WP/20/211

IMF Working Paper

Financial Frictions and Firm Informality: A General Equilibrium Perspective

by Luis Franjo, Nathalie Pouokam, and Francesco Turino

INTERNATIONAL MONETARY FUND

 WP/20/211

IMF Working Paper

Institute for Capacity Development

Financial Frictions and Firm Informality: A General Equilibrium Perspective[1]

Prepared by Luis Franjo [a], Nathalie Pouokam [b], and Francesco Turino [a]
[a]Fundamentos del Análisis Económico (FAE), Universidad de Alicante
[b]International Monetary Fund

Authorized for distribution by Norbert Funke

August 2020

Abstract

In this paper we build a model of occupational choice with informal production and progressive income taxation. We calibrate the model to the Brazilian economy to evaluate the impact of removing financial frictions on informality. We find that financial deepening leads to a drop in the size of the informal sector (from 37 percent to 22 percent of official GDP), to an increase in measured TFP (by 4 percent), to an increase in official GDP (by 27 percent), to a decrease in tax evasion (by 17 percent) and to an increase in fiscal revenues (by 15 percent). When assessing the response of this policy at different levels of financial development, we find a non-linear relationship between the credit-to-GDP ratio on the one hand, and either the size of the informal economy, or GDP per capita on the other hand. We test these features with cross-country data and find evidence in favor of both types of non-linearity. We also investigate changes in the income tax progressitivity as an alternative policy and find it to be more effective in countries with a medium to high level of financial markets development.

Keywords: informality, financial frictions, taxation, entrepreneurship, productivity, misallocation.
JEL Classification: E21, E22, E26, E44, H26, O10, O40.

[1] We would like to thank Lian Allub, M. Dolores Collado, Francesco Serti, Selene Perazzini, Perdo Albarrán, Iñigo Iturbe-Ormaetxe, Pedro Gomes, Carlos Alberto De Resende, Ralph Chami, and seminar participants at the Institute for Capacity Development of the International Monetary Fund in Washington D.C. F. Turino acknowledges support from the Institute for Capacity Development of the International Monetary Fund as part of its visiting scholar program. F. Turino also acknowledges financial support from the MINECO/FEDER (ECO2015- 70540-P). Luis Franjo acknowledges financial support from the Cátedra Germán Bernácer.

Authors' E-Mail Addresses: luis.franjogarcia@ua.es, npouokam@imf.org, fturino@ua.es

1. Introduction

Developing countries have less developed financial markets (Abiad, Detragiache and Tressel (2010)), and larger informal sectors (Antunes and Cavalcanti (2007), Quintin (2008)). A less developed financial market is associated with lower productivity and GDP per capita (Buera, Kaboski and Shin (2011), Allub and Erosa (2018)), while a larger informal sector may imply lower aggregate productivity (D'Erasmo and Boedo (2012), Ulyssea (2018)) and a higher rate of tax evasion (Orsi, Raggi and Turino (2014)). At the same time, advanced economies with developed financial markets may still have a large share of informality (Kuehn (2014), Hassan and Schneider (2016)).

This paper quantitatively assesses the extent to which informality and development in financial markets are related. To this end, we set up a model of entrepreneurship with informality and progressive income taxation in which informal entrepreneurs do not have access to credit and face an endogenous probability of detection by fiscal authorities. Specifically, building on Buera and Shin (2013), we develop a general equilibrium model of occupational choice with credit market imperfections where entrepreneurs decide to produce either in the formal or in the informal sector. Entrepreneurs in the formal sector have access to credit up to a proportional amount of their savings and pay progressive taxes on total income. By contrast, entrepreneurs in the informal sector operate in financial autarky but avoid taxation by concealing production activities. However, as in Allingham and Sandmo (1972), in each period an informal entrepreneur can be caught evading by the government and forced to pay taxes augmented by a penalty surcharge factor, with a probability that increases in the level of capital employed in informal production.

We calibrate our model to Brazilian data, matching the distribution of firms in both the formal and informal sectors. In our model, and consistently with the data (see e.g. La Porta and Shleifer (2014a), Ulyssea (2018)), informal firms are smaller and less productive than formal firms.[2] These findings are driven by two main assumptions: the lack of access to credit and the endogenous probability of detection. While the former assumption sets bounds to the maximum size of informal firms, the latter prevents wealthy entrepreneurs from running large informal businesses and, therefore, turns out to be crucial for capturing the observed dynamics of informal firms. Moreover, because of these properties, we identify three types of informal firms in our model: (i) firms that are sufficiently productive to operate in the formal sector, but are established in the informal sector because of the lack of access to credit;[3] (ii) firms that are still productive enough to profitably operate in the formal sector, but choose nevertheless to hide their production to avoid taxation (*parasite* firms); and (iii) firms that are run by low-skilled agents that are too unproductive to ever become formal entrepreneurs (*survival* firms).[4]

[2] This is a general equilibrium result as firms in the two sectors are ex-ante identical. Differences in productivity are then endogenously triggered by the optimal decisions of heterogeneous entrepreneurs to run their business informally or formally.

[3] As shown in La Porta and Shleifer (2014b), a World Bank survey reports that three quarters of the informal entrepreneurs interviewed see better access to credit as one of the potential benefits of becoming formal.

[4] Parasite and survival firms are definitions typically adopted in the literature on firm informality. The interested reader is refereed to Ulyssea (2018), and the references therein, for further details on this point.

To address the main question of the paper, we first evaluate the effect of financial deepening in our model economy. We find that completely removing financial constraints triggers a sizeable drop in the size of the informal economy (from 37 to 23 percent of official output), and in tax evasion (a drop of 17 percent), together with an important increase in official GDP (by 26.9 percent), an increase in fiscal revenues (by 15 percent), and an increase in measured aggregate TFP (by 4 percent).[5] To investigate the main mechanism behind these results, we further evaluate the potential gains from financial deepening in economies with different degrees of financial development ranging from financial autarky to perfect credit markets. Our main finding is that financial deepening significantly helps lower the size of the informal sector (and of tax evasion) in economies with a credit-to-GDP ratio below 60 percent, but has a virtually negligible effect on the size of the informal sector in economies in which the credit-to-GDP ratio is above this threshold. The mechanism behind this finding stems from a reallocation of resources from the informal sector to the formal sector, resulting in an amplification of the impact of financial deepening on official GDP in countries where the credit-to-GDP ratio is low (below 60 percent). In other words, our model predicts a non-linear relationship between the relative size of the formal economy and the level of financial markets development.

We test the implications of our model with cross-country data, reporting evidence in support of the non-linearities described above. These findings suggest that accounting for the informal economy is crucial for understanding the relationship between financial and economic development. As far as the observed non-linear pattern is concerned, our model offers a clear theoretical explanation. With the three different types of informal firms described above, financial deepening produces two countervailing effects on informal production. On the one hand, less financial frictions bring firms out of informality because highly productive entrepreneurs (first firm type) have larger access to credit. At the same time, a financial deepening reform also results in higher production costs, thereby boosting the relative gains from tax evasion and bringing more parasite firms to the informal sector. The first effect dominates the second one at high levels of financial frictions where a marginal increase in credit availability significantly boosts the gains from formality for talented entrepreneurs. By moving to the formal sector, these entrepreneurs operate at a larger scale, inducing in the aggregate a sizeable increase in official GDP, and a quick decline in the size of the informal sector. The strength of this effect, however, vanishes as the credit-to-GDP ratio gets larger. In fact, with more developed financial markets, highly productive entrepreneurs becomes less financially constrained and, hence, informal production is mostly carried out by small parasite and survival firms that optimally decide to self-finance their businesses to take advantage of tax evasion. In these economies therefore, further relaxing the collateral requirement when borrowing only has little impact on the size of the informal sector.

In terms of policy, the main message from our findings is clear: the role of lowering credit markets frictions in reducing informality and tax evasion is relevant, but only at high levels of credit markets imperfections. In these circumstances, the gains in terms of GDP per capita of financial deepening are amplified by firms switching from informality to formality. These findings contribute to the literature of economic development (Buera et al. (2011) and Allub and Erosa (2018)) by providing an additional argument in favor of financial deepening in developing economies. This argument is further supported by the results of an alternative policy experiment, where we evaluate the effect of changing the degree of progressivity in

[5] With the word measured TFP we refer to total factor productivity in the formal sector.

personal income taxation at different levels of financial development. As in the case of financial deepening, we find that reducing tax progressivity also decreases the size of the informal sector, reduces tax evasion, and, at the same time, boosts total GDP and measured aggregate TFP. However, by quantitatively comparing the effects of the two policies,[6] we draw three important conclusions: first, financial deepening is a more powerful policy for countries with a very low credit-to-GDP ratio; second, the effects of a flat income tax policy on informality and production is amplified by the existence of financial frictions; and third, for countries with a relatively large credit-to-GDP ratio, a flat income tax policy is more effective than financial deepening at curbing informality and tax evasion. Because our mechanism does not account for the role of tax progressivity in reducing income inequality, we do not view our model as supporting a shift to a flat income tax policy. However, one important contribution of our paper is that it highlights an important channel through which tax progressivity affects informality.

Our paper belongs to a branch of the literature that analyzes the role of financial frictions on productivity, aggregate output and entrepreneurship, as for example, Erosa (2001), Fonseca and Doornik (2020), Buera et al. (2011), Buera and Shin (2013), and Allub and Erosa (2018). We add to this literature by studying financial frictions within a framework in which entrepreneurs choose to run their businesses either formally or informally. This allows us to relate firm informality to financial and economic development.

The results of our analysis complement and extend the existing literature that relies on heterogeneous agents models with occupational choice and incomplete financial markets to study sources and implications of firm informality. In this respect, Amaral and Quintin (2006) introduce informality in a model of occupational choice to analyze cross-country differences in labor markets regulation and its interactions with credit markets imperfections. Antunes and Cavalcanti (2007) study the role of regulatory costs and limited enforcement of financial contracts on the size of the informal sector in the U.S. and Peru, whereas Quintin (2008), in a very similar environment, assesses how the size of the informal sector is determined by the degree of enforcement of financing contracts in the formal sector. D'Erasmo and Boedo (2012) also develop a model with informal entrepreneurs but focus on evaluating the costs of firm informality in terms of aggregate TFP. We extend these models to an environment with an endogenous probability of detection and progressivity in personal income taxation. In addition to being more realistic, these assumptions turn out to be crucial to our findings.[7] First, as mentioned earlier, having an endogenous probability of detection is necessary for matching the distribution of firms observed in the informal sector. Second, progressivity in taxation is crucial to generate large shares of informality with perfect capital markets. Absent this effect (i.e flat tax), we get the same conclusion as Quintin (2008): lack of tax enforcement alone (i.e. with perfect credit markets) is not enough to generate a large informal sector.[8] The reason is that progressivity amplifies the incentive of talented entrepreneurs to evade taxation, therefore determining a cutoff value (in terms of managerial ability) that distinguishes between those entrepreneurs that remain informal because of the collateral constraints, and those

[6] More specifically, we compare the effects of a tax flat scheme with those resulting from financial deepening. In the context of models with occupational choice, a flat tax scheme has been analyzed, among the other, by Kitao and Rocio.

