Economics Bulletin

Volume 36, Issue 1

Informality and Productivity

Arturo Anton
Centro de Investigación y Docencia Económicas (CIDE)

Rodolfo Gutierrez
Ministry of Finance, Mexico

Abstract

We examine the role of productivity differences across countries to explain the negative relationship between GDP per capita and informality in the data. Through the lens of a formal-informal two-sector model, we find that differences in overall and sector-specific productivities replicate the relationship between informality and GDP per capita and may explain up to 40 percent of the differences in informality. The dispersion in overall productivity induces most of the variation in output per worker in the informal sector.

We thank Kaniska Dam, Julio Leal, and an anonymous referee for their useful comments and suggestions. The usual disclaimer applies. The views presented in this document are responsibility of the authors and do not represent the views of the Mexican Ministry of Finance.
Citation: Arturo Anton and Rodolfo Gutierrez, (2016) "Informality and Productivity", *Economics Bulletin*, Volume 36, Issue 1, pages 205-217
Contact: Arturo Anton - arturo.anton@cide.edu, Rodolfo Gutierrez - rodolfo_gutierrez@hacienda.gob.mx.
Submitted: February 16, 2015. **Published:** February 11, 2016.

1. Introduction

Taxes, regulation and enforcement policies are usually identified as major determinants of informality.[1] At the same time, it is well known that the data exhibit a negative relationship between GDP per capita and informality (see, for example, Ihrig and Moe, 2004, and La Porta and Shleifer, 2008). However, Antunes and Cavalcanti (2007) and Prado (2011) find that differences in tax rates and in regulatory and enforcement policies do not come close to simultaneously accounting for the large differences in the size of the informal sector and GDP per capita across countries.

In this paper, we examine an alternate but closely related issue: to what extent do productivity differences across countries explain the negative relationship between GDP per capita and informality in the data? To the best of our knowledge, this issue has not been previously examined in the literature. For this purpose, we use a version of the two-sector, formal-informal model of Ihrig and Moe (2004).[2] In particular, output in the formal sector is produced with both capital and labor, whereas informal output is produced with only labor. Labor endowment is fixed and must be allocated optimally to each sector. Following the evidence reported by La Porta and Shleifer (2008) regarding the large productivity differences between formal and informal firms within countries, we also distinguish between *overall* and *sector-specific* productivities. In the model, this distinction is important: both overall and sector-specific productivities have different marginal effects on allocations and, thus, on the relative size of the informal sector. Therefore, productivity differences have the potential to explain the dispersion of informality in the data.

The model suggests that productivity differences across countries can reasonably explain the empirically observed inverse relationship between GDP per capita and informality. The intuition is simple. Suppose productivity in the formal sector increases relative to that in the informal sector so that the marginal productivity of formal labor is relatively higher. Given that labor has decreasing returns, the allocation of labor to the informal sector must fall so that the marginal productivities of labor in both sectors are the same. Consequently, output falls in the informal sector but rises in the formal sector. Consider instead an increase in overall productivity so that the marginal product of capital is higher. As a result, resources are allocated to the formal sector because the informal sector does not use capital to produce goods. The corresponding increase in the marginal product of labor in the formal sector forces informal labor to fall once again according to the mechanism previously discussed. Ultimately, the simultaneous increase in both overall and formal sector productivities leads to an increase in output per capita. This effect explains why a negative relationship between GDP per capita and informality due to changes in productivities may be obtained.

In the model, informality is defined in terms of all unreported income from the production of legal goods and services that would generally be taxable otherwise. Given that informality may be defined alternatively as the lack of a contractual relationship between firms and workers, we also consider a decentralized version of Ihrig and Moe's (2004) model. In this framework, formal firms must cover all the workers' costs derived from a contractual relationship as stated by the law. In contrast, informal firms can avoid these costs but face a probability of being caught and

[1] See Schneider and Enste (2000) for a detailed literature review.

[2] These authors examine the impact of tax rates and enforcement policies on the size of the informal sector. They do not consider how productivity differences affect informal sector output.

forced by the authority to comply with the law. We show that the basic trade-off between productivity and informality discussed above also holds in this model. Therefore, the previous analysis may be extended to a framework where the absence of a contractual bond between employers and workers defines informality.

The model of Ihrig and Moe (2004) is then calibrated to the data. If informality is defined in terms of unreported income for tax purposes, numerical analysis suggests that productivity differences may account for approximately 40 percent of the differences in informality. This result holds for a cross-section and for a panel of countries, and even for an alternative dataset on informality. This value decreases to approximately 24 percent when the alternate definition of informality is used instead. The analysis also suggests that it is the dispersion in overall productivity (rather than sector-specific productivity) that induces most of the variation in informal sector's output per worker.

The next section presents the models. Section 3 describes how the model of Ihrig and Moe is calibrated to fit the data. Section 4 discusses the results. The final section offers conclusions.

