Financial Structure, Informality and Development*

Pablo N D'Erasmo† Hernan J Moscoso Boedo

University of Maryland University of Virginia

February 21, 2012

Abstract

The impact of capital market imperfections and costs of creating and operating formal sector firms on total factor productivity is studied. We propose a firm dynamics model with endogenous formal and informal sectors where firms face a technology adoption opportunity. The model predicts that countries with a low degree of debt enforcement and high costs of formality are characterized by low allocative efficiency and large output shares produced by low productivity, informal sector firms. For frictions parametrized using the *Doing Business* database, the model generates a drop in total factor productivity of up to 25% relative to the US.

Keywords: Financial Structure, Informal Sector, Productivity, Policy Distortions.

JEL Classifications: D24, E26, L11, O16, O17

*We thank an anonymous referee, Cristina Arellano, Rui Castro, Russell Cooper, V.V. Chari, Dean Corbae, John Haltiwanger, Bob Lucas, Ellen McGrattan, Toshihiko Mukoyama, Fabrizio Perri, Diego Restuccia, Nancy Stokey, Vincenzo Quadrini, Erwan Quintin, Pierre-Daniel Sarte, Luis Serven, Eric Young and the seminar participants of the 2009 Midwest Macroeconomics Meetings, the 2009 North American Econometric Society Meetings, NBER Summer Institute 2010, 2010 ITAM Summer Camp and Federal Reserve Banks of Dallas, Minneapolis and Richmond for their very helpful comments. We also thank Laura Alfaro for sharing her data and Javier Miranda for helping with BDS data. All errors are of course ours. We acknowledge the use of the computer cluster of the Economics Department, University of Texas at Austin financed by NSF grant MRI-0521499. Hernan Moscoso Boedo gratefully acknowledges financial support from the Bankard Fund for Political Economy

†Corresponding author: University of Maryland, Department of Economics, 3105 Tydings Hall, College Park, MD 20742, (301)405 3529, derasmo@econ.umd.edu

1 Introduction

A theory of total factor productivity (TFP) based on measured institutional differences across countries is proposed in this paper. In particular, institutional heterogeneity of entry costs to the formal sector, differences in the tax structure and in the efficiency of debt enforcing mechanisms (measured as debt recovery rate and debt enforcing costs at default) are considered. The question that the paper tries to answer is: what is the change in total factor productivity that can be attributed to international differences in measured costs of doing business in the formal sector?

This is done by developing a general equilibrium model of firm dynamics with endogenous entry and exit that incorporates capital financing and bankruptcy decisions. The model allows for the existence of a formal and an informal sector. Entering and operating in the formal sector is costly, but allows firms to choose from an unrestricted set of technologies, while providing access to credit markets with better commitment (given by observed recovery rates and associated costs). The degree of debt enforcement varies across countries and, because there is equilibrium default, it affects the interest rate that firms face. This generates endogenous idiosyncratic prices. In the quantitative exercise, country-specific institutions are imposed, which are based on those measured by the World Bank as reported in its *Doing Business* database. In order to isolate the effects of institutional differences, countries are assumed to all have access to the same production possibilities. By generating large informal sectors and increasing capital misallocation, the frictions explain a drop in total factor productivity of up to 25%.

As Figure 1 shows, informal activity is a feature that seems to be correlated to productivity and output per worker, worldwide.[1] Of the various measures of informal activity, we focus on the fraction of the labor force that participates in the underground economy.[2] The fraction of

[1]Aggregate output measures used to compute productivity and output per worker include all production, formal as well as informal. For more details on the measurement issue, see section 2.2 or the online Appendix

[2]Measured as the fraction of the labor force not covered by a pension scheme. The share of the labor force not covered by pension schemes provides a better direct measure of informality for the US, the benchmark country in the calibration. Schneider and Enste (2000) report various measures of the informal sector across countries (highly correlated with our measure), and is the most comprehensive study regarding informality in

the labor force that is engaged in production outside of the formal sector ranges from around 10% in developed countries to almost 100% at the low end of the income distribution. Even when measures of informal activity are extremely noisy, such a large sector of the economy cannot be ignored if economic development around the world is to be understood.

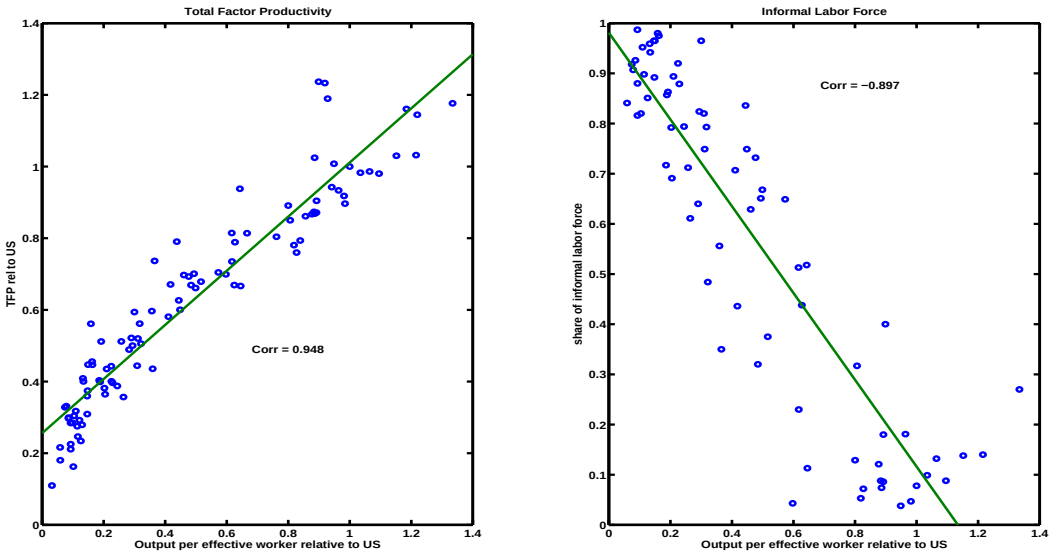

Figure 1: Total Factor Productivity and Size of the informal sector.

Note: Total factor productivity and output per effective worker are author's calculations based on Hall and Jones (1999). See the online Appendix for details. The share of informal labor force corresponds to the share of the labor force not covered by a pension scheme as reported by the *World Development Indicators* 2006. Solid lines indicate OLS regression.

The microeconomic structure of the informal sector displays some salient features. Pratap and Quintin (2008) document that the informal sector is mainly composed of small-scale, self-financed, unskilled intensive activities.[3] Moreover, La Porta and Shleifer (2008) also find that unofficial firms tend to be smaller, less productive (with productivity differences in the 100 300 percent range), and younger than formal firms.[4]

a cross country setting. They include indirect estimates of informal output from energy consumption or money demand or from discrepancies between official and actual employment in household surveys. In the sample all of the countries included do have a pension scheme, alleviating the potential drawback of having countries without formal pension schemes.

[3]Regarding market segmentation, the evidence points towards segmentation in the financial markets but not in the labour markets.

[4]Consistent with micro evidence presented by Bruhn (2008), Bertrand and Kramarz (2002), and McKenzie and Seynabou Sakho (2010), this implies that improvements in the cost of doing business would affect aggregate

The model predictions are consistent with the evidence. More specifically, it delivers a strong negative correlation between the level of informality and income per capita as well as an informal sector with very small, relatively unproductive and young firms compared to those in the formal sector. Moving along the development spectrum, poorer countries display a bimodal distribution of firms, with many small and large ones, but not many middle sized firms (the so called "missing middle" in the empirical literature).

Institutional differences affect average productivity in the formal sector through two forces going in opposite direction: a selection mechanism and the allocation of resources. On one hand, as the cost of doing business increases, the entry threshold to the formal sector also rises (i.e. only the most productive firms decide to operate formally). This effect tends to augment the average productivity of formal firms in economies where there are higher costs. On the other hand, the existence of the informal sector shifts the labor demand curve, increasing the equilibrium wage and changing the schedule of interest rates faced by formal firms. Formal sector incumbents alter their capital accumulation decisions as a function of their individual productivity. Moreover, given that the entry threshold is higher, fewer firms decide to enter and the reduction in the level of competition generates a decrease in the exit threshold (thus, firms that are less productive stay active). These sources of allocative inefficiency become more pronounced as frictions increase. They lead to an increase in the variance of the marginal product of capital within the formal sector (in line with the evidence put forward by Hsieh and Klenow (2009)) that reduces firm-level productivity. The first effect dominates the second, thus generating an increase in output-weighted productivity for formal sector firms as institutions worsen.[5]

The introduction of the informal sector in a firm dynamics model with financial frictions is quantitatively important for productivity differences. In Section 6.2, a counterfactual example

productivity not from the unleashing of informal firms but rather from their displacement by more productive new formal firms.

[5]Although there is no direct access to evidence on this fact, this is most probably a counterfactual prediction of the model. It is important to note that, in order to isolate the effects of institutions, in the quantitative exercise it was assumed that every economy has access to the same technological opportunities. In reality, countries differ along many dimensions, one of which is their distance to the technology frontier. According to Duarte and Restuccia (2010) this explains about 50% of the gains in TFP across countries.[6]

shows that a model with no informal sector generates a reduction in TFP relative to the U.S. that is 20% smaller than the drop produced by the benchmark model. Furthermore, in Section 6.3, the effect of each friction is analyzed to find that in the benchmark case, entry costs account for 3/4 of TFP differences across countries and the financial structure and taxes account for the rest (this implies that the measured variation in taxes has a minimal impact on productivity).

Our approach to firm dynamics started with Hopenhayn (1992) and Hopenhayn and Rogerson (1993), and is close to Cooley and Quadrini (2001) who studied the effects of financial constraints in a similar set up.[7] A related literature on the distributional consequences of frictions in this context started with Restuccia and Rogerson (2008).[8] In all cases, they back up the implied frictions in the firm's environment necessary to generate the observed distribution of firms. As opposed to them, in this paper, frictions that firms face are those observed in the data collected by the World Bank.[9] This paper introduces imperfect capital markets, and along that dimension the most closely related papers include Antunes and Cavalcanti (2007) and Quintin (2008).[10] This paper builds upon this literature by analyzing a model of firm dynamics with idiosyncratic uncertainty and endogenous technology adoption. Also, different financial contracts are introduced here, where default costs are constrained by limited liability.

The paper is organized as follows. Section 2 presents institutional differences across countries as measured by the World Bank. Section 3 and 4 present the theoretical model and its equilibrium. Section 5 is devoted to the calibration of the model to the US data. Section 6 presents the main experiments. Finally, Section 7 concludes.[11]

[7]The modeling assumptions regarding the informal sector follow the steps of Rauch (1991) and Loayza (1996) where the informal activity can be thought of as an optimal response to the economic environment.

[8]Important references are Hsieh and Klenow (2009), Guner, Ventura, and Xu (2008), Arellano, Bai, and Zhang (2010) and Buera, Kaboski and Shin (2011).

[9]Barseghyan and DiCecio (2010) and Moscoso Boedo and Mukoyama (2012) use a similar strategy.