[7] Our framework is similar to the ones developed in Dinola and Rocio, even though they do not explicitly include informal firms in the analysis. Moreover, differently from ours, both these papers assess fiscal policy implications of tax evasion in the U.S. economy.

[8] As we shown in Section 4.3, under a flat tax regime and perfect credit markets, the share of the informal sector in Brazil would be around 6% of official GDP, while with progressive taxation we found 23%.

entrepreneurs that operate *parasite* firms. As such, progressivity turns out to be also crucial to the non-linearities described above.

The rest of the paper is organized as follows. Section 2 describes the theoretical background, while Section 3 provides details on the calibration of the model to the Brazilian economy. The results from a financial deepening reform in Brazil are evaluated in Section 4, where we also discuss the predictions of the model on the relationship between firm informality and financial development. Section 4 also contains empirical evidence in favor of the implications of our model, as well as an analysis of the effects of a flat tax policy at different degree of financial frictions. Section 5 concludes. Details on the numerical procedure to solve the model are deferred to the appendix A.1.

2. Model

Building on Erosa (2001), we consider an overlapping generations model with imperfect financial markets, where the government raises progressive taxes on personal income in order to finance public expenditures. Each generation consists of heterogeneous individuals that live for $J < \infty$ periods, each of them endowed with one unit of time until retirement. Retirement occurs at mandatory age $J_R < J$. Heterogeneity takes the form of individual-specific endowments of managerial ability and endogenously idiosyncratic wealth profiles, the latter resulting from individual savings decisions over the life-cycle. During her working life, an individual decides between becoming a worker or being an entrepreneur on the basis of her managerial ability and her financial wealth. Workers supply their time-endowment inelastically and receive a gross wage, while entrepreneurs decide how much to produce by combining labor and capital with a technology that exhibits decreasing returns to scale.

As in Buera and Shin (2013), imperfections in credit markets take the form of collateral requirements on capital rental that are proportional to the individual's savings. Tax enforcement is also imperfect in that informal transactions − i.e. those carried out by unregistered firms − are detectable by the government only after a monitoring process. Entrepreneurs can therefore escape taxation by running their businesses informally. However, tax avoidance comes at a cost. First, access to credit is precluded to informal entrepreneurs, which therefore operate in financial autarky. Second, as in Allingham and Sandmo (1972) and Orsi et al. (2014), in each period, a number of individuals are audited by the government, so that informal entrepreneurs face a probability of being caught evading, in which case they are convicted for tax evasion and forced to pay the taxes due augmented with a penalty surcharge.

2.1. Households

In each period, a new generation of individuals is born. Time is discrete and each agent discounts the future exponentially with a common discount factor $\beta \in (0,1)$. Preferences over consumption of a newly born individual over her life-cycle are represented by the following inter-temporal utility function

$$\sum_{j=0}^{J} \beta^j u(c_j)$$

where c_j denotes consumption at age j, while $u(c_j) = \frac{c_j^{1-\sigma}}{1-\sigma}$ is a standard CRRA instantaneous utility function with relative risk aversion coefficient equal to $\sigma > 0$. At birth

(i.e. age $j = 0$), each individual is endowed with zero assets and idiosyncratic managerial ability $e \in \Theta$ that is drawn from an invariant distribution with cumulative distribution function $\Phi(e)$. An individual's managerial ability (which is randomly assigned at birth) remains unchanged throughout her life-cycle.

2.2. Firms

Following Quadrini (2000), we assume that the consumption good is produced by small business firms (entrepreneurs) and large corporations. Entrepreneurs combine their managerial abilities, e, with capital, k, and labor, n, to produce output via the following technology

$$e^{\eta}\left(k^{\alpha}n^{(1-\alpha)}\right)^{1-\eta}$$

where α, $\eta \in (0,1)$. This production function exhibits decreasing returns to scale, capturing the idea of *span of managerial control* popularized by Lucas (1978).

Workers are hired in a perfectly competitive labor market at the wage rate w, while capital is financed by the entrepreneurs who combine their own assets, a, and borrowed resources obtained from financial intermediaries. We assume that to have access to credit, entrepreneurs need to comply with business regulations − among them registration with tax authorities − which make their production activities observable by the government. Entrepreneurs who choose to comply with business regulations are referred to as formal entrepreneurs. They demand quantities k_f and n_f of (formal) capital and (formal) labor, respectively. Entrepreneurs who choose not to register with tax authorities are referred to as informal entrepreneurs. These agents run their businesses informally by hiring n_i (informal) workers from the labor market, and by self-financing (informal) capital, k_i. Because they are not registered with tax authorities, informal entrepreneurs may escape taxation by concealing their production activities, which would then be detected by the government only after a monitoring process, as in Orsi et al. (2014).

The corporate sector is assumed made of perfectly competitive firms that have access to a common technology represented by the following aggregate production function

$$A^{\eta}\left(K_c^{\alpha}N_c^{(1-\alpha)}\right)^{1-\eta} - \phi_f \tag{1}$$

where $A > 0$, K_c and N_c respectively denote aggregate capital and aggregate labor in the corporate sector, while ϕ_f stands for a fixed operational cost.

2.3. Financial intermediaries

Perfectly competitive financial intermediaries receive deposits from households at risk-free interest rate, r, and rent capital to firms at rental rate r_k. In equilibrium, a zero-profit condition requires that

$$r_k = r + \delta \tag{2}$$

where $\delta \in (0,1)$ is the capital depreciation rate. As in Buera and Shin (2013), we assume that there is limited contract enforceability for small businesses so that the demand for rented capital by a formal entrepreneur who has accumulated wealth a is subject to the following collateral constraint

$$k_f \leq \lambda a \tag{3}$$

where $\lambda \geq 1$ measures the degree of financial imperfections, with $\lambda = \infty$ corresponding to perfect capital markets, and with $\lambda = 1$ denoting financial autarky. The latter case corresponds to the situation of informal entrepreneurs who have no access to credit, i.e.

$$k_i \leq a. \tag{4}$$

As in Quadrini (2000), we assume that large corporations are not subject to collateral constraints, and therefore operate in a perfect capital markets regime.[9] Consequently, given the production function in (1), the aggregate demands of capital and labor in the corporate sector jointly satisfy the following first order conditions for an interior maximum

$$(1-\eta)(1-\alpha)Y_c/L_c = w \tag{5}$$

$$(1-\eta)\alpha Y_c/K_c = r + \delta. \tag{6}$$

2.4 Government

The government raises taxes on personal income to finance wasteful public expenditures. As in Heathcote, Storesletten and Violante (2017), we assume that taxation is progressive, with tax revenues collected on personal income y described by the following tax function

$$T(y) = y - \lambda_y y^{1-\tau}$$

where $\tau \in (0,1)$ measures the degree of progressivity of the tax system, while $\lambda_y \in (0,1)$ determines the average level of taxation in the economy.

To discourage tax evasion, the government periodically conducts audits. Following an audit, entrepreneurs found conceiling production are convicted of tax evasion and forced to pay the taxes due augmented by a penalty surcharged factor $s > 1$. We assume that the probability of being audited in a given period, namely $p(k)$, is an increasing function of the amount of capital rented by the firm in the given period. This assumption is common in the tax evasion literature and can be rationalized by the fact that large establishments are more visible to tax authorities and therefore find it harder to conceal production than smaller ones (see e.g. Ordonez, 2014).

2.5. Agent's decisions problems

2.5.1. Working age population

During working life (i.e. for $j = 0,1,\ldots,J_r - 1$), at the beginning of each period, knowing her managerial ability, e, and given the amount of assets a accumulated in the previous period, an individual decides whether to be a worker or an entrepreneur. Thereafter, workers choose how much to consume, c, and save, a'. Entrepreneurs instead decide first whether to comply with business regulations, and on the basis of this choice, how much to produce by hiring workers and renting capital (taking into account collateral constraints). After production decisions have taken place, random audits are conducted and fines enforced by the

[9] This assumption captures the idea that it is very unlikely that large corporations would face similar financial imperfections as small business.

government. Informal entrepreneurs can then observe whether or not they have been caught concealing production to evade taxation. On the basis of this, they decide how much to consume and save.

The decision problem of an individual with state variables e and a can be written in a recursive formulation, with the beginning-of-period value function given as follows

$$V(a,e) = max\{V^W(a,e), V_f^E(a,e), V_i^E(a,e)\}$$

The function $V^W(a,e)$ denotes the value function for the agent who chooses to be a worker in the current period, i.e.

$$V^W(a,e) = \max_{c,a'}\{u(c) + \beta V(a',e)\}$$

subject to

$$c + a' = y^w + a - T(y^w)$$

$$a' \geq 0 \tag{7}$$

where $y^w = w + ra$ denotes the worker's personal income, while constraint (7) captures the fact that workers cannot borrow. Functions $V_f^E(a,e)$ and $V_i^E(a,e)$ respectively stand for the value of being a formal entrepreneur and the value of being an informal entrepreneur. The decision problem of a formal entrepreneur takes the following form

$$V_f^E(a,e) = \max_{k_f,n_f,c,a'}\{u(c) + \beta V(a',e)\}$$

subject to (3), (7) and

$$y^E = e^\eta\left(k_f^\alpha n_f^{(1-\alpha)}\right)^{1-\eta} - wn_f - (r+\delta)k_f + ra$$

$$c + a' = y^E + a - T(y^E)$$

In the above equation, y^E denotes the formal entrepreneur's declared income, which amounts to her actual earnings.

By contrast, informal entrepreneurs attempt to escape taxation by concealing their production activities and reporting only their capital incomes. However, in any period, an informal entrepreneur who has rented capital k_i faces a probability $p(k_i)$ of being audited and caught by the government evading taxation. So, let $V_d^E(a,e)$ and $V_{nd}^E(a,e)$ denote the informal entrepreneur's value functions corresponding to the cases of detection and non-detection, respectively. The value of being an informal entrepreneur can then be written as follows

$$V_i^E(a,e) = \max_{k_i,n_i}\{p(k_i)V_d^E(a,e) + (1-p(k_i))V_{nd}^E(a,e)\}$$

subject to equation (4), which describes the collateral constraint in financial autarky. The value function in the case of non-detention is given by

$$V_{nd}^E(a,e) = \max_{c,a'}\{u(c) + \beta V(a',e)\}$$

subject to equation (7) and

$$y^E = ra$$

$$c + a' = y^E + \pi + a - T(y^E)$$

where π represents profits from business activities, i.e.

$$\pi = e^\eta\left(k_i^\alpha n_i^{(1-\alpha)}\right)^{1-\eta} - wn_i - (r+\delta)k_i. \tag{8}$$

Accordingly, concealing production allows the informal entrepreneur to hide profit income π from tax authorities. However, in the case of detection, the government would force the informal entrepreneur to pay the taxes due on the unreported income (i.e. $T(y^E + \pi) - T(y^E)$) scaled up by a penalty surcharge factor s. Consequently, the value function of an informal entrepreneur that has been caught by the government is given by

$$V_d^E(a, e) = \max_{c,a'}\{u(c) + \beta V(a', e)\}$$

subject to equation (7) and

$$y^E = ra$$

$$c + a' = y^E + \pi + a - (1 + s)[T(y^E + \pi) - T(y^E)]$$

where π is defined as in (8).