2. The models

2.1 *A model of unreported income*

The economy is characterized by two sectors, labeled as "formal" (f) and "informal" (i). The informal sector represents the shadow economy as defined by Schneider and Enste (2000). Following Ihrig and Moe (2004), goods in country x's formal sector are produced with capital k and labor l_f, according to the Cobb-Douglas production function $Y_{fx} = \tilde{A}_{fx} k^{\alpha} l_f^{1-\alpha}$. Here, $\tilde{A}_{fx}$ is a productivity parameter and $\alpha \in (0,1)$ is the capital share. In contrast, goods in the informal sector are produced with informal labor l_i only. This specification implicitly assumes that capital is a fixed factor of production in the informal sector.[3] The production function is $Y_{ix} = \tilde{A}_{ix} l_i^{\gamma}$, where $\tilde{A}_{ix}$ is a productivity parameter and $\gamma \in (0,1)$. Note that the subscript x is not included in capital and labor inputs to avoid cluttering notation.

The formal sector pays a tax rate $\tau \in (0,1)$ on output. In contrast, the informal sector may evade it. However, it faces a probability $\rho \in [0,1]$ of being detected and obliged by the authority to pay the tax rate τ.[4] Its expected revenue is

$$(1-\rho)\tilde{A}_{ix} l_i^{\gamma} + \rho(1-\tau)\tilde{A}_{ix} l_i^{\gamma} = (1-\rho\tau)\tilde{A}_{ix} l_i^{\gamma}.$$

There is an infinitely-lived representative agent whose preferences are given by:

$$U = \sum_{t=0}^{\infty} \beta^t \, u(c_t), \qquad (1)$$

where $\beta \in (0,1)$ is the discount factor and $u(c_t)$ is the utility function over consumption c_t. The agent has a time endowment T, which she allocates between the formal and informal sectors. She is also the owner of capital that evolves according to:

$$k_{t+1} = s_t + (1-\delta)k_t, \qquad (2)$$

---

[3] La Porta and Shleifer (2008) report that a typical unregistered, informal firm remains small in size and has insufficient capital (relative to formal firms) despite operating for an average of 7 years. This evidence suggests that informal firms might be particularly subject to some sort of investment constraints.

[4] The term $\rho\tau$ on the right hand side of the expression below may have an alternative explanation: if, in addition to the tax payment the government imposes a penalty as a percentage of informal output, then the value of $\rho\tau$ incorporates both the tax rate faced by formal firms, τ, and the penalty.

where s_t is investment and $\delta \in (0,1)$ is the depreciation rate.

The agent's problem is thus to choose consumption, the next period's capital stock and the allocation of labor in each sector to maximize (1) subject to:

$$c_t + k_{t+1} - (1-\delta)k_t = (1-\tau)\tilde{A}_{fx} k_t^{\alpha} l_{f,t}^{1-\alpha} + (1-\rho\tau)\tilde{A}_{ix} l_{i,t}^{\gamma}, \tag{3}$$

and

$$l_{f,t} + l_{i,t} = T, \tag{4}$$

with $k_0 > 0$ given.

Assuming logarithmic utility for $u(c_t)$, standard optimality conditions yield:

$$(1-\tau)(1-\alpha)\tilde{A}_{fx} k_t^{\alpha} l_{f,t}^{-\alpha} = (1-\rho\tau)\gamma\tilde{A}_{ix} l_{i,t}^{\gamma-1}, \tag{5}$$

$$\frac{c_{t+1}}{\beta c_t} = (1-\tau)\alpha\tilde{A}_{fx} k_{t+1}^{\alpha-1} l_{f,t+1}^{1-\alpha} + 1 - \delta. \tag{6}$$

Equation (5) simply states that the marginal productivity of labor must be the same in each sector so that the agent is indifferent in allocating labor to the formal and informal sectors. Expression (6) is the typical Euler equation for consumption, where the marginal rate of substitution must be equal to the marginal rate of transformation net of taxes.

Steady state

To account for differences between overall and sector-specific productivities, define $\tilde{A}_{fx} \equiv A_x\theta_{fx}$ and $\tilde{A}_{ix} \equiv A_x\theta_{ix}$. A shift in overall productivity A_x affects the productivity levels in both sectors, whereas a shift in either θ_{fx} or θ_{ix} only affects the productivity of a particular sector.[5] We also impose the condition $\gamma \leq 1-\alpha$ so that production in the informal sector does not depend negatively on A_x (see equation (9) below).

Equations (5) and (6) evaluated at the steady state yield:

$$l_{ix} = \left[\frac{(1-\rho\tau)\gamma\theta_{ix}}{(1-\tau)(1-\alpha)\theta_{fx}}\right]^{\frac{1}{1-\gamma}} \left[\frac{\frac{1}{\beta}-1+\delta}{(1-\tau)\alpha A_x\theta_{fx}}\right]^{\frac{\alpha}{(1-\alpha)(1-\gamma)}}, \tag{7}$$

$$k_x = \left[\frac{(1-\tau)\alpha A_x\theta_{fx}}{\frac{1}{\beta}-1+\delta}\right]^{\frac{1}{1-\alpha}} l_{fx}, \tag{8}$$