[10]Antunes and Cavalcanti (2007) and Quintin (2008) study endogenous informal sectors that result from imperfect contract enforcement. Also related, Castro et al. (2008) and Erosa and Hidalgo Cabrillana (2008) study the effects of financial contracts in environments with asymmetric information.

[11]An online Appendix presents a description on how measured aggregate TFP is constructed, how the country level measured institutions are parametrized using the World Banks Doing Business data set and a section where the model is used to address the problem of measured aggregate TFP in the presence of the informal sector.

2 Measured Differences across Countries

For the purposes of this paper, countries differ along two important dimensions. First, they have different measured institutions related to the formal sector (such as entry costs, taxes, and bankruptcy efficiency). Then, they differ in terms of output, both formal and informal. The following subsections explain these differences in detail.

2.1 Institutions

What firms have to do in order to enter, operate in, and exit from the formal sector varies across countries. In order to compare these different costs the World Bank, through its *Doing Business* project, follows a standardized firm across countries and measures regulations to entry, operations, and exit. They measure the costs, in terms of time and resources, along many dimensions affecting the firm, such as starting a business, getting construction permits, employing workers, obtaining credit, protecting investors, paying taxes, trading across borders, enforcing contracts, and closing a business. Of particular interest are the cost of entering the formal sector, the tax rate and the level of tax compliance difficulty (while operating in the formal sector), and the efficiency of the debt enforcement mechanisms if the firm decides to default on its debt. These costs are depicted in Figure 2 against GNI per capita relative to the US in 2009. See the online Appendix for details on the construction of these costs.

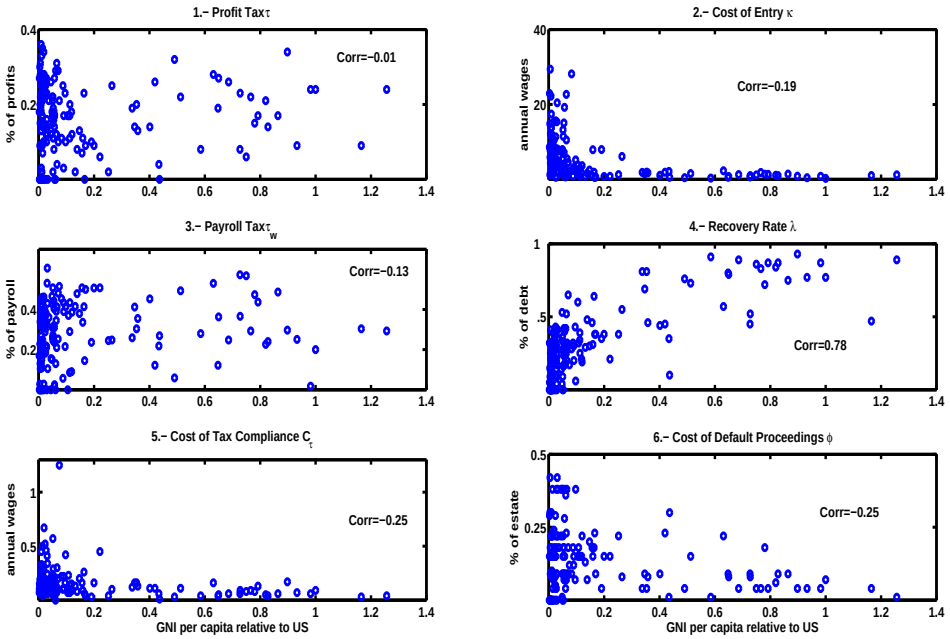

Figure 2: Cost to entry, income tax rate, cost of tax compliance, recovery rate, and cost of default proceedings from the Doing Business Database in 2009.

Entry Cost: The cost of entering the formal sector is constructed as in Moscoso Boedo and Mukoyama (2012). It is the sum of two parts including the costs of registering a business and of dealing with licenses to operate a physical locale.[12] Both have a monetary cost and a time cost (which is translated to monetary units by assuming that one worker has to be employed full time in order for the firm to go through the entry process). The cost of entering the formal sector as a fraction of the wage (denoted by $w\kappa$) varies greatly across countries, with high levels of κ observed only at the low end of the income distribution. Registering a business in the US costs 0.7% of GNI per capita, while in Sierra Leone it is over 1000% of GNI per capita. In terms of time, in the US a business can be started immediately while in other countries it can take a very long time: In Suriname it takes more than two years and in Haiti and Brazil it takes about six months. Dealing with licenses also displays great variation across countries. The cost

[12]The data used to generate the cost of dealing with licenses to operate a physical locale is obtained from the World Bank Doing Business database as "Dealing with Construction Permits". Part of the elements involved in construction permits, such as the cost of connection to basic services, are present when operating a physical locale

is 13% of GNI per capita in the US vs. 600 times per capita income in Liberia and 100 times in Zimbabwe. In terms of time, it takes 40 days to obtain a license in the US compared to up to 1000 days in Haiti.

Tax Structure: The tax rate paid on profits by the firms (τ) and payroll taxes (τ_w) do not seem to exhibit a pattern over the distribution of income per capita, as shown by panels 1 and 3 in Figure 2. What does exhibit a pattern similar to the entry cost is the cost of tax compliance (wc_τ). This cost reflects the time that it takes to pay taxes in each country. It is assumed that during this time there is a full time worker devoted to the tasks related to tax compliance, and therefore translate time into costs as the worker's annual wages. The cost of paying taxes only exhibits levels above 10 weeks for countries below 20% of the US GNI per capita. Paying taxes takes no time in the Maldives, 12 hours in the UAE, 187 hours in the US, and more than 1000 hours in Vietnam, Bolivia, Belarus, Cameroon, and Brazil. This indicates a great deal of variation across countries in the complexity of their tax code. Firms have to bear not only the tax rate per se but also the cost of complying with the tax code, which at the low end of the income distribution is not insignificant.

Bankruptcy efficiency: The efficiency of the system in the event of default has two components: a cost component and a recovery rate. The cost of the system (ϕ), reported as a percentage of the estate's value, includes court fees and the cost of insolvency practitioners, such as legal and accounting fees. It ranges from 1% of the estate's value in countries like Norway and Singapore to more than 40% in Sierra Leone, Liberia, and Ukraine, and above 70% in the Central African Republic. The recovery rate refers to what external lenders obtain once the firm decides to default on its debt (λ). It is effectively zero for many extremely poor countries in sub-Saharan Africa. On the other hand, only in developed countries is it above 75%. Note that this is the return obtained by the external creditor conditional on the borrower defaulting. It measures the cents on the dollar recovered from that point on, and includes different channels to resolve the contract breach such as foreclosure, liquidation, and reorganization, as reported by Djankov et al (2008).

2.2 GDP and the informal economy

This subsection (and the more detailed online Appendix) is devoted to issues regarding the measurement of aggregate output in the presence of an underground economy. The definition and practical measure of GDP is key to generate other measures, such as Total Factor Productivity. Regarding the definition of GDP, the U.N. System of National Accounts 1993 states that informal output should be included in official GDP. Therefore, the question of whether the informal sector is included in aggregate measures of output becomes a practical one. Guidelines on how to measure informal activity were published in OECD (2002). These guidelines direct statistical authorities to confront and analyze data from different sources (for example confrontation and analysis of survey and tax collection data for different dimensions such as production, labor force, or consumption). Then, differences can then be attributed to the informal sector. In 2005-2006, the U.N. conducted a survey across statistical agencies of 45 middle and high income countries to gather information regarding the specific steps taken by each country to include the informal sector in the figures of aggregate output. Countries responding to the survey provided information on the size of the informal economy that had been included in official GDP, as well as what kind of informal activities were accounted for (i.e. not registered, not surveyed, etc). The size of the informal economy (as a fraction of official GDP) included in the official statistics of GDP ranged from zero in Japan and New Zealand to 31.6% in Moldova. Our interpretation is that, in general, official GDP figures include a measure of informal activity. [13]

3 Environment

This is a standard firm dynamics model based on Hopenhayn (1992) that incorporates capital and credit markets as in Cooley and Quadrini (2001). Time is discrete, and the period is set

[13] In the online Appendix, the model is used to address the problem of measured aggregate TFP in the presence of the informal sector. It is shown that, if inputs such as capital and labor are measured with the same or better accuracy as total output, the benchmark results can be thought of as a lower bound of the potential reduction in measured TFP.

to one year. There are three types of entities in the economy: firms, lenders and consumers. Firms can operate in one of the two sectors (formal or informal) and produce the consumption and capital goods used in the economy. They are the capital owners and pay dividends to the consumers. Lenders make loans to the firms. Consumers supply labor to the firms, and receive their profit net of entry costs. The stationary equilibrium is analyzed.

3.1 Consumers

There is an infinitely lived representative consumer who maximizes the expected utility:

$$\mathbb{E}\left[\sum_{t=0}^{\infty}\beta^t U(C_t)\right],$$

where $\mathbb{E}[\cdot]$ is the expectation operator, C_t is consumption (restricted to be nonnegative) and $\beta \in (0,1)$ is the discount factor. The household is endowed with one unit of labor which supplies to firms at the market wage rate w. The consumer is responsible for the creation cost of new firms c_e, consequently owns existing firms in the economy and receives income from the dividends they pay. Moreover, the household has access to a risk free bond B_{t+1} that is in zero net supply and pays r_t units of the consumption good in the following period. Finally, the household receives a lump sum transfer for the total amount of taxes collected.

3.2 Firms and Technology

The unit of production is a single-establishment firm, also understood as a unique investment project. Each project is described by a production function $f(z,k,n)$ that combines productivity z, capital k, and labor n. It is assumed that the production function has decreasing returns to scale. In particular, the production function is defined as $f(z,n,k) = zk^{\alpha}n^{\gamma}$ with $0 < \alpha + \gamma < 1$ and $\alpha, \gamma \in (0,1)$.

There are two processes for z: *high* (h) and *low* (l). The *high* productivity process is given

by

$$\ln(z_{t+1}) = (1-\rho)\ln(\mu_h) + \rho\ln(z_t) + \epsilon_{t+1}$$

with $\epsilon_{t+1} \sim N(0,(1-\rho^2)\sigma^2)$, where σ^2 is the variance of $\ln(z)$, μ_h is the mean, and ρ the autocorrelation parameter of the process. The conditional cumulative distribution of z_{t+1} is denoted by $\eta(z_{t+1}, z_t)$. The use of the *high* productivity process is restricted to the formal sector. To simplify the exposition of the model, the following two assumptions are made. First, it is assumed that the *low* productivity process is a constant given by μ_l and restricted to the informal sector. Second, once operating as either formal or informal, firms are not allowed to switch between sectors. These assumptions imply that formal firms will use the *high* productivity process and that informal firms will use the *low* productivity process. Other potential possibilities would be to allow firms to switch between sectors and to allow formal firms to use the *low* productivity process.[14] The two processes will be calibrated to match the size distribution of formal firms and the size of the informal sector. The difference between the two, is one of the channels that allows the model to generate capital missallocation together with small informal establishments as observed in the data by Bartelsman et. al. (2009) and Perry et. al. (2007). Note that the fraction of firms operating under each process is an endogenous outcome of the model and a function of the country specific frictions.