2.5.2. Retired agents

During retirement (i.e. for $j = J_R, J_R + 1, \ldots, J$), an individual consumes and saves on the basis of the financial wealth accumulated during her working life. Hence, the value function of a retired individual is given as follows

$$V(a, e) = \max_{c,a'}\{u(c) + \beta V(a', e)\}$$

subject to equation (7) and

$$c + a' = y^R + a - T(y^R)$$

where $y^R = ra$ is the retired individual's declared income.

2.6. Equilibrium

For each agent in the economy, let define with $\omega = \{e, a, b(e, a)\}$ the vector containing the individual state variables, e and a, and the occupational choice $b(e, a)$ (i.e. retired, workers, formal entrepreneurs and informal entrepreneurs (detected and undetected)). A stationary equilibrium is given by a price vector $\{r, w\}$, allocations $\{c(\omega), a(\omega)\}$, occupational choices $b(e, a)$, formal and informal workers $\{n_f(\omega), n_i(\omega)\}$, investment in formal and informal capital $\{k_f(\omega), k_i(\omega)\}$, labor and capital in the corporate sector $\{L_c, K_c\}$, a tax function $T(\cdot)$, a penalty surcharge factor s, a probability of being caught evading $p(\cdot)$, and a distribution of individuals over ω, $\xi(\omega)$, such that given the free-risk interest rate r, the wage rate w and the tax system (i.e. s, $p(\cdot)$ and $T(\cdot)$):

• The policy functions $\{c(\omega), a(\omega), k_f(\omega), k_i(\omega), n_f(\omega), n_i(\omega)$,b(e,a)$\}$ solve the agents' decisions problems described in section 2.5.

• Labor and capital in the corporate sector, $\{L_c, K_c\}$, solve optimality conditions (5) and (6).

• Capital and labor markets clear:

$$\int \big(k_f(\omega) + k_i(\omega)\big)d\xi(\omega) + K_c = \int a(\omega)d\xi(\omega)$$

$$\int \big(n_f(\omega) + n_i(\omega)\big)d\xi(\omega) + L_c = \int 1_W d\xi(\omega)$$

where 1_W is an indicator function taking value 1 if the agent is a worker and 0 otherwise.

• The government's budget constraint is balanced, i.e.

$$G = \int [T(y(\omega)) + 1_D s(T(y(\omega) + \pi(\omega)) - T(y(\omega))] d\xi(\omega)$$

where G are public expenditures, $y(\omega)$ is the agent's declared income and 1_D stands for an indicator function that takes value 1 if the individual is an informal entrepreneur that has been audited and 0 otherwise.

- Financial intermediaries earn zero profit, i.e. equation (2) is satisfied.
- The distribution $\xi(\omega)$ is the invariant distribution for the economy.

3. Calibration

We calibrate the model to the Brazilian economy. We begin with specifying the stochastic process for entrepreneurial ability and the probability of being detected while producing informally. As for the former, we follow Allub and Erosa (2018) by assuming that the managerial ability is drawn from a generalized Pareto distribution with CDF

$$\Phi(e) = 1 - \left(1 + \frac{\nu(e-\mu)}{\kappa}\right)^{-1/\nu}$$

where parameters μ, κ and ν determine location, scale and shape of the distribution, respectively. The Pareto distribution is discretized and the first grid point, namely e_{min}, is set to the one with probability mass $\Phi(e_{min})$.

As in Di Nola et al. (2018), the probability of being audited is parametrized by assuming that the detection process follows a logistic distribution with parameters p_1 and p_2, i.e.

$$p(k) = \frac{1}{1+p_1 exp(-p_2 k)}$$

The main advantage of this specification is that it gives a probability of detection which increases quickly as capital becomes larger − with a strength that depends on p_1 and p_2 − so that the benefits from tax evasion decline substantially with the firm's size. This property enables the model to capture the empirical observation that informal firms are typically small scale enterprises (see e.g. Ulyssea (2018)).

We partition the vector of mode's structural parameters in two sub-vectors, and we calibrate them in two different steps. In the first one, we assign values to $\{\beta, \sigma, \delta, \alpha, \phi_f\ \lambda_y\}$, which are calibrated independently, mainly by picking values from external sources. Specifically, in our model, a period of time corresponds to a year, and therefore we set the capital depreciation rate, δ, to 0.05. In addition, the subjective discount factor β is fixed to 0.95, while the relative risk aversion coefficient, σ, is assumed to be equal to 1.5. These numbers are in the range of values typically used in the literature (see e.g. Allub and Erosa (2018) and Ordonez (2014)). The share of labor in production α is set equal to 0.40 as in Sin and Gaglianone (2006), while the operational fixed cost ϕ_f is determined such that in equilibrium, corporate firms earn zero profits. Finally, parameter λ_y in the taxation function is set to 0.70, corresponding to a statutory income tax − without progressivity − of 0.30 as in Ulyssea (2018).

The remaining parameters $\{\lambda, \eta, \mu, \kappa, \nu, \Phi(e_{min}), \tau, s, p_1, p_2, A\}$ are calibrated to match salient features of the Brazilian economy. In particular, we use 11 moments from the data as calibration targets:

- The share of capital in total market output;

- The credit-to-GDP ratio;
- The size of informal economy relative to total formal output;
- The ratio of informal workers to formal workers;
- The share of informal firms in the total number of firms;
- The share of formal firms with up to 5, 5 to 10, 11 to 20, and 21 to 50

Table 1: Calibration Results: Parameter Values

Parameters	Description	Value
β	Subjective discount factor discount factor	0.951
σ	Relative risk aversion coefficient	1.500
δ	Capital depreciation rate	0.050
α	Production function parameter	0.400
ϕ_f	Operational fixed cost	0.506
λ	Tightness of financial frictions	1.367
η	Span of control	0.217
μ	Location Pareto distribution	2.354
κ	Scale in Pareto distribution	0.410
ν	Shape in Pareto distribution	0.989
$\Phi(e_{min})$	Probability mass (Pareto)	0.719
τ	Degree in taxes progressivity	0.144
λ_y	Measure of the average taxation	0.700
s	Surcharge factor	0.257
p_1	Probability of detection parameter	1664
p_2	Probability of detection parameter	1.790
A	TFP in the corporate sector	2.176

employees; and the share of informal firms with less than 2 and less than 5 workers.

The share of capital in total market output and the credit-to-GDP ratio are respectively set to 2.10 and 0.42 as in Allub and Erosa (2018). The target for the size of informal production (relative to formal production) is taken form Medina and Schneider (2018) who estimate an average size of the Brazilian informal economy over the 1991-2015 period of 0.376. The ratio of informal to formal workers and the share of informal firms are set to 0.354 and 0.698, according to the estimates reported in Ulyssea (2018). This paper also provides information of the size distribution (with respect to the number of employees) for both formal and informal firms in Brazil, which we take as targets.

Parameter values are jointly pinned down by minimizing a loss function that computes the distance between the targeted moments from the data and their counterparts in the model.

4. Results

In this section, we first characterize the properties of the general equilibrium in the calibrated model (the benchmark economy) with a specific focus on occupational choices and informality, and then we discuss the implications of removing financial frictions in Brazil The section concludes with a general discussion on the potential gains from financial markets

reforms in developing countries, stressing in particular its effects on the size of the informal economy and tax evasion.

4.1. The Benchmark Equilibrium

Table 2: Calibration Results: Targeted Moments

Moments	Source	Data	Model
Capital-output ratio	Allub and Erosa (2018)	2.100	2.178
Credit-to-GDP ratio	World Bank Database	0.420	0.420
Ratio of informal to formal output	Medina and Schneider (2018)	0.376	0.374
Ratio of informal to formal workers	Ulyssea (2018)	0.354	0.386
Share of informal firms	Ulyssea (2018)	0.698	0.745
Size distribution: informal firms			
$\leq$ 2 workers	Ulyssea (2018)	0.957	0.976
$\leq$ 5 workers	Ulyssea (2018)	0.998	1.000
Size distribution: formal firms			
$\leq$ 5 workers	Ulyssea (2018)	0.701	0.757
$\leq$ 6-10 workers	Ulyssea (2018)	0.141	0.139
$\leq$ 11-20 workers	Ulyssea (2018)	0.083	0.103
$\leq$ 21-50 workers	Ulyssea (2018)	0.048	0.001

In Figure 1, we show how savings and capital decisions (first row), and occupational choices (second row) evolve along the life cycle for individuals with the lowest managerial ability level (first column), middle managerial ability level (second column) and the highest managerial ability level (third row).[10] Age is a key dimension for occupational decisions because it determines wealth accumulation and therefore borrowing limits that are characteristic of our credit-constrained economy. It turns out that two thresholds on the support of the ability distribution are also useful for characterizing the occupational choice decisions. These two points mark three zones (a lower range ability zone, a middle range ability zone, and an upper range ability zone) with unique features when it comes to deciding between being a worker, an informal entrepreneur, or a formal entrepreneur.

In this respect, notice that, at birth, all agents are endowed with zero assets which implies that because of the collateral requirements on capital rental (equation (3)), the only option agents have is to become workers in the first period of their life. The picture illustrates that individuals in the lower range of the ability distribution will remain workers during they

[10] The equilibrium is computed by discretizing the Pareto distribution to pin down 40 different possible managerial abilities. Figure 1 is therefore intended to depict the optimal decisions of few types of individuals (3 out of 40) that are informative enough to characterize the properties of the general equilibrium. Notice furthermore that to compute the stationary distribution of the economy, $\xi(\omega)$, we carry out several simulations of the model with different realizations of a random variable describing audits by the government. In Figure 1, we plot policy functions for a particular simulation.

entire working life. These agents have the unique feature that their ability level is too low to make the operation of a business viable, even at a scale sufficiently low that tax evasion could be afforded with a virtually negligible probability of being detected. These individuals therefore save for retirement only, and never engage in entrepreneurship, even informally.

