

CHAPTER 4

Lagging Behind: Informality and Development

Informality is associated with a wide range of development challenges. Emerging market and developing economies (EMDEs) with greater informality have significantly lower per capita incomes and greater poverty, less developed financial markets, weaker governance and public service provision, poorer human development outcomes, and more limited access to public infrastructure. This wide range of correlates suggests that any policies to address informality need to be embedded in a broader development agenda.

Introduction

Informal activity is widespread in emerging market and developing economies (EMDEs). While informality is often considered a cause of development challenges, informality itself is also a consequence of under-development.¹ EMDEs with more pervasive informality tend to be less developed; rely more on labor-intensive activities that employ unskilled and poorly paid workers; and have limited fiscal resources (World Bank 2019). Life expectancy, maternal mortality and other human-development indicators are, on average, lagging behind in EMDEs with more pervasive informality. Access to public services, such as electricity provision, that are essential to economic development, is limited.

A large informal sector weakens policy effectiveness and the government's ability to generate fiscal revenues.² Government revenues in EMDEs with above-median informality are 5-12 percentage points of GDP below those with below-median informality (World Bank 2019). Limited fiscal resources constrain governments' ability to offer adequate coverage of social protection programs, provide broad access to public sector services, smooth business cycles, and close the productivity gap between the formal and informal sectors (Schneider, Buehn, and Montenegro 2010; World Bank 2020a). In turn, the limited access to public services further discourages firms and workers from engaging with the government, resulting in more participation in the informal sector (Loayza 2018; Perry et al. 2007).

EMDEs with widespread informality score particularly poorly on indicators of development. Many development outcomes are captured and quantified in measures of progress toward the Sustainable Development Goals (SDGs). There are 17 SDGs, each

Note: This chapter was prepared by Franziska Ohnsorge, Yoki Okawa, and Shu Yu. Research assistance was provided by Lorez Qehaja, Arika Kayastha and Jinxin Wu.

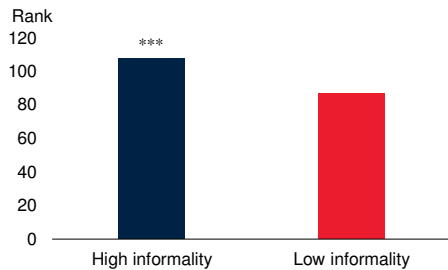
¹ See, for instance, Fields (1975); Harris and Todaro (1970); Loayza (2016); and Ulysees (2020).

² See Dabla-Norris, Gradstein, and Inchauste (2008); Ordóñez (2014); Joshi, Prichard, and Heady (2014); and World Bank (2019).

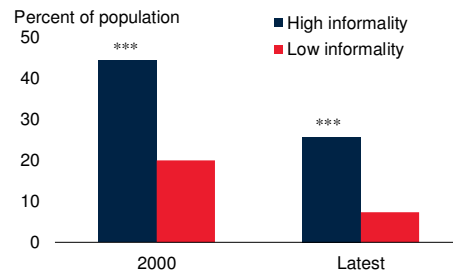
FIGURE 4.1 Development challenges and informality

EMDEs with more pervasive informality face severe development challenges, ranging from extreme poverty to lack of public infrastructure, and lag behind in progress toward the Sustainable Development Goals (SDGs).

A. SDG Global Index rank



B. Extreme poverty headcount



Sources: Sachs et al. (2020); World Bank (World Development Indicators).

Note: "High informality" ("Low informality") are EMDEs with an above-median (below-median) DGE-based informal output measure over the period 1990-2018. EMDEs = emerging market and developing economies; DGE = dynamic general equilibrium model-based estimates in percent of official GDP. Results are robust to regressions using quartile dummies (table 4.14). Based on 132 EMDEs (A) or 155 EMDEs (B). *** indicates that group differences are significant at the 10 percent level.

A. Bars show group averages for the latest year available (that is, 2020). SDG global index rank provides the economy's rank regarding SDG achievement, with a high value suggesting lack of SDG achievement.

B. Bars show group averages for EMDEs with "high informality" and "low informality" in 2000 and the latest year available. Poverty headcount ratio at \$1.90 a day is the percentage of the population living on less than \$1.90 a day at 2011 international prices.

with multiple underlying targets and associated data indicators, to be achieved by 2030. They add specificity to the broad objectives of ending poverty, protecting the planet, and ensuring shared prosperity (Vorisek and Yu 2020). They were adopted in 2015 as a key component of the 2030 Agenda for Sustainable Development. They include human-development-related goals, such as an end to poverty, zero hunger, reduced inequality, high-quality education and health-care services, and also infrastructure-related goals, such as access to clean water, sanitation, and affordable clean energy. Despite recent improvements, progress toward the SDGs has been uneven.

In 2020, EMDEs with above-median informality, on average, ranked around 110 out of 166 in overall SDG achievement, which is significantly worse than EMDEs with below-median informality (figure 4.1).³ In 2018, 26 percent of the population of EMDEs with above-median informality lived in extreme poverty, much more than the 7 percent of the population in EMDEs with below-median informality. In countries with greater informality, income inequality was higher, in part reflecting the wage gap between formal and informal workers and less progressive tax policies (box 4.1; Chong and Gradstein 2007; World Bank 2019).

³Unless otherwise specified, informality refers to dynamic general equilibrium (DGE)-based estimates of informal output in percent of official GDP. Results pertaining to employment informality, proxied by self-employment in percent of total employment, are shown in table 4.12 and table 4.15.

BOX 4.1 Informality and wage inequality

An extensive literature has documented the wage penalty for workers in the informal economy compared to their peers in the formal economy. Estimates of this penalty, however, vary substantially across countries. A comprehensive review of the relevant empirical studies suggests that the wage penalty largely reflects the characteristics of workers who self-select into informal activities.

Worldwide, two billion people, or more than 60 percent of all workers aged 15 and over, have informal jobs (ILO 2018a). Informality is often associated with lower wages than the formal sector. If these lower wages reflect a wage penalty for informality that is independent of worker characteristics, policies that encouraged the movement of workers to the formal sector might be a powerful remedy for income inequality and poverty. If, however, the wage differential largely reflects the characteristics of the workers employed in the informal sector, moving workers to the formal sector would be unlikely to achieve such gains.