where $l_{fx} = T - l_{ix}$. Using equations (7) and (8), production in each sector is given by:

$$Y_{ix} = \psi_i\left(A_x^{1-\alpha-\gamma}\theta_{ix}^{1-\alpha}\theta_{fx}^{-\gamma}\right)^{\frac{1}{(1-\alpha)(1-\gamma)}}, \tag{9}$$

$$Y_{fx} = \psi_f\left(A_x\theta_{fx}\right)^{\frac{1}{1-\alpha}}\left[T - l_{ix}\left(A_x, \theta_{ix}, \theta_{fx}\right)\right], \tag{10}$$

where $\psi_i > 0$ and $\psi_f > 0$ are terms not related to productivity parameters. From (7) - (10), each productivity parameter (A_x, θ_{fx} and θ_{ix}) has a different marginal effect on the variables of

[5] The distinction between overall and sector-specific productivities resembles the "neutral" and "investment-specific" productivities in the two-sector model of Greenwood et al. (1997).

interest l_{ix}, k_x, Y_{ix} and Y_{fx}. Therefore, differences in the levels of these parameters lead to differences in both the allocation and the relative size of the informal sector across countries.

2.2 *A model of non-compliance with labor regulations*

As discussed by Schneider and Enste (2000), Kanbur (2009), and others, there is a lack of consensus to define informality. Some authors define the informal sector in terms of employment relationships that are not legally regulated or protected (see the discussion in Kanbur (2009) and the references therein). In such a case, informality provides employers with flexibility in exchange for the lack of contractual relationships with their employees.

To capture this alternative definition of informality, a decentralized version of the previous model is considered. In particular, the economy is characterized by formal and informal sectors with the same production technologies as before. Firms in the formal sector must comply with all labor regulations imposed by the law. The cost of having a contractual relationship with workers includes the payment of social security contributions and labor taxes, hiring and firing costs, and other employment benefits granted by law. Let $\tau_l \in (0,1)$ to denote the cost of having a contractual relationship per worker as a fraction of the wage rate. Under this setting, a formal firm in country x must choose capital and formal labor to maximize profits Π_{fx} given by:

$$\Pi_{fx} = \tilde{A}_{fx} k^{\alpha} l_f^{1-\alpha} - (1+\tau_l) w_f l_f - rk, \tag{11}$$

where w_f and r represent the wage rate in the formal sector and the rental rate of capital, respectively.

In contrast, firms in the informal sector can avoid compliance with labor regulations. However, these firms face a probability $\rho_l \in [0,1]$ of being detected and obliged by the authority to comply with such regulations. The expected cost per worker in such a case is $\rho_l(1+\tau_l)w_i$, where w_i is the wage rate in the informal sector. With probability $1-\rho_l$, these firms are undetected and earn $\tilde{A}_{ix} l_i^{\gamma} - w_i l_i$. Therefore, expected profits Π_{ix} may be written as:

$$\Pi_{ix} = \tilde{A}_{ix} l_i^{\gamma} - (1+\rho_l\tau_l) w_i l_i. \tag{12}$$

Households are identical, with preferences given by the lifetime utility in (1). They are also the owners of capital and receive the profits from both formal and informal firms. Using equation (4), their budget constraint is:

$$c_t + k_{t+1} - (1-\delta)k_t = r_t k_t + w_{f,t} l_{f,t} + w_{i,t}\left(T - l_{f,t}\right) + \Pi_{fx,t} + \Pi_{ix,t}. \tag{13}$$

The household's problem is thus to choose c_t, k_{t+1} and $l_{f,t}$ to maximize (1) subject to (13) with $k_0 > 0$ given. Assuming, as before, a logarithmic utility for $u(c_t)$, the Euler equation for consumption is given by:

$$\frac{c_{t+1}}{\beta c_t} = r_{t+1} + 1 - \delta, \tag{14}$$

so that the marginal rate of substitution between present and future consumption equals its relative price net of depreciation. Optimality conditions from this problem also yield $w_{f,t} = w_{i,t}$. This is a natural result given the assumption of free mobility of labor between sectors. From firms' optimization problem, the previous condition implies:

$$\frac{(1-\alpha)\tilde{A}_{fx} k_t^{\alpha} l_{f,t}^{-\alpha}}{1+\tau_l} = \frac{\gamma \tilde{A}_{ix} l_{i,t}^{\gamma-1}}{1+\rho_l\tau_l}. \tag{15}$$

Equation (15) indicates that the marginal productivity of labor must be the same in each sector once adjusted for the corresponding expected cost of a unit of labor. Therefore, the household must be indifferent in allocating labor to the formal and informal sectors. This optimality condition resembles equation (5) from the previous model.