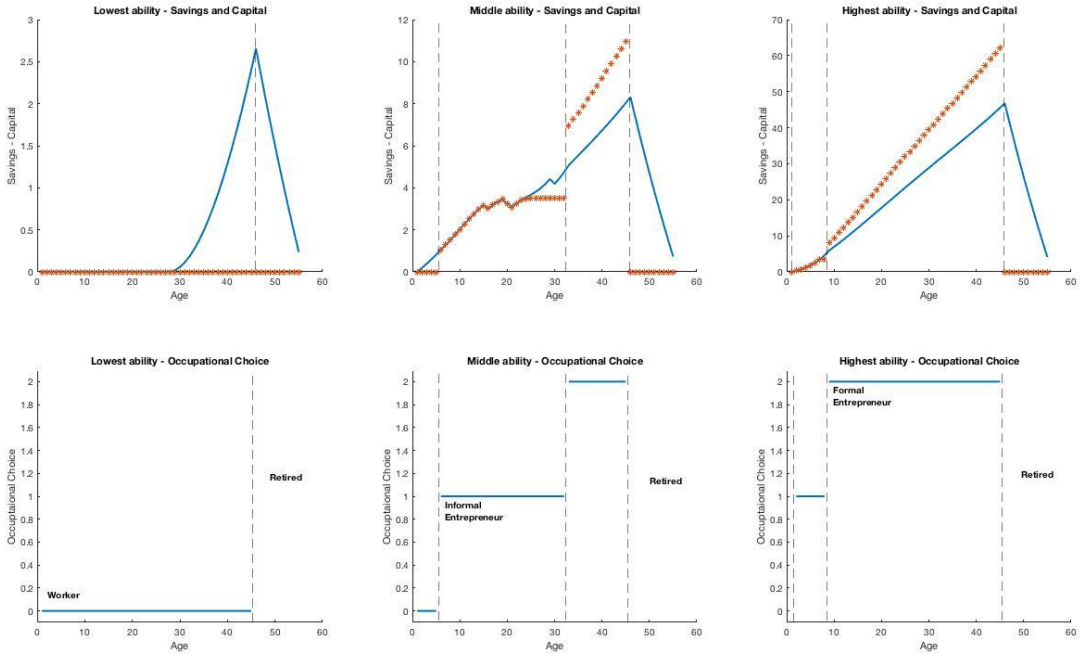

Figure 1: **Inspecting the mechanism.** Note: the picture depicts optimal decisions for agents with the lowest (first column), the middle (second column) and the highest (third column) abilities. The first row depicts the evolution across ages for capital (red discontinuous line) and savings decisions (blue continuous lines). Occupational choices are reported in the second row, where 0 identify workers, 1 informal entrepreneurs and 2 formal entrepreneurs. All parameters are set to their calibrated values.

By contrast, individuals in the middle and in the upper range of the ability distribution switch occupation as soon as they have accumulated enough assets to self-finance informal businesses. These agents have managerial abilities that are sufficiently high to make entrepreneurship a viable option but – because of the collateral constraints – they have accumulated not enough assets to make operating a formal business immediately profitable.[11] Hence, as an alternative to remaining workers, they choose initially to become informal entrepreneurs in order to take advantage of the additional resources coming from tax evasion. In this respect, Figure 1 also illustrates that individuals with a higher managerial ability save more during their working life since they have, in general, more incentives to become entrepreneurs, and hence, in comparison with individuals with middle-range abilities, start an informal business at a younger age.

An important feature highlighted in Figure 1 is that informal firms are small scale enterprises relative to formal firms. This result is consistent with the available empirical

[11] Intuitively, by producing in the formal economy, these individuals would be inefficiently constrained by the borrowing limit and, thus, they would have to produce at a very small scale compared to what would be their optimal scale under perfect credit markets.

evidence for Brazil (see e.g. D' Erasmo (2013) and Ulyssea (2018)) and can be explained intuitively by the combined effect of two main forces. First of all, informal enterprises operate in a financial autarky regime and thus informal entrepreneurs can only use their own savings to finance productive capital. The maximum production scale is then bounded by the total amount of wealth accumulated by the informal entrepreneur. Second, the probability of being audited by the government, $p(k)$, is a increasing function of the amount of capital used in production and thus informal firms have strong incentives to produce at a relatively low scale. The role of these two complementary forces is nicely apparent in the pattern of informal capital for individuals with middle managerial abilities. As Figure 1 illustrates, during the period of time in which these individuals are informal entrepreneurs, capital moves initially in step with savings as a result of the financial autarky constraint (i.e. equation (4)). This pattern continues up to when the capital stock reaches a critical level. After this point, while saving continues to grow, capital stays instead constant as the result of entrepreneurs' desire to keep their firms at relatively small scales to prevent detection.[12]

Another key property of the framework with endogenous occupational choices is that financially constrained entrepreneurs – i.e. those for whom constraints (3) and (4) bind – face a strong incentive to save in order to expand the scale of production through capital accummulation (see e.g. Quadrini (2000)). The reason is that the marginal productivity of capital, for financially constrained firms is higher than the market interest rate,[13] implying that self-financing their own firms is more profitable for these entrepreneurs than lending resources to financial intermediaries. In our framework this effect is amplified in the case of informal entrepreneurs that are not audited by the government. These entrepreneurs, because of tax evasion, are able to accumulate wealth faster as their incomes grow at a higher rate.

As a final remark, notice that once informal entrepreneurs have accumulated enough assets, they find it optimal to comply with business regulations and become formal entrepreneurs. This pattern is common to all entrepreneurs in the benchmark economy, even though the switch from one sector to another occurs at different ages, depending on managerial ability.[14] In sum, as one could expect, agents with ability above a certain threshold would choose to produce informally until they would have accumulated enough assets to relax the strength of existing financial constraints and start producing in the formal sector.

4.2. The effects of a financial markets reform in Brazil

We now assess the impact of removing financial frictions in the Brazilian economy. This assessment is done by comparing the properties of the benchmark model with those of an identical economy with perfect capital markets.[15] Results are reported in Table 3, which provides several key statistics for both the benchmark and the perfect credit markets economy, and in Figure 2 that shows how individual decisions are affected by financial deepening in Brazil.

---

[12] Entrepreneurs in the informal economy face a trade-off in choosing the optimal amount of capital in that if they increase their amount of informal capital by too much, the probability of being detected would increase. Figure 1 shows that this property prevents informal entrepreneurs from being financially constrained since they find it optimal to operate with an informal capital smaller than their total savings.

[13] This is true as the amount of capital employed in production is below the optimal level with perfect capital market.

[14] This is illustrated in Figure 1, which shows that informal entrepreneurs with higher ability accumulate savings at a higher rate, thereby moving to the formal sector earlier in their life-cycle.

[15] The equilibrium for the economy with perfect capital markets is obtained from the benchmark case by setting $\lambda = \infty$ while keeping all of the other parameters at their calibrated values.

4.2.1. The informal economy

Table 3: Experiment Results. Removing financial frictions

	Benchmark	**No financial frictions**
λ	1.37	∞
Size informal economy	37.43%	22.97%
Ghost firms (% Total firms)	74.59%	61.77%
ΔInformal production		-22.12%
ΔFormal production (Official GDP)		26.91%
ΔTotal production		12.72%
ΔCapital		18.65%
ΔTFP (Formal production)		3.68%
ΔTFP (Total production)		2.20%
ΔWage		8.52%
Interest Rate	0.28%	9.61%
Total workers (% Population)	74.98%	80.04%
Total informal entrepreneurs (% Population)	15.13%	12.32%
Total formal entrepreneurs (% Population)	9.89%	7.64%
ΔFiscal revenues		15.44%
ΔTax evasion		-17.30%
Workers (% Total Wealth)	14.42%	63.51%
Informal entrepreneurs (% Total Wealth)	25.46%	25.98%
Formal entrepreneurs (% Total Wealth)	60.12%	10.51%
Constrained formal entrepreneurs	100%	0%
Constrained informal entrepreneurs	59.00%	14.40%

Note: Δ stands for percentage deviations from the benchmark values. Statistics from the no financial frictions case are obtained from the benchmark model by setting $\lambda = \infty$ while holding all the remaining parameters to their calibrated values.

As Table 3 illustrates, relaxing collateral constraints has a substantial impact on all informal sector aggregates. We find in particular that – relative to the benchmark economy – both the percentage of informal (ghost firms) to total firms and the share of informal entrepreneurs decline in the economy with perfect capital markets, leading to an overall decrease of 22.12% in aggregate informal production. By contrast, although the share of formal entrepreneurs declines by about 2 percentage points, we find that aggregate formal production increases by 26.91%. The size of the informal economy decreases by about 15 percentage points, as result of lifting financial constraints. Hence, our model predicts that removing financial frictions would have a huge impact on the size of the informal economy in Brazil.[16] This result is particularly interesting as it shows that around 40% of the actual size of the informal economy in Brazil is explainable by credit markets imperfections.

[16] We measure the size of the informal economy as the ratio of informal production to total formal production, where the latter is computed as the sum of total production by formal entrepreneurs and total production in the corporate sector.

To grasp an intuition for the above results, it is worth pointing out that relaxing the collateral constraint (3) has an asymmetric impact on formal and informal entrepreneurs, in that the removal of financial constraints grants formal entrepreneurs access to all the credit they need to run their businesses, while informal entrepreneurs continue to be in financial autarky. This asymmetry strengthens the incentive to produce in the formal sector, particularly

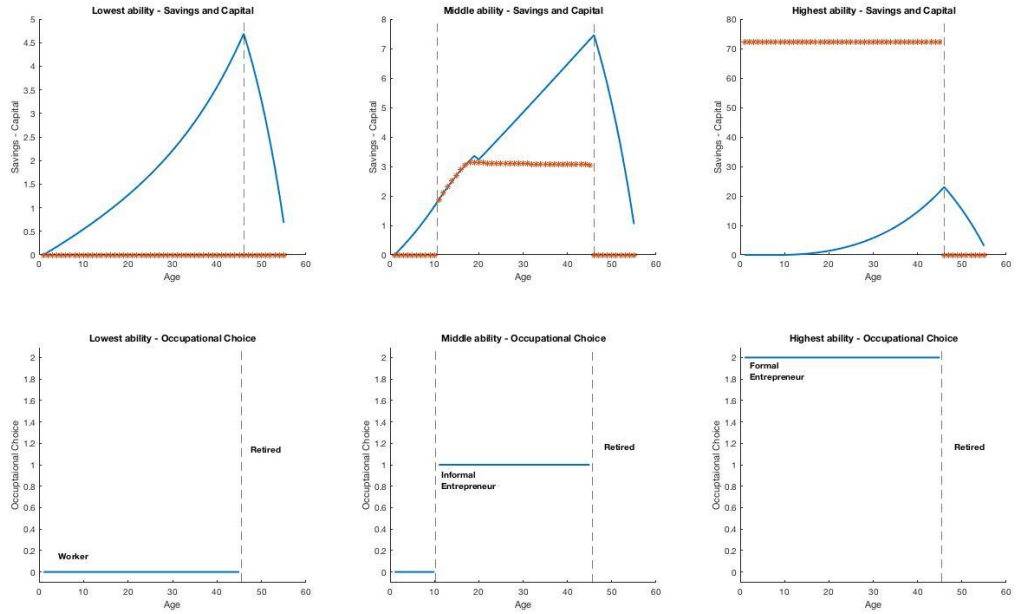

Figure 2: Perfect credit markets and individual decisions. Note: the picture depicts optimal decisions for agents with the lowest (first column), the middle (second column) and the highest (third column) abilities. The first row depicts the evolution of capital (red discontinuous line) and savings (blue continuous lines) across the lifespan. Occupational choices are instead reported in the second row, where 0 identifies workers, 1 informal entrepreneurs and 2 the formal entrepreneurs. Perfect credit markets are obtained by setting $\lambda = \infty$ while holding all the remaining parameters at their calibrated values.