The assumption of different productivity processes is consistent with the evidence provided by La Porta and Shleifer (2008). They document productivity differences between informal firms and small formal firms at the firm level that range from 100% to 300%. They also find that these differences are permanent and not the result of informal sector firms operating at a lower scale in order to avoid detection.[15]

[14] In fact, the version of the model that allows for all of these possibilities was computed and calibrated and delivered that, at the calibrated parameters, the dichotomy between sectors and productivity processes arose endogenously.

[15] For example, differences in sales per worker are much higher (2 to 3 times) than the average entry cost, implying that is not just the barrier to entry that is the main factor affecting scale, productivity or the decision to operate informal. Related to this, they note that in a sample of developing economies approximately 91% of registered firms at the time of the survey started as registered firms and do not come from the informal sector. Moreover, Bruhn (2008), Bertrand and Kamaratz (2001) and McKenzie, David and Sakho (2007) present empirical evidence that shows that improvements in entry costs do not lead to the formalization of previously informal firms and only generate the creation of new businesses.

Firms maximize expected discounted dividends d:

$$\mathbb{E}\left[\sum_{t=0}^{\infty} R^t d_t\right],$$

at the rate R.[16] Firms are created by the consumer paying a cost c_e. Once launched, firms face a technology adoption decision. They draw one investment opportunity from the initial productivity distribution of the h process $\nu(z_0)$. Draws from this distribution are assumed to be i.i.d across firms. Firms then compare z_0 to μ_l and choose between staying out of the market or operating one of the projects as a formal or informal firm, i.e the project choice is non-reversible.[17] Unimplemented projects go back into the pool.

There is a random fixed cost of production c_f, measured in units of output, that is iid across firms and over time with distribution $\xi(c_f)$. A firm that does not pay this fixed cost is not allowed to produce. Firms own their capital and can borrow from financial intermediaries in the form of non-contingent debt $b \geq 0$. They finance investment with either debt or internal funds.

If the firm operates in the formal sector, it is subject to a proportional tax on profits τ, a cost in labor units of filling those taxes $c_\tau w$, and a payroll tax τ_w. Creating a formal sector firm requires an entry cost κw. In the calibration, taxes and the costs are set directly from the corresponding measures in the *Doing Business* database.[18]

3.3 Credit Markets

The credit industry makes loans to the formal and informal sector firms. Creditors are risk-neutral and competitive. They have access to the risk-free bond with return r_t. Asset markets are incomplete. In each period, firms borrow using only one-period non-contingent debt denoted

[16] At the stationary equilibrium, the firm's discount factor is constant.

[17] This is consistent with the evidence presented in Atkeson and Kehoe (2007) who argue that manufacturing plants needed to be completely redesigned in order to make good use of the new technologies.

[18] While government policies can be endogenous, this paper focuses on measuring their effects on aggregates, and policies are taken as exogenous. However, the equilibrium found is consistent with the solution to a model that incorporates a one time political game with full commitment in which the government optimally chooses the taxes and costs reported by the World Bank.

by b. Because there is perfect information, prices depend on firms' characteristics given by their choice of sector (formal or informal), future level of capital, level of borrowing, and current productivity level under each technology. In particular, firms in the formal sector borrow at price $q_j^f(k', b', z)$ and firms in the informal sector borrow at price $q^i(k', b')$. In each period, firms can default on their debt. A default triggers a bankruptcy procedure that liquidates the firm. When making a loan to a formal sector firm, lenders take into account that in the case of default they can recover up to a fraction λ of the original loan. The formal bankruptcy procedure has an associated cost equal to a fraction ϕ of the firm's capital. The values of the recovery rate λ and the bankruptcy cost ϕ are obtained from the *Doing Business* database. Because the capital of the informal firm is not legally registered, the recovery rate of a loan to an informal sector firm that defaults is assumed to be zero. This assumption follows the evidence presented in Pratap and Quintin (2008) where it is suggested that there is segmentation in the financial markets across formal and informal sectors.

Consistent with bankruptcy law across countries, we follow the limited liability doctrine. This limits the owner's liability to the firm's capital.

4 Equilibrium

The stationary equilibrium of the model is analyzed in this section. In this equilibrium the wage rate, the risk free rate and the schedule of loan prices are constant. Every equilibrium function depends on the set of loan prices, the risk-free rate, and the wage rate. For ease of exposition, this dependence is avoided.

4.1 Consumer's Problem

In the stationary equilibrium, all prices and aggregates in the economy are constant. Hence, household maximization implies that the consumer supplies its unit of labor inelastically, $\beta = R = \frac{1}{1+r}$, and that aggregate consumption is:

$$C = w + \Pi + T - E + X, \tag{1}$$

where Π is total dividends from incumbent firms, T is the lump-sum transfer from the income and payroll taxes, E is the aggregate creation cost, and X is the exit value of firms.

4.2 Formal Sector Incumbent

An incumbent firm in the formal sector operating a project with technology h starts the period with capital k, debt b, and previous productivity z_{-1}. Then, the firm draws the fixed cost that is required for continuing the operation, c_f, and decides to either operate the project, exit after repayment of debts, or default and liquidate the firm. The value function of an firm at this stage is denoted as $W^f(z_{-1}, k, b, c_f)$. If it decides to remain in business, it pays c_f and observes the current period's productivity z. The value function of a firm operating in the formal sector is denoted as $V^f(z, k, b, c_f)$.

The incumbent solves the following problem

$$W^f(z_{-1}, k, b, c_f) = \max \left\{ \int V^f(z, k, b, c_f) d\eta(z|z_{-1}), \max\{0, (1-\phi)k - \lambda b\}, k - b \right\} \tag{2}$$

where

$$V^f(z, k, b, c_f) = \max_{n, k', b'} d^f + \beta \int W^f(z, k', b', c_f') d\xi(c_f)$$

s.t.

$$d^f = (1-\tau)\left[zk^\alpha n^\gamma - c_f - w(1+\tau_w)(n+c_\tau)\right] - k' + (1-\delta)k + q^f(k', b', z)b' - b \geq 0$$

The solution to problem (2) provides the exit decision rule $\chi^f(z_{-1}, k, b, c_f)$ that takes the value of 0 if the firm continues to operate, 1 if the firm decides to default, and 2 if the firm decides to exit after repayment. The optimal capital and debt decision rules for a firm in the formal sector are given by $k'^f(z, k, b, c_f)$ and $b'^f(z, k, b, c_f)$, respectively.

4.3 Informal Sector Incumbent

An incumbent firm in the informal sector, after observing the fix operating cost c_f, can choose to stay active or to exit the market after a default. More specifically, the informal incumbent firm solves the following Bellman equation:

$$W^i(k,b,c_f) = \max\left\{V^i(k,b,c_f), k\right\} \tag{3}$$

where the value of remaining in the informal sector is given by

$$V^i(k,b,c_f) = \max_{n,k',b'} d^i + \beta \int W^i(k',b',c_f')d\xi(c_f)$$

s.t.

$$d^i \;=\; \mu_l k^\alpha n^\gamma - c_f - wn - k' + (1-\delta)k + q^i(k',b')b' - b \geq 0$$

The solution to problem (3) provides the exit decision rule $\chi^i(k,b,c_f)$ that takes the value of 0 if the firm continues to operate in the informal sector and 1 if the firm decides to default. The optimal capital and debt decision rules are given by $k'^i(k,b,c_f)$ and $b'^i(k,b,c_f)$.

4.4 Entrants

The value of a potential entrant (net of entry cost) W_e is given by:

$$W_e = \int \max\left\{W^i(0,0,0), \widetilde{V}^f(z_0,0,0,0)\right\}d\nu(z_0) - c_e \tag{4}$$

where $\widetilde{V}^f(z_0,0,0,0)$ is the value of starting as a formal firm given by

$$\widetilde{V}^f(z_0,0,0,0) = \max_{k',b'} \widetilde{d}^f(z_0,0,0,0) + \beta \int W^f(z_0,k',b',c_f')d\xi(c_f)$$

s.t.

$$\widetilde{d}^{f}(z_0, k, b, c_f) = -(1-\tau)\, w(1+\tau_w)\kappa - k^{'} + q^{f}(k^{'}, b^{'}, z_0)b^{'} \geq 0$$

Effectively, an entrant has no capital, no debt, and the cost of production c_f equals zero. The entrant chooses between projects and sectors. The sector and project adoption decisions are made after paying c_e and observing the productivity level z_0, which affects the conditional distribution from which the first productivity parameter will be drawn. Differences in the volatility of the processes together with differences in initial productivity are going to generate variation in the decisions made by the entrants and by the potential lenders. That introduces differences in behavior as a function of volatility and contract enforceability. In equilibrium, under free entry, $W_e = 0$ will hold. The solution to problem (4) provides the entry decision rule $\Xi^{e}(z_0)$ as well as capital and debt decision rules $\widetilde{k}'(z_0, 0, 0, 0)$ and $\widetilde{b}'(z_0, 0, 0, 0)$ for a firm that starts operating in the formal sector.

4.5 Lenders

Lenders make loans to formal and informal firms while taking prices as given. Profit for a loan b' to a firm in the formal sector with future capital k' and, productivity z is

$$\pi^{f}(k^{'}, b^{'}, z) = -q^{f}(k^{'}, b^{'}, z)b^{'} + \frac{1 - p^{f}(k^{'}, b^{'}, z)}{1+r}b' + \frac{p^{f}(k^{'}, b^{'}, z)}{1+r}\min\left\{\lambda b', (1-\phi)k'\right\},$$

where $p^{f}(k^{'}, b^{'}, z)$ denotes the default probability of this borrower.

Profit for a loan b' to a firm in the informal sector with future capital k' is

$$\pi^{i}(k^{'}, b^{'}) \;\; = \;\; -q^{i}(k^{'}, b^{'})b^{'} + \frac{\left[1 - p^{i}(k^{'}, b^{'})\right]}{1+r}b'$$

where $p^{i}(k^{'}, b^{'})$ denotes the default probability of the informal borrower. In equilibrium, the

schedule of prices will adjust so that $\pi^f(k^{'}, b^{'}, z) = 0$ and $\pi^i(k^{'}, b^{'}) = 0$ for all (j, k', b', z).

4.6 Definition of equilibrium

A stationary competitive equilibrium is a set of value functions $\{W^f, W^i, V^f, V^i, \widetilde{V}^f\}$, decision rules (capital, debt, default, exit and sector), a wage rate w, an interest rate r, schedule of lending prices $q^f(k', b', z)$ and $q^i(k', b')$, aggregate distributions $\vartheta(k, b, z; M)$ and $\widehat{\vartheta}(k, b, M)$ of firms in the formal and informal sectors, and a mass of entrants M such that: given prices, firm's value functions and their decision rules are consistent with problems (2), (3) and (4); the free entry condition is satisfied (i.e. $W_e = 0$); lenders make zero profit for every loan type; invariant distributions of firms ϑ and $\widehat{\vartheta}$ are stationary; bond holdings $B = 0$; aggregate consumption satisfies equation (1); and the labor market clears (i.e. $1 = \int n^f(z, k) d\vartheta(k, b, z; M) + \int n^i(k) d\widehat{\vartheta}(k, b; M)$).

