Shades of Grey:
Measuring the Informal Economy Business Cycles[1]

Ceyhun Elgin, Ayhan Kose, Franziska Ohnsorge, and Shu Yu

This draft: October 2019

Abstract: The paper presents a comprehensive database of informality, with a focus on measures that have strong cross-country and over-time coverage. The database includes both model-based and survey-based measures of informality and covers more than 160 economies for as long as 1950-2016. The paper examines two applications of the database. First, it distills stylized facts of the informal economy, including its declining trend and association with weak development outcomes. Second, it documents the main cyclical features of the informal economy. Informal-economy recessions and recoveries do not differ significantly from those in the formal economy. Like formal-economy business cycles, informal-economy business cycles tend to be shallower in advanced economies than in emerging market and developing economies. Informal employment tends to expand slightly during formal-economy expansions only in advanced economies whereas it is often acyclical in emerging market and developing economies.

***Key Words**:* Informal economy, self-employment, employment, business cycle.
***JEL Codes**:* E26, E32, J46, O17

[1] Elgin: Columbia University; and Bogazici University; ce2403@columbia.edu. Kose: World Bank, Prospects Group; Brookings Institution; CEPR; and CAMA; akose@worldbank.org. Ohnsorge: World Bank, Prospects Group; and CAMA; fohnsorge@worldbank.org. Yu (Corresponding Author): World Bank; syu2@worldbank.org. We thank Salvatore Capasso, Kevin Clinton, Norman Loayza, William Maloney, Christopher Towe, and seminar participants at the World Bank, ETH, and University of Heidelberg for useful comments. The findings, interpretations, and conclusions expressed in this paper are those of the authors. They do not necessarily represent the views of the institutions they are affiliated with.

1 Introduction

The livelihoods of the poor in emerging market and developing economies (EMDEs) often depend critically on informal activity. Informal employment accounts for about 70 percent of employment in a typical EMDE. Beyond employment, informal output has been estimated at some 35 percent of GDP in EMDEs, compared with about 15 percent of GDP in advanced economies (World Bank 2019).

While offering the advantage of flexible employment under some circumstances, informality is associated with a wide range of adverse economic outcomes, including low productivity, limited fiscal resources, higher poverty and income inequality. As such, informality is an important consideration in formulating cyclical and structural policies. Despite its policy relevance, the cyclical characteristics of the informal economy remain little explored, in part due to the difficulty of measuring the informal sector.

Our paper makes the following contributions to the literature. First, it compiles a comprehensive database of informality, with a focus on measures that have strong cross-country and over-time coverage. Second, it presents two applications of this new database. In a first application, it distills stylized facts about the informal economy using a wide range of informality measures, such as its size and evolution over time, and tests the consistency among these measures. In a second application, the paper documents the main cyclical features of the informal economy, such as the length and depth of its recessions and recoveries.

First, we construct a comprehensive dataset using various measures of informality, with a focus on cross-country time series. Most of the existing macroeconomic literature on informality rely solely on either survey-based or model-based estimates.

Survey-based measures can cover many dimensions of the informal economy, but they suffer from poor country-year coverage (especially in EMDEs), reporting bias, and lack of consistency in survey methods.[2] Informal employment measures tend to cover only the dimension of labor activities in the informal sector, either the number of hours worked per day in informal employment ("intensity" of participation in informal employment) or, regardless of the number of hours worked per day, the presence of informal employment ("extent" of participation; Meghir et al. 2015). Since the extent of participation in the informal economy and its intensity may react differently during a recession, informal production may be asynchronous with informal employment. For example, during a recession, labor may move from the formal sector to the informal sector and raise participation in the informal economy (Loayza and Rigolini 2011). However, due to the fall in demand during a recession, the intensity of participation, i.e. the number of hours worked in informal employment, may remain the same or even drop, reducing informal output. As a

[2] Survey-based informality measures include discrepancies between income declared for tax purposes and income confirmed in selective audits, labor force or employment surveys (e.g. Binelli and Attanasio 2010; McCaig and Pavcnik 2015), or firm surveys (e.g. Almeida and Carneiro 2012; Putnins and Sauka 2015).

result, measures on informal output are an important complement to measures of informal employment.

In the absence of survey-based measures on informal output, indirect, model-based measures on informal output stand out in their comprehensive country-year coverage and clear economic reasoning, but they rely on strong assumptions (e.g. Orsi et al. 2014; Medina and Schneider 2018). Recent indirect model-based approaches build structural equation models or dynamic equilibrium models to estimate the size of output in the informal economy.[3] Structural models (e.g. the multiple indicators multiple causes [MIMIC] model)—often benchmarked against estimates from other approaches—can be used to estimate the size of the informal economy as a latent factor (Schneider et al. 2010). More recently, dynamic general equilibrium (DGE) models, where households allocate labor between formal and informal sectors, have been used to estimate the size of the informal economy (e.g., Ihrig and Moe 2004; Elgin and Oztunali 2012).

By covering both direct, survey-based indicators as well as indirect, model-based estimates, we overcome the limitations of either. Our dataset combines various cross-country databases and data provided by individual national statistical agencies.[4]

Second, our paper distills the main features of the informal economy. We highlight the circumstances under which individual informality measures could be particularly helpful. This adds to the previous work that focused on the limitations of a confined number of estimation methods.

We identify three different dimensions of informality: output, employment, and perception. Cross-country rankings of informal output or employment are typically consistent with each other while varying overtime. Both output and employment measures show a downward trend since 1990 and a degree of cyclicality, making them more suitable for time-series and cyclical analysis. Among all measures, DGE-based estimates and survey-based estimates on self-employment stand out in their country and time coverage. In contrast, perception-based measures tend to be highly stable over time and could, therefore, be more appropriate for

[3] Some earlier model-based approaches exploit discrepancies between different measures of macroeconomic variables. For instance, Caridi and Passeri (2001) used the forecast error from estimated money demand functions or cash demand functions to measure the size of the informal sector (see Ahumada et al. (2007) for a summary) Studies like Johnson et al. (1997) and Enste and Schneider (1998) estimated the extent of informal production based on the deviation of electricity consumption from its norm, assuming that electric-power consumption is the best physical indicator of overall (informal and informal) economic activities. Tanaka and Keola (2017) assume that the informal sector mainly operates during the night and used nighttime light data to estimate the size of the informal sector.

[4] Official GDP statistics often make an adjustment for informal activity. However, the magnitude of such adjustments is rarely specified. In a survey in 2008, national statistical agencies for about 40 mostly advanced economies or countries in transition reported adjusting their official GDP statistics by 0.8-31.6 percent for the non-observed economy, which is a larger concept than the informal economy. The following activities are included in the non-observed economy: underground, informal (including those undertaken by households for their own final use), illegal and other activities omitted due to deficiencies in the basic data collection program (UN 2008). For all reporting countries, these adjustments were well below those suggested by the measures of informality presented here.

cross-country comparisons. Lastly, for cross-country analyses of narrowly-defined questions, measures from labor, firm, and household surveys may be more suitable, especially when surveys are done consistently.

Third, our paper is the first to document the cyclical features of the informal sector in both advanced economies and EMDEs. We summarize the cyclical features of the informal sector and compare them with those of the formal sector and across advanced economies and EMDEs.

We find that EMDEs and advanced economies differ significantly in their cyclical features in both formal and informal economies. EMDEs experience steeper losses than advanced economies during recessions in both formal and informal economies. Similarly, both informal and formal sectors in EMDEs rebound more forcefully than advanced economies during recoveries. Steeper recessions and stronger recoveries in EMDEs contribute to higher output volatility, confirming findings in former studies (e.g. Aguiar and Gopinath 2007). Recessions and recoveries in the formal economy do not differ statistically significantly from those in the informal economy. In both EMDEs and advanced economies, MIMIC-based informal recessions and recoveries tend to be slightly shallower than those in the formal economy.

Employment in advanced economies (AEs), either formal or informal, moves mildly procyclically. Meanwhile, employment in EMDEs, especially informal employment, appears largely acyclical. This may reflect that wage movements or changes in intensity (i.e. number of hours worked per day) in both formal and informal labor markets bear the brunt of the adjustment over the business cycle (Meghir et al. 2015; Guriev et al. 2016).

The following section describes in detail how various measures are gathered and constructed. In addition to the stylized facts about the features of informal economies across regions, section 3 compares the behavior of various measures systematically and identifies the similarities and differences among them. Section 4 presents the stylized facts about the business cycles of the informal economies, and section 5 concludes.

2 Measurement of Informality: Data

Informality is often defined as market-based legal production of goods and services that are hidden from public authorities for monetary, regulatory, and institutional reasons (Schneider et al. 2010).[5] This general definition encompasses many types of informal activity among workers and firms.[6] Some studies distinguish different types of informality by the

[5] Monetary reasons include avoiding taxes and social security contributions; regulatory reasons include avoiding government bureaucracy or regulatory burdens, while institutional reasons include corruption, the quality of political institutions and weak rule of law. For the purposes of this paper, the informal economy reflects activities that, if recorded, would contribute to GDP, and does not cover criminal activities or home production (Schneider et al. 2010; Medina and Schneider 2018).

[6] See ILO (2018) for detailed definitions on informal workers and informal firms. Informal employment has also been more specifically defined as that of workers without pension coverage, which is a part of social protection (Loayza et al. 2010). See World Bank (2019) for an overview.

motives for participating in the informal economy. For example, some classify informal workers and firms into those that are "excluded" and those that "voluntarily exit" from the formal sector (Perry et al. 2007). Others focus on "subsistence informality", in the absence of which the income of low-skilled workers would fall below subsistence levels (Docquier et al. 2017).

Some others classify informal workers and firms into evaders, avoiders, and outsiders depending on their compliance with regulations and regulations' applicability (Kanbur and Keen 2015). More recent studies distinguish different types of informality by the entities engaged in informal activity, separate from their motivation: within firms, formal and informal workers or activities ("interfirm margin") and, within sectors, informal and formal firms or workers ("intersectoral margin", e.g. Ulyssea 2018). Individual country practices vary widely but typically adhere to these broad principles.

Reflecting the difficulty of measuring informality, the literature has developed a wide range of estimation methods to capture its extent. Our database includes the twelve measures commonly used in the literature, which can be largely categorized into two groups based on their estimation methods. The first group encompasses indirect model-based estimates that mainly focus on the size of the informal sector (i.e. informal output in percent of official GDP). The second group encompasses direct measures gathered from surveys, such as labor force, household, firm, or opinion surveys. In our database, indirect and direct measures together cover up to 196 economies (36 AEs and 160 EMDEs) and for as much as 1950-2016 (see Table A1 for details).

This section describes our comprehensive database of informality measures and the limitations and advantages of each included measure. Indirect measures stand out for their country and year coverage, but they suffer from their narrow focus on economic production and strong reliance on model specification and assumptions. Direct measures of informality capture more dimensions of informality and are independent of model specification and assumptions. However, they tend to have limited country and year coverage, making them less well suited to be used in time-series analyses. Furthermore, indirect measures take a macro perspective when measuring the extent of informality in an economy, while direct measures can provide a micro perspective on how firms and workers behave in the informal sector.

Indirect estimates

Previous studies use various indirect approaches to estimate the size of the informal sector, including the currency-demand approach (e.g. Ardizzi et al. 2014), the electricity-demand approach (e.g. Schneider and Enste 2000), the Multiple Indicators Multiple Causes (MIMIC) model (e.g. Schneider et al. 2010), and the Dynamic General Equilibrium (DGE) model (e.g. Ihrig and Moe 2004; Elgin and Oztunali 2012; Orsi et al. 2014). Among all indirect estimation

methods, the MIMIC model and the DGE model stand out in their year and country coverage.[7] Here we focus on the MIMIC and DGE models to estimate the size of informal economic activity. To make the measures comparable with those in the literature, both DGE-based and MIMIC-based estimates are reported in percent of official GDP.

The multiple indicators multiple causes model (*MIMIC*). The Multiple Indicators Multiple Causes model is a type of structural equations models that can be applied to estimate the size of informal economic activity. Two features of MIMIC make it a particularly attractive estimation approach: first, it explicitly considers multiple causes of informal activity and captures multiple outcome indicators of informal activity; second, it estimates informal activity across countries and over time. Indirect approaches like the currency demand approach or the electricity approach condense the full range of informal activity across product and factor markets into just one indicator. However, the informal sector shows its effects in various markets, which would be captured better in a MIMIC model (Schneider et al. 2010). The data on causes and indicators of informal activity identified in the literature are largely macroeconomic data in a panel setting and can be updated annually.

The limitations of the standard MIMIC model of Schneider et al. (2010) and others has been widely discussed in the literature (e.g. Medina and Schneider 2018; Feige 2016). The limitations include: 1) the use of GDP (GDP per capita and its growth rates) as both cause and indicator variables, 2) its reliance on another independent study's base-year estimates on the informal economy to calibrate the size of informal economy in percent of GDP, and 3) the estimated coefficients are sensitive to alternative model specifications and sample coverage. [8] These limitations can open the MIMIC estimates to manipulation and misrepresentation (Breusch 2005).

Here we replicate the most cited MIMIC study, Schneider et al. (2010), to estimate the size of the informal sector (i.e. in percent of official GDP).[9] Six causes and three indicators are used in the estimation to capture the hypothesized relationships between the informal sector (the latent variable) and its causes and indicators. Once the relationships are identified and the parameters are estimated, the estimation results are used to calculate the MIMIC index, which gives the absolute values of the size of the informal sector after a benchmarking or

[7] The electricity-demand approach and the currency-demand approach suffer from limited data availability and theoretical caveats. The specific caveats concerning the electricity-demand include that it assumes that all informal economic activities require only the use of electricity, the production and use of electricity is assumed to stay constant over time. Similarly, the caveats in the currency-demand approach include assumptions that all transactions in the informal sector are assumed to be paid in cash and that there is no informal sector in the base year (Ahumada et al. 2007). In addition, Schneider et al. (2010) suggest that the currency-demand approach and the electricity-demand approach only use one indicator to capture all effects of the informal sector while the informal sector shows its effects in various markets.

[8] Medina and Schneider (2018) tried to overcome the limitation of using official GDP (which may capture part of the informal economy) by using the night-light data to independently capture economic activity.

[9] See the Appendix for details on the estimation method. The estimation results from the model specification that ensures maximum data coverage (shown in Appendix Table A.2) are used here.

calibration procedure. The MIMIC approach delivers a panel of estimates (*MIMIC*) for 160 countries over the period 1993-2015.

The MIMIC estimates capture both the level of employment and productivity in the informal sector, while measures on informal employment only reflect the level of employment. Despite the comprehensive country and year coverage, the time variation of the *MIMIC* estimates is limited, which makes the estimates less suited for time-series analyses (including the business cycle analysis in section 4).

The DGE model (*DGE*). A Dynamic General Equilibrium (DGE) model (e.g. Ihrig and Moe 2004; Elgin and Oztunali 2012) considers how households allocate labor between formal and informal economies within each period and how the allocation changes over time. In comparison to other methods, the DGE approach stands out in its comprehensive country-year coverage, clear economic reasoning, and its applicability in policy experiments and projection (Loayza 2016).

The DGE approach has some limitations. First, it strongly relies on assumptions (Orsi et al 2014; Schneider and Buehn, 2016). Second, like the MIMIC approach, it requires base-year estimates on the informal economy from another independent study to calibrate the size of informal economy (e.g. Ihrig and Moe 2004; Elgin and Oztunali 2012). Third, a computable DGE model only captures some of the stylized facts of the informal sector. Data availability, especially for EMDEs, presents a challenge to matching DGE models with all aspects of informality.

Here we use a deterministic DGE model proposed by Elgin and Oztunali (2012) to estimate the size of the informal sector. The model captures the essence of labor allocation between the formal and informal sector. In the model, an infinitely-lived representative household is endowed with K_0 units of productive capital and a total of $H_t > 0$ units of time. The household has access to two productive technologies, denoted formal and informal, and maximizes its lifetime utility by solving the following optimization problem:

$$\max_{\{C_t, I_t, K_{t+1}, N_{It}, N_{Ft}\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t U(C_t)$$

$$\text{s.t. } C_t + I_t = (1-\tau_t) A_{Ft} K_t^{\alpha} N_{Ft}^{1-\alpha} + A_{It} N_{It}^{\gamma} \quad (1)$$

$$K_{t+1} = I_t + (1-\delta) K_t \quad (2)$$

$$N_{It} + N_{Ft} = H_t \quad (3)$$

$\beta < 1$ is a discount factor and the instantaneous utility function U(.) is strictly increasing and strictly concave. Eq(1) defines the household's resource feasibility constraint: the sum of consumption C_t and investment I_t should equal the amount produced using the formal and informal technologies. The right-hand side of Eq(1) shows that the formal technology (A_{Ft}) follows a standard Cobb-Douglas specification and is exclusive to the formal sector. K_t is the household's capital stock while N_{Ft} is the number of hours the household devotes to the formal sector. τ_t captures the tax rate imposed on formal output. Informal output depends

on the number of hours the household devotes to the informal sector, N_{It} , and its technology, A_{It}.[10]

The rest of the household's problem is standard: Eq(2) specifies the law of motion for capital, where $\delta \in [0; 1]$ is the depreciation rate. Eq(3) is the household's time constraint. In this simple model, the government's policy τ_t is assumed to be exogenously given and the tax revenue is assumed to be used to finance an exogenous stream of government spending, G_t. Then, given the government policy variable tax burden $\{\tau_t\}$, a competitive equilibrium of the two-sector model is a set of sequences $\{C_t, I_t, K_{t+1}, N_{It}, N_{Ft}, G_t\}_{t=0}^{\infty}$ that maximize expected utility from consumption (i.e. $\sum_{t=0}^{\infty} \beta^t U(C_t)$).