For those agents that are the most productive (the most able agents). To see this, notice that in the benchmark economy with financial frictions, high ability agents without significant financial wealth chose between operating a small business in the formal sector (because of the existence of collateral constraints) and operating a small business in the informal sector in which they could profitably escape taxation (because of financial autarky and because the probability of being audited is a function of the amount of capital rented). Thus, if the probability of detection and the fine are small enough, the gains from tax evasion would easily overdo the benefit of having access to credit, making it likely that these entrepreneurs would decide to be informal producers. By contrast, under perfect credit markets, highly productive entrepreneurs now choose between operating a small informal business, and operating a relatively large formal business. In such circumstances, gains from the potential high scale of formal production can be large enough to induce agents to comply with business regulations

while starting up their own enterprises. This feature is well apparent in Figure 2, which shows that with perfect credit markets, highly productive entrepreneurs open their businesses directly in the formal sector. These entrepreneurs start operating directly at a high scale as they no longer need to accumulate wealth to overcome financial constraints. In the aggregate, this effect induces a redistribution of resources from informal producers to formal producers (redistribution of wealth toward the most productive entrepreneurs), which eventually ends up increasing total formal production and decreasing informal production.

In addition to the above resource reallocation mechanism that involves the most productive individuals, removing financial frictions also dampens the gains from informal entrepreneurship for relative low-skilled agents. The reason is the increase in the wage rate (+8.52%) driven by the larger demand of formal labor (see Table 3). Because of this effect, the opportunity cost of entrepreneurship increases, thereby making it less attractive for workers with relatively low managerial ability to switch occupation and become informal entrepreneurs. This feature is illustrated in Figure 2, which shows that the amount of time in which middle-skilled individuals remain workers in the frictionless economy is twice as large as the amount of time required in the benchmark economy (see Figure 1).

To summarize, there are two main forces behind the decrease in the size of the informal economy after removing financial frictions: first, the increase in the potential scale of formal business operation makes it more attractive to operate in the formal sector to produce at an optimal scale; and second, the increase in wages makes it more profitable for many potential informal entrepreneurs to remain workers. The combined effects of these two mechanisms reduce the share of individuals that optimally decide to become informal entrepreneurs, declining from a high of 15.13% of total working age population in the benchmark economy to a low of 12.72% in the model without financial frictions (see Table 3). As a result, the percentage of individuals under-reporting income to fiscal authorities also declines after relaxing the collateral constraints, inducing a decrease in tax evasion of 17.30%. Table 3 shows that this last effect – jointly with the expansion in formal production – results in quantitatively important gains for the government, with total fiscal revenues that increases by 15.44%.

Overall, our results show that removing all financial frictions is not enough to completely eliminate firms informality. As a matter of fact, we find that the percentage of informal firms is still high in the frictionless economy (61.77%), with an aggregate amount of production that accounts for about 23% of the total market output. These firms are run by low-skilled agents that are producing at a relatively low scale, and find it optimal to become informal entrepreneurs – rather than workers – benefiting from tax avoidance. In this respect, we can think about the case with perfect credit markets as a lower bound to the size of the informal economy, which, being driven by income taxation, can be eliminated only through specific fiscal policy interventions. We will turn to this point in Section 4.3.

4.2.2. Official GDP and aggregate TFP: the amplification effect

One interesting result of the above experiment is that removing financial frictions has a huge impact on total official GDP (+26.91%), which we refer to from now on as aggregate formal production. A related finding is that TFP in the formal sector also increases with the financial markets reform (+3.68%) as shown in Table 3.

There are two main mechanisms behind these results: first, production by entrepreneurs with relatively high ability rises in the presence of perfect capital markets;[17] and second, resources for informal production – i.e. entrepreneurial ability, labor, and capital – are

Table 4: Experiment Results. No Informal Production. Removing financial frictions

	Collateral constraints	**No Financial Frictions**
λ	1.37	∞
ΔGDP		14.47%
ΔCapital		23.52%
ΔWage		7.87%
ΔTFP		1.00%
Total workers (% Population)	79.51%	85.47%
Total entrepreneurs (% Pop.)	20.49%	14.53%
Fiscal Pressure (% Total Income)	33.90%	32.89%
Workers (% Total Wealth)	15.66%	78.67%
Entrepreneurs (% Total Wealth)	84.34%	21.33%
Constrained Entrepreneurs	100%	0%

Note: Δ stands for percentage deviations from the equilibrium with collateral constraints and no informal sector.

partially reallocated to the formal sector when collateral constraints are relaxed.[18] The first channel is a conventional mechanism that is well known in the literature of occupational choice and financial market imperfections (see e.g. Buera et al. (2011) and Buera and Shin (2013)). The second is instead driven by the assumption of imperfect tax enforcement, and represents an unconventional channel that amplifies the effects of the first mechanism.

One way to assess whether this amplification effect is quantitatively important is to compare our benchmark results (i.e. the statistics reported in Table 3), with those obtained in a counterfactual model economy with perfect tax enforceability where, by definition, the informality channel is shut down.[19] Simulated results based on the counterfactual model are reported in Table 4, which provides statistics for the cases with and without the collateral constraint. As the table illustrates, the removal of financial frictions in the counterfactual economy increases GDP by 18.71% and aggregate TFP by 1%. In comparison with the benchmark results, we can then conclude that informality substantially amplifies the effect of a financial markets reform as we find that (i) the increase in official GDP in the benchmark model is almost twice as large as in its counterfactual counterpart, and (ii) TFP gains from removing financial frictions are more than three times higher when informality is accounted for in the model.

[17] As discussed in Section 4.2.1, when collateral constraints are removed, high ability entrepreneurs do not need to accumulate wealth in order to open their own, high scale and more productive firms. Since under the presence of financial constraints richer but less skilled entrepreneurs are running businesses in the economy, this effect ends up increasing not only official GDP but also aggregate TFP.

[18] We show in the previous sections that the incentive to evade taxation are in fact crucially related to the strength of collateral constraints, in that without financial frictions, high-skilled informal entrepreneurs would directly open their businesses in the formal sector.

[19] The counterfactual equilibrium is computed from the benchmark model by setting the penalty surcharge factor $s = \infty$ while keeping all of the remaining parameters at their calibrated values.

4.2.3. Wealth distribution

As it is well known, the removal of financial frictions has important effects in terms of wealth redistribution in model with endogenous occupational choices (see e.g. Quadrini (2000)). Table 3 shows that this property holds true in our framework, where financial wealth held by workers moves from a low of 14.42% of total wealth in the benchmark economy to a high of 63.51% in the model with perfect capital markets. The table further illustrates that this

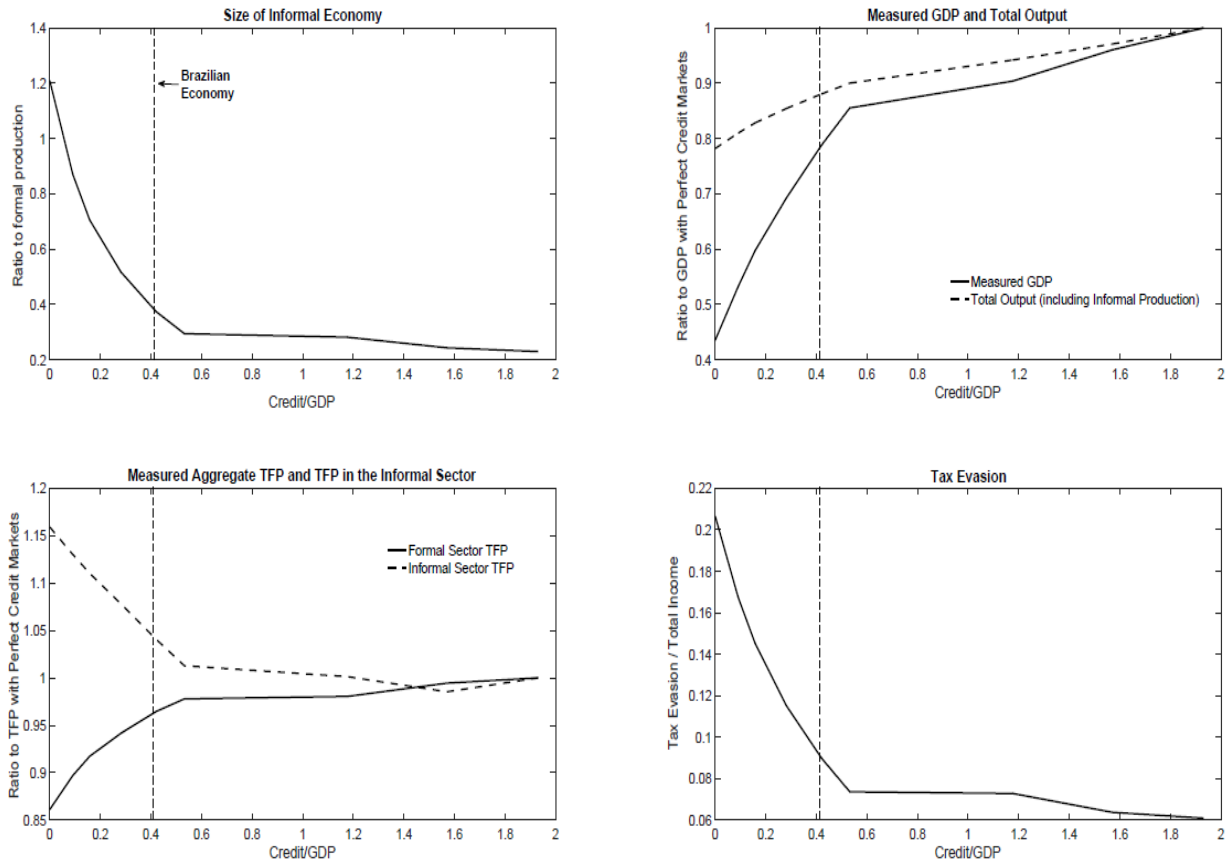

Figure 3: Development and financial frictions in the long-run

effect is entirely driven by a reallocation of resources from formal entrepreneurs to workers, given that the share of wealth held by informal producers is virtually unaffected by relaxing the collateral constraints. Because of this feature, we find that entrepreneurs hold 36.49% of total wealth with perfect capital markets, a value which is about 15 percentage points higher than in its counterpart in the model with complete tax enforcement (see Table 4). Hence, our model predicts that firms informality actually dampens the overall wealth redistribution effect induced by financial markets reforms.