Steady state

Now we solve for steady-state expressions of the variables of interest. For that purpose, the first-order condition for capital from the problem in (11) may be substituted into (14). This new equation, along with (15), solves for the levels of informal labor and capital at the steady-state:

$$l_{ix} = \left[\frac{(1+\tau_l)\gamma\theta_{ix}}{(1+\rho_l\tau_l)(1-\alpha)\theta_{fx}}\right]^{\frac{1}{1-\gamma}}\left[\frac{\frac{1}{\beta}-1+\delta}{\alpha A_x\theta_{fx}}\right]^{\frac{\alpha}{(1-\alpha)(1-\gamma)}}, \tag{16}$$

$$k_x = \left[\frac{\alpha A_x\theta_{fx}}{\frac{1}{\beta}-1+\delta}\right]^{\frac{1}{1-\alpha}} l_{fx}, \tag{17}$$

with $l_{fx} = T - l_{ix}$. Let $\phi_i > 0$ and $\phi_f > 0$ denote terms not related to productivity parameters. After substituting equations (16) and (17) into the production functions of each sector, the following expressions are obtained:

$$Y_{ix} = \phi_i\left(A_x^{1-\alpha-\gamma}\theta_{ix}^{1-\alpha}\theta_{fx}^{-\gamma}\right)^{\frac{1}{(1-\alpha)(1-\gamma)}}, \tag{18}$$

$$Y_{fx} = \phi_f\left(A_x\theta_{fx}\right)^{\frac{1}{1-\alpha}}\left[T - l_{ix}\left(A_x, \theta_{ix}, \theta_{fx}\right)\right]. \tag{19}$$

As in the previous model, an inspection of equations (16) - (19) indicates that each productivity parameter will have a different marginal effect on capital, labor and production. In fact, expressions (16) - (19) closely resemble those listed in (7) - (10). Therefore, differences in productivity parameters also lead to differences in both the allocation and the relative size of the informal sector under this alternate definition of informality.

3. Calibration

In this section, we describe how parameter values are set to replicate some features in the data. Naturally, this approach requires estimates of informality for both developed and developing countries elaborated under a common methodology. To the best of our knowledge, such a dataset is not available if informality is defined in terms of a lack of a contractual relationship between firms and workers. The most comprehensive data for such purposes have been developed by the International Labor Organization (ILO) and Women in Informal Employment: Globalizing and Organizing (WIEGO), but it is only available for 40 developing economies. As will be clear below, the absence of estimates for developed countries like the U.S. leads us to refrain from using this database for calibration purposes. In contrast, estimates for the size of the shadow economy have been elaborated by Schneider (2007) for both developed and developing countries. For this reason, we take the model presented in section 2.1 as our reference case to calibrate parameter values. Given that the model provided in section 2.2 yields similar qualitative predictions as the previous model, later we also use the data from ILO-WIEGO to evaluate the fitness of our exercise.

We assume that each economy in the data may be represented by expressions (7) - (10). Data on GDP per capita for several years and countries are expressed in 2005 international dollars and

adjusted by PPP rates, as reported by the World Bank's World Development Indicators (WDI). As mentioned above, the estimates for the size of the informal economy, $Y_{ix}/(Y_{ix}+Y_{fx})$, are taken from Schneider (2007).

Unless otherwise noted, parameters are calibrated taking the U.S. economy as the benchmark and using 2002 as the reference year. The capital share α is set to 0.36 and $\gamma = 1-\alpha$. The time endowment T is normalized to 100. The depreciation rate δ is set to 6.2 percent, which is consistent with an annual rate of return of capital before taxes of 6.5 percent. The discount factor β is fixed at 0.951, a value similar to those reported elsewhere (cfr., Prescott, 1986; Parente and Prescott, 1992).

We estimate the remaining 5 parameters as follows. Without loss of generality, the productivity parameter A_{us} is set to 1, where the subscript $x \equiv us$ denotes the values for U.S. variables. Parameters τ, ρ, θ_{ius} and θ_{fus} are then simultaneously calibrated to match the following four moments in U.S. data: (1) the size of the informal economy of 8.4 percent; (2) the GDP per capita of \$39,944; (3) the tax revenue of 10 percent of GDP collected by the central government, as reported by the WDI; and (4) the total amount of civil penalties assessed by the fiscal authorities from all types of taxes as a share of GDP (0.17 percent), according to the Internal Revenue Service. Parameter values are reported in Table I.

Table I
Parameter values (U.S. benchmark economy)

α	0.36	A_{us}	1
γ	0.64	τ	0.109
T	100	ρ	0.185
δ	0.062	θ_{ius}	0.813
β	0.951	θ_{fus}	0.384

Note: see the main text for details.