5 Calibration

This section calibrates the model to the US economy. The basis for this calibration can be found in Moscoso Boedo and Mukoyama (2012) and D'Erasmo (2011).

The volatility of the *high* productivity process σ_h is set to 0.2305 and the autocorrelation parameter ρ is set to 0.885 as estimated for the U.S. manufacturing sector by Cooper and Haltiwanger (2006).[19] The process will be discretized to obtain the grid for z and the transition probabilities $\eta(z'|z)$ following the method explained in Tauchen (1986). The number of grid points for z is set to 17. From the transition matrix $\eta(z'|z)$, the unconditional probability $\eta^*(z)$ is derived. The distribution of initial shocks is set to $\nu(z_0) = \eta^*(z)$.

The labor share γ is set to 0.64, a standard value, and the capital share is based on previous estimates of the degree of decreasing returns to scale at the firm level. In particular, $\alpha = 0.21$, so $\alpha + \gamma = 0.85$ as in Restuccia and Rogerson (2008). Following the literature, the risk-free interest rate r is set to 4% per year, which implies that $\beta = \frac{1}{1+r} = 0.9615$. The depreciation

[19] These parameters were estimated from registered manufacturing firms consistent with the firms that use it in the model.

rate δ is set to 7%. The value of the entry cost c_e is calibrated as in Hopenhayn and Rogerson (1993). In particular, the wage rate is normalized to 1 and used to find the value of c_e that, in equilibrium, satisfies the free entry condition with equality. Operating fixed costs are assumed to take values of $\{0, \widehat{c_f}, \infty\}$.

The parameters $\{\tau, c_\tau, \tau_w, \kappa, \lambda, \phi\}$ are taken directly from the values reported in the *Doing Business* database for the U.S. economy in 2009 (see Table 4 below). The parameters are set as follows: the tax rate $\tau = 0.23$, $c_\tau = 0.09$ and $\tau_w = 0.20$; the entry cost $\kappa = 0.26$; and the bankruptcy parameters $\lambda = 0.77$ and $\phi = 0.07$.

Five more parameters are left to calibrate: the mean of the productivity process of the *high* and *low* projects μ_h and μ_l respectively, the operating cost $\widehat{c_f}$, and the associated probabilities $\xi(\widehat{c_f})$ and $\xi(\infty)$. To obtain values for these parameters, the following moments are targeted: the size of the US informal labor force, measured as those workers not covered by a pension scheme (as reported by World Development Indicators 2006), the average size of formal establishments in the U.S. and the exit rates distribution across the size of firms. The data regarding the size distribution of establishments (in the formal sector) and exit rates in the US comes from the *Statistics of US Business* (SUBS) data set for the years 2003-2004. It is the same data used in Moscoso Boedo and Mukoyama (2012).[20]

Table 1 displays the calibrated parameters and a summary of the moments used.

[20] A description of this data set can be found in *http* : *//www.census.gov/epcd/susb/introusb.htm. Statistics of U.S. Businesses* basic data items are extracted from the Business Register, a file of all known single and multi-establishment employers maintained and updated by the U.S. Census Bureau.

Table 1: Model Parameters

Parameter		Value	Moment (US economy)
Discount Factor	β	0.9615	Avg. yearly return 5-year T-Note
Depreciation Rate	δ	0.07	Manufacturing Sector
Labor Share	γ	0.64	Labor Share
Capital Share	α	0.21	Degree of Decreasing Returns
Std Dev	σ	0.2305	Manufacturing Sector
Autocorrelation	ρ	0.885	Manufacturing Sector
Creation Cost	c_e	0.11	Entry Condition
Mean *high* process	μ_h	1.62	Avg. Operating Establishment
Low productivity	μ_l	0.762	Size Informal Sector
Positive Operating Cost	$\widehat{c_f}$	8	Exit Rate Distribution
Distribution Op. Costs	$\xi(\widehat{0})$	0.858	Exit Rate Distribution
	$\xi(\widehat{c_f})$	0.1	
	$\xi(\infty)$	0.042	

Table 2 shows moment values from the data that were used for the calibration, and those produced by the model.

Table 2: Target Moments

Moment	US Data	Model
Average Formal Est.	17.6	17.6
Informal Sector (fraction Labor Force)	7.8%	7.8%
Exit Rate Distribution		
by Employment Size	(%)	(%)
1-4	14.88	13.12
5-9	6.72	7.23
10-19	5.57	6.50
20-49	4.91	4.20
50-99	4.58	4.20
100-249	4.16	4.20
250-499	3.90	4.20
500-	4.22	4.20

Note: the size of the informal labor force is measured as those workers not covered by a pension scheme (World Development Indicators 2006). The data regarding the size distribution of establishments (in the formal sector) and exit rates in the US comes from the *Statistics of US Business* (SUBS) data set for the years 2003-2004 (see Moscoso Boedo and Mukoyama (2012)).

After the calibration exercise is done, the model is tested in different dimensions. In partic-

ular, we ask how the distribution of operating establishments over size and age in the formal sector generated by the model compares with that of the US (obtained from Business Dynamics Statistics for the year 2004). Table 3 shows the joint distribution of age and size of operating establishments as well as the unconditional distributions of establishment size and age (far right column and bottom row respectively).

Table 3: Distribution of US Formal Establishments by Age and Employment Size

Age	Young		Middle		Old		**Total Size Dist.**	
	Data	Model	Data	Model	Data	Model	Data	Model
Employment Size	(%)	(%)	(%)	(%)	(%)	(%)	(%)	(%)
1-4	13.6	0.3	20.6	9.3	14.4	19.9	**48.6**	**29.3**
5-9	2.5	7.2	9.8	10.0	9.5	9.4	**21.8**	**22.2**
10-19	1.2	3.1	6.2	11.4	6.8	5.4	**14.2**	**23.6**
20 - 49	0.7	4.0	3.9	10.5	5.0	2.6	**9.6**	**17.4**
50 - 99	0.2	0.6	1.2	3.8	1.8	0.5	**3.2**	**5.3**
100 - 249	0.1	0.2	0.6	1.6	1.0	0.1	**1.8**	**2.0**
250 +	0.0	0.0	0.2	0.2	0.5	0.0	**0.7**	**0.2**
Total Age Dist.	**18.4**	**15.4**	**42.6**	**46.7**	**39.0**	**37.9**	**100**	**100**

Note: Data corresponds to the distribution of establishments by firm size and age for the year 2004 from Business Dynamics Statistics. "Young" corresponds to 0-1 years in operation,"Middle" corresponds to 2-10 years, and "Old" corresponds to 11 years or more.

The model does a good job of generating the right distributions of operating establishments in the formal sector for both size and age. Regarding size, it generates the right number of small establishments (with less than 19 employees), but misses at the very low end of the distribution (less than 5 employees). With respect to the age distribution of formal establishments, the model is on target when compared to the fraction of young, middle, and old establishments. A deeper look at the joint distribution shows that the model under-predicts the fraction of young establishments in the smallest size category. The reason is that the productivity threshold to enter the formal sector endogenously generates young establishments that are relatively more productive and therefore larger than observed in the data. On the other hand, the model yields a distribution of middle and old establishments across sizes that closely resembles that observed in the data. By construction, the average entry rate and exit rate in the model are identical. Their value equals 8%. The entry and exit rates in the data are 11.1% and 10.2% respectively.

Thus, compared to the US data, the model average entry and exit rates are three and two percentage points lower respectively.[21]

6 The Effects of Country Specific Institutions

What are the changes in total factor productivity that can be attributed to institutional differences across countries? Our focus is on measured differences in the cost of entry to the formal sector, the tax structure, and the efficiency of debt enforcing mechanisms. The experiment can be interpreted as a counterfactual in which the effects of imposing country specific frictions onto the US economy are measured in steady state. Due to the high computational burden of the exercise, the number of observations is limited by grouping countries by income level following the World Bank's definition: High Income Countries (HIC) and Developing Countries, where Developing Countries are classified as Upper Middle Income Countries (UMIC), Lower Middle Income Countries (LMIC) and Low Income Countries (LIC).[22]

The experiment can be described as follows. First, calibrate the model to the US economy by using $(\lambda, \phi, \tau, c_\tau, \tau_w, \kappa)_{US}$.[23] Next, for each income group, $(\lambda, \phi, \tau, c_\tau, \tau_w, \kappa)_g$ are adjusted, where $g \in \{HIC, UMIC, LMIC, LIC\}$ and a new equilibrium is computed.[24] In order to implement it, the *Doing Business* database for the year 2009 is used to obtain the median $(\lambda, \phi, \tau, c_\tau, \tau_w, \kappa)$ for each income group. Table 4 shows parameter values for the US economy (used in the benchmark calibration) and those of High, Upper Middle, Lower Middle and Low

[21] The distance between the model and data in terms of size distribution of young firms, entry and exit rates is partly due to the way the data is collected. In the data, establishments are observed at one point in time. Those establishments that are less than one year old, are considered entrants. However, the model counterpart for entrant establishments is defined as those establishments that are exactly one year old.

[22] Roughly, countries are classified as HIC if their GNI per capita is higher than 25% of the US, UMIC if their GNI per capita falls between 8% and 25% of the US, LMIC if their GNI per capita falls between 2% and 8% of the US and LIC if their GNI per capita is below 2% of the US.

[23] In this case, the wage is normalized to one, and then the set of loan prices $q_j^f(k', b', z)$ and $q^i(k', b')$ are obtained through iteration, until lenders make zero profit on each contract, and the mass of potential entrants M that clears the labor market is found together with the value of entry cost c_e that satisfies the zero entry condition.

[24] More specifically, the wage rate w and loan prices $q_j^f(k', b', z)$ and $q^i(k', b')$ are obtained through iteration until lenders make zero profits and the zero entry condition is satisfied (given c_e obtained for the US). Finally, the mass of entrants M is adjusted in each case to clear the labor market.

Income countries.

Table 4: Frictions across income groups

	λ	ϕ	τ	c_τ	τ_w	κ
US	0.77	0.07	0.23	0.09	0.20	0.26
High (HIC)	0.72	0.08	0.18	0.07	0.28	1.08
Upper Middle (UMIC)	0.30	0.15	0.17	0.10	0.37	1.33
Lower Middle (LMIC)	0.25	0.15	0.17	0.14	0.31	5.08
Low (LIC)	0.15	0.09	0.20	0.13	0.23	7.03

Note: Countries are classified following the World Bank's income groups. Countries are HIC if their GNI per capita is higher than 25% of the US, UMIC if their GNI per capita falls between 8% and 25% of the US, LMIC if their GNI per capita falls between 2% and 8% of the US and LIC if their GNI per capita is below 2% of the US. Median values for each group and friction are reported.

