

Informality Trends and Cycles*

Norman V. Loayza
The World Bank

Jamele Rigolini
The World Bank

Abstract

This paper studies the trends and cycles of informal employment. It first presents a theoretical model where the size of informal employment is determined by the relative costs and benefits of informality and the distribution of workers' skills. In the long run, informal employment varies with the trends in these variables, and in the short run it reacts to accommodate transient shocks and to close the gap that separates it from its trend level. The paper then uses an error-correction framework to examine empirically informality's long- and short-run relationships. For this purpose, it uses country-level data at annual frequency for a sample of developed and developing countries, with the share of self-employment in the labor force as the proxy for informal employment. The paper finds that, in the long run, informality is larger in countries that have lower GDP per capita and impose more costs to formal firms, in the form of more rigid business regulations, less valuable police and judicial services, and weaker monitoring of informality. In the short run, informal employment is found to be counter-cyclical for the majority of countries, with the degree of counter-cyclicality being lower in countries with larger informal employment and better police and judicial services. Moreover, informal employment follows a stable, trend-reverting process. These results are robust to changes in the sample and to the influence of outliers, even when only developing countries are considered in the analysis.

World Bank Policy Research Working Paper 4078, December 2006

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

*We would like to thank Koichi Kume for excellent research assistance. We would also like to thank Omar Arias, Bill Maloney, and Jaime Saavedra for very useful comments and discussions.

1 Introduction

Informality is a fundamental characteristic of underdevelopment. It is shaped both by the modes of production inherent to economies in the transition to modernity and by the relationship that the state establishes with private agents through regulation and monitoring. The received literature finds evidence that the relative size of the informal sector declines with overall development, rises with the burden of regulation, and decreases with the strength of monitoring (see Johnson *et al.*, 1997; Schneider and Enste, 2000; Friedman *et al.*, 2000; and Loayza, Oviedo, and Servén, 2005).

Notwithstanding the importance of these long-run relationships, the behavior of the informal sector is quite dynamic over time. This indicates that informality not only responds to fundamental, long-run forces but also to inter-temporal economic conditions related to the business cycle and transient policies. Thus, for example, the informal sector could act as a buffer that expands in economic recessions or as an adjustment mechanism during temporarily high tax regimes.

This paper studies the behavior of the informal sector – specifically informal employment – over the long and short runs, linking the two aspects both theoretically and empirically. In the long run, informal employment is determined by the trends in the relative costs and benefits of informality. In the short run, informal employment reacts to the temporary conditions created by the business cycle and moves to close the gap that separates it from its equilibrium, long-run trend. In so doing, the paper examines how the variables that drive informality in the long run also determine the way the informal sector reacts to the business cycle.

The literature on informality over the business cycle is relatively scarce and has concentrated on the analysis of the time series of selected countries (see, for instance, Bosch and Maloney, 2005). A further reduced number of papers address the issue of informality cycles in the context of long-run relationships. One of them is Fiess,

Fugazza, and Maloney (2006), which examines the reaction of formal and informal labor markets to permanent and temporary macroeconomic shocks. The paper develops a theoretical model whereby the comovement of relative earnings, sector sizes, and the real exchange rate can indicate the type of shock affecting the economy and whether labor rigidities become binding. It then studies these comovement patterns using time series for Argentina, Brazil, Colombia, and Mexico. Although Fiess, Fugazza, and Maloney study informality over the business cycle as deviation from a steady state, they do not characterize the evolution of the steady state itself. Our paper studies this evolution theoretically and uses cross-country variation to identify it. Once this is done, this paper uses the time-series variation to identify and assess the cyclical properties of informal employment.

The paper first presents a two-sector theoretical model that endogenizes the relative size of informal employment, making it a function of the productivity differential between formal and informal workers. This differential is, in turn, determined by the cost to become and remain formal and the distribution of skills in the labor force. Specifically, the productivity differential has a worker-driven component, given by workers' individual skills, and a sector-related component, given by the relative formal-informal regulatory burden, the strength of enforcement, and the access to productive public services. The size of informal employment is then given by the proportion of workers whose skills fall below a threshold level where the worker is indifferent between the two sectors.

In the model, cycles appear as productivity shocks occur, and informal employment's behavior over the cycle depends on how these shocks affect the productivity differential between the two sectors. The model derives its main results conducting comparative statics exercises on the cumulative function of the skill distribution (i.e., the relative size of informal employment) and the elasticity of this cumulative function with respect to productivity shocks (i.e., the short-run response of informal employment). In the

model, informal employment is predicted to be countercyclical. This happens as shocks affect both sectors homogeneously so that the costs of formality become more binding when negative shocks occur, and conversely in the face of positive shocks. Moreover, the model shows that higher levels of informal employment reduce the counter-cyclical response of informality.

Our theoretical model draws from Rausch (1991), who looks at the relationship between the size of the informal sector and the minimum wage. In his model, people choose to become workers, formal, or informal entrepreneurs depending on their managerial ability; similarly, in our model workers choose to be active in the formal or informal sector depending on their skills. Less directly, our model also draws from previous theoretical papers that have studied the relationship between informality and growth (Loayza, 1996), informality and rent-seeking (Sarte, 2000), inequality and the size of the informal sector (Dessy and Pallage, 2003), and the relationship between informality and the structure of taxation (de Paula and Sheinkman, 2006).

The paper then tests the relevance of the model using cross-country and time-series data. It does so in the context of an error-correction framework: in the long run, informal employment moves in conjunction with the development characteristics that influence and are influenced by informality; in the short run, informal employment moves both in response to production shocks, and in order to close the gap that separates it from its trend, equilibrium level. We also examine how the response of informal employment to production shocks varies with the level of informality itself and its determinants. The empirical findings are broadly consistent with the model's results.

The data consist of annual observations for a collection of developing and developed countries over diverse time periods. The observations are given at the national level, so that the variation in the data (and, thus, its informational value) resides on their cross-country and time-series dimensions. The key piece of information for the paper's empirical analysis (and its main bottleneck) is data on informal employment. For this

purpose, the proxy used in the paper is the share of self-employed in the labor force, as reported in the surveys collected by the International Labor Organization. The final sample consists of 471 observations in 42 countries when we consider all countries, and of 182 observations in 18 countries when we consider developing countries only.

The rest of the paper proceeds as follows. Section 2 presents a stochastic model of employment and production in a dual (formal/informal) economy. Section 3 presents the empirical analysis, introducing first the data and sample and then discussing the estimation results for, respectively, the long- and short-run relationships. The empirical section includes several robustness checks on both the criteria to obtain the sample and the influence of outliers.

2 A Basic Model of Informal Activity

The economy consists of a measure one of workers, and of two sectors, a formal and an informal one. Each worker is the owner of her own firm, hence, in what follows we shall use equivalently the concept of workers and firms. In each sector workers have different productivities depending on their skill level s , which has a distribution $\xi(s)$ over $s \in [0, 1]$. More precisely, productivity in the formal sector is equal to $\vartheta R^\alpha \cdot F(s) - R$, where ϑ represents the overall productivity level of the economy, s the workers' skills, R the level of regulation, $0 < \alpha < 1$, and R^α represents the efficiency with which the government uses regulation (i.e. countries with a higher parameter α make a better use of regulation). We also assume that workers' productivity increases with their skills in a concave manner, so that $F'(s) > 0, F''(s) < 0$.

The way we introduce regulation R wants to capture its dual role. Some degree of regulation can increase firms' productivity by allowing the government to provide better services to firms, such as the ability to solve disputes through an efficient judicial system, standardized procedures, and the ability to solve moral hazard and adverse selection

problems. This “efficient provision of services through regulation” is captured in the productivity function by the term R^α . However, obeying regulations is also costly to firms, in particular when they are used inefficiently and lead to red tape and corrupt practices. De Soto (1989), for instance, estimates that to register a small industry in Peru it takes 289 days and \$1231 to fulfill all the bureaucratic procedures, and Djankov *et al.* (2002) find similar results in a large number of developing countries. Moreover, staying formal can also be very costly. De Soto (1989) finds that, in a sample of 50 small manufacturing firms, the costs of staying formal represent an average of 348 percent of after-tax profits. Interestingly, only 22 percent of such costs are due to taxes, and 5 percent are due to higher public utility rates, while 73 percent of the costs are due to regulation and bureaucratic requirements. To consider the benefits and costs of regulation, we assume therefore that regulation has a constant unit marginal cost, but that the *efficiency* with which it is used can vary: in the model, the varying efficiency is captured by the parameter α , where countries with a higher parameter α perform better in transforming regulation in productivity-enhancing services. Notice that the function $\partial R^\alpha \cdot F(s) - R$ is strictly concave, so that there exists an “optimal” level of regulation $R^*(s)$ that maximizes firms’ productivity, where $R^*(s)$ increases with workers’ skills s .