The model provides a reasonable mapping between the formal economy and informal economy in a dynamic setting. The two key equilibrium conditions are the equilibrium condition that connects the formal and informal economy through labor allocation, and the equilibrium condition that captures the intertemporal substitution. Our calibration and data construction rely on these two conditions to estimate the ratio, $\frac{Y_{It}}{Y_{Ft}}$, which can be further expressed as $\frac{A_{It}N_{It}^{\gamma}}{A_{Ft}K_t^{\alpha}N_{Ft}^{1-\alpha}}$. The model results in estimates on informal output in percent of official GDP for 158 countries (36 AEs and 122 EMDEs) over the period 1950-2016.

The DGE estimates reflect both the level of employment and productivity in the informal sector and stand out in their country and year coverage. The time variation of the DGE estimates is sufficient for time-series analysis, including the business cycle analysis in section 5. However, the time variation of the DGE estimates relies partially on assumptions. For instance, in Elgin and Oztunali (2012), the growth rate of productivity in the informal sector is assumed to be a function of the growth rates of capital and productivity in the formal sector.[11]

Survey-based estimates

Labor force surveys (LFS) and household surveys (HS) on labor related measures. Four existing informality measures are labor related, out of which three are related to employment and one to pension coverage. These measures are mainly gathered from labor force surveys and sometimes covered by household surveys.

Labor related measures have the advantages of not relying on strong assumptions, having no need for base-year estimates for calibration, and having sufficient time variation for time-series analysis. They also have the following limitations: 1) the data are costly to gather, which results in limited country and year coverage; 2) survey methodologies may vary across time and countries, making the measures incomparable; 3) the typical drawbacks of survey-based estimates (such as sample bias) may make the data quality questionable; and

[10] The model also assumes no cost for hiding and the government cannot enforce payment, the household will attempt to hide the income received from the informal sector.

[11] In the case of Elgin and Oztunali (2012), the heavy reliance of DGE estimates on assumptions and base-year estimates on the informal economy for calibration could be reduced by using other sources of information on the informal economy (e.g. survey-based estimates of informal employment; see Elgin et al. 2019).

4) employment measures cannot reflect other changes in the informal sector, such as productivity and number of working hours.

Despite the limitations, survey-based labor-related measures can provide guidance for the construction and use of indirect informality measures. Among all labor related measures, self-employment stands out in its time and country coverage and sufficient level of time variation, making it suitable for time-series analysis and cross-country comparison. When studying labor-related questions (e.g. employment creation and destruction in the informal sector, or social security issues), labor-related measures should be preferred.

The most frequently used measure is the share of self-employment in total employment (labeled *SEMP,* e.g. La Porta and Shleifer 2014, and Maloney 2004). As defined by the 1993 International Classification of Status in Employment, self-employed workers include four sub-categories of jobs (as classified in WDI and ILO): employers, own-account workers, members of producers' cooperatives, and contributing family workers. [12] Self-employed workers are those who, working on their own account (e.g. own-account workers or employers) or with one or a few partners or in a cooperative, hold "self-employment jobs" as defined above. These are jobs where the remuneration is directly dependent upon the profits derived from the goods and services produced.

Two other measures are informal employment (*INF_EMP*) and employment outside the formal sector (*EMP_NF*).[13] These two measures are usually expressed in percent of total employment (or non-agricultural employment) and refer to different aspects of informality.[14] While employment outside the formal sector is an enterprise-based concept, informal employment is a job-based concept and has a broader definition. Informal employment comprises all workers of the informal sector and informal workers outside the informal sector. Almost all persons employed in the informal sector are in informal employment. However, not all informal employment is in the informal sector, e.g. internships in the formal sector without contracts or pension contributions.

For a comprehensive dataset on labor-related measures on informality, we combine the cross-country databases, provided by WDI, ILO and OECD, and gather additional data from various sources. [15] The resulting dataset on self-employment is a panel of 182 countries/regions over the period 1955-2016. The dataset on informal employment covers

[12] Self-employment largely overlaps with informal employment, but not all self-employed workers are informal employment (e.g. the owner of a formally registered firm is formally employed). While contributing family workers are always classified as informal, workers that hold other types of "self-employment jobs" are classified as informal employment when their production units are informal sector enterprises or households. See 17ths ICLS guidelines for details (https://www.ilo.org/public/libdoc/ilo/2013/480862.pdf)

[13] ILO presents a detailed definition of these two measures (http://laborsta.ilo.org/informal_economy_E.html http://www.ilo.org/ilostat-files/Documents/description_IFL_EN.pdf).

[14] ILO reports these two measures both in percent of total employment and in percent of non-agricultural employment. Here, due to space limitation, we focus on these two measures in percent of total employment, which are comparable with the self-employment measure.

[15] See the Appendix for details on how the final data series on self-employment are constructed.

53 countries/regions (all EMDEs) from various years during 2001-2016 while the dataset on employment outside the formal sector contains 57 countries/regions (all EMDEs) from various years during 1999-2016.[16]

Data on pension coverage (*PENSION*) are gathered from various issues of the World Bank's World Development Indicators (book version, reported until 2012). The measure is defined as the fraction of the labor force that contributes to a retirement pension scheme (Loayza et al. 2010). It yields a panel that covers 135 countries from 1990 to 2010. The measures is suitable for analyzing social security issues related to the informal economy.

Firm surveys. Two datasets of firm surveys have outstanding coverage and data quality: World Bank Enterprise Surveys, and Executive Opinion Surveys conducted by World Economic Forum. World Bank Enterprise Surveys cover 139 economies over the period 2006-2016 while Executive Opinion Surveys cover 151 countries over the period 2006-2016.[17]

Both surveys are answered by top managers and business owners, who are business experts and would be familiar with the business climate in a country. The surveys could reflect some dimensions of informality (e.g. the ease of doing business in the informal sector) that are not captured in the output measures of informality or labor-related measures. Similar to labor-related measures, measures from firm surveys also have the advantages of being independent of strong assumptions and base-year estimates for calibration.

There are two drawbacks of informality measures from firm surveys. First, firm surveys tend to have limited year coverage. Second, since perception does not move much over time, these types of measures do not have much time variation. Both drawbacks limit their application in time-series analysis. However, they shed light on the perceived extent of informality in a country and provide guidance for constructing and validating indirect model estimates.

World Bank Enterprise Surveys compile responses on various topics (including informality) from face-to-face interviews with top managers and business owners in over 130,000 companies in 146 countries. The surveys yield the following measures of informality that have been used in the literature (La Porta and Shleifer 2014; World Bank 2019): percent of firms competing against unregistered or informal firms (*WB1*), percent of firms formally registered when they started operations in the country (*WB2*), (average) number of years firms operating without formal registration (*WB3*), and percent of firms identifying practices of competitors in the informal sector as a major constraint (*WB4)*. A higher value of *WB1*, *WB3* and *WB4* and a lower value of *WB2* indicates a higher level of informality. *WB1* and *WB4* also provide some insights into informal firms' competitiveness while *WB2* and *WB3* indicates the constraints imposed by registration requirements.

[16] Data on informal employment and employment outside the formal sector are obtained from ILO.

[17] Due to survey design changes, the data collected by Executive Opinion Survey in year 2004 and 2005 are not comparable with the following years. World Bank has another firm-level survey, called "Productivity and Investment Climate Surveys". Although the surveys occasionally also report measures on informality, those measures are obtained from various sources and use different methodologies.

In comparison to Enterprise Surveys, Executive Opinion Surveys provide a more balanced panel dataset, making them more suitable for business cycle analysis. World Economic Forum has been conducting the Executive Opinion Survey every year since 1979. As reported in the 2014 edition, over 13,000 executives in 144 economies were surveyed. From year 2006, the survey asks the following question, "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1 = Most economic activity is undeclared or unregistered; 7 = Most economic activity is declared or registered)." The average responses at the country-year level constitute a series of informality measures, labeled as *WEF*, with A lower average indicating a larger informal economy.

Household surveys (HS). Household surveys report either the extent of informality in an economy or people's opinions on informal economic activities. Among all, World Value Surveys (WVS) stand out in their country and year coverage while others mainly focus on European countries.[18] It asks whether respondents can justify cheating on taxes in five waves from 1981-1984 to 2010-2014. The responses range from 1 (never justifiable) to 10 (always justifiable). In total, 317,750 respondents from 96 economies participated in the survey. The average responses at the country-year level are used as a measure for attitudes towards informality, labeled as *WVS*. A higher average at the country level implies that people find cheating on taxes more justifiable and thus consider informal activity more acceptable. It is regarded as indirect measures of informality as the lack of tax morality is associated with a higher level of informality (Oviedo et al. 2009).

3 Stylized facts

The various measures differ somewhat, both in the level of informality as well as its variation over time. In general, *MIMIC* estimates indicate lower and less volatile informal sector activity than *DGE* estimates. This partly reflects the assumed underlying drivers of informality in the two approaches: *MIMIC* is based on slow-moving variables such as institutions whereas *DGE* is based on more volatile variables such as employment, investment and productivity. In EMDEs, the share of informal activity in GDP (by either measure) tends to be well below the share of self-employment in total employment, perhaps reflecting lower labor productivity in the informal economy than the formal economy (Loayza 2018).[19] Value- and opinion-based measures tend to be stable over decades, potentially reflecting a profound rigidity in perceptions.

While the individual measures may deviate in some of these specifics, there are important common features across all informality measures. These are highlighted in the remainder of this section.

[18] Those surveys include the Eurobarometer Survey, European Values Survey, and the European Social Survey, which focus on European countries. We do not use those due to their limited coverage over EMDEs. Details on other social surveys are shown in the Appendix (Table A1 and A3).

[19] In this section and below, self-employment is used to proxy for informal employment as in La Porta and Shleifer (2014). In the following sections, "in percent of GDP or output" is used as the equivalent of "in percent of official GDP" in the context of the share of informal output (both DGE-based and MIMIC-based estimates), while "in percent of employment" is used as the equivalent of "in percent of total employment".

3.1 The extent and trend of the informal economy

On average, the informal economy accounts for 32-33 percent of GDP and 31 percent of employment over the period 1990-2016 (Table 1). As shown in previous studies, a higher level of development, e.g. as measured by log per capita income, is associated with lower informality, virtually regardless of the measure of informality other than attitudes-based ones or the year chosen (e.g., La Porta and Shleifer 2014; Figure 1). As a result, informality tends to be considerably more pervasive in EMDEs than in advanced economies: in advanced economies, it accounts for about 19 percent of GDP and 16 percent of employment, on average, whereas in EMDEs, it accounts for 36-37 percent of GDP and 40 percent of employment.

[Table 1 about here]

Within advanced economies, the size of the informal economy varies widely. In 2016, countries such as Greece produced about 25 percent of GDP in the informal sector while hiring about 34 percent of employment informally. Meanwhile, the informal sector in countries ranked higher on governance indicators and Doing Business indicators, such as the United States and Switzerland, generated about 8 percent of GDP and employed about 6-15 percent of total employed.

There is even wider heterogeneity in informal activity among EMDEs (Figure 2). For example, the informal economy ranged from around 10 (in China) to 69 (in Equatorial Guinea) percent of GDP—depending on the measure used—and self-employment ranged from near-zero (Qatar) to 94 (Burundi) percent of employment.

Both informal output and employment have declined since 1990, especially in EMDEs. Between 1990-16, on average, the share of informal output fell by about 7 percentage points of GDP in EMDEs (to 32 percent of GDP) and by about 4 percentage points (to 17 percent of GDP) in advanced economies. Over the same period, the average share of self-employment declined by about 4 percentage points (to 14 percent of total employment) in advanced economies and by about 4.5 percentage points (to 43 percent of total employment) in EMDEs.

In EMDEs, the largest declines in the shares of informal output and employment occurred from the early-2000s onwards in a reversal of a decade of rising informal employment and barely shrinking informal output.[20] In advanced economies, the largest declines in the share of informal employment occurred between the late 1990s and the global financial crisis; they have since partly reversed, amid anemic post-crisis growth (Figure 2).

These declines were broad-based. A country-specific regression of the share of the informal economy in GDP and employment on a time trend is applied to capture this secular decline (Figure 3). In 50 (*WEF*)-100 (*DGE*) percent of advanced economies (depending on the

[20] The persistence of high levels of informality in EMDEs in the early 1990s in part reflects the expanding informal sector in countries in Eastern and Central Europe during their economic transition (Kaufmann and Kaliberda 1996).

measure) and 48 (*WEF*)-81 (*MIMIC*) percent of EMDEs, there has been a statistically significant downward trend in the share of the informal economy in GDP and employment. The trend decline in the share of informal output suggests that growth may be associated with more rapidly rising labor productivity in the formal economy than in the informal economy. As economies grow, formal-sector entrepreneurs' productivity growth may benefit from technological improvements and greater availability of capital that cannot be accessed by their informal-sector peers (e.g. Amaral and Quintin 2006).

Perceptions of informality appear to change much more slowly than actual informal output and employment. [21] Since 2000, perceptions have shifted significantly upwards or downwards into a quartile with higher or lower informality in only 14 percent of all EMDEs, while having no such sizeable shifts in the remaining 86 percent of EMDEs. There were, however, a few exceptions. For instance, between 2006 and 2007, Indonesia's WEF index moved from the median of EMDEs to the quartile of EMDEs with the smallest informal sectors. At the time, Indonesia strengthened tax administration, modernized investment laws, and reduced some restrictions on labor (IMF 2007). This coincided with rapid growth of 6.3 percent, the highest rate since 2000, and reductions in the share of both informal output and employment.

3.2 Correlates of informality

A large informal economy is associated with a range of adverse outcomes, as documented in this section. These correlations, however, do not necessarily reflect causal links from informality to adverse outcomes. Several theoretical models have explored channels through which high informality can cause lower growth, worse governance, and weaker ability of governments to reduce income inequality (Docquier et al. 2017; Loayza 1996; Sarte 2000). Other theoretical studies have focused on the reverse causality, identifying two major reasons for the emergence of informal activity: lack of development (Harris and Todaro 1970; Loayza 2016), and poor governance including burdensome regulations, corruption, or poor public services coupled with weak enforcement (de Soto 1989; Amaral and Quintin 2006; Ordóñez 2014).

The causes and implications of informality predicted by these theoretical models are also confirmed by empirical studies. A large informal economy is associated with lower per capita incomes, less access to credit, limited trade openness, less skilled labor forces, as well as weaker output, investment and productivity growth. Informality is also associated with less effective institutions, weaker governance, more burdensome tax and regulatory regimes, and lower government revenues and expenditures (e.g. Enste and Schneider 1998; Dabla-Norris et al. 2008, World Bank 2019).

[21] For instance, Guiso et al. (2009) demonstrate that the perception of trustworthiness is largely historically determined with limited time-variance.

Per capita incomes. Figure 4 shows these economic and institutional indicators for countries with high and low output and employment informality.[22] For both output and employment informality, GDP per capita in countries with below-median ("low") informality is about 2-3 times of those in countries with above-median ("high") informality.[23] The lower productivity and resource misallocation associated with higher informality may also be reflected in slower output growth (e.g. Hsieh and Klenow 2009; Figure A2).

Access to credit and capital accumulation. A larger informal economy is associated with limited access to credit and slower accumulation of physical or human capital (Ovedio et al. 2009). At the firm level, informality can limit access to conventional bank credit, because of a lack of documentation for assets and inadequate financial statements. Investment activity in the informal sector may also be subdued because informal firms may be unwilling to adopt technologies that would make them more visible to tax and other authorities (Dabla-Norris et al. 2008; Gandelman and Rasteletti 2017). In the long run, the tendency to hire less skilled workers in the informal sector may slow human capital accumulation. Indeed, countries with below-median informality tend to have significantly higher levels of human capital and access to credit (Maloney 2004; Docquier et al. 2017; Figure 4).

Trade openness. A smaller informal sector is associated with greater economic openness, especially to trade.[24] On average, the trade-to-GDP ratio is lower by 17 percentage points in countries with a greater share of self-employment than in countries with a smaller share of self-employment (Figure 4).[25] That said, the impact of major trade liberalization episodes on informality varies across countries and differs between the short- and the long-term (Goldberg and Pavcnik 2003; Dix-Carneiro and Kovak 2017).

Regulatory burdens. Heavier regulatory (or administrative) burdens may encourage informality as workers and firms join the informal sector to avoid regulatory and administrative compliance costs.[26] The *Doing Business* distance-to-frontier scores for countries with below-median informality (by DGE estimates) is 60 points—three-fifths of a standard deviation—higher in countries with below-median output informality (Figure 4). Similarly, the Business Freedom index is about half of a standard deviation higher in countries with below-median output informality than in countries with high (above-median) informality.

---

[22] Additional correlates are shown in Figure A2. Figure A1 redoes the analyses in Figure 4 using MIMIC-based and WEF-based informality measures. The results are in line with Figure 4.

[23] Median informality amounts to about 32 percent of GDP for DGE-based informal output and 34 percent of total employment for self-employment.

[24] Empirical studies, such as Goldberg and Pavcnik (2004 and 2007), Boly (2018), and McCaig and Pavcnik (2018), show that informality declined following some trade liberalization episodes. Conversely, a short-term increase in informality has been attributed to trade liberalization amid rigid labor markets in studies such as Goldberg and Pavcnik (2003), Attanasio et al. (2004), and Bosch et al. (2012).

[25] However, the trade-to-GDP ratio is similar in countries with different output informality. The positive association between employment informality and trade openness could be driven by the level of development or the fact that workers in the tradeable sector need to comply with the regulations and are more viable.

[26] Perry et al. (2007); Ulyssea (2010); Buhn (2011); De Mel et al. (2013); Rocha et al. (2018).