Persistent informality frequently overlaps with poverty as many working poor remain employed in the informal sector (ILO 2018b; box 4.3). Lower wages in the informal sector can result from different worker characteristics, possibly reflecting comparative advantage, or reflect non-wage benefits that might accrue to work in the informal sector (Heckman and Li 2003; Maloney 2004). Wage differences may also reflect subjective well-being or job satisfaction of workers in the formal and informal sectors, with informal jobs offering more flexibility and independence (e.g., Blanchflower, Oswald, and Stutzer 2001; Falco et al. 2011; Sanfey and Teksoz 2007). Alternatively, wage differentials could stem from rigidities and other factors that create a wedge in wages between similar workers in informal and formal employment (Harris and Todaro 1970).

This box sifts through a large body of empirical evidence on informal wage differentials to explore the following questions:

- What factors can create wage differentials between formal and informal sectors?
- How large is the wage gap between formal and informal jobs?
- What accounts for the wage gap between formal and informal jobs?

A comprehensive review of empirical models, identification strategies, estimation methods, and data sources delivers mixed results. Some studies detect a substantial formal wage premium over informal employment; other estimates, however, do not find a significant wage gap after controlling for individual and firm-specific characteristics. In light of these different findings, a Meta-Regression

Note: This box was prepared by Sergiy Kasyanenko.

BOX 4.1 Informality and wage inequality (*continued*)

Analysis (MRA) is used to aggregate multi-study estimates of the formal wage premium and obtain a quantitative assessment of the sources of cross-study variation in research outcomes.

Causes of wage differentials

Wage differentials between formal and informal employment reflect a confluence of worker-, job- and country-specific characteristics. Broadly speaking, differentials can reflect inefficiencies caused by labor market frictions or self-selection of workers into their most productive employment—with diametrically opposed policy implications.

Segmented labor markets. Lower informal wages may result from workers being rationed out of better-paying formal jobs. For example, García and Badillo (2018) find that formal job rationing may affect over 60 percent of the workers in the informal sector of Colombia. Rigidities in the formal job market induced by, for example, labor regulations, unions, tax laws and labor regulations or efficiency wages may restrict competitive access to formal jobs. This creates a wedge in wages between formal and informal employment for workers of equal productivity (Harris and Todaro 1970). Formal wage premiums may also reflect better job matches in formal activities, particularly in denser and larger urban areas (Matano, Obaco, and Royuela 2020).

Self-selection into informal employment. A wage differential can also arise because of worker preferences. Workers may self-select into informal employment, either because of desirable non-wage benefits or amenities attached to informal jobs or because they have a comparative advantage—that is typically unobserved in research studies—in informal sector activities (Heckman and Li 2003; Maloney 2004). A worker may stay at a lower-paying informal job simply because the opportunity cost of foregone wages in the formal sector does not offset non-monetary benefits of informal employment, such as greater autonomy and more flexible working hours relative to a formal, salaried job.

Characteristics of informal workers. In EMDEs, self-employed workers constitute the core of informal employment; they typically lack registration at the national level, do not contribute to social security, and are not entitled to paid annual or sick leave.^a However, not all informal workers are self-employed, and the informal sector itself may be divided into tiers such as informal self-employed entrepreneurs or professional workers and informal non-professional employees.^b In EMDEs, about half of informal workers are non-professional self-employed

a. According to ILO (2018b), nearly 90 percent of all own-account workers—the largest component of self-employed in EMDEs—are in the informal sector, accounting for over 45 percent of all informal jobs.

b. See Cunningham and Maloney (2001); Fields (1990, 2005); and Günther and Launov (2006).

BOX 4.1 Informality and wage inequality (*continued*)

workers—who migrate to formal employment as per capita incomes grow—and the majority of the remainder are informal employees (Gindling, Mossaad and Newhouse 2016). Depending on the restrictiveness of regulations and the quality of education systems, the composition of informal employment varies across countries and EMDE regions. For example, contributing family members (predominantly women) and the self-employed are the majority of informal workers in developing Asia and Africa, where public education systems can be rudimentary, whereas informal employees and employers dominate the informal sector in Europe and Central Asia and in Latin America where tax and business regulations can be burdensome (ILO 2018b).

Interpreting the literature

Selection of studies. The representative sample of studies on informality and wage inequality used here follows the selection guidelines in Stanley et al. (2013) and is broadly similar to criteria applied by van der Sluis, van Praag, and Vijverberg (2005). An initial search was conducted in the major English language repositories of academic articles and working papers.^c A study was included in the database if it: (i) provided a quantitative estimate of the informal-formal wage gap and a corresponding standard error or t-statistic; (ii) used data from micro-level household or labor surveys to obtain these estimates; (iii) analyzed an EMDE or group of EMDEs as defined by the World Bank classification and (iv) was published no earlier than 1990.^d The resulting database included 18 studies with a total of 83 individual coefficient estimates covering 20 EMDEs (annex 4A, table 4.1).

Definitions matter. Differences in estimates of the incidence of informal employment and the wage differentials between formal and informal workers in part reflect differences in data coverage and definitions of informal workers.^e Studies typically find that self-employed informal workers earn the same or more than formal workers, but employed informal workers earn less than formal workers (figure B4.1.1; Abraham 2019).^f In EMDEs with more restrictive business regulations, however, self-employment may be associated with a higher wage penalty because less favorable business climates may deter more productive

c. Covered online databases include EBSCO, EconLit, Google Scholar, JSTOR, International Monetary Fund Working Paper series, IZA Working Papers, the National Bureau of Economic Research (NBER), RePEc, Social Science Research Network (SSRN), and World Bank Policy Research Working Paper series.

d. Prior to 1990, reliable and comparable individual or household level survey data, which is used to estimate wage gaps between the formal and informal sectors, are very limited for developing countries.

e. See Hussmanns (2004); ILO (2013); and Perry et al. (2007).

f. See Arias and Khamis (2008); Kahyalar et al. (2018); Lehmann and Pignatti (2007); Maloney (1999); and Nguyen, Nordman, and Roubaud (2013).

BOX 4.1 Informality and wage inequality (*continued*)

workers, particularly women, from transitioning to non-professional self-employment (Gindling, Mossaad, and Newhouse 2020).