In order to understand the effects of country-specific frictions on firm-level productivity, we start by analyzing aggregate Total Factor Productivity following cross-country studies such as Klenow and Rodriguez-Clare (1997) or Hall and Jones (1999). They compute the following equation:

$$TFP = \frac{Y}{K^{\alpha} H^{(1-\alpha)}}$$

where Y denotes aggregate output, K denotes aggregate capital, H denotes some aggregate for labor (usually adjusted for human capital), and α is the capital share. In the model, aggregate output is the sum across both formal and informal establishments, aggregate capital is the sum of capital across establishments in both sectors and our aggregate measure of labor equals one.[25] The same parameter share as in Hall and Jones (1999) is used, which equals 1/3.

[25] In the online Appendix, the model is used to analyze the measurement issues previously discussed.

Table 5: Differences Across Countries

Panel (a): Main Results

			Developing Countries					
	HIC		UMIC		LMIC		LIC	
	Data	Model	Data	Model	Data	Model	Data	Model
	Data	Model	Data	Model	Data	Model	Data	Model
TFP	0.91	0.90	0.70	0.82	0.44	0.75	0.31	0.75
Informal labor force (%)	8.8	43.2	45.0	71.8	71.7	95.8	95.0	95.0
Output per eff. Worker	0.91	0.85	0.49	0.76	0.25	0.66	0.11	0.66

Note: TFP and Output per effective worker are reported relative to the US value. Data is from author's calculations based on Hall and Jones (1999), see the online Appendix. One unit of effective worker equals one unit of human capital. The size of the informal labor force is taken from the *World Development Indicators*(2006) as the share of the labor force not covered by a pension scheme.

Panel (b): Other Moments

			Developing Countries					
	HIC		UMIC		LMIC		LIC	
	Data	Model	Data	Model	Data	Model	Data	Model
Avg. employment formal	11.1	37.1	129.8	30.9	175.0	80.5	386.4	96.8
ln(Var employment formal)	10.5	8.4	12.7	8.0	12.7	9.8	13.6	9.4
Capital per eff. Worker	1.05	0.84	0.38	0.78	0.18	0.70	0.04	0.69
Formal Entry Rate	0.81	0.63	0.65	0.64	0.62	0.60	0.47	0.58
Business Density	1.62	0.29	0.93	0.17	0.31	0.01	0.03	0.01
Domestic Credit to Private Sector (% GDP)	54.9	65.9	21.3	30.3	16.0	3.9	7.5	4.0

Note: Capital per effective worker, Formal Entry Rate, Business Density and Domestic Credit to Private Sector are reported relative to the US value. Data on average employment and variance of employment is taken from Alfaro et. al. (2009). Capital per effective worker is from author's calculations based on Hall and Jones (1999), see the online Appendix. One unit of effective worker equals one unit of human capital. Data on the Formal Entry Rate and Business Density are taken from the 2008 World Bank Group *Entrepreneurship Survey and Database*. The model counterpart is obtained as total formal labor force over the average size of formal establishments which equals the measure of formal establishment to total population. Domestic Credit to GDP is also taken from the *World Development Indicators* (average 2004-2007). Domestic credit to private sector in the model is computed as the ratio of formal debt to total output.

6.1 Quantitative Results

Panel (a) of Table 5 displays the main results for each income group and compares the model to the data for the median country in each group. The most important result of the paper is that measured institutional differences between Developing Countries and the U.S. generate a decline in TFP of up to 25%. In particular, it generates a reduction in TFP of 18%, 25%, and 25% when the model is solved using institutional parameters corresponding to the median Upper Middle, Lower Middle and Low Income Country respectively (see the first row in Panel (a) of Table 5).[26] The sources of observed productivity differences and the role of each friction will be analyzed in what follows. In short, allocative efficiency and the share of output produced by firms in the informal sector play a key role.

In terms of informal activity, the model generates sizable informal sectors that are negatively correlated with GDP per worker, as observed in the data. The model delivers an informal labor force that is on target across income levels, ranging from around 10% in the US to 95% at the low end of the income distribution. However, the model overshoots the data in the middle of the income distribution.[27]

The model also generates important changes in output per effective worker (up to 34% for developing countries) but the portion not explained by the model is more significant than in the case of total factor productivity. This discrepancy comes from differences of the same order of magnitude in capital per effective worker that result from the fact that lenders in each country have access to the same risk free rate (see Panel (b) of Table 5).

Panel (b) of Table 5 presents other moments across income groups that provide a natural test for our model. One of the main channels affecting capital reallocation is the process of entry into and exit out of the formal sector. In the model, as frictions increase, the exit rate (and the entry rate) decreases. For example, the exit rate in the US is approximately 180%

[26] In order to isolate the effects on total factor productivity that result from institutional differences, the technology possibility frontier was kept constant across countries. This and other factors such as differences in human capital accumulation, while important to understand the gap between the model generated values and the data, are beyond the scope of this paper.

[27] The model can generate the right size of the informal sector by including tax enforcement. This extension was considered and the broad nature of the results changes little but complicates the analysis.

of that of LMIC as observed in the data. In Low Income Countries, firms stay in business for much longer, preventing the natural process of churning of unproductive firms. Lower wages and higher productivity thresholds to enter the formal sector generate larger and more dispersed firms in countries with large frictions. The model is on target both for both average size and variance of size (measured by employees) as reported by Alfaro et. al. (2009). As in the data, the business density drops to 1% of the US's for Low Income Countries.

As in the data, the model generates a sharp decrease in the stock of domestic credit to private sector as a percentage of GDP. In the data for developing economies, domestic credit to private sector ranges from 21% (UMIC) to 7.5% (LIC) relative to the US, whereas the model counterpart goes from 30% to 4%. The model moment includes only the stock of formal credit because the data contains loans from formal entities, and to our knowledge there is no accurate measure of the stock of informal credit across countries.[28]

6.2 Dispersion of marginal product of capital

Understanding how capital is allocated across establishments in the formal sector is central to the analysis, because all measured institutional differences across countries relate to firms in this sector. In their influential paper, Hsieh and Klenow (2009) observe that the standard deviation of the log of the marginal product of capital is greater in India and China that in the US, and conclude that TFP would increase by between 30% and 60% if the dispersion was to adjust to US levels.[29] This paper generates dispersion in the marginal product of capital that is comparable to that documented by Hsieh and Klenow (2009).

[28]The data on private domestic credit includes not only business loans but also personal loans, so these values should be taken as an approximation of the observed relationship between firms' credit and country income.

[29]Note that more dispersion in the marginal product of capital reflects a lower degree of allocative efficiency because a planner would distribute capital to equalize it across firms.

Table 6: Dispersion of Marginal Product of Capital - Formal sector

	US	HIC	UMIC	LMIC	LIC
$var(ln(MPK))$	0.23	0.24	0.28	0.31	0.32

Note: $var(ln(MPK))$ corresponds to the variance of log marginal product of capital.

Table 6 shows that the variance of the marginal product of capital in the LMIC and LIC is 23% and 29% higher than in the US. In the case of China and India (countries that are close to our LMIC and LIC respectively), Hsieh and Klenow (2009) find that the dispersion of marginal product of capital is 13% and 46% higher than in the US. In this case it is the country-specific institutions that endogenously generate this dispersion and point towards inefficiencies in the allocation of capital and labor, even within the formal sector.[30]

Still in the formal sector, the allocation of capital and labor across establishments with heterogeneous productivity can be analyzed against a frictionless benchmark economy with commitment. In a frictionless world with commitment, the Modigliani-Miller theorem applies and optimal allocations can be derived from a static problem. Conditional on surviving, firms solve:

$$\max_{k,n} \left\{(1-\tau)(zk^{\alpha}n^{\gamma} - w(1+\tau_w)n) - (r+\delta)k\right\}$$

The solution to this problem implies that the capital-labor ratio for each country is constant across firms and depends only on factor prices, i.e independent of productivity. More specifically, $\widetilde{(k/n)} = \frac{\alpha}{\gamma}\frac{w(1-\tau)(1+\tau_w)}{(r+\delta)}$. Using a notion of efficiency that is similar to the one used below to study productivity, a measure of the capital to worker ratio in the formal sector is defined as follows: $\widehat{(k/n)} = \overline{(k/n)} + cov((k/n)_s, \omega_s^f)$. This measure captures differences in prices, and as before can be decomposed into a "mean" effect ($\overline{(k/n)}$) and a "variation" effect ($cov((k/n)_s, \omega_s^f)$). An

[30]To maintain consistency with the other tables in this paper, the dispersion of the marginal product of capital across income groups was reported. However, the exercise was also conducted for the China and India specific measured frictions and the values are similar to those reported by Hsieh and Klenow (2009). More specifically, the $var(ln(MPK))$ for China equals 0.257 (7.4% higher than U.S.) and the $var(ln(MPK))$ for India equals 0.3267 (36.13% higher than U.S.). The model generates the right distribution of marginal product of capital while also being consistent with the size of the informal labor force in both China and India.

efficient allocation will imply a covariance equal to zero. Table 7 displays the values of $\widehat{(k/n)}$ and its decomposition for each income group.

Table 7: Capital Per Worker Decomposition in the Formal Sector

Group	$\widehat{(k/n)}$	$\overline{(k/n)}$	$cov((k/n)_s, \omega_s^f)$	$\frac{\widehat{(k/n)}}{\overline{(k/n)}}$
US	2.29	2.63	-0.34	0.84
HIC	2.19	2.89	-0.70	0.81
UMIC	2.09	2.82	-0.73	0.76
LMIC	1.71	2.62	-0.90	0.69
LIC	1.50	2.35	-0.85	0.66

Note: $\widehat{(k/n)}$ corresponds to output weighted capital to labor ratio. $\overline{(k/n)}$ is the arithmetic mean of capital to labor ratio. $cov((k/n)_s, \omega_s^f)$ is the covariance between the capital to labor ratio and output shares. $\widetilde{(k/n)}$ is the "efficient" level capital to output ratio (derived from the solution to a frictionless problem).

The output-weighted capital to labor ratio decreases in the formal sector as institutional parameters are set to those of less developed economies. Part of this decrease is related to the reduction in after-tax wages across countries that affect the efficient ratio, $\widetilde{(k/n)}$, as well as the model average, $\overline{(k/n)}$. But as the last column of Table 7 shows, the departure from optimal levels increases as we move towards lower income countries. This suggests the existence of higher effective average interest rates in the poorer countries, completely generated by heterogeneity in financial frictions. Also note that most of this decrease in $\widehat{(k/n)}$ comes from the covariance term. Low income countries display a larger covariance term (in absolute value), implying that large firms substitute away from capital and towards labor more than small firms. This is the result of differences in the endogenous firm specific schedule of loan prices.

To continue with the productivity analysis, but now moving to all firms in the economy, Figure 3 shows how output is distributed over firms' productivity z.