Governance. Research points to the contribution of poor governance to the pervasive informality in some EMDEs, especially in LAC and ECA.[27] On average, countries with above-median informality have had weaker government effectiveness (by about three-quarters of a standard deviations) than countries with below-median informality (Figure 4). Similar differences are found in the case of control of corruption and rule of law. For example, in Georgia, during the period 1996-2016, the transition to a market economy brought significant improvements in government effectiveness, control of corruption, and rule of law. With output growth averaging about 6 percent per year, the share of informal output fell from 66 percent to 57 percent of GDP, and the share of informal employment in total employment fell by a similar magnitude.

Other economic outcomes. At the macroeconomic level, a higher level of informality is linked with a lower level of labor productivity, a higher level of extreme poverty and income inequality, and a lower level of government revenues and expenditures (Figure A2; World Bank 2019). On average, informal labor productivity below formal-sector productivity in EMDEs, although not in advanced economies (Loayza 2018).[28] Regardless of the measure of informality, on average, government revenues in EMDEs with the most pervasive informality have been 5-10 percentage points of GDP below those with the least pervasive informality. Similarly, in EMDEs with the most pervasive informality, government expenditures were 4-10 percentage points of GDP lower than in those with the lowest informality. The composition of tax revenues is also tilted towards trade taxes in economies with more pronounced informality, making the tax system less progressive.

3.3 Consistency among the various measures of informality

Overall, the various measures of informality appear to capture three distinct aspects of informality: output (DGE and MIMIC estimates), employment (e.g. self-employment and pensions), and perception (e.g. the WEF and WVS surveys). While the common trends and correlates illustrated above suggest a considerable degree of commonality among some of these aspects, they also move slightly differently from time to time. For instance, in some cases, noticeable drops in informal output translated into only moderate falls (e.g., Bulgaria or Sri Lanka) or even expansion in employment in the informal sector (e.g., Bolivia or Paraguay). This section explores the consistency among the various measures of informality.

We illustrate this by examining correlations among various informality measures (both levels and first-differenced measures). [29] First, we check whether various measures are

[27] Sarte (2000) suggests that firms operating in the informal sector to avoid rent-seeking bureaucrats. Choi and Thum (2005) and Dreher and Schneider (2010) report an association between higher informality and weaker law and order and control of corruption. Dabla-Norris et al. (2008) show that the quality of the legal framework is important in determining the size of the informal sector.

[28] Many studies have shown that informal firms tend to be less productive than their formal counterparts; although this productivity differential in part reflects the characteristics of informal firms (e.g., McKenzie and Sakho 2010; Fajnzylber et al. 2011; De Mel et al. 2013).

[29] Various additional correlations are calculated (e.g. both pairwise and rank correlations for a panel setting, cross-country within a specific year, and cross-year within a certain country, and with a longer time period).

consistent with each other in terms of capturing an economy's level of informality, or at least its position in cross-country rankings. Since the informality measures have different units and definitions, Spearman rank correlation, which checks the statistical dependence between the rankings of two data series, is preferred here. The median rank correlations between different variables across countries within a single year are shown in Table 2. Second, we investigate whether various measures share the same movements over time. To do so, the share of country-pairs in which first differences in two measures have the same sign are calculated and reported in Table 3.

[Insert Table 2 about here]

The various measures for informality are generally positively correlated with each other, with the correlations within each block (output, employment, perception) being stronger than correlations between blocks (Table 2). The rank correlation between the estimates of informal output (*DGE* and *MIMIC*) are positive, close to 1, and significant at the 1 percent level. In addition, the rank correlations between *DGE* or *MIMIC* and employment measures or some perception measures are also significant. The correlations among the various measures of informal employment range from 0.31-0.93 and are mostly significant at the 1 percent level. On average, the correlation between an estimate of informal *output* and *SEMP* or *PENSION* is around 0.60, significant at the 1 percent level.

In contrast, perception-based estimates of informality tend to be more correlated with each other than with estimates of informal output or employment. The WVS is an exception: it tends to be uncorrelated or little correlated with all other measures, both perception-based ones and other measures. The weak linkage suggests that a large informal sector reflects more than citizens' weak tax morality, which is captured by WVS. In contrast, the WEF measure among the perception-based measures is the one most correlated with the other output-based measures (about 0.70 with *DGE* and *MIMIC*) and employment-related measures (about 0.5-0.7 with *PENSION* and *SEMP*) since, conceptually, it captures perceived informal economic activities.

To further examine consistency of movements over time among various measures, we check the coincidence of the direction of movements over time in different variables. Table 3 shows the share of country-pairs in which first differences in two measures have the same sign.[30] This is the case in surprisingly few instances—at most in 82 percent of country-year pairs (for *INF_EMP* and *EMP_NF*). The coincidence between the direction of changes in output measures and employment measures is 50-60 percent of country-year pairs, suggesting that

---

The results are largely in line with Table 2 and will be provided upon request. We also conduct a simple factor analysis using measures with sufficient coverage (i.e. *DGE*, *MIMIC, SEMP*, and *WEF*). The results show that these measures share a common factor, which explains 62-87 percent of the variation of the measures (87 percent for *DGE*, 87 percent for *MIMIC*, 62 percent for *SEMP*, and 69 percent for *WEF*).

[30] As a robustness check, we also calculate the pairwise correlations of first-differenced informality measures over time for each country and compute their medians across countries. The results are in line with Table 3 and will be provided as online Appendix. While significant and positive correlations are observed among self-employment, informal employment and employment outside the formal sector, no significant correlations between informal employment (or perception) measures and informal output measures are found.

output measures capture important additional factors to employment measures (e.g., changes in labor productivity or intensity).

[Insert Table 3 about here]

3.4 Two country examples: Mexico and Czech Republic

Two country examples illustrate the perspectives provided by different measures of informality (Figure 6). Mexico illustrates a case of shrinking informality across all measures of informality over the past two decades, whereas the Czech Republic illustrates a divergence between output-based and employment- or perception-based measures.

In *Mexico*, the share of the informal economy in output has declined from 32 percent of GDP in 1993 to 27-30 percent of GDP in 2015. The decline by the *DGE* measure was somewhat stronger than the decline in the *MIMIC* measure, partially reflecting slow moving feature of institutional quality in comparison to capital accumulation. Both *DGE* and *MIMIC* movements, however, were broadly in line with declines that would have been expected based on per capita income growth, supported by regulatory streamlining for startups in 2011, comprehensive labor law reforms in 2012, and some judicial reforms in 2014 (OECD 2011 and 2015). The fall in the share of informal output was accompanied by a 12-percentage point decline in the share of informal employment to 32 percent of employment in 2015—from well above to now broadly in line with informality in peers with similar per capita income. Between 1994 and 2015, acceptance of tax evasion declined from the 80th percentile to the median among EMDEs. Even over the more recent period of 2006-16, Mexico's informal economy has been perceived to shrink.

In the *Czech Republic*, the share of informal output also declined, although not as strongly as in Mexico. Both measures of informal output experienced similar declines from 18-20 percent of GDP in 1994 to 16-17 percent of GDP in 2015. While the share of informal output remains below that suggested by income level, the decline in the share of informal output was less than may have been expected based on per capita income growth. In contrast to informal output, the share of informal employment has steadily increased almost 7 percentage point since 1994 to 18 percent in 2014. The OECD has attributed the rise in self-employment to efforts at circumventing tax and social security obligations (OECD 2016 and 2010). Doing Business reports also suggest that various changes in labor regulations have been implemented between 2008-2010 and 2013-2015. During 2008-2010, these changes in labor regulations were accompanied with rising self-employment. The reverse occurred during 2013-2015, in part reflecting a second round of changes in labor regulations to reduce informal employment. However, the second round of changes in labor regulations raised concerns about informality (reflected in *WEF* index).

4 Cyclical Features of the Informal Economy

4.1 Features of informal economy business cycles

Formal and informal output and employment are significantly more volatile in EMDEs than in advanced economies, possibly reflecting larger shocks to, or lesser resilience to shocks in, EMDEs (Neumeyer and Perri 2005, Aguiar and Gopinath 2007, Restrepo-Echavarria 2014; Table 4). In addition, in both EMDEs and advanced economies, self-employment (as a proxy for informal employment) is somewhat more volatile than formal employment where the labor market is more rigid (Djankov and Ramalho 2009).

Both formal and informal employment alone have higher volatility than the sum of formal and informal employment (i.e. total employment), since formal and informal employment move marginally in opposite directions.[31] This supports former findings that the informal sector may help stabilize total employment over business circles (Loayza and Rigolini 2011; Fernandez and Meza 2015). The informal sector's stabilization role in the labor market is more prominent in EMDEs where 40-60 percent of total employment is informal.

4.2 Dating Business Cycles

We follow Claessens et al. (2012) and apply the Bry and Boschan (1971)'s algorithm to date the business cycles of formal and informal sectors. To apply Bry and Boschan (1971)'s algorithm to annual data, we follow Berge and Jorda (2011) and Harding and Pagan (2002) and define peaks (troughs) as occurring at time t whenever $\{\Delta y_t > (<) 0, \Delta y_{t+1} < (>) 0\}$. As the censoring rule, if there are additional peaks/troughs within a 5-year period around a peak, the one with the deepest contraction/expansion is picked. When calculating characteristics of business cycles, the closest peaks (troughs) before troughs (peaks) are used when there are several peaks (troughs) in a row.

To illustrate our dating method, we show the case of *Mexico* in Figure 7. Between 1980 and 2016, Mexico's official GDP experienced five cyclical peaks (in 1981, 1985, 1994, 2000, and 2008) and five cyclical troughs (in 1983, 1986, 1995, 2001, 2009). The first peak in 1981 was right before the 1982 debt crisis, which tipped the Mexican economy towards the trough in 1983. After 1983, Mexico quickly rebounded and started its trade liberalization process. It joined GATT in 1986 and NAFTA in 1994, which coincided with turning points in official GDP. These turning points suggested that the formal sector in Mexico initially experienced shocks from trade liberalization but subsequently rebounded. The peaks and troughs around 1995 and 2000 corresponded to the 1995 Tequila crisis and the 2000 dot-com bubble bust. The 2008 global financial crisis also affected Mexico's output in the formal sector, reflected by the turning points around 2008.

Turning points of formal and informal output business cycles coincided in the case of Mexico. Informal output expanded as formal output grew, possibly because the income generated by

[31] The correlation between formal and informal employment growth rates is about -0.2 and significant at 1 percent level.

trade liberalization created demand for informal sector activity. However, the speed of expansion was considerably faster (four times as fast) in the formal economy than in the informal economy, possibly due to the formal sector's better access to (or utilization of) capital and advanced technology (World Bank 2019).

[Insert Figure 7 about here]

During 1980-2016, total employment expanded as formal output grew and shrank as formal output contracted. However, the same co-movement did not happen between formal output and self-employment. Self-employment dropped sharply, from 11.7 million to 9.1 million, between 1980 and 1990. while formal employment expanded from 8.9 million to 19.5 million. The expansion in formal employment partially reflected the expanding tradable sector that created new formal employment to absorb self-employed workers. In addition, Mexico conducted a reform to simplify business registration between 2002 and 2006 (i.e. Rapid Business Opening System, Bruhn 2011 and 2013), which is also followed by an expansion in formal employment and a stabilization in self-employment.

4.3 Characteristics of informal output through the business cycle

We study the main characteristics of the recession and recovery phases of both formal and informal business cycles. A recession is defined as the period from peak to trough, while an expansion is the converse, the period from trough to the following peak. A recovery, the early part of an expansion, is defined as the time it takes for output to rebound from the trough to its pre-recession peak. The main characteristics of the recession and recovery phases, including duration, amplitude, and slope, are defined as in Claessens et al. (2012).

- The *duration* captures, for a recession, the period from peak to trough and, for a recovery, the period it takes for output to return to its pre-trough peak.
- The *amplitude* of a recession measures the change in output from a peak to the next trough. The amplitude of a recovery measures the change in output during the first year of an expansion, which is the period between a trough and its following peak.
- The *slope* measures the speed of a given cyclical phase. It is defined as the ratio of amplitude over duration for a recession phase and the ratio of the change from the trough to the last peak divided by the duration for a recovery phase.
- For recessions only, another widely used measure, *cumulative loss*, is calculated. It captures the overall cost of a recession. The cumulative loss is defined as the difference between the sum of annual changes in output and half of the amplitude during a recession.

[Insert Table 5a about here]

The results here are in line with earlier studies (Birinci and Elgin 2013; Bajada 2003) on informal business cycle recessions and expansions in advanced economies.[32] In contrast to these studies, here we focus on recessions and recoveries. Since recovery phases are the

[32] A comparison between our findings and former studies will be provided upon request.

early parts of expansions, they reflect more of an economy's short-term cyclical movements rather than its long-term growth path.

Recessions. Overall, informal economy recessions did not differ statistically significantly from formal economy recessions. However, the recessions of MIMIC-based informal output are slightly shallower and slower in adjustment than those of formal output and DGE-based informal output. Based on DGE estimates, the average informal economy recession lasted 1.6 years with a contraction by, on average, 4 percent per year, 6 percent from the peak to its following trough, and 6.4 percent cumulatively—broadly in line with formal economy recessions. MIMIC-based informal output recessions last, on average, 1.4 year with output contractions of 2.5 percent per year, 3.7 percent from peak to trough, and 4.9 percent cumulatively. The shallower recessions of MIMIC-based informal output could be due to the slow-moving institutional measures embedded in MIMIC's estimation methods (e.g., government effectiveness, see Table A2 for details).

Both formal and informal economy recessions were significantly shallower in advanced economies than in EMDEs, and both DGE-based and MIMIC-based informal recessions are accompanied by more severe output contractions in EMDEs (Table 5a). Formal output contracted by 2.5 percent per year and 4.0 percent from the peak to its following trough during an average recession in advanced economies. It contracted by 4.6 percent per year and 6.6 percent from the peak to its following trough in EMDEs. Similarly, DGE-based informal output contracted by 4.5 percent per year in EMDEs during informal economy recessions, while it contracted by 2.7 percent per year in advanced economies. During an average MIMIC-based informal recession, output shrank by 1.4 percent per year in advanced economic and by 2.9 percent per year in EMDEs. The shallower recessions in advanced economies are consistent with the low volatility of formal and informal business cycles presented in Table 4. The more severe output contractions in EMDEs further confirm EMDEs' vulnerability to shocks. When DGE estimates are used, informal economy recessions lasted somewhat longer (2 years) in advanced economies than in EMDEs (1 year). The same difference is not found when MIMIC-based estimates are used.

[Insert Tables 5b about here]

Recoveries. Informal-economy recoveries also did not differ statistically significantly from formal-economy recoveries (Table 5b). On average, both formal and informal economy took about 2 years to reach the levels of their former peaks after a recession, with output expanding by 2-6 percent in the first year into a recovery and by 2-5 percent per year during the entire recovery phase. MIMIC-based informal recoveries were slightly shorter, occurred less frequently, and were less pronounced than DGE-based informal recoveries and formal recoveries. The finding is consistent with the lower volatility of MIMIC estimates shown in Table 4.

Both formal- and informal-economy recoveries in advanced economies were significantly shallower than in EMDEs (Table 5b). First, formal-sector output recoveries lasted significantly longer in EMDEs (2.2 years on average) than in advanced economies (1.7 year

on average), whereas the duration of informal-sector recoveries did not differ significantly. Second, both informal and formal recoveries were accompanied by significantly larger output gains in EMDEs than in advanced economies. For example, in the first year of a recovery, DGE (MIMIC)-based informal output increased, on average, by 4.9 (3.8) percent in EMDEs and by 2.2 (1.6) percent in advanced economies. Over an average recovery phase, the slope of DGE (MIMIC)-based informal output is 4.3 (2.3) percent in EMDEs and 2.0 (1.5) percent in advanced economies.

Combining the results from Table 5a and 5b, we observe both shallower recessions and recoveries in advanced economies than in EMDEs, contributing to the higher volatility of output in EMDEs. The latter is a feature well documented in the literature (e.g. Aguier and Gopinath 2007). One of the reasons could be EMDEs' tendency to follow procyclical fiscal policy, which exacerbates the underlying business cycle (Frankel et al. 2013).

4.4 Characteristics of informal employment through the business cycle

In this section, the evolution of employment is examined during formal and informal business cycles. We consider both recovery phases and expansion phases, with the latter being defined as the periods from a trough to the next peak (Claessens et al. 2012). On average, expansions lasted about 2-6 year longer than recoveries. Following Kose et al. (2003), employment is log first-differenced and demeaned to remove trend employment growth.

[Insert Table 6 about here]

Formal employment growth dropped below its longrun trend during recessions and recoveries and rose above it during expansions in both advanced economies and EMDEs (Table 6). However, only formal employment growth in advanced economies differs statistically significantly from its longrun trend in all these business cycle phases. Annual formal employment growth in advanced economies dropped by 0.7 percentage points below its longrun trend during the average formal-economy recession and by 1.8 percentage points during the average formal-economy recovery. Formal employment growth in advanced economies, on average, also exceeded its longrun trend by 0.2 percentage points during informal (DGE-based alone) output expansions, in advanced economies and EMDEs alike.

In EMDEs, formal employment growth did not deviate from its longrun trend to a statistically significant degree in most of these formal and informal business cycle phases (except DGE-based informal-economy recession). In particular, the contraction in formal employment growth in EMDEs is significantly less severe than in advanced economies during both formal and informal recoveries. The absence of a significant movement of formal employment to output fluctuations could be due to the high level of informal employment or labor market rigidity in EMDEs (e.g. minimum wages, severance payments, licensing; Oviedo et al. 2009).

Informal employment growth (proxied by self-employment) contracted during both formal and informal **recessions**, especially in advanced economies. On average, informal employment growth dropped below its longrun trend by around 1 percentage point in

advanced economics (significantly below the longrun trend) and by about 0.2-0.7 percentage points (not significantly different from the longrun trend) in EMDEs during formal or informal output recessions. The changes in informal employment growth during recessions in advanced economies do not significantly differ from those in EMDEs.