Next, we turn the attention to firms in the informal sector, which in our model are firms that choose not to obey regulations. Factual evidence suggests that, aside from the fact that they do not have to bear the cost of regulation, informal firms tend to be less productive than formal ones. Penalties for engaging in informal activity can be stiff: detected firms may have to surrender a considerable part of their output or physical capital stock, and to avoid being caught, firms scale down the size of their informal operations, becoming therefore less productive. At least part of the penalties informal firms pay are in the form of bribes. De Soto (1989), for instance, finds that informal entrepreneurs pay between 10 to 15 percent of their gross income in bribes to corrupt government officials, whereas formal entrepreneurs pay an average of only 1 percent

of gross income in bribes (without counting bribes used to become formal). Another cost of informality is the inability to take full advantage of government provided goods, such as the legal and judicial system, and the police. In order to represent the lower productivity of the informal sector, we assume that informal workers have productivity $\phi(E)\vartheta R^\alpha \cdot F(s)$, where E represents enforcement of regulation compliance, $\phi(E)$ the level with which informal firms can benefit from the government's provision of services without paying for them, and $\phi(E) < 1$, $\phi'(E) < 0$. The productivity differential depends therefore on the intensity with which the government fights informal activity by constraining firms to obey regulation and to pay for the services it provides.

In the model, regulation implies a fixed cost that workers have to pay irrespective of their productivity. Therefore, in equilibrium there exists a “skill threshold” $s_I \in [0, 1]$ such that workers with skills $s < s_I$ choose to be informal, while workers with skills $s > s_I$ choose to be formal. This perspective is consistent with the view of informality as a voluntary, equilibrium choice, as proposed in Maloney (2004) and Loayza (1996). Although depending on the parameters ϑ, α, R, E , a fully formal or informal economy can exist, in what follows we focus attention on the case where formal and informal activities coexist. Under this situation, the skill threshold s_I is defined by:

$$F(s_I) = \frac{R^{1-\alpha}}{\vartheta(1 - \phi(E))} \quad (1)$$

and the size of the informal sector is given by $\Xi[s_I(\vartheta, \alpha, R, E)]$, where $\Xi[s]$ represents the cumulative function of the skill distribution $\xi(s)$. The next proposition describes how the size of the informal sector varies with the parameters of the economy:

PROPOSITION 1 *The size of the informal sector increases with the regulatory burden R , while it decreases with overall productivity ϑ , with the efficiency of the provision of services α , and with enforcement E .*

The intuition runs as follows. Firms weight the benefits of being formal against the costs; thus, when regulation decreases or enforcement increases, the formal sector becomes relatively more attractive for them, and more firms join it. Moreover, regulation is a fixed cost that all formal firms have to bear. Therefore, when overall productivity increases, the cost of regulation becomes proportionally smaller, so that more firms join the formal sector.

Informality Cycles

The previous section describes how informal activity (measured by the size of the informal sector) reacts in the long run to changes in overall productivity, the quality of public services, regulation, and enforcement. Next, we shall see how the same setup also explains how informal activity reacts to the business cycle. To this end, we presume that the cycle is caused by changes in overall productivity ϑ , so that the relationship between informal activity and the business cycle is described by the elasticity of informal activity with respect to ϑ , which is equal to:

$$\varepsilon_{\vartheta} = \frac{d\Xi[s_I]}{d\vartheta} \frac{\vartheta}{\Xi[s_I]} = \frac{d\Xi[s_I]}{ds_I} \frac{ds_I}{d\vartheta} \frac{\vartheta}{\Xi[s_I]} = -\frac{F(s_I)}{F'(s_I)} \frac{\xi(s_I)}{\Xi[s_I]} \quad (2)$$

In our model, informal activity reacts therefore counter-cyclically to the business cycle, as regulation being a fixed cost, more workers are willing to join the formal sector during a high cycle.¹ Several other factors also affect the elasticity of informal activity (and, therefore, the size of the cycle). Equation (2) shows that, everything else being equal, *counter-cyclicality is smaller in countries with large informal economies*. The reason is

¹If our model had allowed for different productivity shocks to informal and formal firms, the possibility of pro-cyclical informality would have arisen. Fiess, Fugazza, and Maloney (2006) considers this possibility, obtaining pro-cyclical informal behavior when shocks to the non-tradable sector dominate.

simple: a productivity shock affects the *absolute* number of workers switching sectors, so that in *relative* terms, countries with larger informal economies have cycles of smaller magnitude. Similarly, the skill density at the threshold level $\xi(s_I)$ also influences the elasticity of informal activity, as, for equal productivity shocks, if more workers have skills close to the threshold level s_I , the magnitude of the cycle becomes larger. Finally, the term $F(s_I)/F'(s_I)$ takes into account how much workers' productivity varies around the skill threshold s_I . More precisely, if productivity varies much around the skill threshold s_I (i.e., $F(s_I)/F'(s_I)$ is small), then, everything else being equal, under a shock few workers are going to switch sectors as the required change in s_I necessary to adjust to the shock is small; in contrast, if productivity varies little around the skill threshold s_I (i.e., $F(s_I)/F'(s_I)$ is large), then, everything else being equal, under a shock more workers are going to switch sectors. Notice, also, that the skill threshold s_I represents a sufficient statistics in describing the elasticity ε_θ , so that to study how informal activity reacts to the cycle it is sufficient to look at how the elasticity varies with s_I . In doing the comparative statics, however, we shall make two assumptions that will be justified by the empirical analysis. First, notice that the skill distribution $\xi(s)$ varies from country to country, and that it is empirically not possible to measure it. Hence, in this section we consider it as constant, and in the empirical section we shall consider it as an exogenous random error term. Similarly, as the size of the informal sector is empirically measurable, and across countries there is not necessarily a relationship between s_I and Ξ , in doing the comparative statics we shall consider the size of the informal sector Ξ as an exogenous variable, and in the empirical section we shall then control for it. The next proposition summarizes the behavior of the elasticity:

PROPOSITION 2

1. *The informal sector behaves counter-cyclically with respect to the business cycle.*

2. *Everything else being equal, informality is less counter-cyclical in countries with larger informal economies.*
3. *Everything else being equal, informality is less counter-cyclical in countries with higher overall productivity, stronger enforcement, and in countries which provide better public services.*
4. *Everything else being equal, informality is more counter-cyclical in countries with higher regulation.*

PROOF OF PROPOSITION 2 The derivative of the elasticity with respect to Ξ is equal to:

$$\frac{\partial}{\partial \Xi} \varepsilon_{\vartheta}(s_I, \Xi) = \frac{F(s_I)}{F'(s_I)} \frac{\xi}{\Xi^2} > 0 \quad (3)$$

Similarly, the derivative of the elasticity with respect to s_I is equal to:

$$\frac{\partial}{\partial s_I} \varepsilon_{\vartheta}(s_I, \Xi) = - \left(1 - \frac{F(s_I)F''(s_I)}{F'(s_I)^2} \right) \frac{\xi}{\Xi} < 0 \quad (4)$$

where the negative sign stems from the concavity assumption on $F(s)$. Equation (1) concludes the proof. END OF PROOF.

Propositions 1 and 2 have clear implication on how overall productivity, regulation, enforcement, and the efficient provision of public services affect informality trends and cycles. Next, we shall look at how these implications compare with empirical evidence.

3 Empirics

The empirical section of the paper has a dual objective. First, it will analyze the long-run relationship between informal employment and its main correlates suggested by the theory. This will serve to validate the long-run component of the model and to derive a measure of disequilibrium in informal employment that will be used in the next part of the analysis.

Second, the empirical section will study the short-run movements in informal employment. In particular it will examine how informality reacts to changes in aggregate production, that is, whether informal employment behaves counter- or pro-cyclically with respect to the business cycle. The analysis will allow the informality response to be heterogeneous across countries and will consider whether it varies systematically with the level of informality itself, as suggested by the model.

The connection between the two sections of the empirical analysis is given by an error-correction framework. It is this framework which best fits our theoretical model: In the short run, informal employment moves both in response to production shocks and in order to close the gap that separates it from its trend, equilibrium level; in the long run, informal employment moves in conjunction with the development characteristics that influence and are influenced by informality.