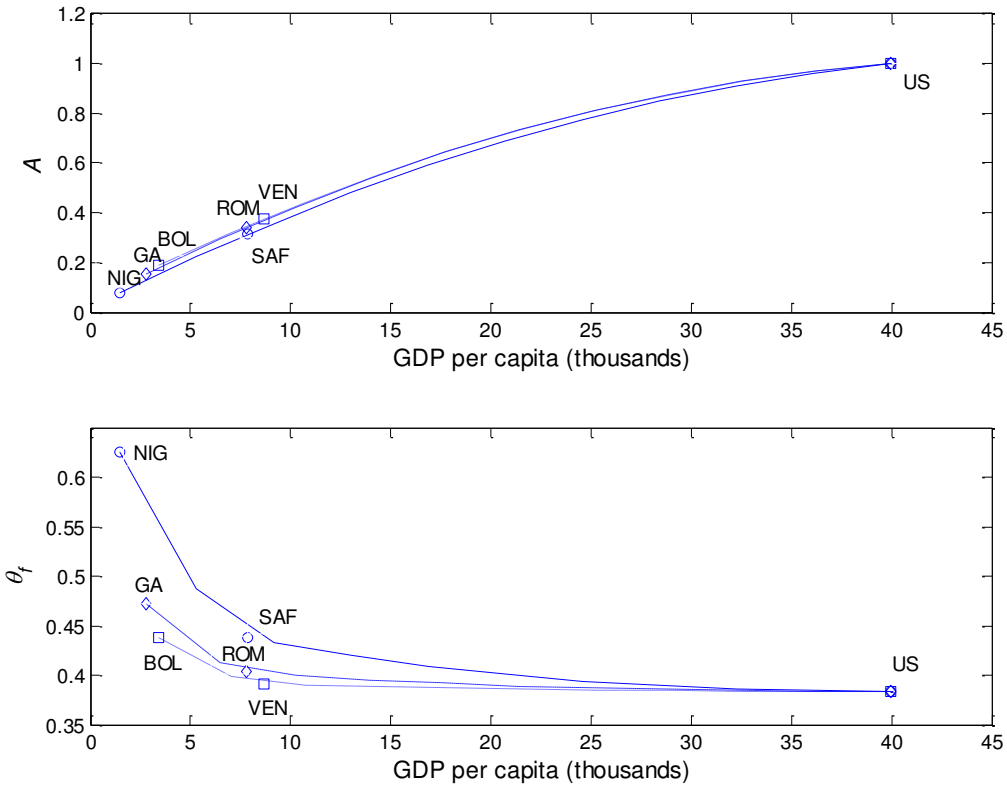
To identify productivity levels in the data, we consider two additional countries with different sizes of informal economies and GDP per capita levels. In particular, we take a lower-middle-income and an upper-middle-income country, following the World Bank's classification. The identification assumption is that these economies share the same parameter values as the U.S. economy, with the exception of technology levels. Without loss of generality, the parameter θ_{ix} is fixed at the U.S. value θ_{ius}. Parameters A_x and θ_{fx} are then calibrated to replicate the size of the informal economy and the level of GDP per capita for each country.

The countries chosen are Nigeria (NIG) and South Africa (SAF). These countries had estimated informal economies of 59 and 30 percent of GDP, and GDP per capita levels of \$1,452 and \$7,864, respectively. The corresponding estimates for the lower-middle-income country are $A_L = 0.08$ and $\theta_{fL} = 0.63$ and for the upper-middle-income country are $A_M = 0.31$ and $\theta_{fM} = 0.44$.

Once we have the 3 estimates of parameter A_x, an interpolation is made on A_x for a given level of GDP per capita. A similar approach is followed for parameter θ_{fx}. The results are presented in Figure 1. Given the values in Table I (except those for A and θ_f), the fitted pairs

$\{A_x, \theta_{fx}\}$ are used to compute the steady-state variables in (7) – (10). This allows us to estimate the fraction of output produced in the informal sector, $Y_{ix}/(Y_{ix} + Y_{fx})$, and the share of workers in the informal sector, l_{ix}, for different pairs $\{A_x, \theta_{fx}\}$. The results derived from this interpolation will be referred to as the "Benchmark case".

Figure 1
Fitting of productivity parameters



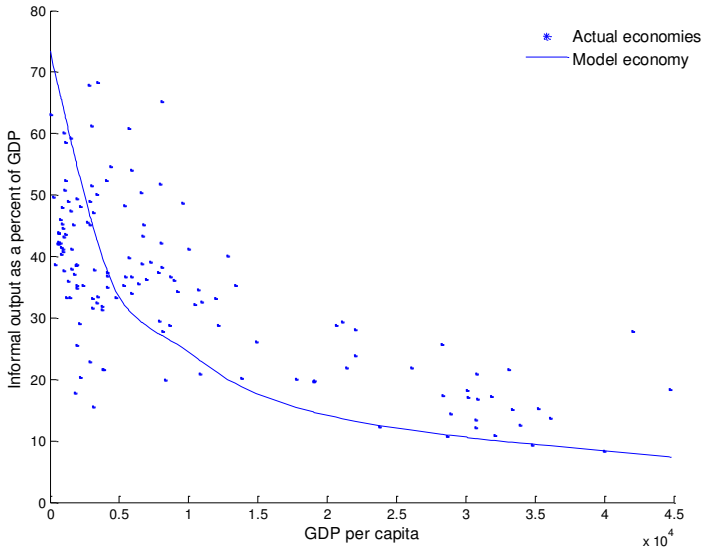
Given that the results might be sensitive to the set of countries chosen, we also consider two alternative combinations of lower-middle-income and upper-middle-income countries. These combinations are Georgia (GA) and Romania (ROM) as one case and Bolivia (BOL) and Venezuela (VEN) as the other case. Interestingly, Bolivia and Georgia are the countries with the highest informal economies in the sample (68.3 and 68 percent of GDP, respectively). Given that the U.S. is the country with the lowest informal economy in the sample, the calibration in these two cases captures the extreme values of informality in the data. The interpolation results from each case are shown in Figure 1. The information that emerges from this figure is that the identification of parameter A is relatively robust. However, parameter θ_f seems more sensitive to the combination of countries chosen. The results that use Georgia and Romania for identifying productivity parameters are henceforth referred to as "Robustness check 1". The remaining case where Bolivia and Venezuela are used instead is labeled as "Robustness check 2".