Methodology matters. Empirical research on the wage differential between informal and formal workers has largely relied on estimating “Mincerian” wage regressions conditional on the observed characteristics of workers, although more recent studies have used quantile regressions to assess sector wage gaps along the wage distribution.^g Such cross-sectional wage regressions are biased when workers’ unobserved characteristics affect both their occupational choice and wages. For example, several studies find workers transitioning from the formal sector into the informal sector after spending several years accumulating experience and knowledge in the formal sector (Gong, van Soest, and Villagomez 2004; Maloney 2004). Hence, studies that rely on panel data to control for time-invariant unobserved worker characteristics find smaller informal-formal wage differentials (Botelho and Ponczek 2011; Cho and Cho 2011; El Badaoui, Strobl, and Walsh 2008; Tansel, Keskin, and Ozdemir 2020). Similarly, semiparametric matching models, such as propensity score matching and difference-in-difference estimators that are immune to the misspecification of the wage regressions, find modest or insignificant wage differentials between formal and informal jobs (Pratap and Quintin 2006).

Empirical estimates of wage differentials

Wage differentials. The estimates of the wage differential between informal and formal workers in the 18 studies selected here range from a formal sector wage penalty of 50 percent in Tajikistan (Huber and Rahimov 2014) to a formal sector wage premium of 113 percent in South Africa (El Badaoui, Strobl, and Walsh 2008) with a median formal wage premium of about 18 percent.^h On average, the random-effects meta-analysis estimate of the wage premium in the formal sector obtained from pooling all studies is 19 percent (figure B4.1.1 and table 4.2).ⁱ

Explaining wage differentials. The wage premium largely disappears in studies using worker fixed effects, which are controlling for unobserved characteristics of

g. See Bargain and Kwenda (2014); Lehmann and Zaiceva (2013); and Tansel and Kan (2012).

h. A formal-sector wage premium indicates that formal-sector wages exceed those in the informal sector, whereas a formal sector wage penalty indicates formal-sector wages are below informal-sector wages. Huber and Rahimov (2014) attribute a large formal wage penalty in Tajikistan to self-selection and find no evidence of labor market segmentation. Meanwhile, El Badaoui, Strobl, and Walsh (2008) find that a formal wage premium in South Africa disappears once they controlled for unobserved worker characteristics.

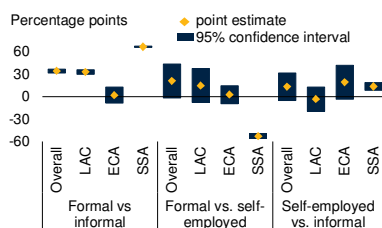
i. See annex 4A for technical details.

BOX 4.1 Informality and wage inequality (continued)

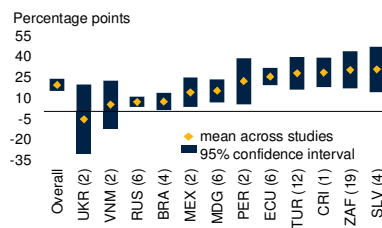
FIGURE B4.1.1 EMDEs: Estimates of informal-formal wage gaps

Estimates of informal-formal wage gaps vary considerably across countries and definitions of informality. Countries in Latin America and the Caribbean and Sub-Saharan Africa tend to exhibit both a higher incidence of informality and a larger wage premium in the formal sector.

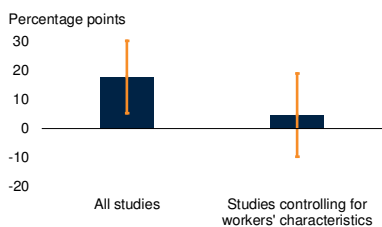
A. EMDEs: Informal-formal wage gaps



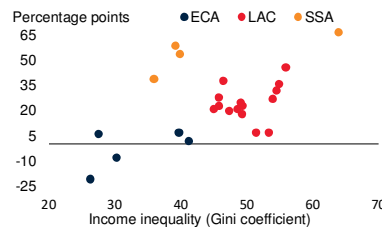
B. Informal-formal wage gaps: meta-analysis



C. Informal-formal wage gaps



D. Informal-formal wage gap and income inequality by EMDE regions



Sources: Gindling, Mossaad, and Newhouse (2016); World Bank.

Note: A positive wage gap indicates a penalty for working informally—a lower wage for informal workers than for comparable formal workers. A negative wage gap indicates a premium for working informally—a higher wage for informal workers than for comparable formal workers. EAP = East Asia and Pacific; ECA = Europe and Central Asia; LAC = Latin America and the Caribbean; MNA = Middle East and North Africa; SAR = South Asia; SSA = Sub-Saharan Africa. Wage gap between wage employees in the informal and formal sectors is displayed on the vertical axis.

A. Formal vs. informal = a wage gap between wage employees in the formal and informal sectors; Formal vs. self-employed = a wage gap between workers with formal jobs and self-employed workers; Self-employed vs. informal = a wage gap between self-employed workers and wage employees in the informal sector.

B. BRA = Brazil; CRI = Costa Rica; ECU = Ecuador; MDG = Madagascar; MEX = Mexico; PER = Peru; RUS = Russian Federation; SLV = El Salvador; TUR = Turkey; UKR = Ukraine; VNM = Vietnam; ZAF = South Africa. The number of studies or estimates for each economy is shown in parenthesis; economy means are calculated using a random-effects meta-analysis model.

C. The wage premium (shown in bars) is obtained from 18 empirical studies of the wage gap between formal and informal workers. The whiskers show the 90 percent confidence intervals. See box 4.1 for details.

D. Income inequality is measured as Gini coefficient provided by the World Bank's World Development Indicators.

BOX 4.1 Informality and wage inequality (*continued*)

workers.^j It turns into a statistically insignificant 8 percent penalty in studies that compared wages of self-employed informal workers with formal-sector employees.^k Informal employment tends to be associated with low levels of education and a U-shaped age profile and is more prevalent in rural areas, where there are fewer job alternatives in the formal sector (Gasparini and Tornarolli 2007; Hazans 2011).^l In general, low productivity attributes of workers in the informal sector may limit their earnings potential in comparable formal jobs.

Country characteristics. Even after controlling for study and sample-specific attributes, most of the cross-study variation in the estimates of the wage differential remains unexplained. That said, wage premia tend to be higher where informality is more widespread (figure B4.1.1). Differentials are particularly wide in Latin America and the Caribbean and Sub-Saharan Africa but below-average in Europe and Central Asia and South Asia (figure B4.1.1). Overall, the data do not offer strong evidence of a significant relationship between the size of the formal sector wage premium and the level of development or the quality of institutions.