The data consist of annual observations for a collection of developing and developed countries over diverse time periods.² The observations are given at the national level, so that the variation in the data (and, thus, its informational value) resides on their cross-country and time-series dimensions. In principle, if we had sufficiently large time series per country, we could estimate both long- and short-run relationships using individual country data. Since this is not the case, we need to make some identifying assumptions, which are either directly supported by the data or validated by our results. The first is

²The next subsection provides detailed information on sample composition, variable definitions, and data sources.

that the long-run relationship can be estimated from the cross-country variation in the data. Although for efficiency purposes time-series data are also used, the within-country variation is quite small relative to the cross-country variation. Therefore, identification of long-run parameters comes from the comparison across countries. Figure 1 illustrates this point for the relationship between informal employment and GDP per capita.

The second assumption is that the speed of adjustment to the (equilibrium) trend relationship is homogeneous across countries. This implies a stable dynamic process for informal employment that is common to all countries, including those whose limited time-series data would not clearly reveal such a process. Finally, the third assumption is that the short-run response of informal employment to cyclical movements in aggregate output is heterogeneous across countries but in a systematic way that links this response to specific country characteristics. Next we describe how these assumptions affect the long- and short-run specifications of the empirical model.

LONG RUN

First we examine whether in the long run informal employment is determined by the flexibility of the formal business environment, the quality of public services available to formal enterprises (e.g., the police and judicial system), and the enforcement of taxes and business regulations. The first two factors should determine the opportunity cost of informality, while the last one would represent its direct cost. Since informal employment is also related to other features of underdevelopment – such as the lack of education, rudimentary infrastructure, and laggard technology – we also relate the level of informal employment to a country’s per capita GDP. In light of these considerations, a straightforward representation of the long-run regression equation is given by:

$$I_{ct} = \alpha_0 + \alpha_1 Y_{ct} + \alpha_2 Flex_c + \alpha_3 Law_c + \alpha_4 Govex_c + \delta_{ct} \quad (5)$$

where the subscripts c and t represent country and time period, respectively. I represents informal employment, proxied by the ratio of self to total employment, as reported by the International Labor Organization. Y denotes the average level of income, as measured by the log of per capita GDP. $Flex$ represents business flexibility, proxied by the Fraser Institute index of credit, labor, and regulatory flexibility. Law measures the enforcement of contracts and the prevalence of the rule of law and the efficiency of the police and judicial systems, proxied by the International Country Risk Guide index of law and order. Finally, $Govex$ is the ratio of government expenditures to GDP and attempts to measure the government's ability to monitor and enforce formal taxes and regulations; we assume, therefore, that this ability is linked to the availability of government's financial resources. The variables $Flex$, Law , and $Govex$ are measured as country averages due to their stability over time and, for $Flex$ and Law , to the incompleteness of their data in the time dimension.

SHORT RUN

As mentioned above, the short run is modeled as an error-correction process, where informal employment changes in response to output shocks and in order to close the gap that separates it from its long-run equilibrium level. A simple formulation of the short-run process is given by:

$$\Delta I_{ct} = \beta_0 + \beta_{1c}\Delta Y_{ct} + \beta_2 LRdev_{ct-1} + \varepsilon_{ct} \quad (6)$$

where $LRdev$ represents the deviation of informal employment from its trend value, as derived from the long-run equation; and Δ is the difference operator denoting the (proportional) change with respect to the previous year's value.³ In this formulation the

³In the case of informal employment, the proportional difference is computed as the absolute difference divided by the previous year's value: $(I_t - I_{t-1})/I_{t-1}$. In the case of GDP per capita, the proportional difference is simply computed as the difference of the logs.

response of informal employment to output changes can vary from country to country (which is indicated by the subscript c attached to β_1). We consider whether this heterogeneity is systematic by allowing β_1 to be a (linear) function of the (average) levels of informal employment and, potentially, per capita GDP, business flexibility, rule of law, and government expenditures. Thus, the short-run regression equation can be written as:

$$\begin{aligned} \Delta I_{ct} = & \beta_0 + (\beta_{10} + \beta_{11} I_c + \beta_{12} Y_c + \beta_{13} Flex_c + \beta_{14} Law_c \\ & + \beta_{15} Govec) \Delta Y_{ct} + \beta_2 LRdev_{ct-1} + \varepsilon_{ct} \end{aligned} \quad (7)$$

Estimation is conducted in two sequential steps: first, we estimate the long-run relationship, analyze it, and use it to derive the deviations from equilibrium. Second, we estimate and analyze the short-run equation. Before discussing the results, the next subsection presents our data set, providing information on sample composition, variable definitions, and data sources.

Data and Sample

Our measure of informal employment corresponds to the percentage of the active labor force that is self employed. In most developing countries, there is a strong association between self-employment and informal activity, as most self employed workers tend to be low-skilled, unregistered workers. In fact, self employment correlates well with other estimates of informal activity such as the Schneider (2005) measure of informal production: the correlation among non-Eastern European countries equals 0.75. In addition, it presents some advantages with respect to traditional estimates of informal production

based on excess currency demand or latent variable methodologies (e.g., the MIMIC model). The first advantage is conceptual: what self-employment measures is clear and well defined (although we can argue that self-employment does not comprise all informal activity). In contrast, methodologies based on estimated residuals or derived latent variables render data whose meaning is subject to multiple interpretations. The second advantage is practical: self-employment data are available not only for a cross section of countries but also for several consecutive years per country. Conversely, most other measures are either only available across countries or have very limited time-series. Needless to say, for the type of research we undertake here both cross-country and time-series dimensions are crucial.

Data on self-employment are obtained from the International Labor Organization, which on its website (<http://laborsta.ilo.org>) publishes yearly employment statistics for most countries. Self-employment is measured as the percentage of self employed workers with respect to the total active population, and the full dataset contains 783 observations in 93 countries. We drop, however, some observations for the following reasons.

First, we drop countries from Eastern Europe, as self employment in these countries appears to be still in transition to market-economy levels. In particular, self-employment levels in Eastern Europe remain substantially lower than in non-Eastern European countries, and the gap persists after correcting for their level of per capita GDP and their institutional characteristics. This gap is likely the result of their socialist past, when employment took place exclusively in state enterprises. Since the gap is large and changing over time, including Eastern European countries would bias both the long and short run estimates.

Second, we drop countries that do not have at least four consecutive pairs of observations, which corresponds to the minimum threshold of observations for the short-run regression. This choice is a compromise between having sufficient time-series observa-

tions per country and not eliminating too many countries from the sample. At any rate, as a robustness check we shall vary the threshold from a minimum of two to a maximum of eight consecutive pairs of observations.

The variable *Flex* measures the regulatory environment, and consists of the Fraser Institute index of credit, labor, and regulatory flexibility (<http://www.fraserinstitute.ca>). *Flex* measures how much the regulation of credit and labor markets, and of the business environment, “restricts entry into markets and interferes with the freedom to engage in voluntary exchange” (Economic Freedom of the World, 2005). The index considers factors such as the presence of foreign Banks, interest rates controls, minimum wages, and firms’ entry costs. It varies from zero to ten, where ten represents the highest degree of flexibility.

The variable *Law* measures the degree to which contracts are enforced and the efficiency of the police and judicial systems. It consists of the International Country Risk Guide index of law and order (<http://www.icrgonline.com>), which measures the strength and impartiality of the legal system, and the popular observance of the law. The index varies from one to six, where six represents the highest degree of law and order.⁴ Finally, the variables *Y* and *Govex* represent, respectively, the log of GDP per capita in constant 2000 US Dollars, and general government expenditure as a percentage of GDP. Both variables are from the World Development Indicators (2005). Appendix 2 shows univariate summary statistics and bivariate correlations for all variables included in the empirical analysis.

We also consider two different samples of countries: all countries, and only low and middle income countries with a per capita GDP below 9000 US dollars. We work with

⁴A main caveat of the *Flex* and *Law* indexes is that they are based to a large extent on subjective assessments, such as firms’ surveys (for instance, 10 out of 15 components used to create the *Flex* index are based on survey data). Thus, the indexes can change from one year to the next if people’s perceptions changes (because, for instance, of a corruption scandal), even if structurally little has changed in the country. It is also to avoid these noisy fluctuations that we take countries’ averages of *Flex* and *Law*.

the two samples partly to check the robustness of the results and partly to consider the possibility that self-employment may have a different worker composition in rich as in developing countries. The final sample consists therefore of 471 observations in 42 countries when we consider all countries, and of 182 observations in 18 countries when we only consider low and middle-income countries. Appendix 1 lists the countries and corresponding years of coverage for each sample under study.

Results

Table 1 presents the results on the estimation of the basic long-run relationship. The connection between informal employment and GDP per capita is quite strong: higher informal employment is associated with lower GDP per capita, and this association is not only statistically significant but also economically meaningful (see Col. 1). In fact, 80% of the variation in informal employment can be explained by the variation in GDP per capita. The close connection between self-employment and GDP per capita has already been documented by, among others, Blau (1987), Maloney (2001), and Gollin (2002). Using Schneider's measures of informal production, Loayza, Oviedo, and Servén (2005) also find a strong relationship between informality and national income.