4. Results

To validate the model's predictions, we take advantage of the estimates on the size of the informal economy (as a percent of GDP) for 145 countries in 2002 provided by Schneider (2007). We only consider those countries for which we also have the information on GDP per capita. This leaves us with a sample of 140 countries. Due to space constraints, only the benchmark case in Figures 2 – 5 is discussed in detail below. The results under robustness checks 1 and 2 are qualitatively similar.

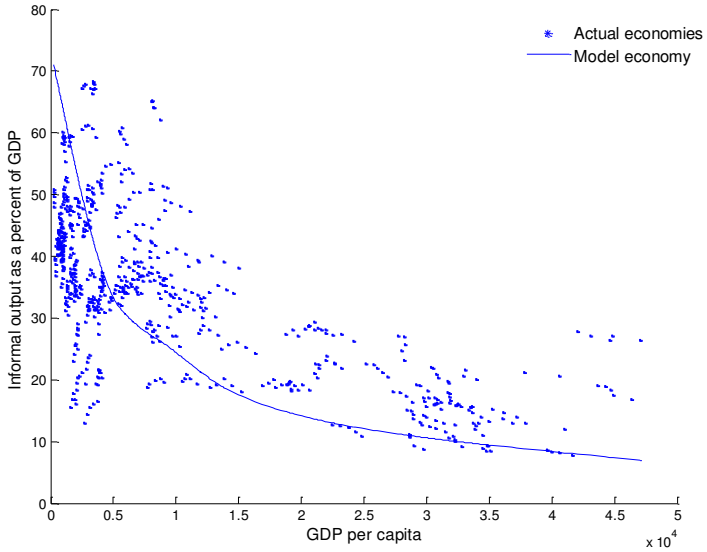
Figure 2 presents the relationship between GDP per capita and the percent of output produced in the informal sector for a cross-section of countries in 2002. The data depict a negative convex relationship between these two variables, which confirms the findings reported by Ihrig and Moe (2004) and La Porta and Shleifer (2008). Interestingly, the model is able to replicate this convex relationship, although it under-predicts the share of output in the informal sector for high-income countries. To evaluate the model's fitness, an OLS regression is estimated, with the actual values for the share of informal output as the dependent variable and a constant term and the predicted values of the model as regressors. The adjusted R-squared is 0.37, suggesting that the model's fit is relatively good. Under robustness checks 1 and 2, the adjusted R-squared value is 0.40 in each case.

Figure 2
GDP per capita and informal output as a percent of GDP
(cross-section data)



Taking advantage of the informal economy estimates of Schneider (2007) for several years, we also perform a panel data analysis. In particular, we consider 5 years of data for the 140 countries above. Therefore, we are left with a panel of 700 observations. The results are presented in Figure 3. Once again, the model replicates the negative relationship between GDP per capita and the size of the informal economy in the data. An OLS regression with a similar specification as above yields an adjusted R-squared of 0.36. For robustness checks 1 and 2, the corresponding value is 0.39 in each case.

Figure 3
GDP per capita and informal output as a percent of GDP
(panel data)



An alternative way to evaluate the predictions of the model is to consider a different variable related to informality, as well as a completely different dataset. This approach is especially important given that the productivity parameter estimates used are those obtained in the previous section. In particular, we take the estimates for l_{ix} under the benchmark case and compare them against the data on the share of workers employed in the informal sector, as reported by ILO-WIEGO. For this dataset, the informal sector is defined in terms of private unincorporated enterprises that are unregistered or small in terms of the number of employed persons. The cross-country data refer to non-agricultural employment in 40 developing countries. We only consider those economies for which we also have information on GDP per capita. This restriction leaves us with a sample of 37 countries.

The results are presented in Figure 4. Remarkably, the model is able to replicate the negative relationship in the data relatively well. To evaluate the model's fitness, an OLS regression is estimated with the share of workers in the informal sector from the data as the dependent variable. The regressors include a constant term and the predicted values of l_{ix} from the model. The adjusted R-squared is 0.39, which also suggests a good fit. When robustness checks 1 and 2 are considered instead, the corresponding R-squared value is 0.40 in each case. Overall, what emerges from all of these exercises is that productivity differentials across countries may account for approximately 40 percent of the differences in informality observed in the data.