Conclusion

Despite years of declining poverty, many working poor remain employed in the informal sector where they face significantly lower wages than workers in the formal economy. Estimates of the formal sector wage premium vary widely but, in the meta-analysis of the 18 studies conducted here, amount to just under 20 percent of informal wages. However, among studies controlling for worker characteristics, there is no statistically significant evidence of a formal sector wage premium.

This suggests that any formal-informal wage differential is largely a reflection of the characteristics of the types of workers who self-select into informal and formal employment. Workers in the informal sector tend to be less-skilled, younger or older, and more agricultural than workers in the formal sector. This points to the need for policies to lift these workers' human capital and, thus, allow them to switch into productive formal employment.

j. In the regression in annex 4A, the informal-formal wage gap for studies using fixed worker effects is estimated as the sum of the coefficients for μ and *FE*, tested for joint significance: the test statistic $F(1,76) = 0.41$ indicates that the null hypothesis of a zero sum of two coefficients cannot be rejected at any conventional significance level.

k. In the regression in annex 4A, the wage premium for self-employed is estimated as the sum of the coefficients for μ and *Self-employed*, tested for joint significance: the test statistic $F(1,76) = 0.33$ indicates that the null hypothesis of a zero sum of two coefficients cannot be rejected at any conventional significance level.

l. Younger and older workers are typically less productive, while older retired workers may be choosing informal employment to supplement their social security benefits.

It would take substantial additional financial resources to meet the SDGs by 2030, even if they were accompanied by big strides in policy improvements (UN SDSN 2019; Vorisek and Yu 2020). The World Bank estimates that low- and middle-income countries face additional investment needs of \$1.5-2.7 trillion per year between 2015 and 2030 to meet infrastructure-related SDGs alone (Rozenberg and Fay 2019). The IMF estimates that additional spending of about \$1.3 trillion per year during 2019–30 is required to make meaningful progress toward infrastructure-related SDGs in EMDEs, and another \$1.3 trillion for human-development-related SDGs (Gaspar et al. 2019). The UN estimates that an additional \$400 billion per year is needed in lower-income developing countries between 2019 and 2030, mainly for social protection, health, education, and climate change mitigation and adaptation (UN SDSN 2019).

Against this backdrop, this chapter addresses the following questions:

- What are the development challenges associated with the informal economy?
- What are the correlates of widespread informality?
- What are the correlates of changes in the informal sector over time?

Contributions. The chapter makes the following contributions to the literature on informality. First, it provides a systematic and comprehensive overview of developmental challenges facing countries with large informal sectors, highlighting their association with a wide range of development weaknesses and shortfalls from the SDGs.⁴ Previous studies have focused on the economic or institutional correlates of informality—such as per capita income (for instance, La Porta and Shleifer 2014; Loayza, Servén, and Sugawara 2010) or control of corruption (for instance, Choi and Thum 2005; Dreher and Schneider 2010)—and largely disregarded the linkages between informality and other aspects of sustainable development, ranging from life expectancy to lack of access to public infrastructure.

Second, the chapter is the first published study to empirically and systematically examine a broad range of correlates of informality in a large group of EMDEs, numbering around 130. Previous studies have tended to focus on one dimension of informality, rely on a more limited range of correlates, and only examine the correlates of cross-country differences in informality without focusing on EMDEs. To identify the robust correlates of informality, the chapter is also the first to use a Bayesian Model Averaging (BMA) approach, which is designed to take account of model uncertainty (Fernandez, Ley, and Steel 2001).

Third, in three boxes, this chapter illustrates how informality can pose developmental challenges to EMDEs. Box 4.1 conducts the first extensive meta-analysis of studies that

⁴ See Medina and Schneider (2019); Oviedo, Thomas, and Karakurum-Özdemir (2009); and Schneider, Buehn, and Montenegro (2010).

documented wage differences for workers in formal and informal sectors. Box 4.2 utilizes a unique firm-level dataset to show how the productivity gap between formal and informal firms in EMDEs can be narrowed by improvements in business climates.⁵ Box 4.3 empirically tests whether there is a strong relationship between declines in informality and poverty reduction (or income inequality).

Main Findings. The chapter demonstrates that EMDEs with pervasive informality face a wider range of greater development challenges than other EMDEs. First, informality is associated with poor economic outcomes. Countries with larger informal sectors have lower per capita incomes, greater poverty, less financial development, and weaker growth in output, investment and productivity. Informal firms are less productive than their formal counterparts (box 4.2).

Second, more pervasive informality is associated with significantly lower government revenues and expenditures, less effective policy institutions, more burdensome tax and regulatory regimes, and weaker governance. Weaknesses in governance and revenue collection constrain the provision of public services in EMDEs with more pervasive informality, contributing to poorer human development outcomes. People living in EMDEs with more widespread informality suffer from a greater prevalence of hunger, poorer health and education, and greater gender inequality. Countries with more widespread informality offer poorer access to, and lower-quality, infrastructure.

Third, the results from the BMA approach suggest that economic development, human capital, and governance are particularly robust correlates of informality. That said, other correlates such as infrastructure, for instance, are also relevant.

Fourth, while informality is linked to a host of developmental challenges, formalization alone is unlikely to offer an effective path out of under-development. For instance, while declines in informality were associated with poverty reduction, they were not systematically linked to declining income inequality (box 4.3). This may reflect the fact that informality itself is a symptom of underdevelopment, in line with the meta-analysis of the literature that finds that the wage penalty largely reflects the characteristics of informal workers (box 4.1).

The following section summarizes the transmission mechanisms underlying the linkage between informality and development challenges. Here informality is regarded as both a cause and a consequence of underdevelopment: there are reasons to expect causation potentially to run in both directions. The subsequent sections examine the economic and institutional correlates of informality, followed by sections that describe the link between informality and various SDGs. The penultimate section summarizes the finding of the Bayesian Model Averaging approach, followed by a conclusion in the final section.

⁵ Existing studies, such as Meghir, Narita, and Robin (2015) and Ulyssea (2018), explore the productivity gap between formal and informal firms in individual countries.

Links between informality and development challenges

EMDEs with widespread informality face relatively large development challenges. Informality may be linked with these challenges through several channels. For the purposes of the discussion here, informality refers to output informality, but the results are robust to using employment informality.