If we replace GDP per capita by the determinants of the opportunity and direct costs of informality, all of them carry coefficients with the expected negative sign and high statistical significance. Thus, informal employment is more prevalent when business flexibility, the rule of law, and government resources are weaker. The variation in these three variables explains 72% of the variation in informal employment, slightly less than the explanatory power of GDP per capita alone.

When we consider the most comprehensive model (which adds GDP per capita to business flexibility, the rule of law, and government resources as determinants of informal employment), we find that all variables retain their negative sign and three

of the four remain statistically significant –the exception being government resources. In all cases, the size of the coefficients is somewhat reduced, indicating that GDP per capita captures some of the effects of the variables measuring the costs of informality, and *vice versa*. Government resources is in fact so much associated with GDP per capita that its relationship with informal employment appears to be embedded in the informality-income relationship. Figure 2 compares the average informal employment per country with the corresponding predicted level according to this model. With an R-squared of 85%, the accuracy of the model’s fit is shared by most countries in the sample, with no discernable bias for countries in different income levels (or those in Latin America in particular).

As the following exercises show, the basic results are quite robust to different samples and to the influence of outliers. Table 2 presents the results for the samples obtained using different criteria for the minimum number of time-series observations per country. The estimation of the error-correction model relies on having sufficient consecutive observations for each country. Using a large time-series dimension improves the model’s ability of capturing informality’s dynamic process but comes at the cost of eliminating countries with few annual observations. For our basic specification, we chose 4 consecutive pairs of annual observations as the minimum threshold. In this robustness exercise, we first relax this criterion (applying a minimum of 2 consecutive pairs and, thus, allowing more countries in the sample) and then restrict it further (applying a threshold of 8 pairs and, therefore, including fewer countries). Moreover, for each of the three cases, we consider not only OLS estimation but also a weighted least squares (WLS) procedure that reduces the influence of outliers. The results are remarkably robust, despite the considerable variations in the data and the potential presence of outlying observations.

Also, our basic results are obtained using a sample of countries comprising both developed and developing countries. Table 3 examines the robustness of these results

to the use of a sample of developing countries only. In this case the variation in all variables is significantly more limited and the sample size is reduced to fewer than half observations and countries. Not surprisingly, the R-squared falls to 66%, indicating that the informational value that developed countries bring about is rather important. Despite this large change, most results are robust. GDP per capita and the rule of law remain negatively and significantly related to informal employment. The coefficient on government resources retains its negative sign and now becomes statistically significant; on the other hand, business flexibility loses relevance. The simpler models (not shown in the tables) are more robust. When only GDP per capita is included in the regression, it carries a negative and highly significant coefficient. Likewise, when we exclude GDP per capita from the regression— replacing it by the variables that measure the opportunity and direct costs of informality— business flexibility, the rule of law, and government expenditures all carry negative and significant coefficients, just as it happened for the full country sample. The table also shows the results obtained with the WLS procedure that limits the influence of extreme observations. They are basically the same as those obtained under OLS, confirming the results' robustness to potential outliers.

Table 3 also presents the long-run estimation using only Latin American and Caribbean countries. Despite the fact that the sample of observations and countries is smaller than one-third of the full sample, the results are quite similar; in fact, the only exception is that business flexibility does not carry a significant coefficient when GDP per capita is also considered as an explanatory variable. In the simpler model —where only the variables representing the costs of informality are included— business flexibility, the rule of law and government expenditures are negative and significantly related to informal employment, as it happened with the full sample of countries.

In summary, informal employment is more pervasive in countries having lower GDP per capita and imposing more costs (or generating less advantage) to formal enterprises, in the form of more rigid business regulations, less valuable public services, and weaker

monitoring of informality. The results are robust to changes in the sample and the influence of outlier observations.

SHORT RUN

We now examine how informal employment reacts to deviations from its equilibrium level and to fluctuations in GDP per capita growth. The basic results are presented in Table 4. Column 1 shows the simplest error-correction model, where full cross-country homogeneity is assumed. The coefficient on the deviation from the long run (also known as “adjustment” coefficient) is negative and statistically significant. This supports the assumption of dynamic stability in the sense that informal employment moves, at least in part, to return to its long-run equilibrium. The coefficient on the GDP per capita growth is also negative and statistically significant, indicating that in average informal employment behaves counter-cyclically.

In column 2, we allow the coefficient on GDP per capita growth to vary with the average level of informal employment. We find the interaction coefficient to be significantly positive, meaning that in countries with larger informal sector, informal employment tends to be less counter-cyclical. Figure 3 simulates the change in informal employment growth (and corresponding 90% confidence bands) due to a 5 percentage-point increase in GDP per capita growth as a function of the average level of informal employment. Figure 3 also identifies the predicted values for the Latin American countries in the sample. For the majority of countries, we estimate that informal employment moves counter to the business cycle. The exceptions are the countries with the largest informal sectors (such as Peru and Bolivia), for which informal employment appears to be a-cyclical.

In column 3 of Table 4, we consider the possibility that the determinants of informality have an independent effect on the counter-cyclicity of informal employment.

For this purpose, we enlarge the set of interactions with GDP per capita growth. We find that neither the level of income nor business flexibility affect the cyclical response of informality. Larger government expenditure appears to induce a more pro-cyclical response, but, as we will see below, this result is not robust. On the other hand, the interaction with law and order does appear to be robust, indicating that stronger rule of law reduces the counter-cyclical of informal employment.

Since the level of informality and the measure of the rule of law tend to go in opposite directions, their effects on the cyclical of informal employment could cancel each other. To consider this possibility, we estimate the elasticity of informal employment growth to changes in GDP per capita growth for all countries in the sample, taking into account only the significant interactions –that is, a model where GDP per capita growth is interacted with informal employment and rule of law only (not shown in the table). The results are presented in Figure 4. For ease of presentation, the elasticities (and corresponding 90% confidence bands) are ordered by country income and smoothed out by fitting a cubic spline. For 83% of the sample, the response of informality to economic growth is significantly negative, that is, counter-cyclical. For 15%, this response is not statistically different from zero, and only in one case (Vietnam) informality is estimated to be pro-cyclical.

To check how sensible the elasticities derived from the empirical model are, we match them with the elasticities obtained on a country-by-country basis (that is, through a set of country dummy variables interacted with GDP per capita growth in the error-correction model). The latter are subject to large imprecision due to the small size of the time-series dimension of most countries. However, in spite of this, the two sets of elasticities are reasonably in line with each other, having a correlation coefficient of the order of 0.65.

Next, we conduct robustness checks similar to those in the long-run analysis. In Table 5 we consider the robustness of the results to the presence of potential outliers

and changes in the minimum threshold of time-series coverage. In Table 6 we examine the robustness of the results to the samples of all developing and only Latin American countries, as well as to potential outliers in these samples. In both cases, results remain basically the same.

To summarize the robust results, the deviation from the long-run (or “error-correction” term) always carries a negative and highly significant coefficient, indicating informal employment’s tendency to trend reversion. The growth rate of GDP per capita also carries a negative and significant coefficient, which has to be considered jointly with the significant interaction coefficients to establish the cyclicity of informal employment. Only two interactions are significant in all robustness exercises: they are the interactions of GDP per capita growth with the level of informal employment and with the rule of law index. Informality’s counter-cyclicity decreases with the level of informal employment and, independently, decreases with the rule of law. Since improvements in the rule of law have the effect of reducing informal employment in the long run, the two interactions with GDP per capita growth would tend to go in opposite directions. However, using the actual values in the sample under study, we find that for the large majority of countries (and not only for the average), informal employment is significantly and robustly countercyclical. In brief, the short-run results are robust to changes in the sample and the influence of outliers, including the cases where the informational value of the variables is reduced as when only developing and Latin American countries are considered.

4 Conclusions

This paper studies the trends and cycles of informal employment. It first presents a theoretical model where the size of informal employment is determined by the relative

costs and benefits of informality (in terms of regulatory burden, enforcement, and access to public services) and the distribution of workers' skills. In the long run informal employment varies with the trends in these variables, and in the short run it reacts to accommodate transient shocks and also to close the gap that separates it from its trend level.

The paper then examines empirically informality's long- and short-run relationships. It does so in the context of an error-correction framework, using country-level data at annual frequency for a sample of developed and developing countries. Using the share of self-employment in the labor force as proxy for informal employment, the paper finds that in the long run informality is larger in countries that have lower GDP per capita and impose more costs to formal firms, in the form of more rigid business regulations, less valuable public services, and weaker monitoring of informality. The results are robust to the criteria used to obtain the sample and to the influence of outlier observations.