As mentioned earlier, an alternate definition of informality is the absence of a contractual relationship between firms and workers that is reflected in the lack of social security and other benefits for workers. To account for this phenomenon, we take advantage of the data on the share of informal jobs in total employment elaborated by ILO-WIEGO. According to ILO-WIEGO, informal jobs include all employees not covered by social security or not entitled to other employment benefits. These data are thus in accordance with the alternate definition of informality. We have shown in section 2 that the effects of productivity parameters on informal labor are qualitatively similar irrespective of how informal labor is defined (see equations 7 and

16). Therefore, we also compare the predicted values of l_{ix} in the benchmark case against the share of informal jobs in total employment as reported by ILO-WIEGO. The cross-country data for which we also have information on GDP per capita decreases the original sample to 34 countries.

Figure 4
GDP per capita and share of workers in the informal sector

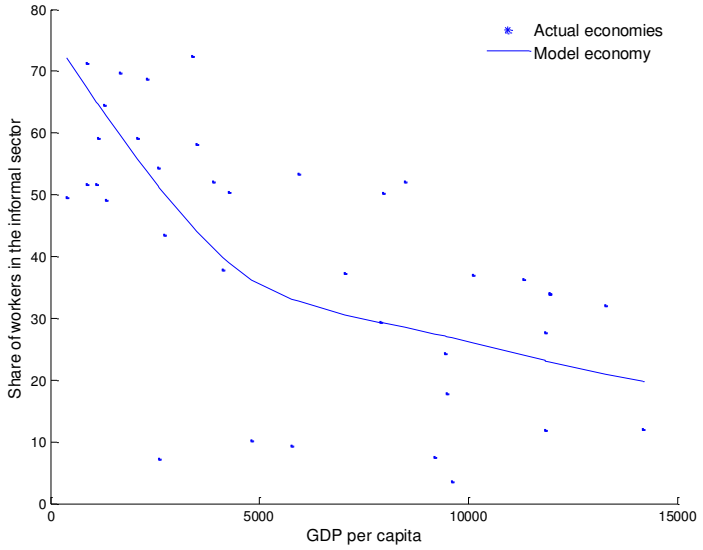


Figure 5 illustrates how the model fits the data under this alternate definition of informality. The large dispersion of informality for a given level of GDP per capita is noteworthy. Nonetheless, the model does a fair job of explaining the negative pattern in the data. An OLS regression with the predicted values of l_{ix} as the explanatory variable yields an adjusted R-squared value of 0.21. This number increases to 0.24 when robustness checks 1 and 2 are considered instead. These results show that differences in productivity are still capable of explaining a sizable fraction of the differences in informality across countries.

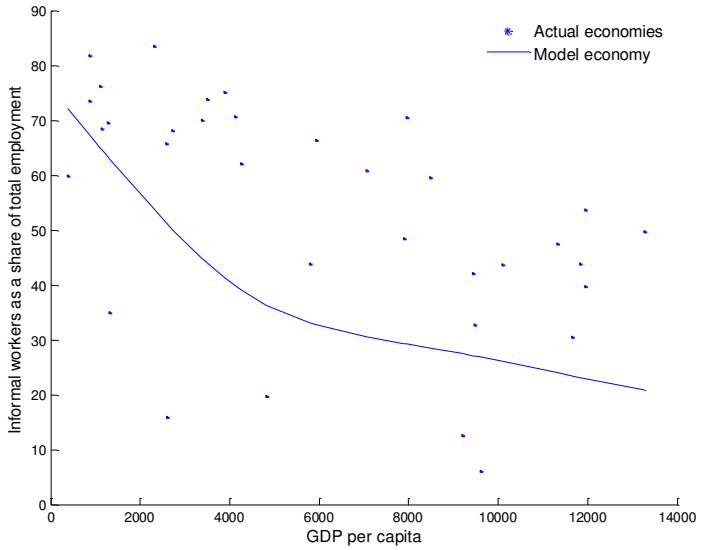
The final exercise is to evaluate the relative importance of overall and sector-specific productivity levels to account for the differences in informality across countries. In particular, we ask the following: if the dispersion in production per worker in the informal sector across countries is only attributable to changes in overall and sector-specific productivity levels, what is the relative contribution of each term? The exercise thus assumes that all parameters for a given country are kept fixed with the exception of A_x and θ_{fx}. The dispersion in informal production per worker is measured with respect to its corresponding U.S. level.

Let $\hat{y}_{ix} \equiv y_{ix}/l_{ix}$ denote production per worker in country x's informal sector. From the production function Y_{ix} and (7), it may be shown that:[6]

$$\frac{\hat{y}_{ix}}{\hat{y}_{ius}} = \left(\frac{A_x}{A_{us}}\right)^{\frac{1}{1-\alpha}} \left(\frac{\theta_{fx}}{\theta_{fus}}\right)^{\frac{1}{1-\alpha}}. \qquad (20)$$

[6] Expression (20) also holds for the model described in section 2.2. For that case, use equation (16) instead of (7) to substitute into the production function of the informal sector.