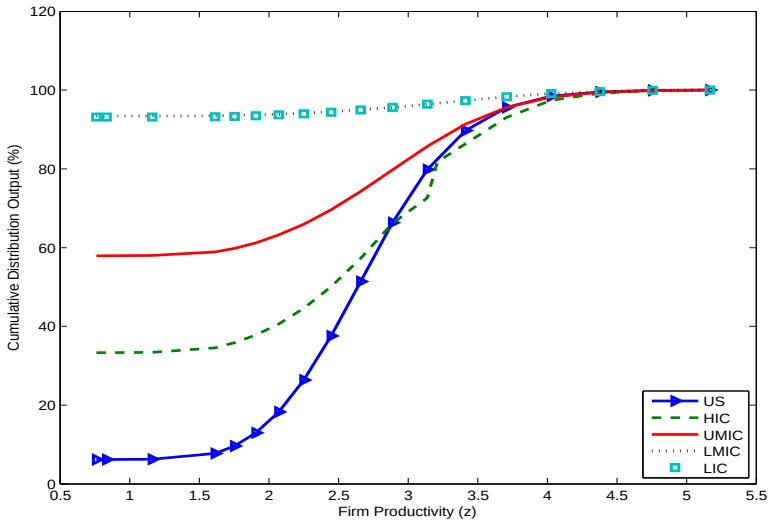

Figure 3: Distribution of Output over Productivity across Income Groups

The increase in the entry cost κ raises the entry threshold to the formal sector[31], generating a reallocation towards more productive firms in that sector. However, as this cost rises the share of production in the informal sector (less productive firms) also increases. The latter effect dominates, that is when distortions are higher the share of output produced by low productivity firms increases. For example, firms with productivity less than or equal to 1.5 ($z \leq 1.5$) account for about 10% of output in the US, 35% in HIC, 60% in UMIC and around 90% in LMIC and LIC.

6.3 Firm Level Productivity and Misallocation

It is crucial to provide a measure that captures how efficiently resources are allocated in the economy. To address this issue, weighted average plant level productivity is decomposed following Olley and Pakes (1996) (also used by Bartelsman et al (2009) for example):

$$\widehat{z} = \int z_s \omega_s ds = \Omega \mu_u + (1 - \Omega)[\overline{z} + cov(z_s, \omega_s^f)]$$

[31]The productivity entry threshold to the formal sector goes from 1.32 standard deviations from the log-mean of the h productivity process in the US to 1.64 and 5.78 in the UMIC and LIC respectively.

where $\widehat{z}$ is the average of plant level productivity weighted by output share, Ω is the informal share of output, ω_s is the output share of each establishment, ω_s^f is the output share of each establishment in the formal sector, and $\overline{z}$ is the unweighted mean productivity in the formal sector. Therefore, the output-weighted productivity can be decomposed into three terms. First is the effect of informal activity given by Ω, and then the formal-weighted productivity which can be decomposed into the unweighted average of firm-level productivity plus a covariance between output share and productivity. The covariance captures allocative efficiency within the formal sector because it reflects the extent to which firms with higher than average productivity have a greater market share. Table 8 displays the values of this decomposition across income groups.

Table 8: Firm Productivity Decomposition

Group	$\widehat{z}$	Ω	$\overline{z}$	$cov(z_s, \omega_s^f)$
US	2.69	0.06	2.22	0.61
HIC	2.27	0.37	2.17	0.85
UMIC	1.70	0.65	2.18	0.80
LMIC	0.93	0.94	2.26	1.08
LIC	0.93	0.94	2.24	1.06

Note: $\widehat{z}$ corresponds to output weighted firm level productivity. $\overline{z}$ is the arithmetic mean of firm level productivity. Ω is the informal share of output. $cov(z_s, \omega_s^f)$ is the covariance between firm level productivity and its output share.

Note that the value of output-weighted productivity correlates with the value of measured TFP. As distortions increase, the value of $\widehat{z}$ decreases. This effect is generated by massive shifts to the informal economy, which goes from 7% in the US to 94% in the median Low Income Country. However, the output-weighted productivity of the formal sector increases in the poorer countries (the sum of $\overline{z}$ and $cov(z_s, \omega_s^f)$). This is the direct result of higher productivity entry thresholds to the formal sector together with lower wages in the poorer countries. As observed before, misallocation of capital increases with the cost of doing business and poorer countries are less efficient in assigning resources (see Tables 6 and 7). However, the selection effect into the formal sector dominates the misallocation effect producing this counterfactual result (i.e. an increase in the output-weighted productivity of the formal sector as frictions increase).

6.4 How Important is the Informal Sector for Productivity?

This section analyzes how much the existence of the informal sector contributes to the main results. More specifically, it compares the benchmark economy with a model in which firms do not have the option of operating in the informal sector. As in the standard Hopenhayn (1992) model, potential entrants choose only between activating their technology or staying out of the market. If they decide to enter, firms are created as formal firms.[32] Table 9 presents a comparison between the data, the benchmark economy and the simplified version of the model for the case of LMIC.[33]

Table 9: Counter Factual: No Informal Sector Model

	Data	Benchmark	No Informal Sector
TFP	0.44	0.75	0.80
Informal labor force (%)	71.7	95.8	-
Output per eff. Worker	0.25	0.66	0.70
Avg. Employment Formal	175.0	80.5	115.8
Formal Entry Rate	0.62	0.60	0.54
Var(ln(MPK)) Formal Sector	-	1.35	1.20
Entry Threshold (z)	-	4.38	3.71
Wage Rate	-	0.78	0.64

Note: TFP, Output per effective worker, Capital per effective worker, Formal Entry Rate and Var(ln(MPK)) are reported relative to the US value. See also notes in Tables 5 and 6.

This table shows that the introduction of an informal sector is quantitatively important for explaining productivity differences across countries. The benchmark experiment generates a decline in TFP of about 25% when the model is parameterized with measured institutions for LMIC, whereas a model with no informal sector generates a drop in TFP that is 20% smaller (1-0.80 vs 1-0.75). Both the benchmark model and the model with no informal sector do not generate enough differences in output per effective worker when compared to the data. Output per effective worker relative to the U.S. is 6% lower in the benchmark economy than in the

[32] In order to make the comparison easier to understand, the simplified model is not "re-calibrated " and solved using the same parameters obtained from the benchmark calibration. However, the predictions of this simplified model are not far from the true data values used as original targets. For example, the average of the formal establishment in the US in the simplified model equals 17.95 workers (vs 17.60 in the data).

[33] This section presents only the LMIC case to simplify the analysis. All other cases produce similar results.

counterfactual case.

In line with what other studies in the literature have found (see for example Buera, Kaboski and Shin (2011)), the model with no informal sector generates a decrease in aggregate measured productivity and output per worker as frictions increase. The introduction of a technology choice and the fact that firms are allowed to operate in the informal sector improves upon the previous results. There are two effects going in opposite direction that make the introduction of the informal sector non-trivial. On one hand, in countries where the costs of doing business are high, a large fraction of total output is being produced by lower productivity firms in the informal sector, which is reflected in a lower measured TFP. On the other hand, the existence of the informal sector shifts the labor demand curve and keeps the equilibrium wage rate in the benchmark economy higher than in the model with no informal sector (22% higher). This is a direct result of introducing an additional choice at the entry point. In the model with an informal sector, all potential entrants effectively produce (either formally or informally), whereas in the case without an informal sector the choice boils down to producing formally or not entering at all, and therefore the mass of firms that effectively produce is smaller.

A higher wage rate makes the substitution between workers and capital more expensive (note the differences in average firm size across models), resulting in a significantly higher variance of log marginal productivity of capital (12.5% increase in the relative variance of log-MPK). Moreover, a higher wage rate also results in a higher entry threshold into the formal sector and a relatively lower exit rate that implies that at the calibrated parameters, on average, productivity in the formal sector is higher once the informal sector is in place than when is not.

6.5 Effects of each friction

In order to understand the individual effects of each institution (entry costs, tax structure and bankruptcy efficiency), isolated changes from US institutions to LMIC institutions are analyzed.[34] Each friction affects aggregate moments through different channels. In summary,

[34]LMIC parameter values provide a natural benchmark for understanding the results because they are noticeably different from those of the US (the calibrated economy) but are within the observed range.

by moving from US to LMIC parameters, the entry costs are responsible for roughly 3/4 of the gap in total factor productivity generated by the model and the bankruptcy efficiency parameters and taxes are responsible for 1/4 of that gap.

Table 10: Effect of Each Friction

	US	κ_{LMIC}	$\{\lambda, \phi\}_{LMIC}$	$\{\tau, c_\tau, \tau_w\}_{LMIC}$	LMIC
TFP	1	0.79	0.94	0.97	0.75
Informal labor force (%)	7.8	78.4	27.1	17.9	95.8
Output per eff. Worker	1	0.71	0.87	0.96	0.66
$var(ln(MPK))$	0.24	0.257	0.268	0.238	0.31

Note: TFP and Output per effective worker are reported relative to the US value. Model TFP is calculated as $TFP \equiv \frac{Y}{K^{\widetilde{\alpha}}}$ where $\widetilde{\alpha} = 1/3$ is taken from Hall and Jones (1999). See Table 4 for specific parameters.

Entry cost: The second column of Table 10 reports the effects of only changing the formal sector entry costs. Changes in the entry cost directly affect the productivity threshold that make firms indifferent between the formal and informal sectors. This change (and the implied effect on equilibrium functions) produce a large shift to the informal sector (from 7.8% of the labor force to 78.4%) and a relatively small effect on the efficiency of the formal sector's resource allocation (an increase in 7% in the marginal product of capital). These effects are responsible for the 21% and 29% decreases in TFP and output per worker, respectively.

Bankruptcy efficiency: The third column of Table 10 reports the effects of only changing the bankruptcy efficiency parameters. TFP declines 6% and output per worker declines 13% (approximately 24% and 38% of the total fall in TFP and output per worker respectively). These effects are due to an increase in informal activity (the informal labor force increases to 27.1%) and also to an inefficient resource allocation within the formal sector. The variance of the log of the marginal product of capital increases 12% (approximately double the effect when compared to the effects of entry costs). This is responsible for almost 1/3 of the gap between the US and the LMIC in terms of the dispersion of the marginal product of capital in the formal sector.

Tax structure: The fourth column of table 10 reports the effects of only changing the tax

structure parameters. Although there are large changes in parameters (26% decrease in the tax rate on profits and 55% increase in the payroll tax rate and in the cost of tax compliance), the results in terms of total factor productivity are smaller than the effects coming from changes in entry costs and bankruptcy efficiency parameters. Informal activity increases from 7.8% to 17.9% while output and productivity decrease by only 3 and 4 percent, respectively (approximately 10% of the total decrease). In the formal sector, the dispersion of the marginal product of capital is also unaffected.

7 Conclusion

A theory of total factor productivity based on measured capital market imperfections and costs of creating and operating formal sector firms was presented. The model predicts that countries with a low degree of debt enforcement and high costs associated with the formal sector are characterized by low allocative efficiency and a large share of output produced by low-productivity firms in the informal sector.