The short-run results indicate that informal employment follows a stable, trend-reverting process. Moreover, informal employment is found to be counter-cyclical for the majority of countries. Informality's counter-cyclicity decreases with the level of informal employment and, independently, decreases with the quality of policy and judicial services (and less significantly with GDP per capita, business regulatory flexibility, and strength of enforcement). These results are robust to changes in the sample and to the influence of outliers, even when only developing countries are considered in the analysis.

References

- Blau, D. (1987), "A Times-Series Analysis of Self-Employment in the United States," *The Journal of Political Economy*, 95(3), 445-67.
- Bosch, M., and W. Maloney (2005), "Labor Market Dynamics in Developing Countries: Comparative Analysis Using Continuous Time Markov Processes," World Bank Policy Research Working Paper 3583.
- De Paula, A., and J. A. Scheinkman (2006), "The Informal Sector," mimeo, University of Pennsylvania.
- De Soto, H. (1989), *The Other Path*, Basic Books.
- Dessy, S., and Pallage, S. (2003), "Taxes, Inequality, and the Size of the Informal Sector," *Journal of Development Economics*, 70, 225-33.
- Djankov, S., R. La Porta, F. Lopez de Silanes, and A. Shleifer (2002), "The Regulation of Entry," *The Quarterly Journal of Economics*, 117(1), 1-37.
- Fiess, N., M. Fugazza, and W. F. Maloney (2006), "Informal Labor Markets and Macroeconomic Fluctuations," mimeo, The World Bank.
- The Fraser Institute (2005), *Economic Freedom of the World 2005 Annual Report*.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton (2000), "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries," *Journal of Public Economics*, 76, 459-93.
- Gollin, D. (2002), "Getting Income Shares Right," *The Journal of Political Economy*, 110(2), 458-74.
- Johnson, S., Kaufmann, D., Shleifer, A., Goldman, M.I., and Weitzman, M.L. (1997), "The Unofficial Economy in Transition," *Brookings Papers on Economic Activity*, 2, 159-239.

- Loayza, N. (1996), “The Economics of the Informal Sector: A Simple Model and Some Evidence from Latin America,” *Carnegie-Rochester Conference Series on Public Policy*, 45, 129-62.
- Loayza, N., A. M. Oviedo, and L. Serven (2005), “The Impact of Regulation on Growth and Informality – Cross-Country Evidence,” World Bank Policy Research Working Paper 3263.
- Maloney, W. (2001), “Self Employment and Labor Turnover in Developing Countries: Cross-Country Evidence,” in *World Bank Economists’ Forum*, Shantayanan Devarajan, F. Halsey Rogers, and Lyn Squire Eds., The World Bank.
- Maloney, W. F. (2004), “Informality Revisited,” *World Development*, 32(7): 1159-1178.
- Rausch, J.E. (1991), “Modelling the Informal Sector Formally,” *Journal of Development Economics*, 35, 33-47.
- Sarte, P.-D. (2000), “Informality and Rent-Seeking Bureaucracies in a Model of Long-Run Growth,” *Journal of Monetary Economics*, 46, 173-97.
- Schneider, F., and D. H. Enste (2000) , “Shadow Economies: Size, Causes, and Consequences,” *Journal of Economic Literature*, 38, 77-114.
- Schneider, F. (2005), “Shadow Economies of 145 Countries all over the World: What do we really know?,” mimeo, Johannes Kepler University of Linz.
- The World Bank (2005), *World Development Indicators*, www.worldbank.org.

Table 1. Long-Run Informality Relationships

Dependent variable: Self employment rate (as ratio to total workers)			
	OLS (1)	OLS (2)	OLS (3)
GDP per capita (in logs, annual)	-0.0759 *** 0.0043		-0.0516 *** 0.0060
Business flexibility (index from Fraser Institute, range:0-10, country average)		-0.0293 ** 0.0111	-0.0167 * 0.0092
Law and order (index from ICRG, range:0-6, country average)		-0.0457 *** 0.0072	-0.0191 *** 0.0050
Government expenditure (as % of GDP, country average)		-0.0050 ** 0.0022	-0.0015 0.0015
Constant	0.9065 *** 0.0388	0.6954 *** 0.0666	0.9030 *** 0.0424
Observations/Countries	525/42	525/42	525/42
R-squared	0.80	0.72	0.85

Numbers in parentheses are robust, (country) clustered standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

Table 2. Robustness to Outliers and Minimum Time-Series Coverage, Long-Run

Dependent variable: Self employment rate (as ratio to total workers)						
	Min.Obs. = 4		Min.Obs. = 2		Min.Obs. = 8	
	OLS (1)	WLS (2)	OLS (3)	WLS (4)	OLS (5)	WLS (6)
GDP per capita (in logs, annual)	-0.0516 *** 0.0060	-0.0529 *** 0.0052	-0.0515 *** 0.0058	-0.0530 *** 0.0051	-0.0573 *** 0.0072	-0.0580 *** 0.0066
Business flexibility (index from Fraser Institute, range:0-10, country average)	-0.0167 * 0.0092	-0.0158 * 0.0082	-0.0169 * 0.0092	-0.0158 * 0.0081	-0.0189 ** 0.0094	-0.0178 ** 0.0083
Law and order (index from ICRG, range:0-6, country average)	-0.0191 *** 0.0050	-0.0184 *** 0.0045	-0.0185 *** 0.0046	-0.0177 *** 0.0041	-0.0127 ** 0.0052	-0.0125 ** 0.0048
Government expenditure (as % of GDP, country average)	-0.0015 0.0015	-0.0013 0.0014	-0.0017 0.0014	-0.0016 0.0013	-0.0013 0.0018	-0.0011 0.0016
Constant	0.9030 *** 0.0424	0.9047 *** 0.0380	0.9036 *** 0.0421	0.9049 *** 0.0375	0.9390 *** 0.0520	0.9329 *** 0.0464
Observations/Countries	525/42	525/42	546/47	546/47	457/33	457/33
R-squared	0.85	0.87	0.85	0.87	0.83	0.85

Numbers in parentheses are (robust) standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes:1. Min.Obs. Refers to the minimum number of consecutive pairs of annual self-employment observations that a country must have in order to be included in the sample. Consecutive pairs are needed to obtain observations in growth rates as required by the short-run analysis.

2.WLS is a weighted least squares procedure that reduces the influence of outliers in the estimation.

Table 3. Robustness to Sample of Countries, Long-Run

Dependent variable: Self employment rate (as ratio to total workers)						
	Full		Developing		LAC	
	OLS (1)	WLS (2)	OLS (3)	WLS (4)	OLS (5)	WLS (6)
GDP per capita (in logs, annual)	-0.0516 *** 0.0060	-0.0529 *** 0.0052	-0.0579 *** 0.0086	-0.0575 *** 0.0080	-0.0607 *** 0.0126	-0.0597 *** 0.0113
Business flexibility (index from Fraser Institute, range:0-10, country average)	-0.0167 * 0.0092	-0.0158 * 0.0082	0.0045 0.0111	0.0043 0.0096	0.0036 0.0121	0.0022 0.0108
Law and order (index from ICRG, range:0-6, country average)	-0.0191 *** 0.0050	-0.0184 *** 0.0045	-0.0301 *** 0.0079	-0.0290 *** 0.0070	-0.0243 ** 0.0096	-0.0229 ** 0.0087
Government expenditure (as % of GDP, country average)	-0.0015 0.0015	-0.0013 0.0014	-0.0062 * 0.0035	-0.0063 * 0.0030	-0.0054 0.0052	-0.0051 0.0046
Constant	0.9030 *** 0.0424	0.9047 *** 0.0380	0.9155 *** 0.0849	0.9110 *** 0.0765	0.9215 *** 0.1045	0.9115 *** 0.0946
Observations/Countries	525/42	525/42	205/18	205/18	149/12	149/12
R-squared	0.85	0.87	0.66	0.69	0.65	0.67

Numbers in parentheses are (robust) standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: 1. "Full" sample includes all high-income and developing countries with available data.

"Developing" sample includes only developing countries --those with avg. per capita income lower than \$9,000 (at 2000 prices).

"LAC" sample includes only countries in Latin America and the Caribbean.