The reason for this finding lies in the mechanism that strengthens – through the marginal productivity of capital – the incentive of financially constrained entrepreneurs to accumulate wealth. While switched off for formal entrepreneurs, this mechanism still hold trues for informal entrepreneurs in the presence of perfect capital markets. The reason is

because informl entrepreneurs have no access to credit regardless of the degree of financial markets imperfections. In this respect, Table 3 shows that 14.40% of total informal entrepreneurs continue to be financially constrained once we remove the collateral requirements on capital rental in the formal sector. Compared with the unconstrained entrepreneurs, these individuals accumulate financial wealth at a higher rate to self-finance capital, a mechanism that is further amplified by the increase in the interest rate (see Table 3).

4.3. Development and financial frictions: a discussion

A large strand of the literature has highlighted the important role of financial markets in economic development (see e.g. Buera et al. (2011), Buera and Shin (2013) and Allub and Erosa (2018)). In this section, we analyze this linkage through the lens of our model, focusing particularly on the role of informality. To this end, Figure 3 illustrates how the size of the informal economy, official and total GDP, TFP in the two sectors, and tax evasion respond to changes in the degree of financial frictions. The reported results are computed for the steady state equilibrium of the model, and for different values of the parameter controlling the tightness of collateral constraints, λ, while keeping all other parameters fixed at their calibrated values. Since λ does not have a direct empirical counterpart, the results are presented in a way that lends itself to data comparisons. Specifically, we plot steady state statistics against the implied credit-to-GDP ratio, which is monotonically increasing in λ and observable in the data. Furthermore, figures for GDP and TFP are normalized by the perfect-credit-markets level and can be read as a measure of costs (or gains) induced by financial frictions. For the sake of comparison, we also report GDP in the economy with perfect tax enforceability.

Two main results are worth emphasizing. First of all, as Buera and Shin (2013), we find that financial frictions may entail quantitatively important costs in terms of both GDP and TFP losses. In our model, worsening the tightness of collateral constraints to reach financial autarky brings official GDP and formal TFP down to respectively 43.52% and 86.11% of their counterpart levels with perfect credit markets. Comparison with the economy with perfect tax enforceability confirms that such costs are indeed amplified by firms informality. Second, consistently with data, the model predicts a negative correlation between financial frictions (as inversely related to the credit to GDP ratio) and informality. As parameter λ rises from 0 (pure self-financing economy) to ∞ (perfect credit markets), the size of the informal economy monotonically declines, moving from a maximum of 120% of official GDP to a minimum of about 23%. This pattern is mirrored by the response of tax evasion, which also decreases monotonically when we relax the collateral constraints, declining from 20% of total output in the pure self-financing economy to around 6% with perfect capital markets.

The above results complement the findings of Antunes and Cavalcanti (2007) and Quintin (2008) who also report a negative correlation between firms informality and financial development in models with imperfect contract enforceability. In terms of policy implications, however, our findings extend the existing literature by showing that the role of lowering credit markets frictions in reducing informality and tax evasion is only relevant (and is also extremely important in quantitative terms) at high levels of credit market imperfections. As Figure 3 illustrates, there is a non-linear relationship between financial development and informality: easing the collateral constraints has a significant impact on the size of the informal economy for countries with credit to GDP ratio below 60%, while no quantitatively important effect is found above this threshold. As a consequence of this feature, we also find a non-linear effect

of financial frictions on tax evasion, official GDP and (to a lesser extent) TFP, with the implication that poor countries may have sizeable gains from financial deepening both in terms of growth in official GDP and in fiscal revenues (because of the significant impact on tax evasion).

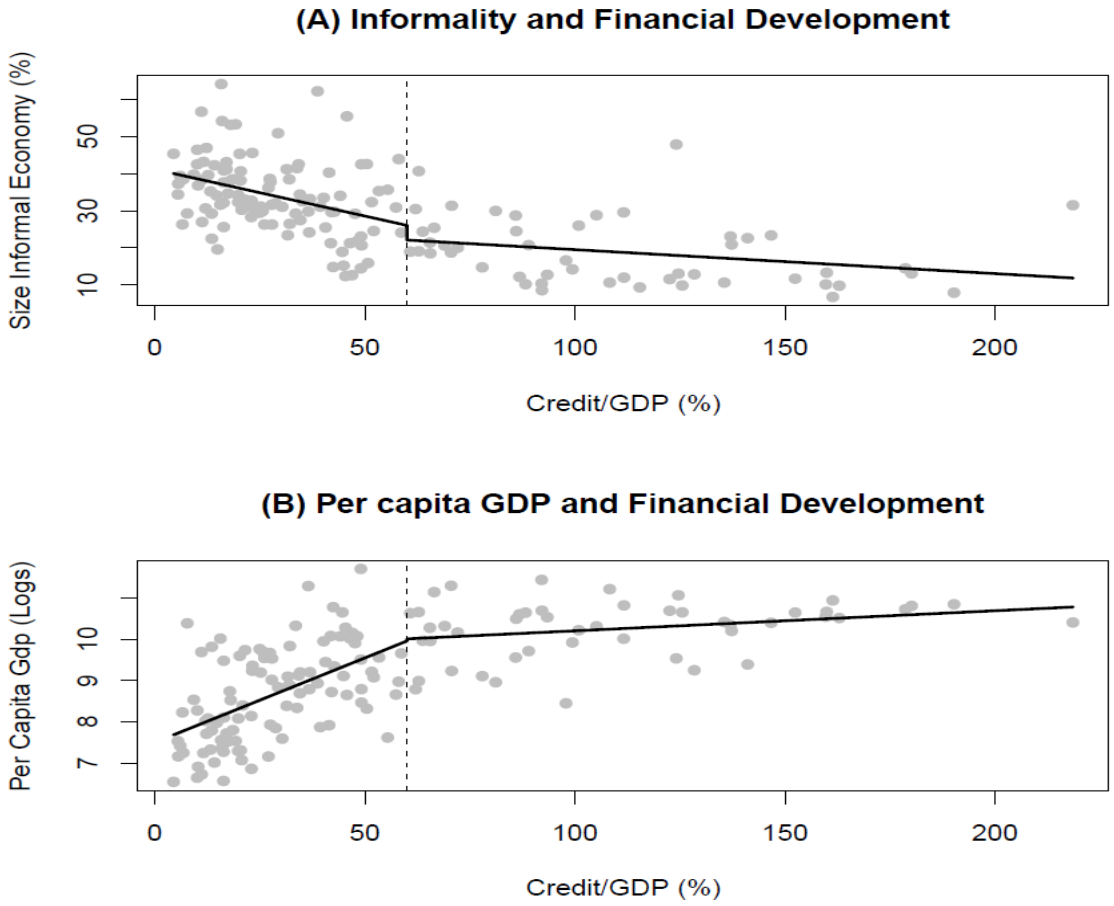

Figure 4: Testing for the non-linear effect of fiscal market frictions on informality (panel A) and GDP per capita (panel B). The vertical dashed line refers to the threshold level for the credit to GDP ratio of 60.8%.

To understand the mechanism behind the non-linear effect, notice that in our model there are three different types of informal firms. First of all, we have businesses that are run by highly productive but poor entrepreneurs, who choose informality because of the collateral constraints. As shown in Figure 2, individuals with the highest level of managerial ability would never produce informally if credit market were perfect.[20] A second group involves the so-called *parasite* firms, which are productive enough to produce formally but decide to remain informal to take advantage of the higher earnings induced by tax evasion. Finally, we have firms that are run by low-skilled agents that are too unproductive to ever become formal

[20] This type of firms can be somehow related to the so-called De Soto's view, which refers to the informal sector as an untapped reservoir of entrepreneurial energy, held back by government regulations. The interested reader is referred to Ulyssea (2018) and the references therein for further details on this point.

entrepreneurs.[21] With these different types of firms, relaxing the collateral constraints produces two countervailing effects on informal production. On the one hand, less financial frictions bring firms out of informality because (i) highly productive entrepreneurs have larger access to credit so they can produce formally at a profitable higher scale; and (ii) given that the wage rate increases, low-skilled individuals exit informal entrepreneurship to become workers. On the other hand, the induced higher production costs also increase the relative gains from tax evasion, thereby bringing more parasite firms to the informal sector.[22] As Figure 3 illustrates, the first effect dominates the second one at high level of financial frictions where an increase in credit availability boosts the gains from formality for talented entrepreneurs. By moving to the formal sector, these entrepreneurs operate at a larger scale, inducing in the aggregate a sizeable increase in official GDP and a quick decline in the size of the informal sector. The strength of this scale effect, however, decreases as the credit-to-GDP ratio gets larger. In fact, with more developed financial markets, highly productive entrepreneurs become less financially constrained and, hence, informal production is mostly carried out by small parasite firms that, optimally, decide to self-finance their businesses to take advantage of tax evasion. In these economies, therefore, further relaxing the collateral constraints only has a little impact on the size of the informal sector. This is the reason why we find a threshold level for the credit-to-GDP ratio above which further increasing λ has a rather small effect on informality and official GDP.

We test the empirical plausibility of the non-linear effects by running the following cross countries regression

$$y_i = \gamma_0 + \gamma_1 D_i + \gamma_2 credit_i + \gamma_3 D_i credit_i + \epsilon_i \qquad (9)$$

where y_i denotes either the ratio of informal to formal production or the logs of per capita GDP, $credit_i$ stands for credit to GDP ratio, and D_i is a dummy variable that takes value 1 if $credit_i$ of country i is below 60.8%. Data for the size of informal production are taken from Medina, while the source for GDP per capita and the credit-to-GDP ratio is the World Bank database. The dummy variable and its interaction with the credit-to-GDP ratio in equation (9) allow regression coefficients to differ across regions characterized by countries with advanced and less developed financial markets. To determine the threshold level on the credit to GDP ratio that distinguishes the two regions, we have complemented our data with information contained in Abiad et al. (2010) who provide a database of financial reforms covering 91 economies over the 1973-2005 period. The threshold level is then determined according to the minimum value for the credit-to-GDP ratio among the identified advanced economies, which corresponds to the Belgium economy with a ratio of 60.83%.[23]

Estimated results are reported in Figure 4 that displays scatter plots of the data along with the regression lines, and in Table 5, which provides regression coefficients and other statistics. We find that the coefficients for the interaction term are statistically significant in both regressions. The sign of these two coefficients is also coherent with the model predictions: negative when informality is included in the regression as dependent variable and positive in the case of per capita GDP. These findings and the fact that the coefficients on credit are not statistically significant imply that, in the data, a higher credit to GDP ratio is associated with lower informally and higher GDP per capita only in countries with less developed financial

[21] The literature usually refers to this case as *survival* firms.