We find that this mechanism is quantitatively important. When institutions (entry costs, debt enforcement and tax structure) are parameterized using the World Bank *Doing Business* database, the model generates a decline in total factor productivity of up to 25% (relative to the US economy). As in the data, the model delivers a strong negative correlation between income per-worker and the size of the informal sector. Incorporating the informal sector into the model increases its explanatory power by approximately 25%.

8 References

Alfaro, L., Charlton, A. and Kanczuk, F., 2009. Plant-Size Distribution and Cross-Country Income Differences. In: NBER International Seminar on Macroeconomics 2008. NBER, Cambridge, Mass.

Antunes, A., Cavalcanti, T., 2007. Start up costs, limited enforcement, and the hidden economy. European Economic Review 51, 203-224.

Arellano, C., Bai Y., Zhang, J., 2010. Firm Dynamics and Financial Development. Manuscript, University of Minnesota.

Atkeson, A., Kehoe P., 2007. Modeling the transition to a new economy: lessons from two technological revolutions. American Economic Review 97(1), 64-88

Barseghyan, L., DiCecio R., 2010. Entry Costs, Misallocation, and Cross-Country Income and TFP Differences. Working Paper 2009-005A, Federal Reserve Bank of St. Louis.

Barro, R., Lee, J., 2000. International Data on Educational Attainment: Updates and Implications. CID Working Paper No. 42

Bartelsman, E., Haltiwanger J., Scarpetta, S., 2009. Cross-country differences in productivity: The role of allocative efficiency. NBER Working paper #15490. Cambridge, Mass.

Bertrand, M., Kramarz, F., 2002. Does Entry Regulation Hinder Job Creation? Evidence From The French Retail Industry. The Quarterly Journal of Economics 117(4), 1369-1413.

Bruhn, M., 2008. License to sell : the effect of business registration reform on entrepreneurial activity in Mexico. Policy Research Working Paper Series 4538, The World Bank, Washington.

Buera, F., Kaboski, J., Shin, Y., 2011. Finance and Development: A Tale of Two Sectors. American Economic Review, forthcoming.

Castro, R., Clementi, G., MacDonald, G., 2009. Legal Institutions, Sectoral Heterogeneity, and Economic Development. Review of Economic Studies 76(2), 529-561.

Cooley, T., Quadrini V., 2001. Financial Markets and Firm Dynamics. American Economic Review 91(5), 1286-1310.

Cooper, R., Haltiwanger J., 2006. On the Nature of Capital Adjustment Costs. Review of Economic Studies 73(3), 611-633.

D'Erasmo, P., 2011. Investment and Firm Dynamics. Mimeo. University of Maryland.

Djankov, S., La Porta R., Lopez-De-Silanes F., Shleifer A., 2002. The Regulation Of Entry. The Quarterly Journal of Economics 117(1), 1-37.

Djankov, S., Hart, O., McLiesh, C., Shleifer, A., 2008. Debt Enforcement Around the World. Journal of Political Economy 116(6) 1105-1149.

Djankov, S., McLiesh, C., Ramalho, R., Shleifer, A., 2010. The Effect of Corporate Taxes on Investment and Entrepreneurship. American Economic Journal: Macroeconomics 2(3), 31-64.

Duarte M., Restuccia D., 2010. The Role of the Structural Transformation in Aggregate Productivity. The Quarterly Journal of Economics 125(1), 129-173.

Erosa, A., Hidalgo Cabrillana, A., 2008. On Finance as a Theory of TFP, Cross-Industry Productivity Differences, and Economic Rents. The International Economic Review 49(2), 437-473.

Foster, L., Haltiwanger, J., Krizan C., 2001. Aggregate Productivity Growth: Lessons from Microeconomic Evidence. In: Hulten, C., Dean, E., Harper, M. (Eds.), New Developments in Productivity Analysis. University of Chicago Press, Chicago.

Guner, N., Ventura, G., Xu Y., 2008. Macroeconomic implications of size-dependent policies. Review of Economic Dynamics 11, 721-744.

Hall, R., Jones, C., 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114(1), 83-116.

Hallward-Driemeier, M., Khun-Jush, G., Pritchett, L., 2010. Deals Versus Rules: Policy Implementation Uncertainty and Why Firms Hate It. NBER Working paper #16001. Cambridge, Mass.

Heston, A., Summers, R., Aten, B., 2009. Penn World Table Version 6.3, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

Hopenhayn, H., 1992. Entry, exit, and firm dynamics in long run equilibrium. Econometrica 60(5), 1127-1150.

Hopenhayn, H., Rogerson, R., 1993. Job Turnover and Policy Evaluation: A General Equilibrium Analysis. Journal of Political Economy 101, 915-938.

Hsieh, C., Klenow, P., 2009. Misallocation and Manufacturing TFP in China and India. Quarterly Journal of Economics 124, 1403-1448.

Klenow, P., Rodriguez-Clare, A., 1997. The Neoclassical Revival in Growth Economics: Has It Gone Too Far? In: Bernanke, B., Rotemberg, J., (Eds.) NBER Macroeconomics Annual 1997, Cambridge, MA, pp. 73-102.

La Porta, R., Shleifer, A., 2008. The Unofficial Economy and Economic Development. Brookings Papers on Economic Activity.

Loayza, N., 1996. The economics of the informal sector: a simple model and some empirical evidence from Latin America. Carnegie-Rochester Conference Series on Public Policy 45, 129-162

McKenzie, D., Seynabou Sakho, Y., 2010. Does it pay firms to register for taxes? the impact of formality on firm profitability. Journal of Development Economics, 91(1), 15-24.

Moscoso Boedo, H., Mukoyama, T., 2012. Evaluating the Effects of Entry Regulations and Firing Costs on International Income Differences. Journal of Economic Growth, *forthcoming*.

OECD, 2002. Measuring the Non-Observed Economy: A Handbook. OECD, Paris.

Olley, S., Pakes, A., 1996. The Dynamics of Productivity in the Telecommunications Equipment Industry. Econometrica 64(6), 1263-1297.

Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., Saavedra-Chanduvi, J., 2007. Informality: Exit and Exclusion. The World Bank, Washington.

Pratap, S., Quintin, E., 2008. The Informal Sector in Developing Countries: Output, Assets, and Employment. In: Davies, J. (Ed) Personal Wealth from a Global Perspective. Oxford University Press, New York, USA, 373-394.

Quintin, E., 2008. Contract Enforcement and the Size of the Informal Economy. Economic Theory 37, 395-416.

Rauch, J., 1991. Modelling the informal sector formally. Journal of Development Economics 35(1), 33-47

Restuccia, D., Rogerson, R., 2008. Policy Distortions and Aggregate Productivity with Heterogeneous Establishments. Review of Economic Dynamics 1124, 707-720.

Schneider, F., Enste, D., 2000. Shadow Economies: Size, Causes, and Consequences. Journal of Economic Literature 38, 77-114.

Schneider, F., Buehn, A., Montenegro, C., 2010. Shadow Economies all over the World. New Estimates for 162 Countries from 1999 to 2007. World Bank, Policy Research Working Paper 5356. Washington

Syverson, C., 2004. Substitutibility and Product Dispersion. Review of Economics and Statistics 86(2), 534-550.

Tauchen, G., 1986. Finite state Markov-chain approximation to univariate and vector autoregressions. Economics Letters 20, 177-81.

Tybout, J., 2000. Manufacturing Firms in Developing Countries: How Well Do They Do, and Why? Journal of Economic Literature 38, 11-44.

United Nations, 1993. System of National Accounts 1993. United Nations, New York.

United Nations, 2008. Non-Observed Economy in National Accounts. United Nations, New York.

United Nations, 2008b. National Accounts Statistics: Analysis of Main Aggregates. United Nations, New York.

Appendix

This appendix presents a description on how measured aggregate TFP is constructed, how the country level measured institutions are parametrized using the World Bank's Doing Business data set and an explanation on how aggregate data is affected by the existence of an underground economy is presented. In the final section, the model is used to address the problem of measured aggregate TFP in the presence of the informal sector.

Aggregate Data

The data on physical capital, human capital and total factor productivity was constructed following Hall and Jones (1999). This is essentially an update on Hall and Jones (1999) using Heston, Summers and Aten (2009) and Barro and Lee (2000). This allows for the construction of series of physical capital up to 2007 and human capital up to 1999.

Total output and output per worker are both net of mining as in Hall and Jones (1999). The mining share was obtained from UN (2008b) for the years 1970 to 2008 for all the countries except for Taiwan. In the case of Taiwan it was obtained from the Taiwanese statistical authority for the years 2002 to 2008. For the years not available in the data, average for the observed years were imputed.

Total output is obtained by multiplying real output per capita times total population (variables rgdpl and pop in Heston, Summers and Aten (2009))net of mining share.

Output per worker refers to real GDP per worker in Heston, Summers and Aten (2009), variable rgdpwok, net of mining share for the year 2007 for all the countries other than Bahrain and Iraq. The data for Bahrain is for the year 2006 and for Iraq it is for the year 2003

Physical capital stock in 2007 is constructed using the perpetual inventory method. The sample was limited to countries with investment share (ci in Heston, Summers and Aten (2009)) starting on or before 1990 and all available investment data was used. The initial value of capital stock was computed as investment in the initial year over depreciation rate plus the average geometric growth rate of investment for the first 10 years of data. Following Hall and Jones (1999), a depreciation rate of 6% was assumed.

Effective units of labor (H in Hall and Jones (1999)), are constructed using average educational attainment, measured in 1999 for the population aged 25 and over, as reported by Barro and Lee (2000)

Total factor productivity is generated after assuming the following production function

$$TFP = \frac{Y}{K^{\alpha} H^{(1-\alpha)}}$$

Where Y denotes aggregate output, K denotes aggregate capital, H denotes the units of effective labor and α is the capital share (1/3 as in Hall and Jones (1999)).

101 countries remain in the sample, with data on physical capital, effective labor, and output.

Data on informal labor force is obtained from the World Development Indicators 2006. It refers to the fraction of the labor force not covered by a pension scheme. Of the 101 countries with full data explained above, 75 are left with information regarding informal labor force.

[Locate Table 11 about here]

Frictions

The data on the frictions used in the model is obtained from the *Doing Business* database maintained by the World Bank for the year 2009.[35]

The model has six frictions. The recovery rate in the event of a default (λ), the default proceedings costs (ϕ), the profit tax rate (τ), the payroll tax rate (τ_w), the cost of paying taxes (c_τ), and the entry cost to the formal sector (κ).

The recovery rate (λ) is obtained directly from the *closing a business* section of *Doing Business* and measures the cents on the dollar recouped by creditors through the bankruptcy, insolvency or debt enforcement proceedings.

The default proceedings costs (ϕ) are also obtained directly from the *closing a business* section of *Doing Business* and measure the cost of proceedings as a percentage of the estate's value (we assume it to be equal to the stock of capital of the firm at the moment it declares default).

The profit tax rate (τ) and the payroll tax rate (τ_w) are obtained from the *paying taxes* section of *Doing Business*. The payroll tax rate (τ_w) comes from the column *Labor tax and contributions* in *Doing business*. Because both tax rates are expressed as a function of profits, they need to be adjusted and the labor tax rate expressed as a function of payroll. To do that, the standardized balance sheet and income statements was used to construct the exercise as explained in table 1 of Djankov et. al. (2010).