Low productivity, low incomes. Informal workers tend to be less skilled and lower paid than their formal counterparts (Loayza 2018; Perry et al. 2007; World Bank 2019). A meta-analysis of worker-level empirical studies shows that informal workers are, on average, paid 19 percent less than formal workers (figure 4.2; World Bank 2019). In part, this reflects lower productivity on account of lower skill and experience levels than formal workers have. The meta-analysis suggests that, when controlling for worker characteristics, the wage gap is no longer statistically significant (box 4.1).⁶ In 2020, these features made participants in the informal sector particularly vulnerable during lockdowns associated with the COVID-19 pandemic (World Bank 2020a; box 2.1).

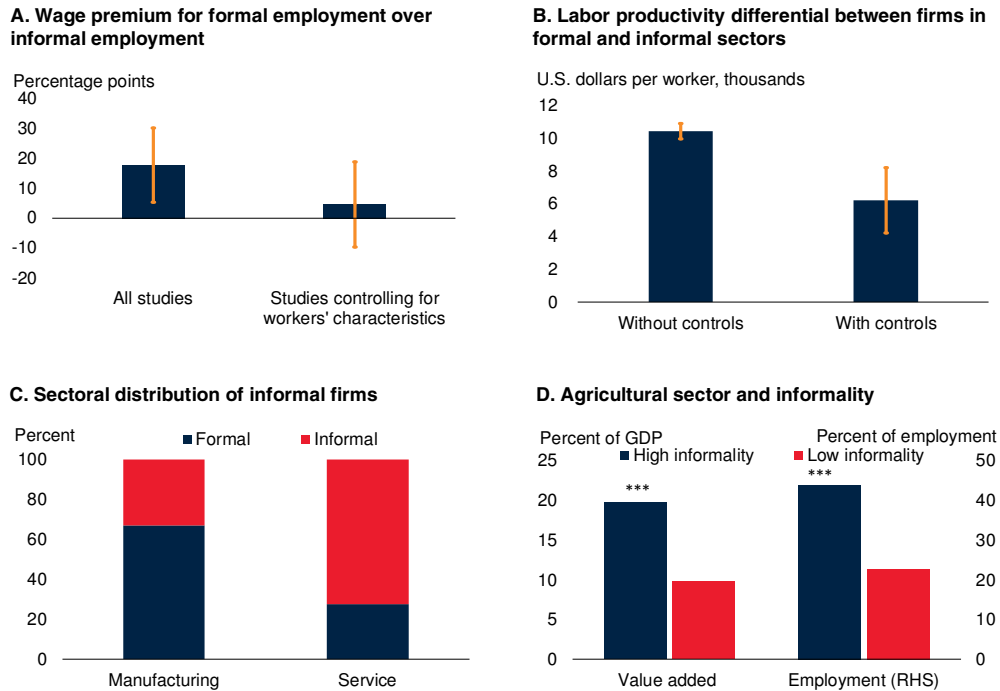
Similarly, informal firms tend to be small, lack funds, and operate in labor-intensive sectors, and, as a result, are less productive than formal firms (figure 4.2; Fajnzylber, Maloney, and Montes-Rojas 2011; Farazi 2014; McKenzie and Sakho 2010). They tend to invest less, possibly in an effort to avoid adopting technologies that would make them more visible to tax and other authorities (Dabla-Norris, Gradstein, and Inchauste 2008; Gandelman and Rasteletti 2017). For example, in about 11,600 firms that participated in Enterprise Surveys in 18 economies during 2007-2014, the fraction of formal firms that invested was significantly higher than that of informal firms. Low productivity in the informal sector can also cast a shadow over formal firms: a sizeable informal sector that competes with the formal sector for low-skilled workers reduces the incentives to invest in human and physical capital and new technologies and slows growth in the long run (box 4.2; Amin, Ohnsorge, and Okou 2019; Distinguin, Rugemintwari, and Tacneng 2016; Docquier, Müller and Naval 2017; Loayza 1996; Perry et al. 2007; Sarte 2000).

Sectoral distribution. Informal workers in EMDEs tend to be concentrated in the agricultural and services sectors (figure 4.2). Agricultural employment in EMDEs is about 90 percent informal. In EMDEs with above-median informality, the agricultural sector, on average, accounts for about 20 percent of GDP and for nearly 40 percent of total employment—almost twice as much as in EMDEs with below-median informality. In addition, informal firms tend to be smaller, less productive and concentrated in labor-intensive sectors, such as low-value-added services. In EMDE service sectors, about 72 percent of firms are informal, compared with 33 percent in manufacturing (World Bank

⁶This lower productivity may also account for the inability of the formal sector in cities to absorb rural migrants during the urbanization process (Fields 1975; Harris and Todaro 1970; Loayza 2016).

FIGURE 4.2 Features of informal firms and workers

Informal workers tend to be less well-paid and employed in the agricultural or services sectors. Informal firms are less productive than their formal-sector peers.



Sources: Amin, Ohnsorge, and Okou (2019); World Bank (Enterprise Surveys, World Development Indicators); World Bank.

A. The wage premium (shown in bars) is obtained from 18 empirical studies of the wage gap between formal and informal workers. The whiskers show the 90 percent confidence intervals. See box 4.1 for details.

B. Difference between labor productivity of formal and informal firms, without controlling for firm characteristics and with controlling for firm characteristics. Firm productivity is measured as sales per worker in 2009 U.S. dollars. Whiskers show the corresponding ± 2 standard errors.

C. See Amin, Ohnsorge, and Okou (2019) for data coverage.

D. Bars are group means from the latest year available calculated for EMDEs with "high informality" (that is, above-median DGE-based informal output in percent of GDP) and those with "low informality" (EMDEs with below-median DGE-based informal output measure) over the period 1990-2018. DGE stands for dynamic general equilibrium model-based estimates. EMDEs = emerging market and developing economies. *** indicates group differences are not zero at 10 percent significance level.

2020a).⁷ Large productivity gaps exist between manufacturing, agricultural, and service sectors in EMDEs (Dieppe 2020).

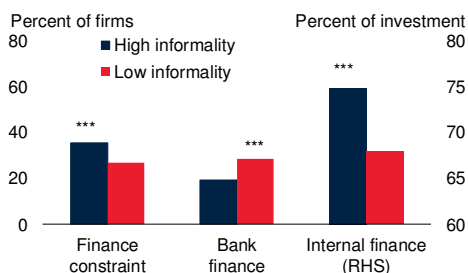
Access to resources. In EMDEs with widespread informality, access to finance and public resources is limited. Informal firms struggle to access conventional banks because of lack of property rights, lack of documentation for assets, and inadequate financial statements (Bose, Capasso, and Wurm 2012; de Soto 1989). In the World Bank's

⁷As these economies grow richer, households tend to shift away from agricultural and informal sector goods (Saracoğlu 2008).