[22] This feature is well apparent in Figure 2, which shows that − in sharp contrast with the benchmark case (see Figure 1) − agents in the middle range of managerial ability never engage in formal entrepreneurship if credit markets are perfect. These are parasite firms, which decide optimally to operate at a small size to reduce the probability of being caught and take advantage of tax evasion.

[23] This value corresponds to the 70th percentile in the cross-country distribution of credit-to-GDP ratio in the data.

markets. As nicely apparent in Figure 4, this result is clearly consistent with a non-linear effect of financial frictions.[24] We thus take the estimated results provided in Table 5 as suggestive evidence in favour of our theory.

Table 5: Testing for the non-linear effects. Regression coefficients.

Dependent variable:	**(A) Size of informal sector**	**(B) GDP per capita (logs)**
constant	25.97^{***} (3.96)	9.72^{***} (0.38)
credit	-0.06 (0.03)	0.00 (0.00)
credit * dummy	-0.19^{**} (0.07)	0.04^{***} (0.01)
dummy	15.16^{***} (4.41)	-2.21^{***} (0.43)
Adj. R^2	0.41	0.50
Num. obs.	155	148
RMSE	9.26	0.89

Note: Standard Errors in parenthesis below the coefficients. Asterisks denote significance levels (i.e. $^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$). Coefficients are estimated with OLS using sample averages over the 2001-2015 periods.

One of the main policy conclusion that can be drawn from the above analysis is that developing countries may effectively fight against firms informality and tax evasion by deepening financial markets. This conclusion is further corroborated by Figure 5, where, for different values of λ, the gains from financial deepening are compared with those resulting from an alternative policy that introduces a flat tax on personal income (i.e $\tau = 0$). As the figure illustrates, in economies with very tight borrowing constraints, financial deepening strongly dominates the flat tax policy not only in terms of lower informality and tax evasion but also in terms of higher total GDP, TFP and fiscal revenues. These results are instead generally reversed in the case of financially developed economies where the role of financial frictions becomes of secondary importance.

Another interesting finding from the above experiment is that the efficiency gains from adopting a flat tax are higher the lower the level of financial development in the country. To get an intuition for this result, notice that a decrease in tax progressivity results in lower income taxation for the highly productive individuals, and this in our model affects the economy through two main margins. First of all, the scale of business operation for the most productive entrepreneurs that are already producing in the formal sector increases (*intensive margin*). Second, the relative gains from tax evasion decline for the richest individuals, thereby pushing talented agents that run informal firms to switch sectors and become formal entrepreneurs (*extensive margin*). The impact of this last margin is magnified when the policy is adopted

[24] In the empirical appendix A.2, we provide evidence that these results remain qualitatively unaffected when performing a difference-in-difference analysis exploiting data on financial reforms to control for possible endogeneity in the regression analysis presented in this section.

under the presence of financial frictions given that – as we have seen in the previous section – with tight borrowing constraints, high-ability entrepreneurs with no enough wealth decide to operate in the informal sector. By contrast, with more developed financial markets, the effect of a flat tax reform are entirely driven by the intensive margin as the more talented entrepreneurs do not need to operate informally to overcome the borrowing constraints. This

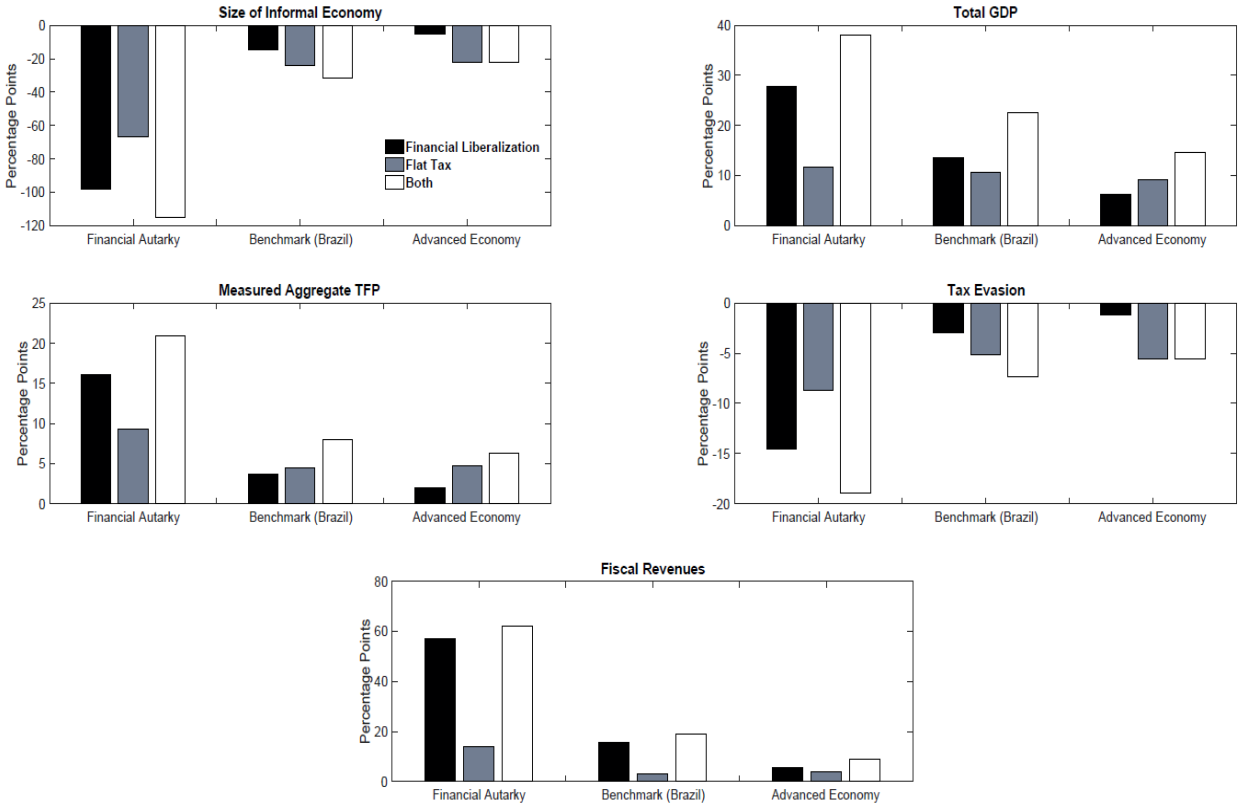

Figure 5: The effects of a flat tax reform for different degrees of financial frictions. Note: financially developed economies are identified by setting parameter $\lambda = 2.5$, implying a credit to GDP ratio of around 117%. Results from a flat tax reform are obtained by setting parameter $\tau = 0$ while keeping all other parameters at their calibrated values. For comparison, the picture reports results from financial deepening (black columns), and the effects when both policies are taken simultaneously (white columns).

is the reason why we find that the gains of a flat tax are amplified when financial markets are less developed.

To conclude, notice that the above experiment also sheds light on the importance of assuming progressive income taxation in our analysis, and more generally in analyzing firm informality. In fact, absent this effect, we reach the same conclusion as Quintin (2008): in the model, the lack of tax enforcement alone (i.e. with perfect credit markets) is not enough to drive a large informal sector. This can be seen in Figure 5 in the case of financially developed economies where financial markets are assumed to be nearly perfect. As the picture illustrates, in this case the size of the informal economy is 17 percentage points lower (moving from 22% of official GDP to around 6%) when taxation is flat instead of progressive. Among other things, this finding suggests that the effect of financial deepening on the size of the informal economy

may be potentially over-predicted in models where income taxation is assumed to be flat. As such, this also highlights the crucial role of progressive income taxation in shaping the non-linear relationship we found between the credit-to-GDP ratio and, both, the size of the informal sector and GDP per capita.

5. Conclusion

In this paper we build a model of occupational choice with progressive income taxation and informal production in which informal entrepreneurs have no access to credit and face an endogenous probability of detection by fiscal authorities. We calibrate our model to the Brazilian economy and evaluate the impact of removing financial frictions in our benchmark economy. We find that after removing financial frictions, the size of the informal economy decreases from 37% to 23%, together with important gains from this policy on official GDP (+27%) and measured aggregate TFP (+4%), and a drop in tax evasion (-17%) leading to an increase in fiscal revenues (+15 %).

We discuss the implications of removing financial frictions for any level of financial development and find a non-linear relation between the credit-to-GDP ratio on the one hand, and either the size of the informal economy, or official GDP, on the other hand. In particular, we find that, after a certain threshold value of the credit-to-GDP ratio, financial deepening has a more limited effect on the size of informality and official output. We test this hypothesis in the data and find evidence for the two non-linearities. This evidence suggests that accounting for the informal sector is crucial for understanding the non-linear relationship between financial and economic development. We then evaluate the effects of implementing a flat tax given different levels of financial development, as an alternative policy, and find that decreasing tax progressivity decreases the size of the informal sector, reduces tax evasion, and, at the same time, boosts total GDP and measured aggregate TFP. Moreover, we find an amplification affect of this policy in the presence of financial frictions.

The results on this paper suggest that financial markets interact with informality with important aggregate consequences up to a certain degree of financial development. In order to fight informality, financial market reforms seem to be very effective in developing countries, while a flat tax reform of personal income taxation seems to be more effective in countries with medium to high levels of financial market development.

References

Abiad, Abdul, Enrica Detragiache, and Thierry Tressel, "A New Database of Financial Reforms," IMF Staff Papers, June 2010, 57 (2), 281–302.

Allingham, Michael G. and Agnar Sandmo, "Income tax evasion: a theoretical analysis," Journal of Public Economics, 1972, 1 (3-4), 323–338.

Allub, Lian and Andrs Erosa, "Financial frictions, occupational choice and economic inequality," Journal of Monetary Economics, 2018.

Amaral, Pedro S and Erwan Quintin, "A competitive model of the informal sector," Journal of monetary Economics, 2006, 53 (7), 1541–1553.

Antunes, Antonio R and Tiago V de V Cavalcanti, "Start up costs, limited enforcement, and the hidden economy," European Economic Review, 2007, 51 (1), 203–224.

Buera, Francisco J. and Yongseok Shin, "Financial Frictions and the Persistence of History: A Quantitative Exploration," Journal of Political Economy, 2013, 121 (2), 221–272.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin, "Finance and development: A tale of two sectors," American economic review, 2011, 101 (5), 1964–2002.