Informal employment growth remained broadly around its longrun trend during formal and informal **recoveries** (i.e. not statistically different from its longrun trend). However, on average, informal employment growth exceeded its longrun trend by about 0.7 percentage points in advanced economies during both formal and informal recoveries. As the informal labor market is more flexible than the formal labor market in advanced economies, informal employment responds to the recovery of output while formal employment does not. Informal employment growth fell slightly (0.4 percentage points) below its longrun trend in EMDEs during formal recoveries, partially reflecting a labor outflow from the informal sector to the formal sector. However, the differences between advanced economies and EMDEs were not statistically significant.

During informal and formal economy **expansions**, informal employment growth rose statistically significantly above its longrun trend only in advanced economies (not in EMDEs). Informal employment growth grew above trend by 0.5 percentage points in advanced economies during both formal and DGE-based informal economy expansions, which was significantly higher than in EMDEs. Informal employment growth did not differ from its longrun trend in EMDEs during both formal and informal expansions. The acyclicality of informal employment growth during informal expansion in EMDEs suggests that the labor is more likely to adjust in wages or number of hours worked per day rather than employment (Meghir et al. 2015; Guriev et al. 2016).

5 Conclusion

Using a comprehensive database of model-based and survey-based estimates of informal activity, we identify a rich set of measures available for cross-country analysis and a more limited set of measures available for time-series or panel analysis. Using the widest possible range of measures, we illustrate the trend decline in informal output and employment and its positive correlation with indicators of economic and institutional development. We identify three somewhat distinct aspects of informality: output, employment, and perceptions.

We illustrate two applications of our database. First, we document the correlates for informality that are consistent across these measures. In particular, higher informality is associated with lower per capita income, weaker governance and business climates, more restrictive regulations, poorer access to credit, higher poverty and inequality, and more constrained government operations.

Second, we document the stylized facts of informal-economy business cycles. Informal-economy business cycles comove with formal-economy output cycles, in that output is highly correlated and turning points tend to coincide. Like formal-economy output cycles, informal-

economy output cycles tend to be shallower in advanced economies than in EMDEs. Informal employment tends to expand procyclically during formal-output expansions only in advanced economies whereas it tends to be acyclical in EMDEs. In contrast to distinct cyclical movements in informal output and employment, perceptions of informality are highly persistent.

6 References

Aguiar, M. and Gopinath, G., 2007. "Emerging Market Business Cycles: The Cycle is The Trend." *Journal of Political Economy* 115(1): 69-102.

Ahumada, H., F. Alvaredo, and A. Canavesa. 2007. "The Monetary Method and The Size of the Shadow Economy: A Critical Assessment." *Review of Income and Wealth* 53 (2): 363–371.

Almeida, R., and P. Carneiro. 2012. "Enforcement of Labor Regulation and Informality." *American Economic Journal: Applied Economics* 4 (3): 64-89.

Amaral, P. S., and E. Quintin. 2006. "A Competitive Model of the Informal Sector." *Journal of Monetary Economics* 53 (7): 1541-1553.

Ardizzi, G., C. Petraglia, M. Piacenza, and G. Turati. 2014. "Measuring The Underground Economy with The Currency Demand Approach: A Reinterpretation of The Methodology, with An Application to Italy." *Review of Income and Wealth* 60 (4): 747-772.

Attanasio, O., P. Goldberg, and N. Pavcnik. 2004. "Trade Reforms and Wage Inequality in Colombia." *Journal of Development Economics* 74 (2): 331-366.

Bajada, C., 2003. "Business Cycle Properties of The Legitimate and Underground Economy in Australia." *Economic Record* 79(247): 397-411.

Berge, T., and Ò. Jordà. 2011. "Evaluating the Classification of Economic Activity into Recessions and Expansions." *American Economic Journal: Macroeconomics* 3.2: 246-277

Binelli, C., and O. Attanasio. 2010. "Mexico in the 1990s: The Main Cross-Sectional Facts." *Review of Economic Dynamics* 13(1) : 238-264.

Birinci, S., and C. Elgin. 2013. "Shadow Economy over the Business Cycle: How do Formal and Informal Cycles Interact?" Unpublished manuscript.

Boly, A. 2018. "On the Short- and Medium-Term Effects of Formalisation: Panel Evidence from Vietnam." *The Journal of Development Studies* 54 (4): 641–656.

Bosch, M., E. Goñi-Pacchioni, and W. Maloney, 2012. "Trade Liberalization, Labor Reforms and Formal–Informal Employment Dynamics." *Labour Economics* 19 (5): 653-667.

Breusch, T. 2005. "Estimating The Underground Economy, Using MIMIC Models", Working Paper, National University of Australia, Canberra.

Bruhn, M. 2011. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico." *The Review of Economics and Statistics* 93 (1): 382-386.

Bruhn, M. 2013. "A Tale of Two Species: Revisiting the Effect of Registration Reform on Informal Business Owners in Mexico." *Journal of Development Economics* 103 (2): 275-283.

Bry, G. and C. Boschan.1971. *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*. NBER.

Caridi, P. and P. Passerini. 2001. "The Underground Economy, The Demand for Currency Approach and The Analysis of Discrepancies: Some Recent European Experience." *Review of Income and Wealth* 47: 239-250.

Choi, J., and M. Thum. 2005. “Corruption and the Shadow Economy.” *International Economic Review* 46 (3): 817-836.
Claessens, S., M. A. Kose, and M. E. Terrones. 2012. "How Do Business and Financial Cycles Interact?" *Journal of International Economics* 87 (1): 178-190.
Dabla-Norris, E., M. Gradstein, and G. Inchauste. 2008. “What Causes Firms to Hide Output?” *Journal of Development Economics* 85 (1-2): 1-27.
De Mel, S., D. McKenzie, and C. Woodruff. 2011. “What is the Cost of Formality? Experimentally Estimating the Demand for Formalization.” *American Economic Journal: Applied Economics* 5 (2): 1221-50.
Dix-Carneiro, R., and B. Kovak. 2017. “Trade Liberalization and Regional Dynamics” *The American Economic Review* 107 (10): 2908-2946.
Djankov, S. and R. Ramalho. 2009. “Employment Laws in Developing Countries.” *Journal of Comparative Economics* 37(1): 3-13.
Docquier, F., T. Müller, and J. Naval. 2017. “Informality and Long-Run Growth.” *The Scandinavian Journal of Economics* 119 (4): 1040-1085.
Dreher, A., and F. Schneider. 2010. "Corruption and The Shadow Economy: An Empirical Analysis." *Public Choice* 144 (1-2): 215-238.
Elgin, C., and O. Oztunali. 2012. “Shadow Economies Around The World: Model Based Estimates.” Working Paper No. 2012/05, Bogazici University, Department of Economics, Istanbul.
Fajnzylber, P., W. Maloney, and G. Montes-Rojas, 2011. “Does Formality Improve Micro-Firm Performance? Evidence from the Brazilian SIMPLES Program.” *Journal of Development Economics* 94 (2): 262–276.
Feige, E. L. 2016. “Reflections on the Meaning and Measurement of Unobserved Economies: What Do We Really Know about the “Shadow Economy”?” *Journal of Tax Administration* 2 (1): 1-37.
Fernández, A., and F. Meza. 2015. “Informal Employment and Business Cycles in Emerging Economies: The Case of Mexico.” *Review of Economic Dynamics* 18 (2): 381–405.
Frankel, J., Vegh, C. and G. Vuletin. 2013. “On Graduation from Fiscal Procyclicality.” *Journal of Development Economics* 100(1): 32-47.
Gandelman, N., and A. Rasteletti. 2017. “Credit Constraints, Sector Informality and Firm Investments: Evidence from a Panel of Uruguayan Firms.” *Journal of Applied Economics* 20 (2): 351–72.
Goldberg, P., and N. Pavcnik. 2003. “The Response of the Informal Sector to Trade Liberalization.” *Journal of Development Economics* 72 (2): 463-496.
———. 2004. “Trade, Inequality, and Poverty: What Do We Know? Evidence from Recent Trade Liberalization Episodes in Developing Countries.” Working Paper 10593. National Bureau of Economic Research, Cambridge, MA.
———. 2007. “Distributional Effects of Globalization in Developing Countries.” *Journal of Economic Literature* 45 (1): 39–82.
Guiso, L., P. Sapienza, and L. Zingales. 2009. “Cultural Biases in Economic Exchange?” *The Quarterly Journal of Economies* 124(3): 1095-1131.
Guriev, S, B Speciale, and M Tuccio. 2018. “How do Regulated and Unregulated Labor Markets Respond to Shocks? Evidence from Immigrants During the Great Recession.” *The Journal of Law, Economics, and Organization (*forthcoming*).*
Harris, J., and M. Todaro. 1970. “Migration, Unemployment, and Development: A Two Sector

Analysis." *The American Economic Review* 60 (1): 126-142.
Hsieh, C., and P. Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4): 1403–48.
International Labor Organization (ILO). 2018. *Informal Economy*. Geneva: International Labor Office, https://www.ilo.org/ilostat-files/Documents/description_IFL_EN.pdf
Ihrig, J., and K. S. Moe. 2004. "Lurking in The Shadows: The Informal Sector and Government Policy." *Journal of Development Economics* 73 (2): 541-557.
Johnson, S., D. Kaufmann, and A. Shleifer. 1997. "The Unofficial Economy in Transition." *Brookings Papers on Economic Activity (*Fall): 159-240.
Kaufmann, D., and A. Kaliberda. 1996. "Integrating the Unofficial Economy into the Dynamics of Post-Socialist Economies: A Framework of Analysis and Evidence." Policy Research Working Paper 1691, World Bank, Washington DC.
Kanbur, R., and M. Keen. 2015. "Rethinking Informality." Voxeu (blog), June 15, https://voxeu.org/article/rethinking-informality.
Kose, M. A., C. Otrok, and C. Whiteman. 2003. "International Business Cycles: World, Region, and Country-Specific Factors." *The American Economic Review* 93 (4): 1216-1239.
Lackó, M. 2000. "Hidden Economy—An Unknown Quantity? Comparative Analysis of Hidden Economics in Transition countries 1989-95." *Economics of Transition* 8 (1): 117-149.
La Porta, R., and A. Shleifer. 2014. "Informality and Development." *Journal of Economic Perspectives* 28 (3): 109-126.
Loayza, N. 1996. "The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America." *Carnegie-Rochester Conference Series on Public Policy* 45: 129-162.
———. 2016. "Informality in the Process of Development and Growth." *The World Economy* 39 (12): 1856-1916.
———. 2018. "Informality: Why Is It So Widespread and How Can It Be Reduced?" *Research & Policy Brief,* World Bank, Kuala Lumpur.
Loayza, N., and J. Rigolini. 2011. "Informal Employment: Safety Net or Growth Engine?" *World Development* 39 (9): 1503-1515.
Loayza, N., L. Servén, and N. Sugawara. 2010. "Informality in Latin America and the Caribbean." In *Business Regulation and Economic Performance*, edited by N. Loayza and L. Servén. Washington, DC: World Bank.
Maloney, W. 2004. "Informality Revisited." *World Development* 32 (7): 1159-1178.
McCaig, B., and N. Pavcnik. 2015. "Informal Employment in a Growing and Globalizing Low-Income Country." *American Economic Review* 105 (5): 545-550.
McKenzie, D., and Y. Sakho. 2010. "Does it Pay Firms to Register for Taxes? The Impact of Formality on Firm Productivity." *Journal of Development Economics* 91 (1): 15-24.
Medina, M., and F. Schneider. 2018. "Shadow Economies around the World: What Did We Learn Over the Last 20 Years?" IMF Working Paper 18/17, International Monetary Fund, Washington, DC.
Meghir, C., R. Narita, and J. Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review* 105 (4): 1509-46.
Neumeyer, P.A. and Perri, F., 2005. "Business Cycles in Emerging Economies: The Role of Interest Rates*." Journal of Monetary Economics* 52(2): 345-380.
OECD. 2010. *OECD Economic Surveys: Czech Republic 2010*. OECD: Paris.
———. 2011. *OECD Economic Surveys: Mexico 2011*. OECD: Paris.

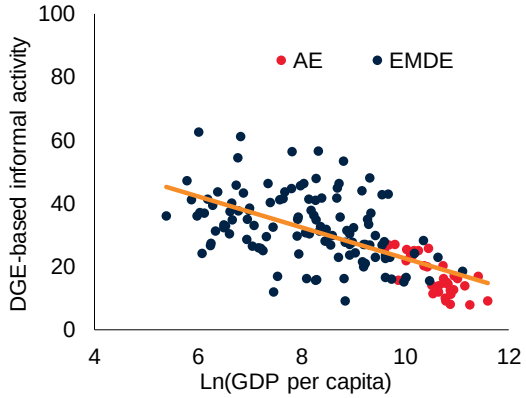
———. 2015. *OECD Economic Surveys: Mexico 2015*. OECD: Paris.
———. 2016. *OECD Economic Surveys: Czech Republic 2010*. OECD: Paris.
Ordóñez, J. 2014. "Tax Collection, the Informal Sector, and Productivity." *Review of Economic Dynamics* 17 (2): 262-286.
Orsi, R., D. Raggi, and F. Turino. 2014. "Size, Trend, and Policy Implications of the Underground Economy." *Review of Economic Dynamics* 17 (3): 417-436.
Oviedo, A., M. Thomas, and K. Karakurum-Özdemir. 2009. "Economic Informality: Causes, Costs, and Policies–A Literature Survey." Working Paper 167, World Bank, Washington, DC.
Perry, G. E., W. F. Maloney, O. S. Arias, P. Fajnzylber, A. D. Mason, and J. Saavedra-Chanduvi. 2007. *Informality: Exit and Exclusion*. Washington, DC: World Bank.
Putnins, T. and A. Sauka. 2015. "Measuring the Shadow Economy Using Company Managers." *Journal of Comparative Economics* 43 (2): 471-490.
Restrepo-Echavarria, P. 2014. "Macroeconomic Volatility: The Role of The Informal Economy." *European Economic Review* 70: 454-459.
Rocha, R., G. Ulyssea, and R. Rachter. 2018. "Do Lower Taxes Reduce Informality? Evidence from Brazil." *Journal of Development Economics* 134 (3): 28-49.
Sarte, P. 2000. "Informality and Rent-Seeking Bureaucracies in a Model of Long-Run Growth." *Journal of Monetary Economics* 46 (1): 173–97.
Schneider, F., and D. Enste. 2000. "Shadow Economies: Size, Causes, and Consequences." *Journal of Economic Literature* 38 (1): 77-114.
Schneider, F., A. Buehn, and C. E. Montenegro. 2010. "Shadow Economies All over the World: New Estimates for 162 Countries from 1999 to 2007." Policy Research Working Paper 5356, World Bank, Washington, DC.
Schneider, F. and A. Buehn. 2016. "Estimating the Size of the Shadow Economy: Methods, Problems and Open Questions." IZA Discussion Paper 9820, IZA-Institute of Labor Economics, Bonn, Germany.
Tanaka, K., and S. Keola. 2017. "Shedding Light on The Shadow Economy: A Nighttime Light Approach." *Journal of Development Studies* 53 (1): 32-48.
Ulyssea, G. 2010. "Regulation of Entry, Labor Market Institutions and the Informal Sector." *Journal of Development Economics* 91 (1): 87-99.
———. 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review* 108 (8): 2015-47.
World Bank. 2019. *Global Economic Prospects 2019 January: Darkening Skies.* Washington, DC: World Bank.