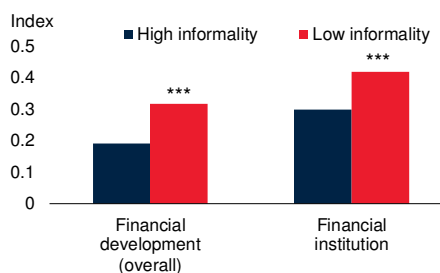
FIGURE 4.3 Access to finance and public services

In EMDEs with more pervasive informality, access to external finance and public services is more limited, constraining productivity and entrepreneurship.

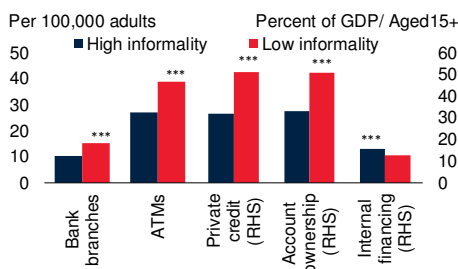
A. Firms' financing conditions and informality



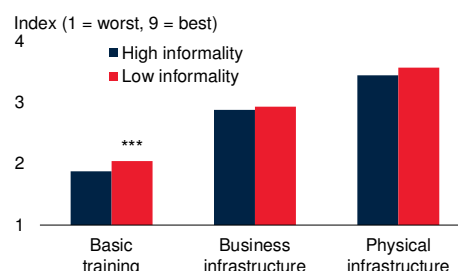
B. Financial development and informality



C. Household access to finance and informality



D. Entrepreneurial framework conditions and informality



Sources: International Monetary Fund (Financial Development Database); World Bank (Enterprise Surveys, Global Financial Development Database, World Development Indicators).

Note: "High informality" ("Low informality") are emerging market and developing economies (EMDEs) with above-median (below-median) dynamic general equilibrium (DGE)-based informal output measures. *** indicates group differences are not zero at 10 percent significance level.

A. Bars are group means for EMDEs over the period 2000-18. "Finance constraint" measures the percent of firms identifying access to finance as a major constraint. "Bank finance" measures the percent of firms using banks to finance investment. "Internal finance" measures the average proportion of investment financed internally.

B. Bars show simple averages for EMDEs over the period 2010-18. "Financial development (overall)" is the aggregate financial development index from the IMF's Global Financial Development Database. It purports to measure the overall level of financial development, in both "financial institutions" and "financial markets." The latter refers to the accessibility, depth, and efficiency of an economy's stock and debt markets, which is less relevant for informal participants in EMDEs. The "Institutions" index measures how developed financial institutions are in terms of their depth (size and liquidity), accessibility (ability of individuals and companies to access financial services), and efficiency (ability of institutions to provide financial services at low cost and with sustainable revenues).

C. Bars are unweighted group averages for EMDEs over the period 2010-18. "Bank branches" measures the number of commercial bank branches per 100,000 adults. "ATMs" measures the number of automated teller machines (ATMs) per 100,000 adults. "Private credit" measures domestic credit to private sector in percent of GDP. "Account ownership" is the percentage of survey respondents (aged 15 and above) who report having an account (by themselves or together with someone else) at a bank or other financial institution, or report personally using a mobile money service in the past 12 months. "Internal financing" refers to the percentage of respondents (aged 15 and above) who report saving or setting aside any money in the past 12 months to start, operate, or expand a farm or business.

D. Bars are group means for EMDEs over the period 2000-18. The score is based on the Global Entrepreneurship Monitor's National Expert Survey. Ranging from 1 to 9, a higher score represents better perceived conditions. Basic training = extent to which training in creating or managing small and medium-sized enterprises (SMEs) is incorporated within the education and training system at primary and secondary levels; Business infrastructure = presence of property rights that support or promote SMEs, including commercial, accounting, and other legal and assessment services and institutions; Physical infrastructure = ease of access to physical resources, communications, utilities, transportation, land, or space at a price that does not discriminate against SMEs. Outliers (Saudi Arabia and Kuwait), are not included.

BOX 4.2 Casting a shadow: Productivity in formal and informal firms

Labor productivity in the average informal firm in emerging market and developing economies (EMDEs) is only one-quarter of that of the average firm operating in the formal sector. Moreover, firms in the formal sector that face informal competition are, on average, only three-quarters as productive as those that do not. This suggests that competition from the informal sector can erode formal firms' market share and resources available to boost productivity as they shoulder the costs of regulatory compliance. More effective governance and stronger control of corruption can help mitigate these effects.

Introduction

The differential in labor productivity between formal and informal firms is well established in the literature (Loayza and Rigolini 2006; Oviedo 2009). However, there is mixed evidence on the impact of a large informal sector on formal firms' labor productivity.^a Some studies suggest that the informal and formal sectors operate independently so that there are no productivity spillovers (La Porta and Shleifer 2014). Others report that competition from the informal sector may erode the profitability of firms in the formal sector, limiting their resources to enhance firm productivity. The aggregate effect varies with country characteristics (Amin and Okou 2020).

Against this backdrop, this box documents the productivity gap between formal and informal firms and their interactions. Specifically, it addresses the following questions:

- How large is the differential in labor productivity between formal and informal firms?
- To what extent are formal firms exposed to informal competition?
- How does informal competition affect the labor productivity of formal firms?

Productivity differential between formal and informal firms

Literature review. The literature documents that informal firms in EMDEs are less productive than formal firms, with labor productivity gaps ranging between 30 and 216 percent (La Porta and Shleifer 2008; Perry et al. 2007). The productivity gap between informal and formal firms is attributed to more backward technologies in informal firms, their greater reliance on unskilled labor,

Note: This box was prepared by Mohammad Amin and Cedric Okou. The box closely follows box 3.3 in the January 2019 *Global Economic Prospects* report.

a. See Gonzalez and Lamanna (2007); Heredia et al. (2017); and Mendi and Costamagna (2017).

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

their more limited economies of scale, and their more restricted access to services, markets, and funding.^b Labor productivity has also been found to vary within the informal sector along different dimensions such as firm size and type of activity (Amin and Huang 2014; Amin and Islam 2015).