Figure 5
GDP per capita and informal workers as a share of total employment



From (20), the variance of the (log of) relative output per worker predicted by the model may be estimated as a function of the dispersion in overall and sector-specific productivities. Namely,

$$\text{Var}\left[\ln\left(\frac{\hat{y}_{ix}}{\hat{y}_{ius}}\right)\right] = \mu^2\text{Var}\left[\ln\left(\frac{A_x}{A_{us}}\right)\right] + \mu^2\text{Var}\left[\ln\left(\frac{\theta_{fx}}{\theta_{fus}}\right)\right] + 2\mu^2\text{Cov}\left[\ln\left(\frac{A_x}{A_{us}}\right), \ln\left(\frac{\theta_{fx}}{\theta_{fus}}\right)\right],$$

where $\mu \equiv 1/(1-\alpha)$.

The results for all the pairs of lower-middle-income and upper-middle-income countries considered are presented in Table II. The numbers for the benchmark case suggest that the dispersion in overall productivity levels accounts for most of the variation in output per worker in the informal sector. The same results are obtained under robustness checks 1 and 2.

Table II
Variance of informal output per worker (relative to the U.S. economy)

Specification	Variance term related to:		Covariance term	$\text{var}\left[\ln\left(\frac{\hat{y}_{ix}}{\hat{y}_{ius}}\right)\right]$
	Ratio of overall productivities	Ratio of sector-specific productivities		
	(A)	(B)	(C)	(A) + (B) + (C)
Benchmark case	1.511	0.053	-0.559	1.005
Robustness check 1	0.847	0.009	-0.164	0.692
Robustness check 2	0.685	0.004	-0.090	0.599

5. Conclusions

This study contributes to the literature by evaluating the role of productivity differences in determining the negative relationship between GDP per capita and informality in the data. The results suggest that variations in productivity levels across countries significantly explain the dispersion in informal output and in both the share of workers in the informal sector and the share of informal workers in total employment. Additionally, variations in sector-specific productivities seem to play a small role in explaining differences in output per worker in the informal sector. Therefore, these results are a complementary step towards understanding why the relative size of the informal sector decreases as the economy develops.

One of the crucial assumptions in both models is that capital is a fixed factor of production in the informal sector. As mentioned previously, this might be the result of investment constraints faced by such firms. Given this friction, an increase in overall productivity leads to a reallocation of resources away from the informal sector and towards the formal sector. In a more general setting, one might think of a framework in which formal sector firms face lower investment constraints than informal sector firms and where such constraints are explicitly modeled. At the same time, the results of this paper note the need for a better understanding of the fundamental differences in sector-specific and overall productivities across countries. A recent study along these lines is presented by Buera et al. (2011). The authors show that financial frictions distort the allocation of resources across heterogeneous production units in a two-sector model with manufacturing and services. This misallocation, in turn, lowers aggregate and sector-level productivities. Given our previous discussion, an extension of this framework to a formal-informal setting would be an interesting research topic.

References

Antunes, A. R. and T. V. de V. Cavalcanti (2007) "Start up costs, limited enforcement, and the hidden economy" *European Economic Review* **51**, 203-224.

Buera, F. J., J. P. Kaboski and Y. Shin (2011) "Finance and development: a tale of two sectors" *American Economic Review* **101**, 1964-2002.

Greenwood, J., Z. Hercowitz and P. Krusell (1997) "Long-run implications of investment-specific technological change" *American Economic Review* **87** (3), 342-362.

Ihrig, J. and K. Moe (2004) "Lurking in the shadows: the informal sector and government policy" *Journal of Development Economics* **73**, 541–557.

Internal Revenue Service, IRS Tax Statistics Table 17-2002.

International Labor Organization (ILO) and Women in Informal Employment: Globalizing and Organizing (WIEGO), Statistics on Women and Men in the Informal Economy, available at: http://laborsta.ilo.org/informal_economy_E.html.

Kanbur, R. (2009) "Conceptualising informality: regulation and enforcement" *Indian Journal of Labour Economics* **52** (1), 33-42.

La Porta, R. and S. Shleifer (2008) "The unofficial economy and economic development" *Brookings Papers on Economic Activity* **39** (2), 275-352.

Parente, S. L. and E. C. Prescott (1992) "Technology adoption and the mechanics of economic development" in *Political Economy, Growth, and Business Cycles* by A. Cukierman, Z. Hercowitz and L. Leiderman, Eds., The MIT Press: Cambridge, MA, 199-224.

Prado, M. (2011) "Government policy in the formal and informal sectors" *European Economic Review* **55**, 1120-1136.

Prescott, E. C. (1986) "Theory ahead of business cycle measurement" *Carnegie-Rochester Conference Series on Public Policy* **25** (1), 11-44.

Schneider, F. (2007) "Shadow economies and corruption all over the world: new estimates for 145 countries" *Economics: The Open-Access, Open-Assessment E-Journal* **1** (9), 1-47.

Schneider, F. and D. Enste (2000) "Shadow economies: size, causes and consequences" *Journal of Economic Literature* **38** (1), 77–114.

World Bank, World Development Indicators, available at: http://data.worldbank.org/data-catalog/world-development-indicators.