Figures

Figure 1. Informality and Development

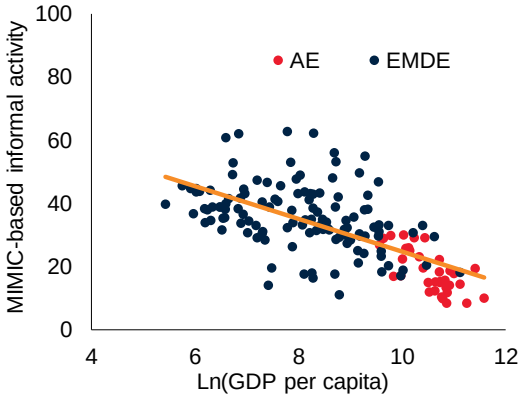
A. DGE-based informal activity

(Percent of official GDP)



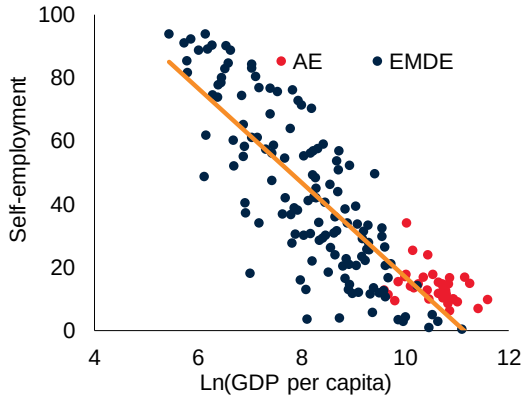
B. MIMIC-based informal activity

(Percent of official GDP)



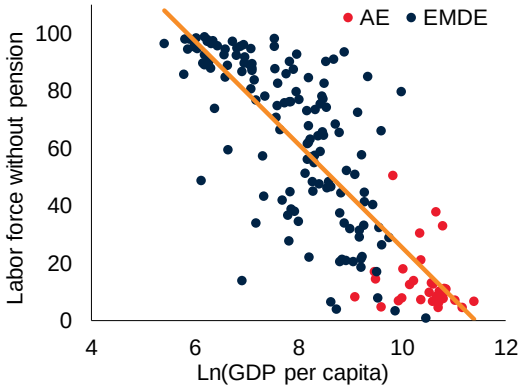
C. Self-employment

(Percent of total employment)



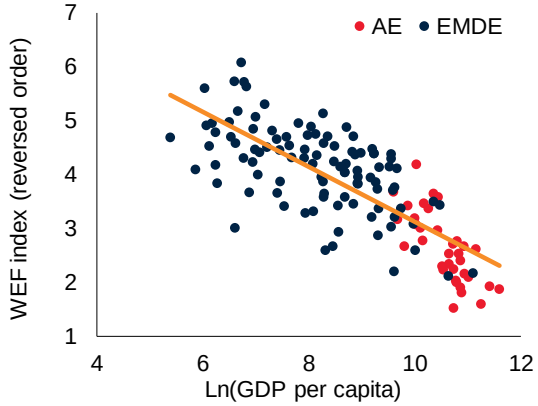
D. Labor force without pension

(Percent of labor force)



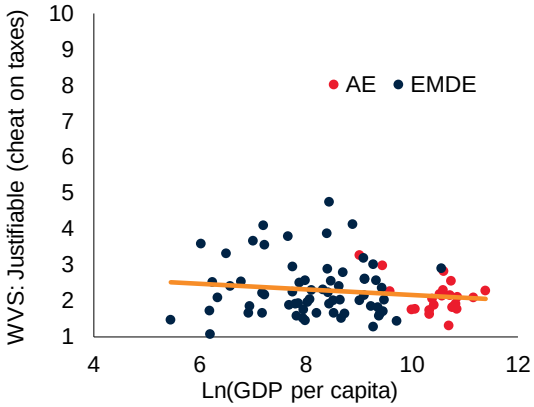
E. Perceived informal activity

(Average response, 1=low, 7=high)



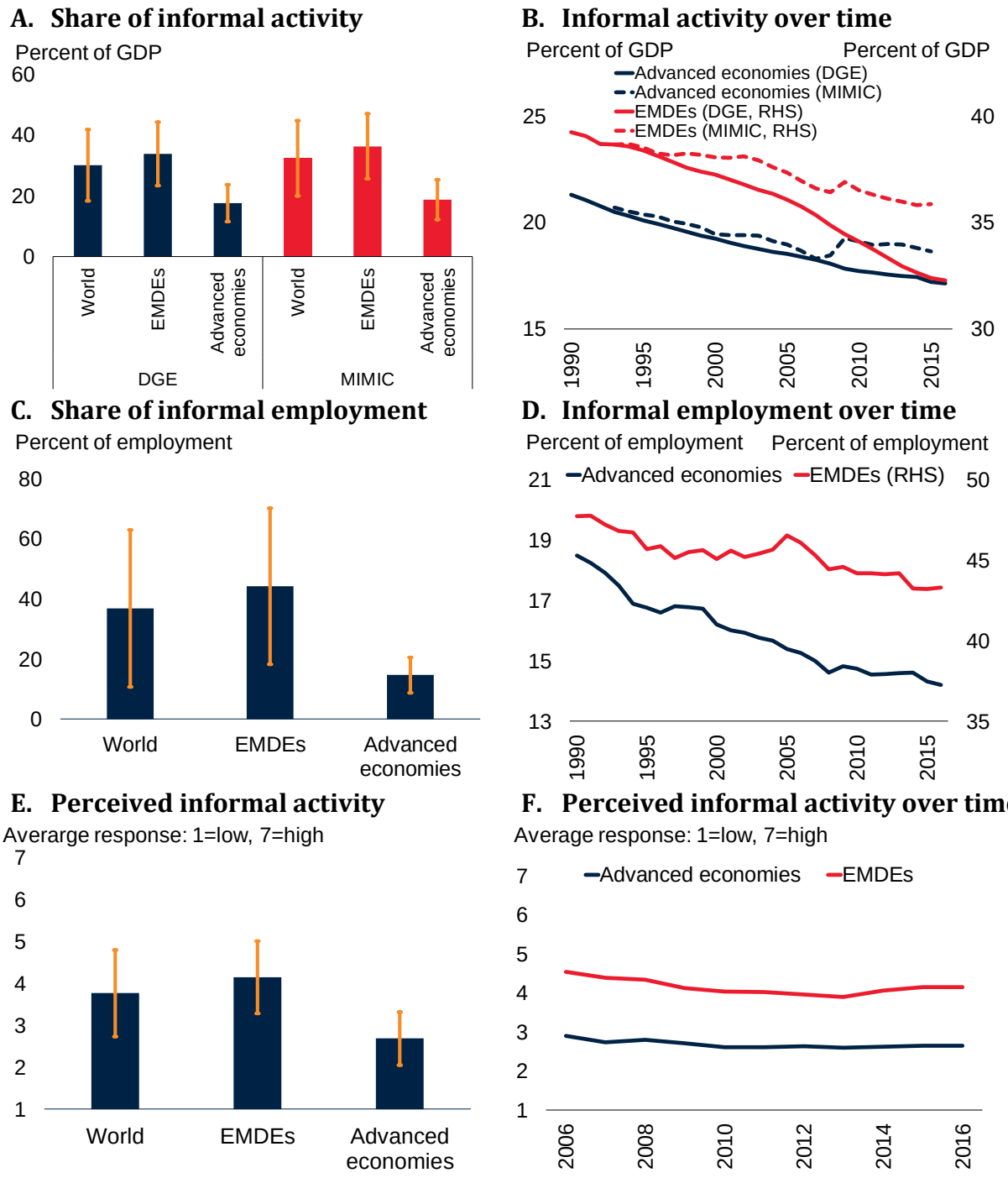
F. Attitudes to informality

(Average response, 1=never justifiable, 10=always justifiable)



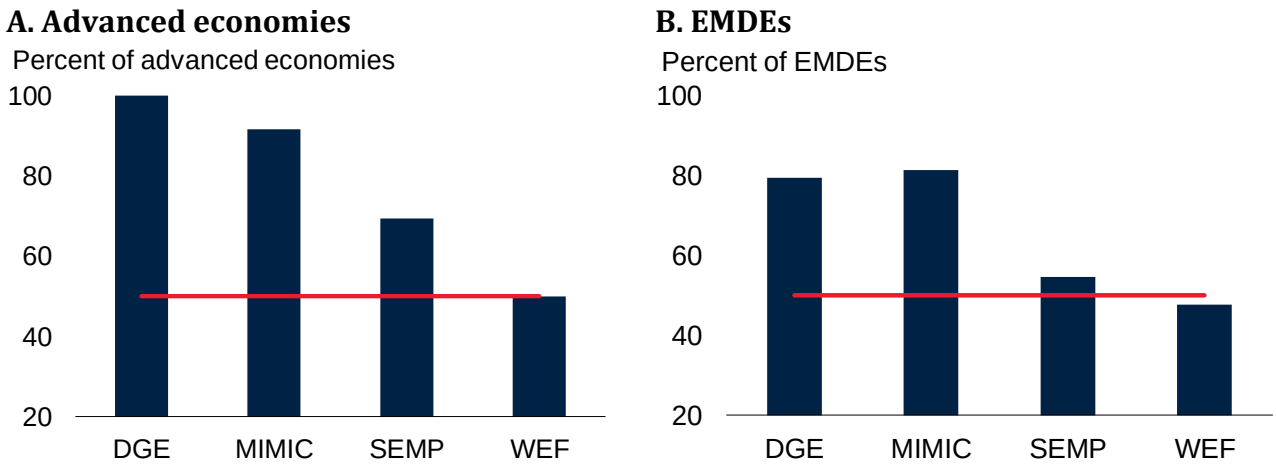
Note: See Section 2 for detailed information on data sources. Data for latest available year. 2016 for DGE model (A), 2015 for MIMIC estimates (B), 2016 or the latest available number for self-employment and labor force without pension (C and D), 2016 for perception of pervasiveness informality using WEF data (E) and the most recent response for WVS questions on justification for cheating (F). Advanced economies (AE) are shown in red and emerging and developing economies (EMDE) are shown in blue, while the fitted lines are shown in orange. "Ln (GDP per capita)" is GDP per capita (constant 2010 US$, in logs), taken from World Development Indicators (WDI).

Figure 2. Informality in Advanced Economies and EMDEs



Note: See Section 2 for details. Notes: In A, C and E, group means for the period 2006-2016 are shown in blue bars (MIMIC in A shown in red bars) with their -1 and +1 SD shown in orange whiskers. The group statistics are calculated for world, advanced economies, and emerging and developing economies (EMDEs). In B, D, and F, group means are calculated for advanced economies (in blue; dashed lines for MIMIC estimates in B) and emerging and developing economies (EMDEs, in red; dashed lines for MIMIC estimates in B). In D, missing data for informal employment are extrapolated in EMDEs for earlier years and filled using the latest available observation in recent years. In E and F, the WEF (World Economic Forum) index of informality is used.

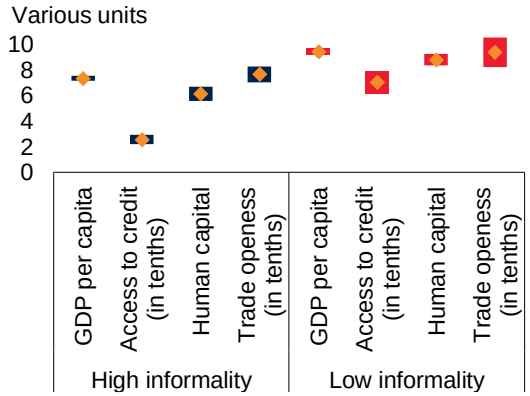
Figure 3. Countries with downward time trend in informality



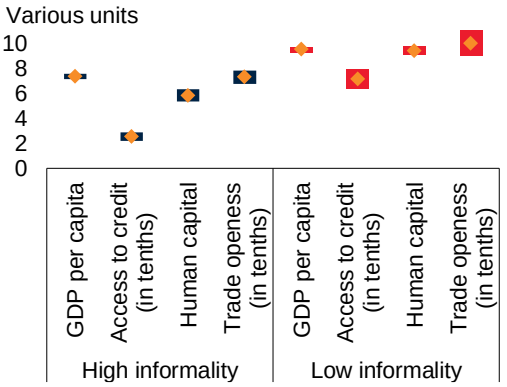
Note: Data for the period 1990-2016. Based on country-specific linear regressions of the share of informality by each of the four measures of informality with a sufficiently long-time dimension. Figures show the share of advanced economies (A) and EMDEs (B) for which the time trend is statistically significantly negative (at least at 10 percent level). In (B), missing values for self-employment are interpolated. Red line indicates 50 percent.

Figure 4. Correlates of Informality in EMDEs

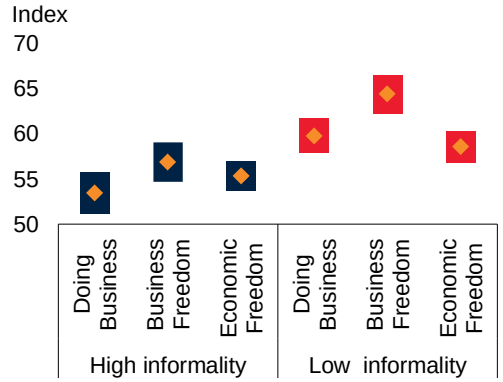
A. Economic correlates in EMDEs with high and low output informality



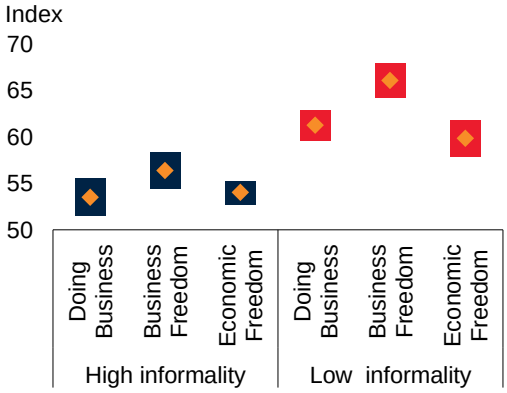
B. Economic correlates in EMDEs with high and low employment informality



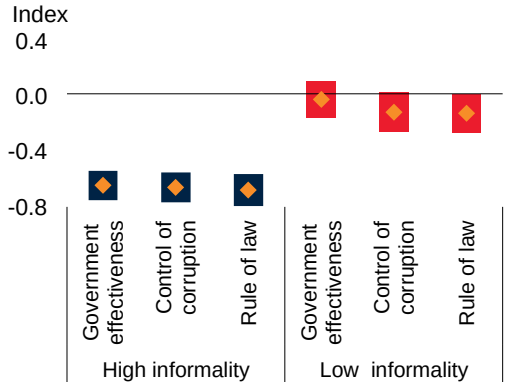
C. Regulatory burdens in EMDEs with high and low output informality



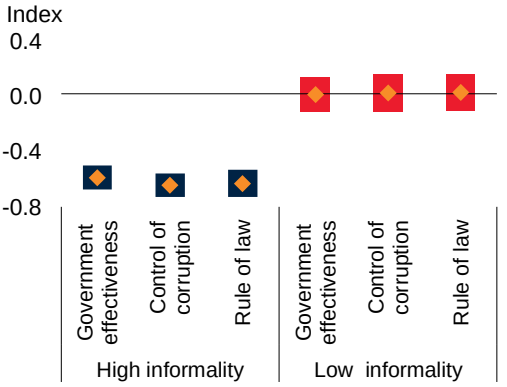
D. Regulatory burdens in EMDEs with high and low employment informality



E. Governance in EMDEs with high and low output informality



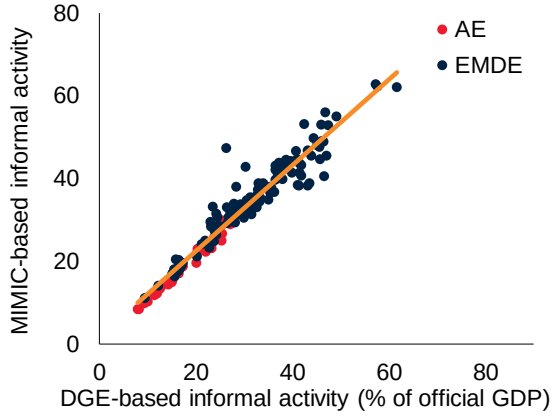
F. Governance in EMDEs with high and low employment informality



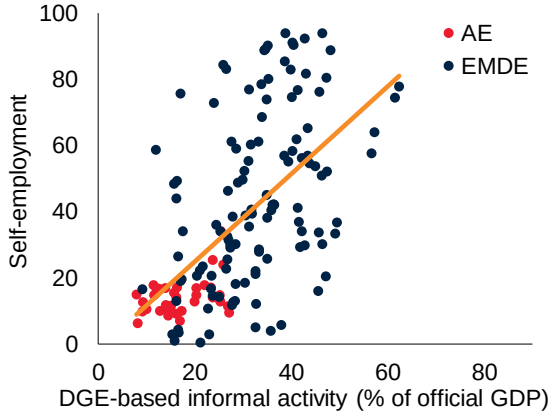
Source: See Section 2 for details. Notes: Data for the period 1990-2016 and EMDEs. The diamonds show the unweighted group averages, with the corresponding 90 percent confidence intervals shown in bars. "High-informality" are EMDEs with above median DGE-based informal output measures (self-employment shares) in A, C, E (B, D, F), while "Low informality" are EMDEs with below median DGE-based informal output measures (self-employment shares). In A-B, the correlates include GDP per capita (in logs, constant 2010 USD, WDI), access to credit (i.e., domestic credit to private sector in percent of GDP, WDI), human capital (average years of schooling, Barro and Lee 2013), trade openness (the sum of imports and exports in percent of GDP, WDI). In C-F, the correlates include doing business (measured as the overall distance to frontier with 100 being the frontier, Doing Business); business freedom and economic freedom (Heritage Foundation; the scores are between 0 and 100, with 100 being the freest economic / business environment); government effectiveness, control of corruption, and rules of law (World Governance Indicators, with higher values corresponding to better outcomes).