Methodology. In this box, the labor productivity gap between formal and informal firms is estimated using the World Bank’s Enterprise Surveys data collected over the period 2007-14 for a cross-section of 4,036 informal firms and 7,558 formal firms in 18 EMDEs. Formal firms are those that register with the relevant authorities; unregistered firms belong to the informal sector. To estimate the productivity gap, a measure of labor productivity—log annual sales in 2009 U.S. dollars per worker—is regressed on a dummy variable that takes the value 1 for informal firms and 0 otherwise and a set of control variables capturing additional firm characteristics (employment size, time in business, location, sector, economy).^c

Lower productivity in informal than formal firms. Virtually across the board, firm-level labor productivity is much lower in the informal sector than in the formal sector (table 4.3).^d The productivity differentials vary widely in this sample, from 48 (Côte-d’Ivoire) to 93 percent (Argentina). On average across the whole sample, labor productivity in informal firms is only one-quarter of labor productivity in formal firms.

Drivers of productivity gap between informal and formal firms. Firm size, age, location in the capital city and manager experience are associated with significantly larger productivity gaps between informal and formal sectors (figure B4.2.1; table 4.4). Formal firms appear to be better equipped to reap productivity benefits from large size, advanced age, and urban location than informal firms.

- *Firm age.* As firms grow older, they are either sufficiently productive to survive or they disappear (“selection effect”; Brandt, Van Biesebroeck, and Zhang 2012). In addition, learning from experience may have taught older firms lessons that deliver productivity gains (“learning effect”; Luttmer 2007). These effects appear to be much more pronounced among formal firms than among informal firms. As a result, the productivity differential

b. See Amaral and Quintin (2006); Galiani and Weinschelbaum (2012); and Jovanovic (1982).

c. Commonly used revenue-based measures of productivity may conflate efficiency and price effects. Disentangling these effects, by using physical productivity measures, may shed new light on productivity patterns, especially at the firm level (Cusolito and Maloney 2018; Jones and Nordhaus 2008).

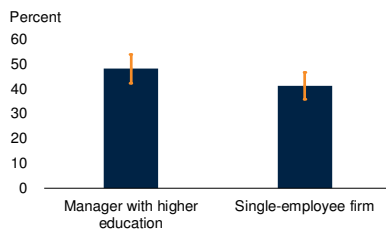
d. Exceptions are the Democratic Republic of Congo and Cabo Verde, possibly due to low productivity of formal firms.

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

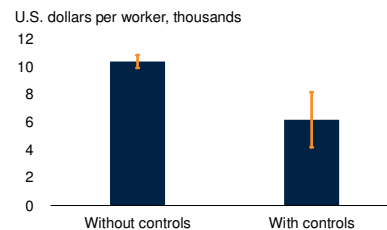
FIGURE B4.2.1 Labor productivity in informal firms

In EMDEs, labor productivity is significantly higher in informal firms that have managers with higher education and in those without any employees other than the owner. This labor productivity differential between formal and informal firms is particularly pronounced among larger and older firms that operate in the capital city and are led by experienced managers.

A. Labor productivity differential between different types of informal firms



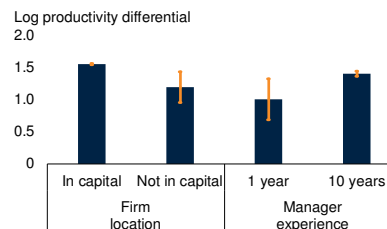
B. Average labor productivity differential between formal and informal firms



C. Labor productivity differential between formal and informal firms, by informal firm's age and size



D. Labor productivity differential between formal and informal firms, by informal firm's location and manager's experience



Source: World Bank.

Note: World Bank's Enterprise Surveys data for 135 countries (2008-2018). Labor productivity is proxied by annual sales per worker in 2009 U.S. dollars. Whiskers show the corresponding ± 2 standard errors.

A. Cross-country average of percent difference between labor productivity in the median informal firm with a manager with higher education or without any employees other than the owner, and the median informal firm with a manager without higher education or with more employees than the owner. Estimates from table 4.3.

B. Difference between productivity of formal and informal firms, with and without controlling for other firm characteristics. Labor productivity in the average formal and average informal firm, controlling for firm characteristics (firm size and age, manufacturing sector activity, location in the capital city, and country fixed effects) as shown in column (1) in table 4.4.

C.D. Difference in log of labor productivity between the average formal and average informal firm in each group, as estimated in coefficient estimates of table 4.4. "Not in capital" stands for "not located in capital city"; "In capital" stands for "located in capital city."

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

between formal and informal firms widens as the age of firms increases. Among one-year-old firms, informal firms have about half the labor productivity of formal firms. Among ten-year-old firms, informal firms have less than one-quarter the labor productivity of formal firms.

- *Firm size.* Larger firms can reap economies of scale that raise their labor productivity compared to smaller firms. Again, in this sample, this effect appears to be stronger among formal firms than among informal firms. Among firms with one employee, informal firms have just under one-third the labor productivity of formal firms; among firms with ten employees, informal firms have less than one-quarter the productivity of formal firms.
- *Firm location.* Capital cities are typically among countries' largest economic centers and so can offer agglomeration benefits, such as larger markets, better infrastructure to access markets and operate, a larger pool of workers, and greater technology spillovers (Duranton and Puga 2004; Rosenthal and Strange 2004). Again, formal firms appear to be better able to benefit from these locational advantages, though the difference is economically modest (although statistically significant). Among firms operating inside the capital city, informal firms' productivity is 31 percent that of similar formal firms; outside the capital city, informal firms' productivity is 30 percent that of similar formal firms.
- *Manager experience.* Managerial ability has been associated with higher labor productivity, through a variety of channels including hiring decisions and input choices (Fernandes 2008). Again, managerial experience appears to benefit formal firms' productivity more than informal firms' productivity. Among firms managed by managers with one year of experience, informal firms' labor productivity is just over one-third that of formal firms; among firms with managers with ten years of experience, informal firms' labor productivity is less than one-quarter that of formal firms.