2. WLS is a weighted least squares procedure that reduces the influence of outliers in the estimation.

Table 4. Short-Run Informality Relationships

Dependent variable: Annual change in the self employment rate			
	OLS (1)	OLS (2)	OLS (3)
Deviation from Long-Run	-0.1752 *** 0.0460	-0.2014 *** 0.0552	-0.2880 *** 0.0862
$\Delta \ln(\text{GDPpc})$	-0.3132 *** 0.0791	-0.8603 *** 0.2008	-5.0215 ** 2.2080
Interactions:			
$\Delta \ln(\text{GDPpc}) * \text{Self}$		2.1899 *** 0.6247	6.5559 *** 2.3540
$\Delta \ln(\text{GDPpc}) * \text{GDPpc}$			0.1613 0.1265
$\Delta \ln(\text{GDPpc}) * \text{Business}$			0.0864 0.0765
$\Delta \ln(\text{GDPpc}) * \text{Law and Order}$			0.1897 *** 0.0552
$\Delta \ln(\text{GDPpc}) * \text{Gov. Expenditure}$			0.0232 * 0.0134
Constant	0.0073 ** 0.0026	0.0087 ** 0.0027	0.0073 ** 0.0027
Observations/Countries	475/42	475/42	475/42
R-squared	0.08	0.11	0.13

Numbers in parentheses are (robust) standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: 1. The deviation from the long-run relationship is the difference between the actual and the projected self-employment rate, where the latter is given by the estimated long-run relationship in the benchmark case (see Table 1, Col. 3)

2. Variables interacted with the change in GDP per capita correspond to their country (time invariant) averages.

Table 5. Robustness to Outliers and Minimum Time-Series Coverage, Short-Run

Dependent variable: Annual change in the self employment rate						
	Min.Obs. = 4		Min.Obs. = 2		Min.Obs. = 8	
	OLS (1)	WLS (2)	OLS (3)	WLS (4)	OLS (5)	WLS (6)
Deviation from Long-Run	-0.2880 *** 0.0862	-0.2021 *** 0.0406	-0.2930 *** 0.0855	-0.2092 *** 0.0422	-0.2724 *** 0.0979	-0.1926 *** 0.0476
$\Delta \ln(\text{GDPpc})$	-5.0215 ** 2.2080	-3.2063 *** 1.1178	-5.3415 ** 2.2471	-3.5869 *** 1.2361	-5.5050 ** 2.5333	-4.0404 *** 1.2647
Interactions:						
$\Delta \ln(\text{GDPpc})$ * Self	6.5559 *** 2.3540	4.2601 *** 1.1441	6.8706 *** 2.3916	4.7160 *** 1.2765	6.7256 *** 2.6529	4.9347 *** 1.3130
$\Delta \ln(\text{GDPpc})$ * GDPpc	0.1613 0.1265	0.1090 0.0688	0.1632 0.1329	0.1224 0.0773	0.2160 0.1491	0.1605 ** 0.0737
$\Delta \ln(\text{GDPpc})$ * Business	0.0864 0.0765	0.0515 0.0394	0.0947 0.0769	0.0322 0.0425	0.1099 0.0871	0.1255 ** 0.0545
$\Delta \ln(\text{GDPpc})$ * Law and Order	0.1897 *** 0.0552	0.1566 *** 0.0336	0.2223 *** 0.0572	0.2217 *** 0.0433	0.1419 *** 0.0553	0.0545 ** 0.0442
$\Delta \ln(\text{GDPpc})$ * Gov. Expenditure	0.0232 * 0.0134	-0.0025 0.0087	0.0246 * 0.0136	-0.0055 0.0087	0.0236 0.0146	0.0011 0.0120
Constant	0.0073 ** 0.0027	0.0038 ** 0.0019	0.0072 ** 0.0026	0.0041 ** 0.0018	0.0076 *** 0.0029	0.0030 0.0023
Observations/Countries	475/42	471/42	489/47	485/47	419/33	415/33
R-squared	0.13	0.13	0.15	0.17	0.13	0.12

Numbers in parentheses are (robust) standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

- Notes: 1. The deviation from the long-run relationship is the difference between the actual and projected self-employment rate, where the latter is given by the estimated long-run relationship in the benchmark case (see Table1, Col.3)
2. Variables interacted with the change in GDP per capita correspond to their country (time invariant) averages.
3. Min.Obs. Refers to the minimum number of consecutive pairs of annual self-employment observations that a country must have in order to be included in the sample. Consecutive pairs are needed to obtain observations in growth rates as required by the short-tun analysis.
4. WLS is a weighted least squares procedure that reduces the influence of outliers in the estimation.

Table 6. Robustness to Sample of Countries, Short-Run

Dependent variable: Annual change in the self employment rate						
	Full		Developing		LAC	
	OLS (1)	WLS (2)	OLS (3)	WLS (4)	OLS (5)	WLS (6)
Deviation from Long-Run	-0.2880 *** 0.0862	-0.2021 *** 0.0406	-0.2391 ** 0.0967	-0.2198 *** 0.0682	-0.3090 *** 0.0774	-0.2941 *** 0.0590
$\Delta \ln(\text{GDPpc})$	-5.0215 ** 2.2080	-3.2063 *** 1.1178	-2.4896 1.5949	-2.6317 * 1.4368	-5.1583 *** 0.9309	-5.5224 *** 0.9558
Interactions:						
$\Delta \ln(\text{GDPpc}) * \text{Self}$	6.5559 *** 2.3540	4.2601 *** 1.1441	3.2541 * 1.5838	2.8307 * 1.5579	6.5929 *** 0.9214	7.3597 *** 0.8482
$\Delta \ln(\text{GDPpc}) * \text{GDPpc}$	0.1613 0.1265	0.1090 0.0688	0.0036 0.1199	0.0692 0.1040	0.2149 ** 0.0733	0.2649 *** 0.0705
$\Delta \ln(\text{GDPpc}) * \text{Business}$	0.0864 0.0765	0.0515 0.0394	-0.0520 0.0817	-0.0954 0.0642	0.0090 0.0785	0.0242 0.0613
$\Delta \ln(\text{GDPpc}) * \text{Law and Order}$	0.1897 *** 0.0552	0.1566 *** 0.0336	0.2449 *** 0.0841	0.2746 *** 0.0754	0.2517 *** 0.0332	0.2892 *** 0.0362
$\Delta \ln(\text{GDPpc}) * \text{Gov. Expenditure}$	0.0232 * 0.0134	-0.0025 0.0087	0.0521 0.0420	0.0413 0.0393	0.0171 0.0254	-0.0244 0.0283
Constant	0.0073 ** 0.0027	0.0038 ** 0.0019	0.0079 ** 0.0031	0.0090 *** 0.0024	0.0080 ** 0.0031	0.0090 *** 0.0025
Observations/Countries	475/42	471/42	182/18	182/18	133/12	133/12
R-squared	0.13	0.13	0.11	0.16	0.15	0.26

Numbers in parentheses are (robust) standard errors.

* significant at 10%; ** significant at 5%; *** significant at 1%

Notes: 1. The deviation from the long-run relationship is the difference between the actual and projected self-employment rate, where the latter is given by the estimated long-run relationship in the benchmark case (see Table 1, Col. 3)

2. Variables interacted with the change in GDP per capita correspond their country (time invariant) averages.

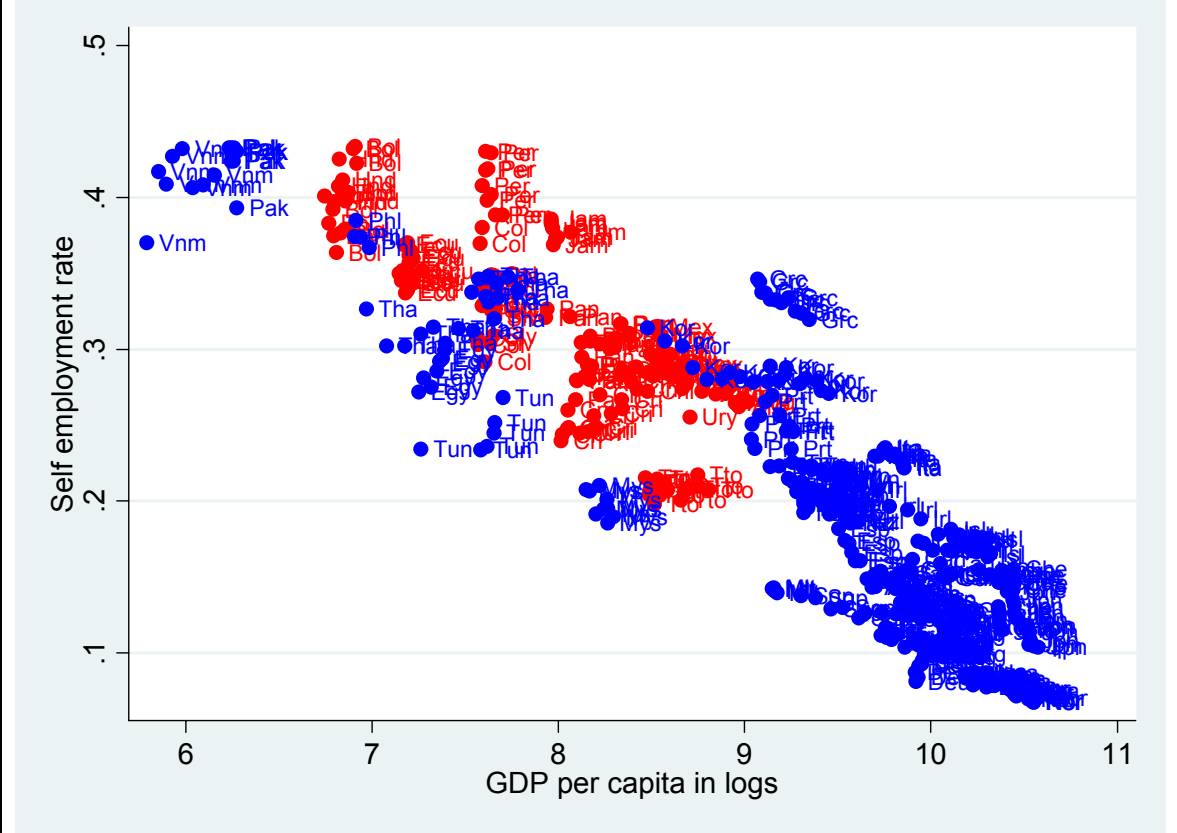
3. "Full" sample includes all high-income and developing countries with available data.