The cost of paying taxes (c_τ) is obtained from the *paying taxes* section of *Doing Business*. It is generated from the *Time (hours per year)* column. It is re-scaled to annual wages assuming one worker has to be employed full time during those hours. For example, if it takes 40 hours to pay taxes in country x, then our measure of c_τ equals 1/52 (assuming that a full time worker works 40 hours per week).

Finally the entry cost (κ) is constructed from the *Starting a Business* and *Dealing with Construction Permits* sections of *Doing Business*. From *Starting a Business*, columns *Time (Days)* and *Cost (% of income per capita)* are used and from *Dealing with Construction Permits*, columns *Time (Days)* and *Cost (% of income per capita)* are used. Both time measures are in terms of calendar days. Both are transformed to wages by assuming that a worker has to be fully employed during those days. Both time dimensions are added to both costs dimensions assuming that income per capita and wages are perfectly correlated. For example, if in country x it takes 365 days to start a business and costs 50% of GNI per capita, and it takes 730 days to deal with permits and it costs 100% of GNI per capita, our measure of κ equals 4.5.

[Locate Table 12 about here]

Aggregate GDP and informal sector measurement

This section is devoted to understanding how aggregate data is affected by the existence of an underground economy. In particular we concentrate the attention on the effects of the Non-Observed Economy on the official GDP figures published by the statistical authorities.

Regarding the GDP definition and measure (key to generating other related measures such as TFP), the UN System of National Accounts 1993 manual states in its point 6.36: *"Because certain kinds of producers try to conceal their activities from public authorities it does not*

[35]Data freely available online at www.doingbusiness.org.

follow that they are not included in national accounts in practice", and continues, *"Because the underground economy may account for a significant part of the total economy of some countries, it is particularly important to try to make estimates of total production which include it, even if it cannot always be separately identified as such."*

This clearly indicates that from a definition point of view, informal activities are within the production boundary and should be included in the GDP measures.

Given that the agents involved in the underground economy may conceal their activity, it could be very hard for the statistical authorities to actually include the output of the informal economy in the aggregate GDP. To alleviate the problem and help agencies adhere to the UN guidelines, in 2002 the OECD released its Handbook for Measurement of the Non-Observed Economy (OECD (2002)) as the basis for the work in this area. It provides guidelines for the estimation of the Non-Observed economy through multiple approaches (Production, Expenditure and Income approaches).

In order to actually estimate the Non-Observed economy, data from different sources must be confronted and analyzed. OECD (2002) gives examples of data confrontation that can be used to estimate the size of the Non-Observed Economy, which include:

- *"Enterprise survey data versus taxation data; wages paid versus taxes raised; sales of goods and services due to VAT versus VAT raised, and production versus production related taxes.*
- *Enterprise survey data about the production of commodities versus enterprise survey data about purchases of commodities; supply of goods and services versus the use of them.*
- *Expenditure survey data versus retail trade survey data; household expenditures versus retail trade.*
- *Expenditure survey data versus income or taxation data; household expenditures versus available income.*
- *Enterprise survey data versus labour force survey data; use of labour versus supply of labour; turnover, value added, intermediate consumption, etc. versus the use of labour."*

Following the OECD (2002), the UN conducted a survey across statistical agencies to gather information regarding the specific steps each country follows in order to include the Non-Observed Economy in their measure of GDP. The results were published in UN (2008). Forty five countries responded to the survey, with only two (Japan and New Zealand) not estimating the Non-Observed Economy.

In addition to providing the size of the Non-Observed Economy included in the official GDP figures, countries also reported what kind of activities within the Non-Observed economy were accounted for. Broadly speaking, there are four categories of producers susceptible to be part of the Non-Observed Economy. These are the not registered, not surveyed, those misreporting and other deficiencies. Table 1 in UN (2008) reports what kind of adjustment each country conducts and, if available, the size of the estimate for the Non-Observed Economy included in the official GDP figure.

Of the forty three countries that do estimate the Non-Observed Economy and include it in the official number of GDP, four countries did not report the size of the estimate. The estimates for the remaining thirty nine range from 0.8% for the US (in 1997) to 31.6% for Moldova (in

2001). The correlation between GDP per worker (from Heston et al (2009) for the year 2000) and the fraction of the Non-Observed Economy in the official GDP figure is -0.73. It is clear that most of the countries do correct their official statistics by adding the informal sector in their GDP numbers.

The size of the informal sector as a fraction of total output has also been estimated independently. Schneider et. al. (2010) estimate the size of the informal economy for 162 countries from 1999 to 2007. Their empirical method is based on the statistical theory of unobserved variables, which considers multiple causes and indicators of the shadow economy, as well as multiple effects of its existence. In particular they use the Multiple Indicators Multiple Causes (MIMIC) model, which is a type of structural equations model, based on aggregate data for each country at each point in time. They use freedom indicators (Business, Economic and fiscal) from the Heritage Foundation, size of the money supply, labor force participation rate, official GDP per capita, unemployment rate, government size among others to estimate the size of the informal economy. The fraction of the non-observed economy included in the official statistics is a lower bound when compared to Schneider et. al. (2010), which (for the same group of countries) range from 8.8% in the US to 65.5% in Georgia.

TFP and Informal Sector Measurement

Section 2.2 presented evidence that shows that reported GDP includes output generated by the informal sector. However, by construction, informal activities are difficult to measure and some fraction of it could go missing when constructing the values of aggregate inputs and output at the country level, affecting observed total factor productivity. With the help of the model one can understand this relationship and show that the fall in measured TFP is in fact amplified (and thus increases the predictive power of the model) as long as measurement problems in official statistics are smaller with respect to informal labor and capital than informal output.

Let Y_j^o denote total output observed for country j. Moreover, let Y_j^f and Y_j^i denote total output effectively produced by the formal and informal sector, respectively, in country j. Then, $Y_j^o = Y_j^f + \omega_j^y Y_j^i$, where $\omega_j^y \in [0,1]$ corresponds to the fraction of informal output included in official statistics in country j. Similarly, let K_j^o and N_j^o denote total capital and labor observed respectively and let K_j^f, K_j^i, N_j^f and N_j^i denote capital in the formal sector, capital in the informal sector, labor in the formal sector and labor in the informal sector, respectively. Then, $K_j^o = K_j^f + \omega_j^k K_j^i$ and $N_j^o = N_j^f + \omega_j^n N_j^i$ where $\omega_j^k \in [0,1]$ and $\omega_j^n \in [0,1]$ corresponds to the fraction of informal capital and informal labor included in official statistics, respectively. Using these expressions, measured total factor productivity in country j becomes

$$TFP_j^o = \frac{Y_j^o}{(K_j^o)^\alpha (N_j^o)^{(1-\alpha)}}.$$

To simplify the exposition of the results in what follows, it is assumed that the U.S. correctly measures output, capital and labor (i.e. $\omega_{US}^y = \omega_{US}^k = \omega_{US}^n = 1$). Using the model, Table 13 shows how measured relative TFP (between an economy parameterized with LMIC institutions and the US) reacts as a function of ω_{LMIC}^y, ω_{LMIC}^k and ω_{LMIC}^n.

[Locate Table 13 about here]

This table shows that as long as inputs (capital and labor) are measured with the same or better degree of accuracy as output, the model's figures for total factor productivity can be thought of as a conservative measure. The main results of the paper are presented using $\omega^y_{LMIC} = \omega^k_{LMIC} = \omega^n_{LMIC} = 1$ and, as Table 5 shows, relative TFP falls approximately 25%. If, for example, capital and labor are fully captured in official statistics but only 75% of informal output is included, (i.e. $\omega^n_{LMIC} = \omega^k_{LMIC} = 1$ and $\omega^y_{LMIC} = 0.75$), relative measured TFP would fall by 42.9%. If only 75% of informal capital and labor while only 50% of informal output is observed (i.e. $\omega^y_{LMIC} = 0.50$ and $\omega^n_{LMIC} = \omega^k_{LMIC} = 0.75$), the drop in relative TFP would be equal to 48.2%, explaining most of the observed difference in the data (equal to 56%). Thus, the benchmark case ($\omega^y_{LMIC} = \omega^k_{LMIC} = \omega^n_{LMIC} = 1$) can be interpreted as a lower bound on the effects of institutions on measured TFP. Note that the model would generate an increase in relative TFP, a counterfactual prediction, only if output is measured more accurately than inputs (capital and labor).

Figures and Tables Appendix

Table 11: Descriptive Statistics

	Mean	Median	SD	Min	Max	N
h	0.630	0.597	0.172	0.324	1.000	101
y	0.350	0.213	0.307	0.013	1.007	101
$(K/Y)^{\alpha/(1-\alpha)}$	0.869	0.885	0.222	0.378	1.548	101
TFP	0.621	0.597	0.284	0.110	1.237	101
K/H	0.460	0.289	0.444	0.012	1.635	101
Y/H	0.484	0.366	0.357	0.032	1.334	101
$InfLF(\%)$	0.570	0.668	0.331	0.038	0.987	75

Note: h refers to effective units of labor per worker, y is output per worker, $(K/Y)^{\alpha/(1-\alpha)}$ refers to the capital-output ratio as in Hall and Jones (1999), TFP is our total factor productivity, K/H is the capital per effective unit of labor ratio and Y/H is the output per effective unit of labor. All variables are relative to US levels, with the exception of the informal labor force.

Table 12: Descriptive Statistics

	Mean	Median	SD	Min	Max	N
λ	0.31	0.28	0.25	0.00	0.93	180
ϕ	0.13	0.09	0.12	0.00	0.76	181
τ	0.18	0.18	0.15	0.00	1.77	181
c_τ	0.14	0.11	0.14	0.00	1.25	181
τ_w	0.27	0.28	0.15	0.00	0.61	181
κ	10.76	3.16	48.56	0.26	611.84	181

Table 13: TFP and Informal Sector Measurement

ω^n_{LMIC}	ω^k_{LMIC}	ω^y_{LMIC}				
		1	0.75	0.5	0.25	0
1	1	0.748	0.571	0.394	0.217	0.040
	0.75	0.819	0.625	0.432	0.238	0.044
	0.5	0.928	0.709	0.489	0.269	0.050
	0.25	1.137	0.868	0.599	0.330	0.061
	0	2.078	1.586	1.095	0.603	0.112
0.75	1	0.898	0.685	0.473	0.261	0.048
	0.75	0.983	0.751	0.518	0.285	0.053
	0.5	1.114	0.851	0.587	0.323	0.060
	0.25	1.364	1.042	0.719	0.396	0.073
	0	2.494	1.904	1.314	0.724	0.134

Note: TFP reported relative to the US value. Model TFP is calculated as $TFP^o = \frac{Y^o}{(K^o)^{\widetilde{\alpha}}(N^o)^{(1-\widetilde{\alpha})}}$ where $\widetilde{\alpha} = 1/3$ is taken from Hall and Jones (1999).