciones de política." In *Mercado de trabajo en Colombia: Hechos, tendencias e instituciones*, edited by Luis Eduardo Arango and Franz Hamann. Bogotá: Bank of the Republic.

Mubiru, Alex. 2010. "Domestic Resource Mobilization across Africa: Trends, Challenges, and Policy Options." Committee of Ten Policy Brief 2. Tunis: African Development Bank.

Perry, Guillermo E., W. Maloney, O. Arias, and others. 2007. *Informality: Exit and Exclusion.* Washington: World Bank.

Schneider, Friedrich. 2005. "Shadow Economies of 145 Countries All over the World: What Do We Really Know?" Working Paper 2005-13. Basel: Center for Research in Economics, Management, and the Arts.

Torgler, Benno, and Friedrich Schneider. 2009. "The Impact of Tax Morale and Institutional Quality on the Shadow Economy." *Journal of Economic Psychology* 30(2): 228–45.