Figure 5. Scatterplots with DGE-based informal activity

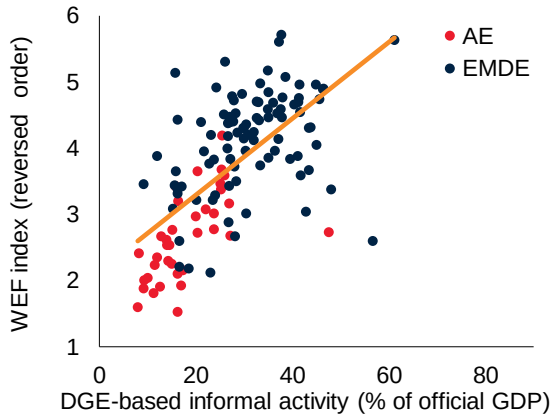
A. MIMIC-based informal activity
(Percent of official GDP)



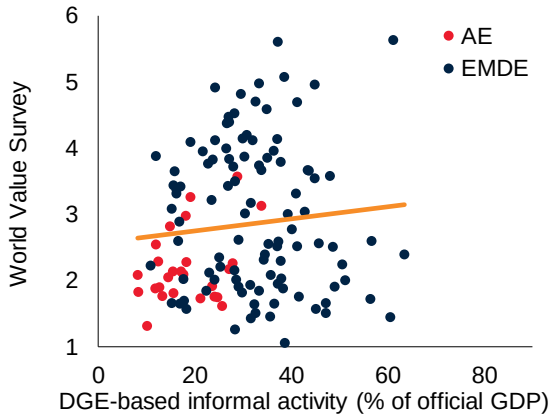
B. Informal employment shares
(Percent of total employment)



C. Perceived informal activity
(Average response, 1=low, 7=high)

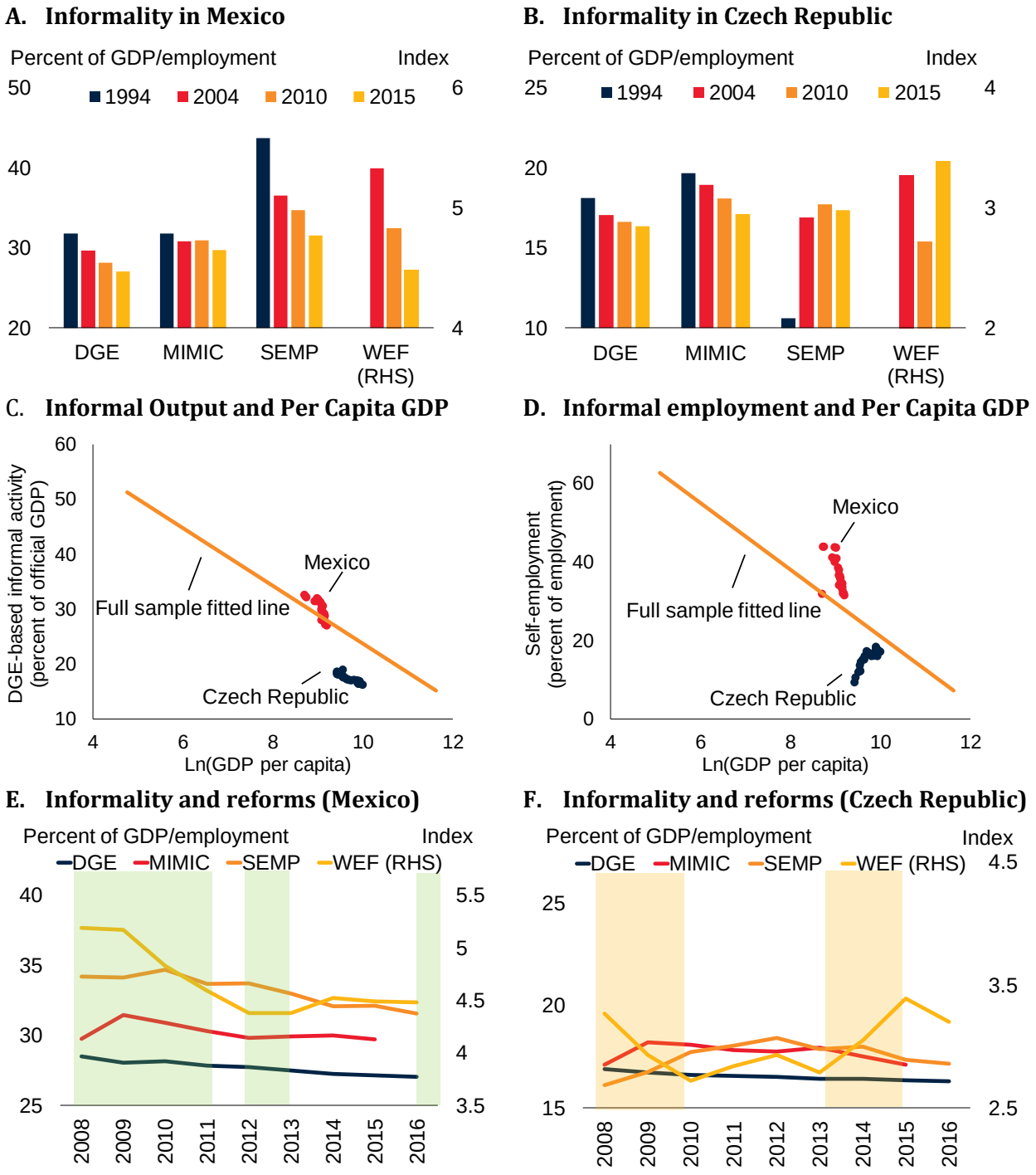


D. Attitudes to informality
(Average response, 1=never justifiable, 10=always justifiable)



Notes: Data for latest available year. 2015 for DGE-based and MIMIC-based estimates (A), latest available years for DGE-based estimates and self-employment shares (B), 2016 for DGE and perception of pervasive informality using WEF data (C) and the most recent response for WVS questions on justification for cheating and DGE in the same year (D). Advanced economies (AE) are shown in red and emerging and developing economies (EMDE) are shown in blue, while the fitted lines are shown in green.

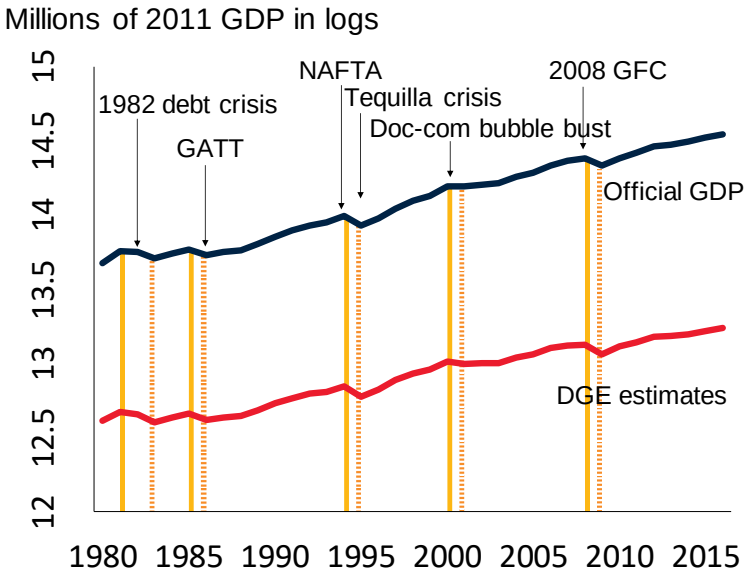
Figure 6. Informality in Mexico and Czech Republic



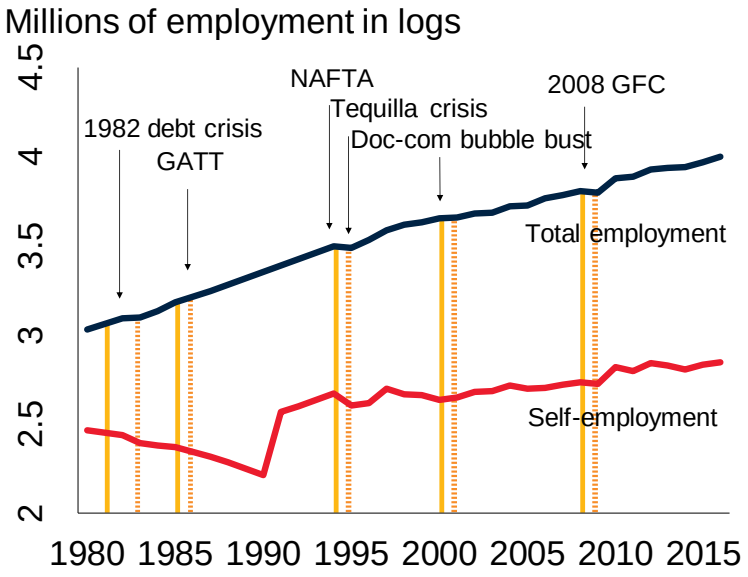
Notes: Data for the period 1990-2016. World Economic Forum (WEF) index has been reordered to make "7= Most economic activity is undeclared or unregistered; 1= Most economic activity is declared or registered" where a higher level suggesting a larger informal sector in the country. Detailed information on the informality measures is listed in Table A.1. A. B. Data for 2004 and 1994 not available for perceptions. Hence, for perceptions, the earliest available data (2006) shown in series labelled 2004. E.F. Years with tax reforms are in shaded green while years with labor reforms are in shaded orange. Information on reforms are obtained from Doing Business database.

Figure 7. Formal and informal business cycles in Mexico

A. Formal GDP and informal DGE



B. Total employment and self-employment



Note: The blue lines depict official GDP (A) and total employment (B). The red lines depict DGE estimates (A) and self-employment (B). Peaks of official GDP are labelled in green lines while troughs are labelled in orange dash lines. BBQ dating method is applied here. See main text for details about the dating method.

Tables

Table 1a. Summary statistics

Estimation method		Aspect		Measures	# of obs	# of ctry	Time period	Mean	Median	Min	Max
Indirect		Output (a)		DGE (% of GDP)	4,239	158	1990-2016	32.2	32.1	7.9	76.1
				MIMIC (% of GDP)	3,680	160	1993-2015	33.3	33.5	8.0	69.0
Direct (Survey-based)	Labor force Surveys	Employment		Pension coverage (% of labor force)	359	135	1990-2010	44.4	36.0	1.1	99.0
				Self-employment (% of total employment)	2,523	179	1990-2016	31.2	26.1	0.0	95.5
				Informal employment (% of total employment)	234	53	2001-2016	65.5	69.6	18.9	99.7
				Employment outside the formal sector (% of total employment)	249	57	1999-2016	57.7	58.1	11.2	94
	Firm surveys	Perception	(a)	WEF (1-7=Most informal)	1,527	151	2006-2016	3.8	3.9	1.5	6.4
			Firms	WB: % firms competing against informal firms	229	139	2006-2016	54.9	55.7	7.2	95.2
				WB: % firms formally registered when founded	214	136	2006-2016	89.1	91.9	29.1	100.0
				WB: Number of years operated without registration	214	136	2006-2016	0.7	0.5	0.0	6.8
				WB: % firms that found competitors in the informal sector as a constraint	230	138	2006-2016	29.5	28.6	0.0	76.0
	HS		(b)	WVS: Justifiable (Cheating on taxes)	200	94	1994-2010	2.3	2.1	1.0	4.7

Notes: Data for the period 1990-2016. DGE is benchmarked to Schneider et al. (2010). World Value Survey (WVS) asks whether cheating on taxes is justifiable (1 is "never justifiable" and 10 is "always justifiable") and reports average responses at the country-year level, with a higher level suggesting that the country is more tolerant towards the informal sector. World Economic Forum (WEF) asks "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1= Most economic activity is undeclared or unregistered; 7= Most economic activity is declared or registered)" and reports average responses at the country-year level. Here the average responses have been reordered to make "7= Most economic activity is undeclared or unregistered; 1= Most economic activity is declared or registered" where a higher level suggesting a larger informal sector in the country. The WEF data for year 2004 and 2005 are dropped since different ordering were used before 2006, which makes the numbers incomparable over time. WB shows the results for World Bank Enterprise Surveys. HS stands for "Household surveys", (a) stands for "Output", and (b) stands for "Opinions/Tax Morality". Detailed information is listed in Table A.1. Since the data on self-employment for Equatorial Guinea is only for year 1983, the data on self-employment are available for 179 countries (instead of 180 economies) here.

Table 1b. Summary statistics by country groupings

	World		AEs		EMDEs	
	Mean	**Median**	**Mean**	**Median**	**Mean**	**Median**
Output						
DGE (% of GDP)	32.2	32.1	18.9	17.3	36.2	36.2
MIMIC (% of GDP)	33.3	33.5	19.4	18.2	37.4	36.8
Employment						
Pension coverage (% of labor force)	44.4	36.0	86.5	90.0	30.5	24.1
Self-employment (% of total employment)	31.2	26.1	15.8	13.9	39.6	37.0
Informal employment (% of total employment)	65.5	69.6			65.5	69.6
Employment outside the formal sector (% of total employment)	57.7	58.1			57.7	58.1
Perception						
WEF (1-7=Most informal)	3.8	3.9	2.7	2.6	4.1	4.2
WB: % firms competing against informal firms	54.9	55.7	35.8	35.9	56.2	57.7
WB: % firms formally registered when founded	89.1	91.9	98.2	98.5	88.6	91.3
WB: Number of years operated without registration	0.7	0.5	0.2	0.1	0.8	0.5
WB: % firms that found competitors in the informal sector as a constraint	29.5	28.6	17.5	16.1	30.2	29.2
WVS: Justifiable (Cheating on taxes)	2.3	2.1	2.2	2.1	2.3	2.1

Notes: Data for the period 1990-2016. DGE is benchmarked to Schneider et al. (2010). World Value Survey (WVS) asks whether cheating on taxes is justifiable (1 is "never justifiable" and 10 is "always justifiable") and reports average responses at the country-year level, with a higher level suggesting that the country is more tolerant towards the informal sector. World Economic Forum (WEF) asks "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1= Most economic activity is undeclared or unregistered; 7= Most economic activity is declared or registered)" and reports average responses at the country-year level. Here the average responses have been reordered to make "7= Most economic activity is undeclared or unregistered; 1= Most economic activity is declared or registered" where a higher level suggesting a larger informal sector in the country. WB shows the results for World Bank Enterprise Surveys. Detailed information is listed in Table A.1. Country groupings follow the method used by Global Economic Prospects (2019).

Table 2. Spearman rank correlations (across countries within individual years)

	DGE	MIMIC	PENSION	SEMP	IF_EMP	EMP_NF	WEF	WB1	WB2	WB3	WB4	WVS
Output												
DGE (% of GDP)	1.00											
MIMIC (% of GDP)	0.98***	1.00										
Employment												
Pension coverage (% of labor force)	-0.59***	-0.60***	1.00									
Self-employment (% of total employment)	0.63***	0.62***	-0.86***	1.00								
Informal employment (% of total employment)	0.38	0.38	-0.90***	0.79***	1.00							
Employment outside the formal sector (% of total employment)	0.31*	0.34**	-0.60	0.71**	0.93***	1.00						
Perception												
WEF (1-7=Most informal)	0.69***	0.70***	-0.46***	0.72***	0.54***	0.53***	1.00					
WB: % firms competing against informal firms	0.40***	0.40***	-0.07	0.37***	0.18	0.37	0.60***	1.00				
WB: % firms formally registered when founded	-0.25***	-0.23	0.67***	-0.54***	-0.62**	-0.60**	-0.57***	-0.60***	1.00			
WB: Number of years operated without registration	0.28**	0.26**	-0.30	0.39***	0.31	0.42	0.45***	0.38***	-0.73***	1.00		
WB: % firms that found competitors in the informal sector as a constraint	0.37***	0.36***	0.08	0.20*	-0.05	0.09	0.48*	0.73***	-0.36***	0.25**	1.00	
WVS: Justifiable (Cheating on taxes)	0.21	0.27*	0.31*	-0.21	-0.50	-0.16	0.14	-0.20	0.33	-0.07	-0.21	1.00

Note: Data for the period 1990-2016. Medians of rank correlation of data across countries within each year. All survey-based measures are interpolated. DGE is benchmarked to Schneider et al. (2010). World Value Survey (WVS) asks whether cheating on taxes is justifiable (1 is "never justifiable" and 10 is "always justifiable") and reports average responses at the country-year level, with a higher level suggesting that the country is more tolerant towards the informal sector. World Economic Forum (WEF) asks "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1= Most economic activity is undeclared or unregistered; 7= Most economic activity is declared or registered)" and reports average responses at the country-year level. Here the average responses have been reordered to make "7= Most economic activity is undeclared or unregistered; 1= Most economic activity is declared or registered" where a higher level suggesting a larger informal sector in the country. "WB" here stands for "World Bank Enterprise Surveys". Detailed information is listed in Table A.1. "***" implies significance at 1% level, "**" implies significance at 5% level, "*" implies significance at 10% level.

Table 3. Coincidence of signs of first-differences

	DGE	MIMIC	PENSION	SEMP	INF_EMP	EMP_NF	WEF	WVS
DGE (% of GDP)	100							
MIMIC (% of GDP)	59.7	100						
Pension coverage (% of labor force)	45.4	45.1	100					
Self-employment (% of total employment)	58.4	58.1	49.0	100				
Informal employment (% of total employment)	59.3	57.8	47.8	62.0	100			
Employment outside the formal sector (% of total employment)	57.5	54.9	46.2	66.5	81.9	100		
WEF (1-7=Most informal)	53.3	54.2	52.8	49.6	59.3	57.3	100	
WVS: Justifiable (Cheat on taxes)	59.3	57.7	55.7	54.9	70.0	70.4	51.1	100

Note: Data for the period 1990-2016. Shares of country-year pairs in which the first difference in the two measures has the same sign (opposite for PENSION) are shown. Survey-based estimates are interpolated to fill the gaps in data series. DGE is benchmarked to Schneider et al. (2010). World Value Survey (WVS) asks whether cheating on taxes is justifiable (1 is "never justifiable" and 10 is "always justifiable") and reports average responses at the country-year level, with a higher level suggesting that the country is more tolerant towards the informal sector. World Economic Forum (WEF) asks "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1= Most economic activity is undeclared or unregistered; 7= Most economic activity is declared or registered)" and reports average responses at the country-year level. Here the average responses have been reordered to make "7= Most economic activity is undeclared or unregistered; 1= Most economic activity is declared or registered" where a higher level suggesting a larger informal sector in the country. WB shows the results for World Bank Enterprise Surveys. Detailed information is listed in Table A.1.

Table 4. Volatility of formal and informal economies

	Output		
	[1]	[2]	[3]
	Formal output	**DGE-based informal output**	**MIMIC-based informal output**
World	6.67	6.14***	5.27***
AEs	3.92 ^	3.99 ^	2.40*** ^
EMDEs	7.21	6.61***	5.81***
	Employment		
	[4]	[5]	[6]
	Total employment	**Formal employment**	**Self-employment**
World	3.54	5.39***	6.97***
AEs	2.18 ^	2.48*** ^	5.06*** ^
EMDEs	3.9	6.16***	7.62***

Note: Data are for the period 1990-2016. Formal output is captured by official GDP, while informal output uses DGE-based or MIMIC-based estimates. Total employment is the sum of formal employment and self-employment. Volatility shows the standard deviations (SDs) of the concerning variables' annual growth rates. "***" implies significant differences at 1 percent level in the SDs of the annual growth rates of formal output and those of informal output in row [1]-[3] (in the SDs of the annual growth rates of total employment and those of formal/self-employment in row [4]-[6]). The shaded areas indicate that the SDs of the annual growth rates of DGE-based informal output (formal employment) significantly differ from those of MIMIC-based informal output (self-employment). "^" indicates significant differences at 5 percent level between advanced economies (AEs) and emerging markets and developing economies (EMDEs).