"Developing" sample includes only developing countries --those with avg. per capita income lower than \$9,000 (at 2000 prices).

"LAC" sample includes only countries in Latin America and the Caribbean.

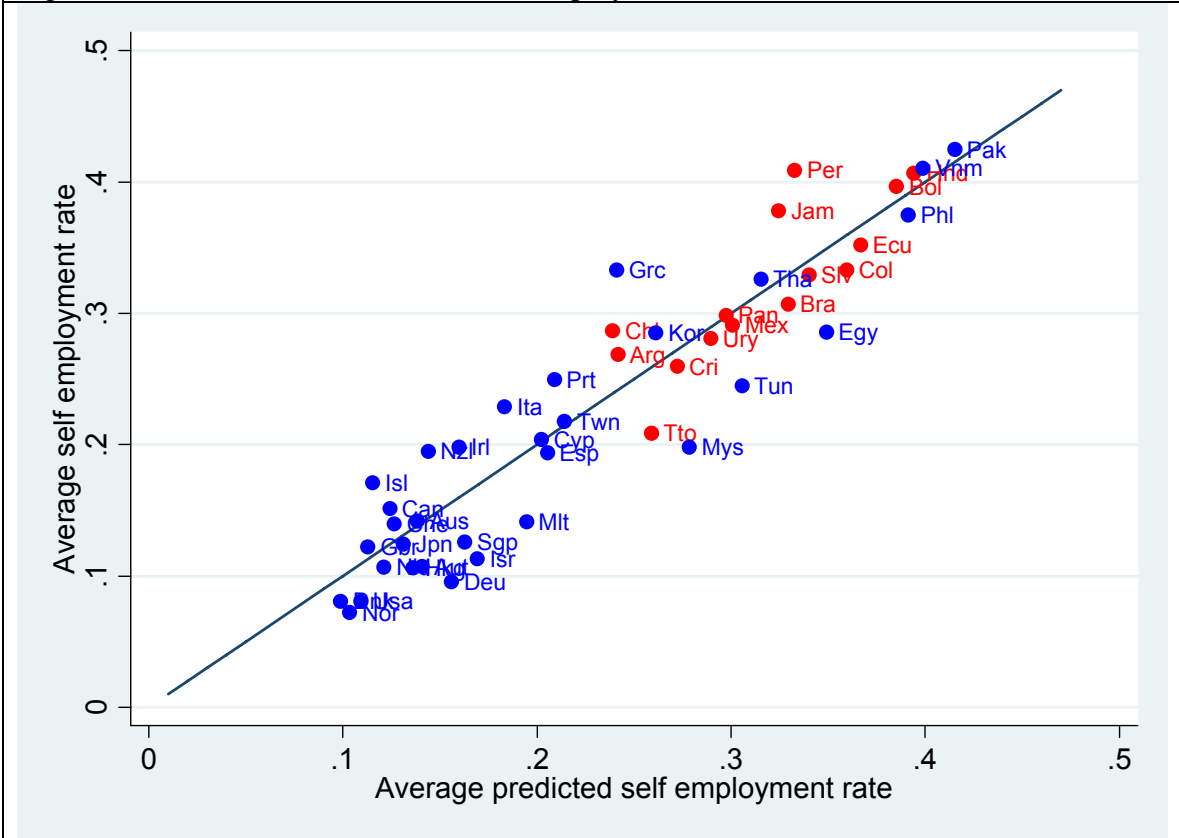
4. WLS is a weighted least squares procedure that reduces the influence of outliers in the estimation.

Figure 1. Informal Employment and GDP per Capita



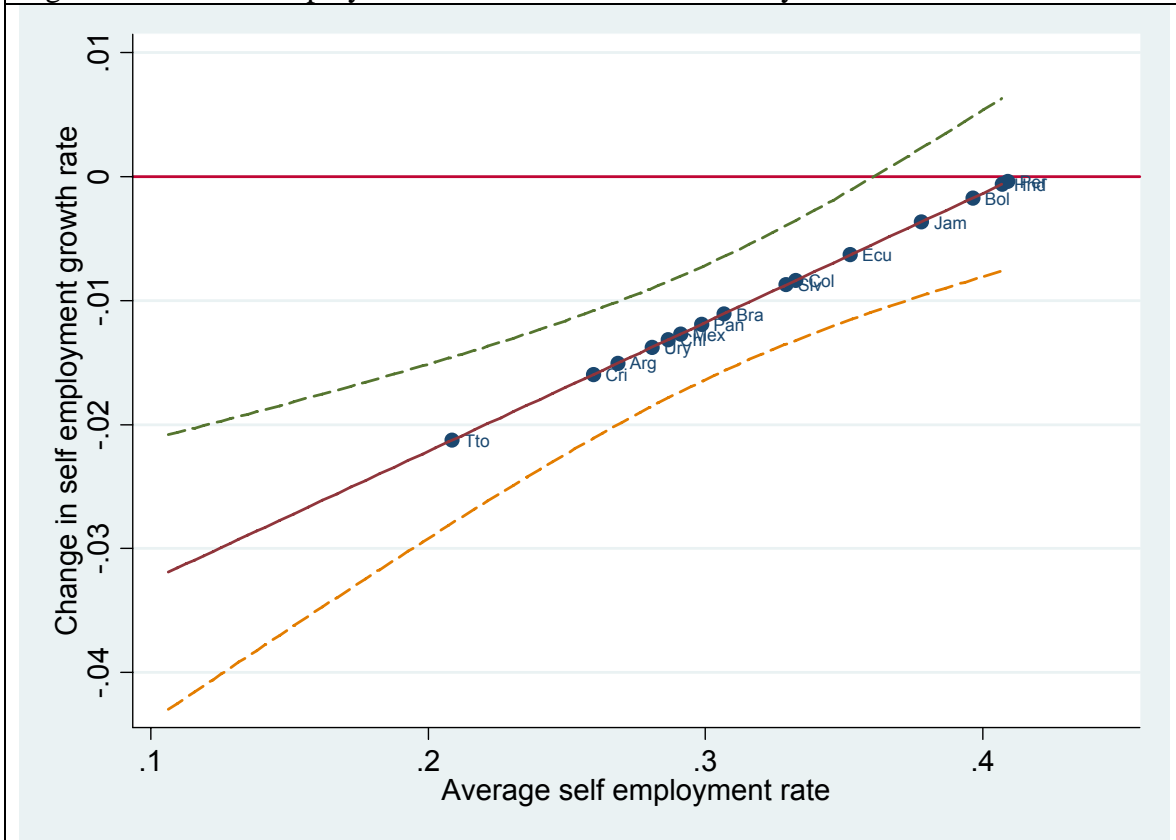
Note: Countries in Latin America and the Caribbean are highlighted in red.

Figure 2. Actual vs. Predicted Informal Employment



Note: 1. The predicted self-employment rate is derived from the long-run regression presented in Table 2, Col. 3 (min.obs.=2).
 2. Countries in Latin America and the Caribbean are highlighted in red.