Carroll, Christopher D, "The method of endogenous gridpoints for solving dynamic stochastic optimization problems," Economics letters, 2006, 91 (3), 312–320.

D'Erasmo, Pablo, "Access to Credit and the Size of the Formal Sector in Brazil," Research Department Publications IDB-WP-404, Inter-American Development Bank, Research Department April 2013.

D'Erasmo, Pablo N and Hernan J Moscoso Boedo, "Financial structure, informality and development," Journal of Monetary Economics, 2012, 59 (3), 286–302.

Erosa, Andres, "Financial intermediation and occupational choice in development," Review of Economic Dynamics, 2001, 4 (2), 303–334.

Fella, Giulio, "A generalized endogenous grid method for non-smooth and non-concave problems," Review of Economic Dynamics, 2014, 17 (2), 329–344.

Fernandez-Bastidas, Rocio, "Entrepreneurship and Tax Evasion," Mimeo, University of Alicante, Department of Economics 2018.

Fonseca, J., & Van Doornik, B. (2020), "Financial Development and Labor Markets: evidence from Brazil" Central Bank of Brazil Working Paper No. 532.

Hassan, Mai and Friedrich Schneider, "Size and development of the shadow economies of 157 countries worldwide: Updated and new measures from 1999 to 2013," 2016.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Optimal Tax Progressivity: An Analytical Framework," The Quarterly Journal of Economics, 2017, 132 (4), 1693–1754.

Kitao, Sagiri, "Entrepreneurship, taxation and capital investment," Review of Economic Dynamics, 2008, 11 (1), 44–69.

Kuehn, Zoe, "Tax rates, governance, and the informal economy in high-income countries," Economic Inquiry, 2014, 52 (1), 405–430.

Lucas, Robert E., "On the Size Distribution of Business Firms," Bell Journal of Economics, Autumn 1978, 9 (2), 508–523.

Medina, Leandro and Friedrich Schneider, "Shadow Economies Around the World: What Did We Learn Over the Last 20 Years?," IMF Working Papers 18/17, International Monetary Fund January 2018.

Nola, Alessandro Di, Georgi Kocharkov, Almuth Scholl, and Anna-Mariia Tkhir, "The Aggregate Consequences of Tax Evasion", Working Paper Series of the Department of Economics, University of Konstanz 2018-06, Department of Economics, University of Konstanz July 2018.

Ordonez, Julio Cesar Leal, "Tax collection, the informal sector, and productivity," Review of Economic Dynamics, April 2014, 17 (2), 262–286.

Orsi, Renzo, Davide Raggi, and Francesco Turino, "Size, trend, and policy implications of the underground economy," Review of Economic Dynamics, 2014, 17 (3), 417–436.

Porta, Rafael La and Andrei Shleifer, "Informality and development," Journal of Economic Perspectives, 2014, 28 (3), 109–26.

Porta, Rafael La and Andrei Shleifer "Informality and Development," NBER Working Paper, 2014, (w20205).

Quadrini, Vincenzo, "Entrepreneurship, Saving and Social Mobility," Review of Economic Dynamics, January 2000, 3 (1), 1–40.

Quintin, Erwan, "Contract enforcement and the size of the informal economy," Economic Theory, 2008, 37 (3), 395–416.

Sin, Hui Lok and Wagner Piazza Gaglianone, "Stochastic simulation of a DSGE model for Brazil," MPRA Paper 20853, University Library of Munich, Germany January 2006.

Ulyssea, Gabriel, "Firms, Informality, and Development: Theory and Evidence from Brazil," American Economic Review, August 2018, 108 (8), 2015–47.

Appendix

A.1. Computational Appendix

In this section we describe the computational procedure we employ to solve for the equilibrium in our model economy. We follow the method proposed in Fella (2014).

In the household problem, savings are continuous choices but occupational choice decisions are discrete choices that make the value function non-concave and thus complicate the use of the endogenous grid method (EGM) proposed in Carroll (2006). Fella (2014) proposes a method that combines the EGM with value function iteration (VFI) in the identified non-concave regions caused by occupational choice decisions.

We first solve for the cash-in-hand and marginal productivity of capital over the state space $\omega = \{e, a, b\}$, where a stands for savings, e is entrepreneurial ability, and b is occupation choice. b is defined over the three occupational choices in our model: worker, entrepreneur in the formal sector, and entrepreneur in the informal sector. Assets, a, are defined over a discrete grid space of 250 points, and e is discretized over 40 entrepreneurial abilities. We define cash-in-hand as all resources available for a household, i.e. the right hand side of budget constraints defined in sections 2.5.1 and 2.5.2, and is composed of: the stock of assets (savings); and income coming from a particular occupation choice. The marginal productivity of capital is defined as the partial derivative of cash-in-hand with respect to assets (savings).

Workers obtain their incomes from the wage they get in the labor market and interests they received on their savings, all of it net of taxes. Then cash-in-hand for workers in a particular period is the sum of income plus the total amount of savings they start with, in a given period.

Income made by formal entrepreneurs comes from the profits they obtain by running their formal businesses plus the interests on their savings, all of it net of taxes. The profits they make depend on the amount of savings they start with. When they decide to produce in the formal sector since, given a particular ability, entrepreneurs are financially constrained by equation (3) until their savings are greater than a certain threshold at which they are able to overcome financial constraints. This constraint will affect the marginal productivity of capital they face, which will be a function of their savings, until that particular threshold is reached, conditional on their ability. Hence, marginal productivity of capital will eventually converge to the market interest rate as they increase their amount of savings.

Informal entrepreneurs solve their maximization problem taking into account that the amount of capital used in production also affects the probability of being audited by fiscal authorities and the fact that they operate under financial autarky. Also their income is going to be affected by the financial constraint defined in equation (4) until a certain threshold in savings is reached. In case they are not caught, their income woyuld be made of the profits made plus income from savings. In case they are caught, they would have to pay the corresponding taxes plus a surcharge fee on the amount of income evaded.

Retirees get their income just from the interests received on their savings. Their cash-in-hand is therefore made of their savings (stock of assets) plus the income received over it.

Once we have the cash-in-hand and the marginal productivity of capital over the state space ω and since we have a life-cycle model, we can solve the model by backward induction

from the oldest generation to the first. We already know the policy functions for the oldest generation and can therefore solve backwards by iterating on the Euler Equation and applying the method proposed by Fella (2014) for the policy function of the generation before the one for which we already know value and policy functions. We do so until we reach the first generation. By this procedure we obtain the optimal policy and value functions for each state variable (occupational choice, assets and ability).

We then simulate the model economy over the probability of informal entrepreneurs being detected and then forced to pay the corresponding surcharge factor on evaded taxes. Under the assumption that households are born with zero assets we obtain the equilibrium variables.

The algorithm we employ can be summarized in these steps:

1. Guess an initial value for the wage, w^0, and the interest rate, r^0, in the economy.
2. Given w^0 and r^0, solve for the amount of capital and labor employed in the corporate sector.
3. Given w^0 and r^0, solve for the policy functions by solving the household's problem.
4. Simulate the economy over the probability of detection in the informal sector and get all the aggregate variables: capital demand, K^d; capital supply, K^s; labor demand, L^d, and labor supply, L^s.
5. Check for convergence in capital and labor markets.
6. Update w^1 and r^1 if necessary until convergence.

A.2. Empirical Appendix

In this section, we perform a difference-in-difference analysis to control for endogeneity in our empirical assessment for the impact of financial frictions on the relative size of the informal economy. To this end, we use the IMF Financial Reforms's database[25] which synthesizes information on credit controls into a credit controls index with values between 0 and 1 (a larger value indicating less credit controls). Summary statistics are presented in Table (6) and Table (7).

We perform our difference-in-difference analysis comparing the cross-country variations observed in 1991 with those observed in 2003. Defining financial deepening as a positive change in the credit controls index between 1991 and 2003 for a given country, we ask whether financial deepening, understood as a policy that is not directly targeted to the informal sector but that positively affects the credit-to-GDP ratio, is associated with a reduction in the size of the informal economy over the period 1991-2003.

Specifically, we estimate:

$$Inf_{i,t} = \beta_0 + \gamma_i + \beta_2 PostTreatment_t + \beta_1 did_{i,t} + \beta_3 credit_{i,t}$$

[25] https://www.imf.org/en/Publications/WP/Issues/2016/12/31/A-New-Database-of-Financial-Reforms-22485.

~~Were~~ where *Inf* is the size of the informal economy as a share of GDP; *PostTreatment* is a dummy variable which takes on value 1 in year 2003 and value 0 in year 2001; and *did* is the difference-in-difference variable which is defined as the product of *PostTreatment* and a *treatment dummy* variable taking on value 1 if the country experienced financial deepening (and value 0 otherwise) .

We find evidence for a strong association between financial deepening and a reduction in the size of the informal economy in percentage of GDP (Table 7, Table 8). We further separate our sample below and above the threshhold of 60 percent for the private credit-to-GDP ratio across countries in 1991, and find that the association between financial deepening and a reduction in the size of the informal economy is even stronger for the least financially developed countries (Table 8).

Table 6 : Summary Statistics

VARIABLES	(1) N	(2) mean	(3) sd	(4) min	(5) max
Inf as share of GDP in 1991	82	33.43	14.67	7.56	68.09
Inf as share of GDP in 2003	82	31.14	14.48	7.78	69.01
Credit Controls Index in 1991	82	1.613	0.978	0	3
Credit Controls Index in 2003	82	2.451	0.750	0	3
Private Credit-to-GDP in 1991	62	62.08	55.62	1.452	209.9
Private Credit-to-GDP in 2003	78	70.55	60.61	4.725	231.6

Table 7 : Summary Statistics for Changes Over 1991-2003

VARIABLES	Control Group Δ Credit Controls=0	Treatment Group Δ Credit Controls>0
Change in Inf as share of GDP	-6.20	-8.51
Change in Credit-to-GDP ratio	40.84	63.32
Change in Credit Controls Index	0	1.563

Table 8 : Estimation Results

Dependent Variable = Inf	All Countries	Least Financially Developed Countries Credit-to-GDP<60	Most Financially Developed Countries Credit-to-GDP>60
Diff-and-diff Estimate	-2.479***	-2.817**	-2.276***
	(0.712)	(1.228)	(0.648)
PostTreatment Dummy	-1.034*	-0.673	-1.337**
	(0.533)	(0.419)	(0.533)
Private Credit-to-GD ratio	-0.0143	-0.0168	-0.0100
	(0.0116)	(0.0334)	(0.0116)
Constant	33.43***	41.52***	24.53***
	(0.644)	(0.755)	(1.060)
R-squared	0.505	0.586	0.664
Observations	140	72	68
Number of ifs	78	36	38
Size of Treatment Group		24	11
Country FE	YES	YES	YES

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1