Table 5a Cyclical features of recessions in formal and informal economy

	# of observations	Duration (years)	Amplitude (%)	Cumulative loss (%)	Slope (%)
Formal Output					
World	208	1.5	-5.9	-6.4	-4.1
		[1.0]	[-3.0]	[-1.9]	[-2.2]
AEs	72	1.5	-4.0**	-4.1	-2.5**
		[1.0]	[-2.3]	[-1.5]	[-1.8]
EMDEs	280	1.5	-6.6	-7.2	-4.6
		[1.0]	[-3.1]	[-2.0]	[-2.4]
DGE-based Informal Output					
World	309	1.6	-6.0	-6.4	-4.0
		[1.0]	[-3.1]	[-2.0]	[-2.3]
AEs	86	1.7**	-4.6*	-5.1	-2.7**
		[2.0***]	[-3.0]	[-2.0]	[-2.0]
EMDEs	223	1.5	-6.5	-6.9	-4.5
		[1.0]	[-3.2]	[-2.0]	[-2.4]
MIMIC-based Informal Output					
World	155	1.4	-3.7	-4.9	-2.5
		[1.0]	[-1.6]	[-0.9]	[-1.4]
AEs	44	1.5	-2.4	-2.7	-1.4***
		[1.0]	[-0.7***]	[-0.3**]	[-0.5**]
EMDEs	111	1.3	-4.2	-5.7	-2.9
		[1.0]	[-2.2]	[-1.1]	[-1.8]

Note: Data for recession episodes starting in the period 1990-2016. Business cycle turning points determined based on formal and informal GDP levels (i.e. official GDP statistics for formal output, DGE and MIMIC estimates for informal output) using the algorithm of Harding and Pagan (2002). Recession is defined as the phase from peak to trough while its corresponding "Duration", "Amplitude", "Cumulative loss" and "Slope" are defined as in Claessens et al. (2012). All statistics correspond to sample means. Medians are in brackets. Asterisks refer to the significant differences in means (or medians) between advanced economies (AEs) and emerging markets and developing economies (EMDEs). "***" implies significance at 1% level, "**" implies significance at 5% level, "*" implies significance at 10% level. Differences between informal and formal economies that are significant at 10 percent level are highlighted in shaded gray.

Table 5b Cyclical features of recoveries in formal and informal economy

	# of observations	Duration (years)	Amplitude (%)	Slope (%)
Formal Output				
World	179	2.1	5.4	4.7
		[2.0]	[3.4]	[2.1]
AEs	37	1.7*	2.9**	2.0*
		[1.0]	[2.3**]	[1.1**]
EMDEs	142	2.2	6.1	5.4
		[2.0]	[3.8]	[2.4]
DGE-based Informal Output				
World	215	2.1	4.2	3.7
		[2.0]	[3.0]	[2.3]
AEs	56	1.9	2.2***	2.0**
		[2.0]	[1.7***]	[1.6***]
EMDEs	159	2.1	4.9	4.3
		[2.0]	[4.0]	[2.6]
MIMIC-based Informal Output				
World	69	1.7	3.2	2.1
		[1.0]	[2.6]	[1.2]
AEs	19	1.4	1.6***	1.5
		[1.0]	[1.5***]	[0.5*]
EMDEs	50	1.7	3.8	2.3
		[1.0]	[3.2]	[1.7]

Note: Data for recovery episodes starting in 1990-2016. Business cycle turning points determined based on formal and informal GDP levels (i.e. official GDP statistics for formal output, DGE and MIMIC estimates for informal output) using the algorithm of Harding and Pagan (2002). Recovery is defined as the time it takes for output to rebound from the trough to the peak level before the recession while its corresponding "Duration", "Amplitude", and "Slope" are defined as in Claessens et al. (2012). All statistics correspond to sample means. Medians are in brackets. Asterisks refer to the significant differences in means (or medians) between advanced economies (AEs) and emerging markets and developing economies (EMDEs). "***" implies significance at 1% level, "**" implies significance at 5% level, "*" implies significance at 10% level. Differences between informal and formal economies that are significant at 10 percent level are highlighted in shaded gray.

Table 6 Employment growth during formal and informal business cycles

	Formal Output Business Cycle								
	Total employment			**Total employment (excl. self-employment)**			**Self-employment**		
	Recession	Recovery	Expansion	Recession	Recovery	Expansion	Recession	Recovery	Expansion
World	-0.5	-0.4	0.2	-0.4	-1.1	0.0	-0.5	-0.0	0.2
	[-0.2]	[-0.2]	[0.2]	[-0.3]	[-1.1]	[0.2]	[-0.5]	[0.0]	[0.2]
AEs	-0.6	-1.4**	0.3	-0.7	-1.8*	0.2	-1.0	0.7	0.5
	[-0.2]	[-1.5***]	[0.3]	[-0.2]	[-2.0***]	[0.3*]	[-0.6]	[0.4]	[0.5*]
EMDEs	-0.4	-0.2	0.1	-0.2	-0.7	-0.1	-0.3	-0.4	0.0
	[-0.2]	[0.1]	[0.2]	[-0.6]	[-0.5]	[0.1]	[0.1]	[-0.1]	[0.0]
	DGE-based Informal Output Business Cycle								
	Total employment			**Total employment (excl. self-employment)**			**Self-employment**		
	Recession	Recovery	Expansion	Recession	Recovery	Expansion	Recession	Recovery	Expansion
World	-0.5	-0.3	0.2	-0.9	-0.7	0.2	-0.5	0.5	0.2
	[-0.3]	[-0.3]	[0.2]	[-0.4]	[-1]	[0.2]	[-0.5]	[0.2]	[0.2]
AEs	-0.6	-1.1***	0.2	-0.7	-1.5**	0.2	-0.9	0.7	0.5
	[-0.3]	[-1.2***]	[0.3]	[-0.3]	[-1.5***]	[0.3]	[-0.6]	[0.4]	[0.5***]
EMDEs	-0.5	-0.1	0.2	-1.1	-0.2	0.2	-0.2	0.4	-0.0
	[-0.4]	[0.0]	[0.2]	[-0.6]	[-0.2]	[0.1]	[0.1]	[0.0]	[0.0]
	MIMIC-based Informal Output Business Cycle								
	Total employment			**Total employment (excl. self-employment)**			**Self-employment**		
	Recession	Recovery	Expansion	Recession	Recovery	Expansion	Recession	Recovery	Expansion
World	-0.5	-0.3	-0.0	-0.6	-0.5	0.1	-0.9	0.3	-0.3
	[-0.4]	[-0.3]	[0.1]	[-0.6]	[-0.6]	[-0.1]	[-0.3]	[0.0]	[0.0]
AEs	-1.1*	-1.2	-0.1	-1.4	-1.6*	-0.2	-1.3	0.7	0.3
	[-0.4]	[-1.0***]	[-0.2*]	[-0.6]	[-1.4***]	[-0.2*]	[-0.3]	[-0.1]	[-0.1]
EMDEs	-0.2	-0.0	-0.0	-0.1	0.1	0.2	-0.7	0.1	-0.6
	[-0.4]	[0.0]	[0.2]	[-0.3]	[0.0]	[0.1]	[0.0]	[0.2]	[0.1]

Note: Data for the period 1990-2016. Business cycle turning points determined based on formal and informal GDP levels (i.e. official GDP statistics for formal output, DGE and MIMIC estimates for informal output) using the algorithm of Harding and Pagan (2002). Recession is defined as the phase from peak to trough while recovery is defined as the time it takes for output to rebound from the trough to the peak level before the recession (Claessens et al. 2012). Expansion is defined as the period from trough to next peak (Claessens et al. 2012). All statistics correspond to the sample medians of demeaned, first differenced, and logged employment. Means are in brackets. Shaded cells represent numbers that significantly differ from zero. Asterisks refer to the significant differences in means (or medians) between advanced economies (AEs) and emerging markets and developing economies (EMDEs). "***" implies significance at 1% level, "**" implies significance at 5% level, "*" implies significance at 10% level.

Appendix

Appendix: Methodologies

The multiple indicators multiple causes model (*MIMIC*).

To estimate the size of the informal sector (i.e., in percent of official GDP) with the MIMIC model, this study closely follows Schneider et al. (2010) and includes six causes and three indicators used in their study.[33]

The six cause variables used here are: (1) size of government (general government final consumption expenditure, as a percent of GDP, obtained from UN, spliced with WDI) as proxy for indirect taxation; (2) share of direct taxation (direct taxes in percent of overall taxation, WDI); (3) fiscal freedom index obtained from Heritage Foundation as a tax burden variable in a wide sense; (4) business freedom index provided by Heritage Foundation; (5) the unemployment rate and GDP per capita to capture the state of the economy (obtained from WDI, the latter is spliced with WEO); and (6) a measure on government effectiveness provided by Worldwide Governance Indicators. The three indicator variables include: (1) growth rate of GDP per capita (WDI, spliced with WEO); (2) the labor force participation rate (people over 15 economically active as a percentage of total population, WDI, spliced with Haver analytics), and (3) currency as a ratio of M0 (currency outside the banks) over M1 (obtained from IMF IFS).

The estimation results are shown in Appendix Table A2. The model specification that ensures maximum data coverage (i.e. Column (5)) is used to generate the MIMIC index of the share of informal output relative to official GDP ($\tilde{\eta}_t$). Then we conduct an additional benchmarking procedure where $\tilde{\eta}_t$t is converted into absolute values of the informal sector ($\widehat{\eta_t}$) using the following equation:[34]

$$\widehat{\eta_t} = \frac{\tilde{\eta}t}{\tilde{\eta}_{2000}} \eta^*_{2000},$$

where t denotes year, $\tilde{\eta}_{2000}$ is the value of the estimated index in the base year 2000, and η^*_{2000} is the exogenous estimate (base value) of the shadow economies in 2000. While the estimates ($\tilde{\eta}_t$) determine the movement of the absolute values of the informal sector over time, the base values η^*_{2000} decide the rankings of the countries' informal sector within the sample in year 2000. The base values η^*_{2000} are taken from Schneider (2007) or, for another 10 countries, from Schneider et al. (2010).

The DGE model (*DGE*).

The calibration follows Elgin and Oztunali (2012) and takes parameter values suggested by the earlier literature (e.g. Ihrig and Moe 2004).[35] α is assumed to be equal to 0.36, and γ takes the value of 0.425. Data are gathered from PWT 9.0 for capital stock (K_t), private consumption (C_t), formal employment (N_{Ft}), depreciation rates (δ, country averages), and tax rates (τ_t). By matching the productivity in the informal sector to the informal economy size in 2007 of the series reported in Schneider et al. (2010) and assuming that A_{It} grows at the average growth rate of K_t and A_{Ft},[36] the DGE estimates are computed for 158 countries over the period 1950-2016.

The estimation results are qualitatively robust to different model specifications such as using alternative values for δ, α, γ, adding labor-leisure choice, tax enforcement parameter to informal sector income (for

[33] MIMIC is a type of Structural Equation Model (SEM). The estimation of a SEM with latent variables can be done by means of LISREL (used by Schneider et al. 2010), SPSS and Stata. Here Stata is used.

[34] Calibration is performed separately for each country. Following Schneider et al. (2010), the MIMIC index has been adjusted to the positive range by adding a positive constant.

[35] Elgin and Oztunali (2012) are not using the model to do a full calibration exercise, where each equilibrium condition is satisfied for every period. Since only two of the equilibrium conditions are utilized, stationarity of empirical data for calibration is a lesser concern. Their approach is followed here.

[36] This assumption implies that growth in the formal sector can spillover to the informal sector via capital accumulation and technological diffusion.

example, using revenue in percent of GDP rather than government spending in percent of GDP for τ_t), see Elgin and Oztunali (2012) for details.

Labor-related Measures on Informality

Several cross-country databases report the survey-based estimates on the share of self-employment in total employment:[37] 1) the World Development Indicators (WDI 2016), which cover 175 countries from 1980 (mainly from 1990s) to 2014; 2) the International Labor Organization (ILO 2016), which covers 109 countries from 1997 to 2014; and 3) OECD (2016 and 2018), which covers 34 OECD countries from 1955 to 2016. When regarding employment outside the formal sector and informal employment, ILO compiled statistics for up to 57 medium and low-income countries for as much as 1999-2016.

For a comprehensive dataset on labor-related measures on informality, we combine the cross-country databases, provided by WDI, ILO and OECD, and gather additional data from the national statistical bureaus (offices), Haver Analytics, the disclosed Living Standards Measurement Studies (LSMS, World Bank), and spliced data from IADB and Eurostat to fill some gaps in years. Data priority is first given to cross-country databases (WDI 2016, ILO 2016 and OECD 2018) and then national statistical bureaus (offices), Haver Analytics, and LSMS, followed by estimates obtained from previous studies, IADB and Eurostat. IADB reports the share of self-employment in total employment (15-64 years old) for 19 Latin American countries between 1990 and 2016, while Eurostat reports the same measure for 29 EU countries and 5 non-EU countries for the period 1983-2016. By focusing on employment of the 15-64 years old groups, their data are systematically lower than those from other cross-country databases. The final step adds 62 more observations to the sample (2 percent of the full sample).

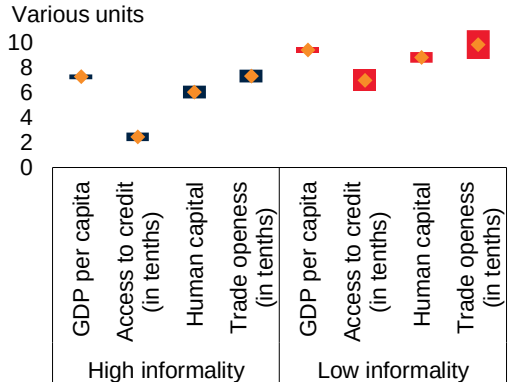
Here lists the national statistical bureaus (offices) where their websites are searched and contacted:

Country	Country	Country	Country	Country
Angola	Cyprus	Kenya	Oman	Togo
Belarus	Democratic Republic of the Congo	Kuwait	Papua New Guinea	Trinidad and Tobago
Belize	Egypt	Kyrgyz Republic	Republic of Congo	Tunisia
Benin	Equatorial Guinea	Lao P.D.R.	Romania	Tunisia
Bhutan	Eritrea	Latvia	Rwanda	United Arab Emirates
Bosnia and Herzegovina	Ethiopia	Lebanon	Senegal	Venezuela
Botswana	Fiji	Libya	Sierra Leone	Vietnam
Brunei Darussalam	FYR Macedonia	Lithuania	Singapore	Yemen
Bulgaria	Gabon	Malawi	Solomon Islands	Macao, China
Burkina Faso	Georgia	Maldives	Sudan	Argentina
Burundi	Guinea	Malta	Suriname	Azerbaijan
Cape Verde	Guinea-Bissau	Mauritania	Swaziland	Bahrain
Cambodia	Guyana	Moldova	Syria	China
Cameroon	Haiti	Mozambique	Taiwan, China	Ghana
Central African Republic	Hong Kong SAR	Myanmar	Tajikistan	Morocco
Chad	Iran	Nepal	Tanzania	Qatar
Comoros	Jamaica	Niger	The Bahamas	Saudi Arabia
Croatia	Jordan	Nigeria	The Gambia	

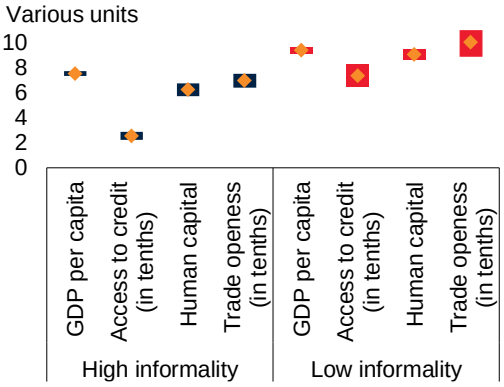
[37] Both ILO and WDI only report model-based estimates from 2018 onwards, which may suffer from caveats such as strong economic assumptions and reliance on other studies' independent estimates to do the benchmarking. Due to the issues related with model-based estimates, we gather historical WDI and ILO reports to obtain survey-based estimates. The model-based estimates from ILO and WDI were used to update the share of self-employment when no other source of information is available.