Figure 3. Informal Employment Reaction to the Business Cycle

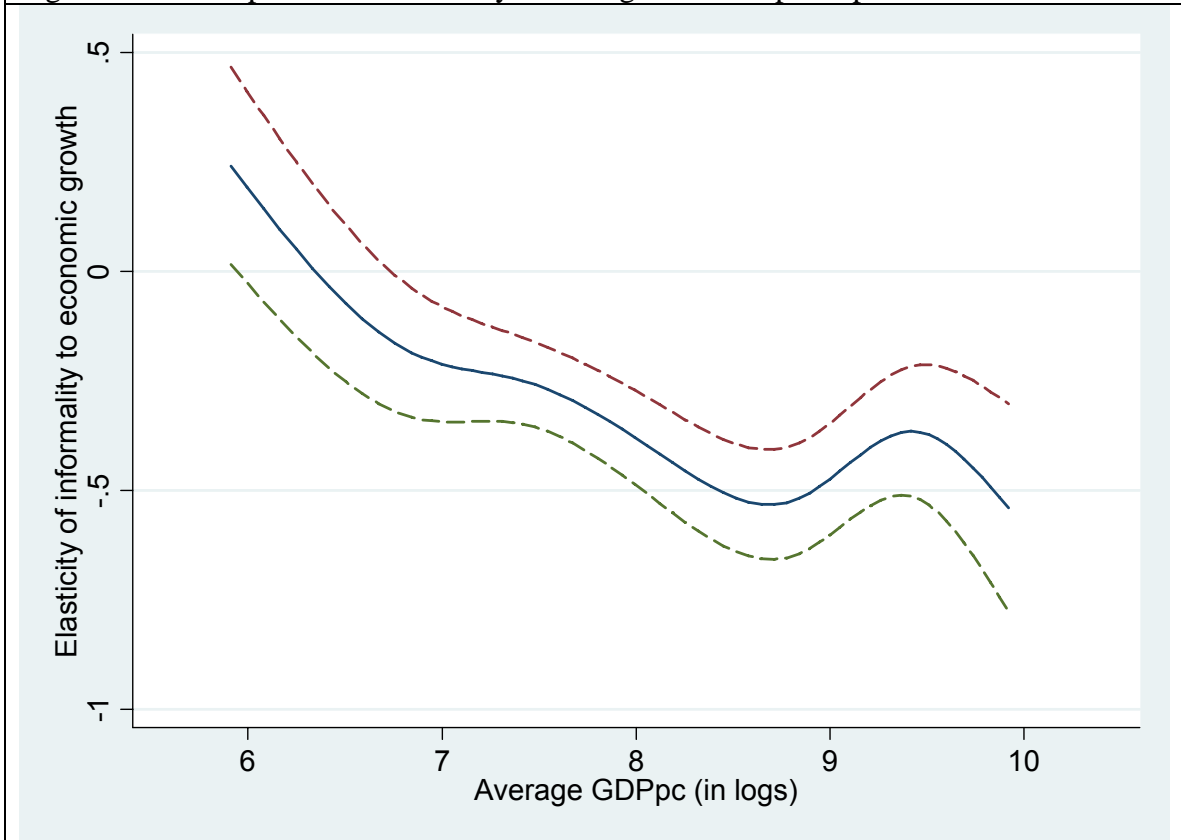


Note: 1. The graph simulates the change in informal employment growth due to a 5 percentage-point change in the GDP per capita growth rate. The simulation is based on a short-run regression like the one shown in Table 4, Col. 2, but with a sample resulting from applying the constraint of min. obs. =2.

2. The dotted lines are the 90% confidence bands.

3. Only countries in Latin America and the Caribbean are highlighted. From left to right they are Trinidad and Tobago, Costa Rica, Argentina, Uruguay, Chile, Mexico, Panama, Brazil, El Salvador, Colombia, Ecuador, Jamaica, Bolivia, Honduras, and Peru.

Figure 4. The Response of Informality to Changes in GDP per capita Growth



Note: 1. The elasticity of informality to economic growth is obtained from the regression:
 $\Delta I_{ct} = \beta_0 + (\beta_{10} + \beta_{11}I_c + \beta_{12}Law_c)\Delta y_{ct} + \beta_2 LRdev_{ct-1} + \varepsilon_{ct}$. For each country, the elasticity is given by $\hat{\beta}_0 + \hat{\beta}_{11}I_c + \hat{\beta}_{12}Law_c$ and the corresponding variance is equal to $Var[\Delta I_{ct}] = Var(\beta_{10}) + Var(\beta_{11})I_c^2 + Var(\beta_{12})Law_c^2 + 2Cov(\beta_{11}, \beta_{12})I_c Law_c + 2Cov(\beta_{10}, \beta_{11})I_c + 2Cov(\beta_{10}, \beta_{12})Law_c$. The sample consists of all countries (47) with min.obs.=2.

2. The point estimates correspond to the 90% confidence bands are smoothed out using the cross medians as knots to fit a cubic spline.

3. Average GDP per capita (in the horizontal axis) is used only as an order criterion for the elasticities.

Appendix 1. Sample of Countries

Country Name	Period	Min.Obs.=2 (47countries)	Min.Obs.=4 (42countries)	Min.Obs.=8 (33countries)	Developing (18countries)	LAC (12countries)
Argentina	1996 - 2003	√	√		√	√
Australia	1987 - 2004	√	√	√		
Austria	1994 - 2004	√	√	√		
Bolivia	1989 - 2000	√	√	√	√	√
Brazil	2001 - 2003	√				
Canada	1987 - 2004	√	√	√		
Chile	1996 - 2004	√	√	√	√	√
Colombia	1992 - 2000	√	√	√	√	√
Costa Rica	1987 - 2004	√	√	√	√	√
Cyprus	1999 - 2004	√	√			
Denmark	1995 - 2004	√	√			
Ecuador	1988 - 2004	√	√	√	√	√
Egypt, Arab Rep.	1997 - 2003	√	√		√	
El Salvador	1995 - 2004	√	√	√	√	√
Germany	1991 - 2004	√	√	√		
Greece	1993 - 2003	√	√	√		
Hong Kong, China	1993 - 2004	√	√	√		
Honduras	1996 - 2004	√				
Ireland	1986 - 2004	√	√	√		
Iceland	1991 - 2002	√	√	√		
Israel	1995 - 2003	√	√	√		
Italy	1993 - 2003	√	√	√		
Jamaica	1997 - 2004	√	√		√	√
Japan	1987 - 2004	√	√	√		
Korea, Rep.	1986 - 2004	√	√	√		
Mexico	1991 - 2004	√	√	√	√	√
Malta	2001 - 2004	√				
Malaysia	1995 - 2003	√	√	√	√	
Netherlands	1995 - 2003	√	√			
Norway	1996 - 2004	√	√	√		
New Zealand	1991 - 2004	√	√	√		
Pakistan	1995 - 2002	√	√		√	
Panama	1982 - 2004	√	√	√	√	√
Peru	1996 - 2004	√	√	√	√	√
Philippines	2001 - 2004	√				
Portugal	1992 - 2003	√	√	√		
Singapore	1986 - 2003	√	√	√		
Spain	1986 - 2004	√	√	√		
Switzerland	1991 - 2004	√	√	√		
Taiwan, China	1994 - 2002	√	√	√		
Thailand	1987 - 2004	√	√	√	√	
Trinidad and Tobago	1987 - 2002	√	√	√	√	√
Tunisia	1999 - 2003	√	√		√	
United Kingdom	1992 - 2004	√	√	√		
Uruguay	2000 - 2003	√				
United States	1987 - 2004	√	√	√		
Vietnam	1996 - 2003	√	√		√	

Appendix 2. Descriptive Statistics for the 42-Country Sample

(a) Univariate

variable	Full				Developing					
	mean	median	min	max	sd	mean	median	min	max	sd
Self employment rate (ratio to total workers)	0.220	0.210	0.067	0.433	0.098	0.312	0.304	0.185	0.433	0.065
GDP per capita (in logs)	9.036	9.362	5.792	10.576	1.158	7.779	7.783	5.792	9.016	0.732
Business flexibility (0-10 index)	6.320	6.300	3.900	8.800	0.916	6.001	6.100	3.900	7.100	0.657
Law and Order (0-6 index)	4.658	5.000	1.000	6.000	1.352	3.475	3.667	1.000	5.000	1.094
Government expenditure (% of GDP)	15.279	14.648	6.233	31.229	4.774	12.371	11.740	6.233	22.734	3.204

(b) Bivariate Correlation: Upper triangle for full sample and lower triangle for developing country sample

variable	Self employment rate	GDP per capita	Business flexibility	Law and Order	Government expenditure
Self employment rate	1.00	-0.89	-0.50	-0.74	-0.49
GDP per capita	-0.73	1.00	0.44	0.73	0.50
Business flexibility	-0.37	0.49	1.00	0.41	0.11
Law and Order	-0.31	0.12	0.13	1.00	0.40
Government expenditure	-0.14	0.22	0.21	-0.31	1.00