Figure A1. Correlates of Informality in EMDEs: MIMIC-based informal activity and WEF index

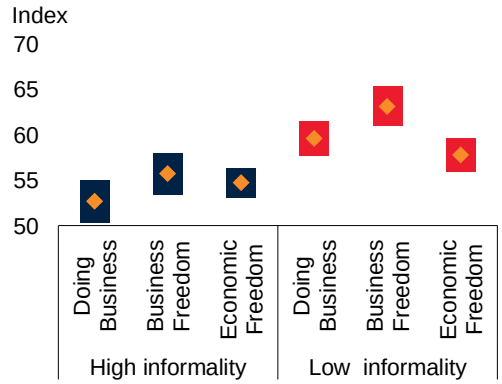
A. Economic correlates in EMDEs with high and low output informality



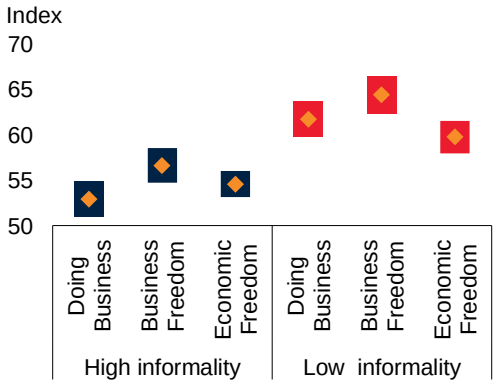
B. Economic correlates in EMDEs with high and low perceived informality



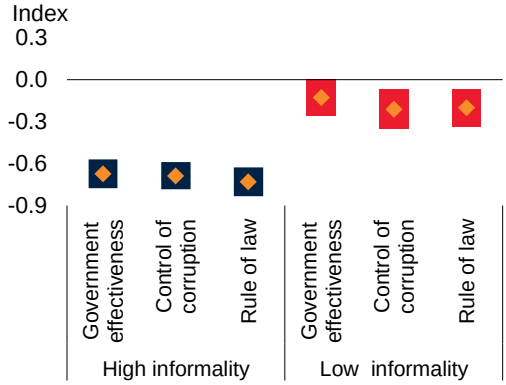
C. Regulatory burdens in EMDEs with high and low output informality



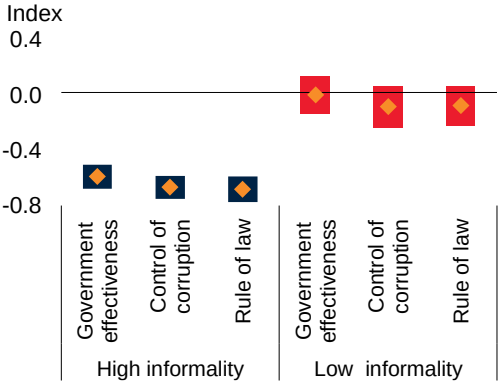
D. Regulatory burdens in EMDEs with high and low perceived informality



E. Governance in EMDEs with high and low output informality



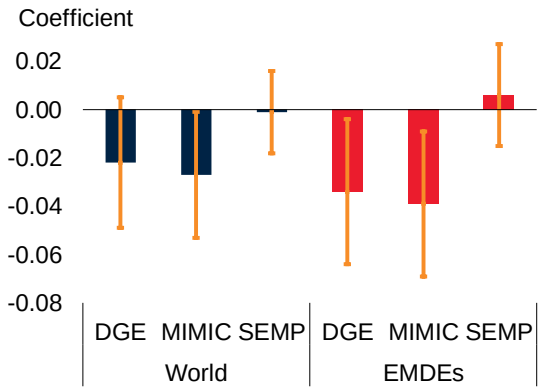
F. Governance in EMDEs with high and low perceived informality



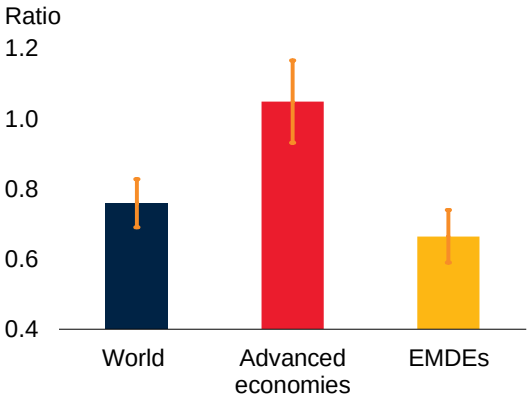
Source: See Section 2 for details. Notes: Data for the period 1990-2016 and EMDEs. The diamonds show the unweighted group averages, with the corresponding 90 percent confidence intervals shown in bars. "High-informality" are EMDEs with above median MIMIC-based informal output measures (WEF-index, reversed order) in A, C, E (B, D, F), while "Low informality" are EMDEs with below median MIMIC-based informal output measures (WEF index, reversed order). In A-B, the correlates include GDP per capita (in logs, constant 2010 USD, WDI), access to credit (i.e., domestic credit to private sector in percent of GDP, WDI), human capital (i.e., average years of schooling, Barro and Lee 2013), trade openness (i.e., the sum of imports and exports in percent of GDP). In C-F, the correlates include doing business (measured as the overall distance to frontier with 100 being the frontier, Doing Business); business freedom and economic freedom (Heritage Foundation; the scores are between 0 and 100, with 100 being the freest economic / business environment); government effectiveness, control of corruption, and rules of law (World Governance Indicators, with higher values corresponding to better outcomes).

Figure A2. Correlates of Informality: Macroeconomic implications

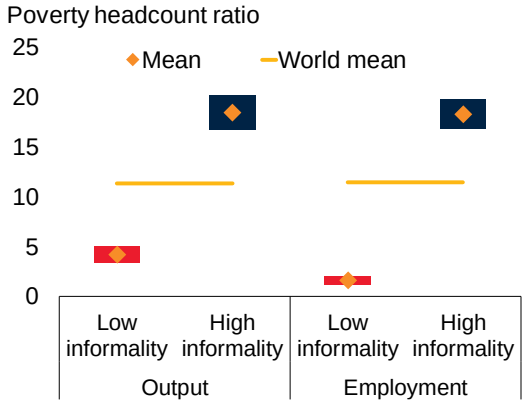
A. Correlation between informality and GDP growth in EMDEs



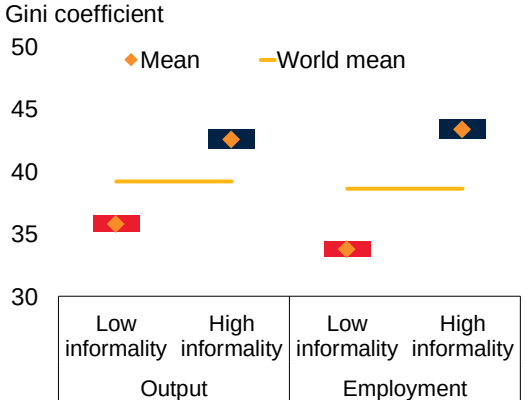
B. Ratio of informal labor productivity to total labor productivity



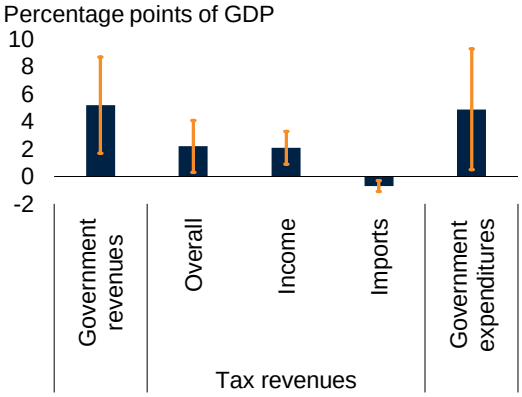
C. Informality and extreme poverty



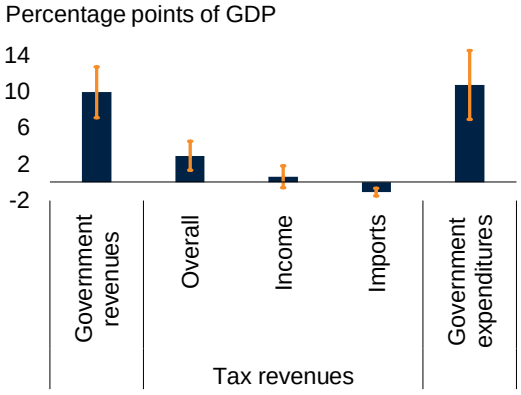
D. Informality and income inequality



E. Differential in fiscal indicators between third of EMDEs with the lowest and highest output informality



F. Differential in fiscal indicators between third of EMDEs with the lowest and highest employment informality



Source: See Section 2 and World Bank (2019) for details. Notes: A. The corresponding 90 percent confidence intervals are shown in whiskers with the coefficient estimates shown in bars. Annual GDP growth is regressed against the various measures of informality while controlling for real GDP per capital (in logs, WDI). B. The average ratios over 1990-2016 (as calculated in Loayza, 2018) are shown in bars with corresponding 95 percent confidence intervals shown in whiskers. C.D. Group averages over 1990-2016 for countries with higher informality (above median) and those with lower informality (below median) are shown in diamonds with corresponding 95 percent confidence intervals shown in bars. The world averages over 1990-2016 are shown in yellow lines. E.F. Differences in percentage points of GDP between the average fiscal indicators among the third of EMDEs with the highest and lowest informality are in bars, with corresponding 90 percent confidence intervals shown in whiskers. All fiscal indicators and informality measures are 2000-16 averages. See World Bank (2019) for details.

Table A1. Summary of Data Coverage

Data	Method	Sources	Measure	# of Ctry	Period	Setup
MIMIC-based informal output	Indirect estimates (MIMIC)	Author's calculations	Size of the informal sector as a percentage of official GDP estimated using the model of Schneider, Buehn, and Montenegro (2010)	160	1993-2015	Balanced panel
DGE-based informal output	Indirect estimates (DGE)	Author's calculations	Size of the informal economy as a percentage of official GDP estimated using the approach of Elgin and Oztunali (2012).	158	1950-2016	Balanced panel
Share of self-employment	Labor force survey and household survey	WDI, ILO, OECD, National Statistical Offices, and LSMS.	The share of self-employment in total employment.	180	1955-2016	Unbalanced panel
Share of informal employment	Labor force survey and household survey	ILO	The share of informal employment in percent of total employment	53	2001-2016	Repeated cross-sections
Share of employment outside the formal sector	Labor force survey and household survey	ILO	The share of employment outside the formal sector in percent of total employment	57	1999-2016	Unbalanced panel
Pension coverage	Labor force survey and household survey	WDI	The fraction of the labor force that contributes to a retirement pension scheme	135	1990-2010	Unbalanced panel
World Bank Enterprises Survey	Firm survey	World Bank Enterprise Survey	Four measures on informality: percent of firms competing against unregistered or informal firms, percent of firms formally registered when they started operations in the country, (average) number of years a firm operated without formal registration, and percent of firms identifying practices of competitors in the informal sector as a major constraint.	139	2006-2016	Repeated cross-sections
World Economic Forum (Executive Opinion Survey)	Firm survey	World Economic Forum	The extent of informal economy based on the question: "In your country, how much economic activity do you estimate to be undeclared or unregistered? (1 = Most economic activity is undeclared or unregistered; 7 = Most economic activity is declared or registered)"	151	2004-2016 (2004-05 unused)	Balanced panel dataset
Informal Market Index (Heritage Foundation)	Firm survey	Heritage Foundation	The subjective perceptions of general compliance with the law, with particular emphasis on the role played by official corruption. The index ranges from 1 to 5 with higher values indicating more informal market activity.	165	1995-2005	Balanced panel dataset
Non-observed activities (% of GDP)	National account approach	UN (2008)	Non-observed activities (% of GDP)	44	Various years	Cross-sectional data
The Eurobarometer Survey: Indirect measure of the informal economy	Household surveys and social opinion surveys	The Eurobarometer Survey	Interviewers ask respondents whether he/she has purchased goods or serviced embodied undeclared work or supplied labor in the informal economy. The survey also includes information on whether the respondents receive all or part of their regular salary or the remuneration for extra work or overtime hours as cash-in-hand and without declaring it to tax or social security authorities. Finally, the survey shows whether respondents find informal economic activities acceptable.	27	2007 and 2013	Repeated cross-sections
World Values Survey: Tax morale	Household surveys and social opinion surveys	World Values Survey	Interviewers ask whether respondents can justify cheating on taxes. Detailed descriptions are reported in Table A3.	94	81-84, 94-98, 99-04, 05-09, 10-14	Repeated cross-sections
European Values Survey: Tax morale	Household surveys and social opinion surveys	European Values Survey	Interviewers ask whether it is justifiable for the respondents or their compatriots to cheat on taxes or pay cash to avoid taxes.	16-47	1981, 1990, 1999, and 2008	Repeated cross-sections
European Social Survey: Indirect measure of the informal economy	Household surveys and social opinion surveys	European Social Survey	Interviewers ask whether respondents paid cash for goods or services with no receipt so as to avoid VAT or taxes over the past five years and whether respondents have a written employment contract.	24	Every two years from 2004-2014	Repeated cross-sections

Table A2. MIMIC Model Estimation Results (1993-2015)

	[1]	[2]	[3]	[4]	[5]
	88 Developing Countries	**98 Developing Countries**	**120 Countries**	**151 Countries**	**161 Countries**
Size of government	0.133***	0.143***	0.157***	0.152***	0.145***
	(0.023)	(0.021)	(0.024)	(0.019)	(0.019)
Share of direct taxation	0.035		0.009		
	(0.023)		(0.022)		
Business Freedom	0.035	0.040**	0.058**		
	(0.021)	(0.020)	(0.024)		
Fiscal Freedom	0.002	-0.010	-0.038		
	(0.023)	(0.020)	(0.025)		
Unemployment rate	0.078***	0.105***	0.055**	0.067***	0.066***
	(0.023)	(0.021)	(0.022)	(0.019)	(0.019)
GDP per capita	-0.342***	-0.324***	-0.393***	-0.381***	-0.385***
	(0.035)	(0.027)	(0.029)	(0.022)	(0.022)
Government effectiveness			-0.069***	-0.043**	-0.042**
			(0.020)	(0.018)	(0.018)
Growth rate of GDP per capita	-0.835***	-0.618***	-0.362***	-0.310***	-0.306***
	(0.119)	(0.085)	(0.079)	(0.064)	(0.064)
Labor force participation rate	-0.321***	-0.219***		-0.167***	-0.155***
	(0.091)	(0.073)		(0.053)	(0.052)
Growth rate of labor force			-0.091		
			(0.064)		
Currency (M0/M1)	1.000	1.000	1.000	1.000	1.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Statistical tests					
RMSEA	0.061	0.057	0.070	0.087	0.089
p(RMSEA<=0.05)	0.097	0.190	0.002	0.000	0.000
Chi^2 (p)	63.922	60.646	124.517	153.29	160.63
	(0.00)	(0.000)	(0.000)	(0.000)	(0.000)
AIC	27388.448	33527.217	41436.305	43231.405	44080.904
BIC	27464.278	33602.241	41522.616	43306.446	44156.205
CFI	0.820	0.852	0.761	0.771	0.764
TLI	0.685	0.734	0.590	0.571	0.558
SRMR	0.033	0.030	0.041	0.046	0.047
CD	0.846	1	1	1	1
Observations	1,159	1,570	1,627	2,374	2,422

Note: Absolute z-statistics in parentheses. ***, **, * denote significance at the 1, 5, and 10% significance levels. All variables are used as their standardized deviations from the mean. Data sources for variables used in the model are listed in Section II footnote 6. Following the MIMIC models' identification rule, the currency (M0/M1) variable is fixed to an a priori value. The currency variable shows the level of money(cash) in circulation. "AIC" stands for "Akaike's information criterion" and "BIC" stands for "Bayesian information criterion. "RMSEA" stands for "Root Mean Square Error of Approximation". "TLI" stands for "Tucker Lewis Index", "CFI" stands for "Comparative Fit Index", "SRMR" stands for "Standardized Root Mean Square Residual", and "CD" shows the coefficient of determination. These are goodness-of-fit statistics.

Table A3. World Value Survey, European Value Surveys and European Social Surveys

Survey	Coverage
World Value Survey (WVS)	**Questions: "Justifiable: cheating on taxes". 1 is "never justifiable" and 10 is "always justifiable"**
WVS 1981-1984	9 countries/regions: Argentina, Australia, Finland, Japan, South Korea, Mexico, South Africa, Sweden, and United States
WVS 1989-1993	16 countries/regions: Argentina, Brazil, Belarus, Chile, China, India, Japan, South Korea, Mexico, Nigeria, Poland, Russia, South Africa, Spain, Switzerland, Turkey.
WVS 1994-1999	52 countries/regions: Albania, Azerbaijan, Argentina, Australia, Bangladesh, Armenia, Bulgaria, Belarus, Chile, China, Taiwan, Colombia, Croatia, Czech Rep., Dominican Rep. , El Salvador, Estonia, Finland, Georgia, Hungary, India, Japan, South Korea, Latvia, Lithuania, Mexico, Moldova, Montenegro, New Zealand, Nigeria, Norway, Peru, Philippines, Poland, Puerto Rico, Romania, Russia, Serbia, Slovakia, Slovenia, South Africa, Spain, Sweden, Switzerland, Ukraine, Macedonia, United States, Uruguay, Venezuela, West Germany, East Germany, Bosnia.
WVS 2000-2004	37 countries/regions: Albania, Algeria, Argentina, Bangladesh, Bosnia, Canada, Chile, China, India, Indonesia, Iran, Japan, Jordan, South Korea, Kyrgyzstan, Mexico, Moldova, Montenegro, Morocco, Nigeria, Pakistan, Peru, Philippines, Puerto Rico, Serbia, Singapore, Viet Nam, South Africa, Zimbabwe, Spain, Uganda, Macedonia, Egypt, Tanzania, United States, Venezuela.
WVS 2005-2009	56 countries/regions: Andorra, Argentina, Australia, Brazil, Bulgaria, Canada, Chile, China, Taiwan, Colombia, Cyprus, Ethiopia, Finland, France, Georgia, Germany, Ghana, Guatemala, Hong, Kong, Hungary, India, Indonesia, Iran, Italy, Japan, Jordan, South, Korea, Malaysia, Mali, Mexico, Moldova, Morocco, Netherlands, New, Zealand, Norway, Poland, Romania, Russia, Rwanda, Viet, Nam, Slovenia, South, Africa, Spain, Sweden, Switzerland, Thailand, Trinidad, and, Tobago, Turkey, Ukraine, Egypt, United, Kingdom, United, States, Burkina, Faso, Uruguay, Serbia, and, Montenegro, Zambia.
WVS 2010-2014	57 countries/regions: Algeria, Azerbaijan, Argentina, Australia, Armenia, Brazil, Belarus, Chile, China, Taiwan, Colombia, Cyprus, Ecuador, Estonia, Georgia, Palestine, Ghana, Hong, Kong, India, Iraq, Japan, Kazakhstan, Jordan, South, Korea, Kuwait, Kyrgyzstan, Lebanon, Libya, Malaysia, Mexico, Morocco, Netherlands, New, Zealand, Nigeria, Pakistan, Peru, Philippines, Poland, Romania, Russia, Rwanda, Singapore, Slovenia, South, Africa, Zimbabwe, Spain, Sweden, Thailand, Trinidad, and, Tobago, Tunisia, Turkey, Ukraine, Egypt, United, States, Uruguay, Uzbekistan, Yemen.