Productivity differentials across informal firms. Labor productivity also differs across different types of informal firms, although the characteristics that are associated with higher labor productivity in informal firms differ across countries.^e In two-fifths of economies, informal firms having a manager with higher education or without any employees other than the owner are significantly more productive than other informal firms (column 1 in table 4.4). Other informal firm characteristics, such as operating in the services sector or being a

e. See Amin and Huang (2014); Amin and Islam (2015); de Mel, McKenzie, and Woodruff (2011); Deininger, Jin, and Sur (2007); Grimm et al. (2012); Haltiwanger, Lane, and Spletzer (1999); Islam (2018); and Maloney (2004).

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

start-up, are accompanied by higher productivity in some countries but lower productivity in others.

Robustness tests. Labor productivity differentials between formal and informal firms remain significant when each control variable is included separately. They are also robust to including interaction terms of informality with firm characteristics that make “hiding” revenue more difficult, such as firm size, exporter status, location in the capital city or having reported fixed asset purchases (Amin and Okou 2020).

Productivity of formal firms amid high informality

Impact of informal competition on formal firms. The extent of competition between formal and informal firms depends partly on the reasons for the existence of the informal firms (Amin and Okou 2020).^f

- *Informality as a survival strategy of unproductive firms.* Low-productivity firms may be forced into informal operations or, if they continue to operate formally, into employing informal workers because this may reduce their costs (Boly 2018; Ulyssea 2018). Operating in the informal sector or employing informal labor may, therefore, be a survival strategy for less-productive firms. “Surviving” informal firms are likely to operate in very different markets and sell different products than formal firms (La Porta and Shleifer 2014). Competition between informal and formal firms and its impact on formal firms may then be limited.
- *Informality as an evasion strategy of productive firms.* Some informal firms may be sufficiently productive to survive in the formal sector yet choose to remain informal to benefit from the cost advantages of noncompliance with (possibly excessive) taxes and regulations (de Mel, McKenzie, and Woodruff 2011; Maloney 2004).^g Such informal firms could have untapped potential for a productivity boost (de Soto 1989). On the other hand, they can create aggressive competition with formal firms that do shoulder the additional cost of tax and regulatory compliance. Credit constraints tend to be higher for formal firms in sectors that host many informal firms (Distinguin,

f. This discussion assumes that firms are either formal or informal. In practice, the degree of informality can vary (Perry et al. 2007; Ulyssea 2018). Firms can operate fully informally, both in product markets and labor markets (“extensive margin”). They sell their output informally and employ informal labor. Or firms can operate semi-formally (“intensive margin”): they sell their output into formal product markets but employ, in part, informal labor, as observed in EMDEs and low-income countries.

g. Such circumstances are likely to be associated with an environment of weak regulatory and tax enforcement (Benjamin and Mbaye 2012; Dabla-Norris, Gradstein, and Inchauste 2008; Quintin 2008; Ulyssea 2010).

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

Rugemintwari, and Tacneng 2016). Such informal competition can reduce the profitability necessary for formal firms to invest in productivity-enhancing new technologies or to innovate, especially in a context of weak property rights enforcement.^h Alternatively, this very competition could force formal firms to increase productivity or, for the lowest-productivity ones, to exit.ⁱ

Extent of informal-firm competition for formal firms. In the World Bank's nationally representative survey data for 75,137 formal (registered) firms in 135 economies between 2008 and 2018, about 55 percent of formal firms reported facing competition from informal firms.^j The share of informal firms competing against formal firms was about 60 percent in EMDEs, 13 percentage points higher than in advanced economies. The level of competition varied widely across economies, ranging from about 7 percent in Bhutan to 95 percent in Uganda. Smaller firms were significantly more likely to be exposed to informal competition than larger firms but there is little evidence of any other systematic difference between firms that were exposed and those that were not (figure B4.2.2).

Impact of informal competition on the productivity of formal firms

Methodology. Ordinary least squares regressions are used to estimate the difference in labor productivity between formal firms that compete against informal firms and those that do not. In the baseline specification, the dependent variable is again labor productivity measured by the (log of) annual sales per worker in 2009 U.S. dollars. The main explanatory variable is the informal competition indicator proxied by the proportion of other formal firms in a subgroup that report facing competition from informal firms. Such a subgroup is defined as a group of firms of similar size and in the same region and sector.^k

h. This has been documented for Latin America countries, India, Panama, Poland, Portugal, Russian Federation, and Turkey. For evidence, see Allen and Schipper (2016); Capp, Elstrodt, and Jones (2005); Distinguin, Rugemintwari, and Tacneng (2016); Farrell (2004); Friesen and Wacker (2013); Gonzalez and Lamanna (2007); Heredia et al. (2017); Iriyama, Kishore, Talukda (2016); Perry et al. (2007); and Vargas (2015).

i. This was documented for the Arab Republic of Egypt. See Ali and Najman (2017); Melitz (2003); and Schipper (2020).

j. In the World Bank's Enterprise Surveys, formal firms are asked the following question: "Does this establishment compete against unregistered or informal firms?"

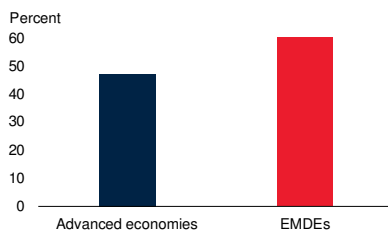
k. As a caveat, the informal competition faced by a specific firm may also be driven by its low productivity, thus generating endogeneity concerns. To address this possible endogeneity issue, the proportion of formal firms facing informal competition in a group of firms of similar size in the same region and sector (a "cell") is used rather than a firm dummy. A cell proportion should be much less correlated with the productivity of a specific firm, and therefore, should be more robust to endogeneity concerns.

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

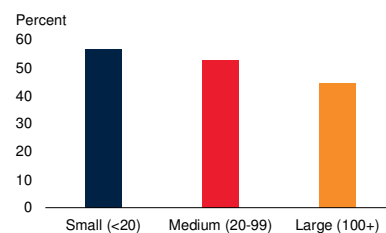
FIGURE B4.2.2 Formal firms facing informal competition

A larger share of formal firms in EMDEs—about three-fifths—reported facing informal competition in 2008-18 than in advanced economies—about half. The degree of informal competition reported by formal firms was higher for smaller than for larger firms.

A. Formal firms reporting competition from informal firms: By country group



B. Formal firms reporting competition from informal firms: By firm size



Source: World Bank (Enterprise Surveys).

Note: World Bank's Enterprise Surveys data for 135 economies (2008-18). Figures show the shares of formal firms.

Productivity gap between formal firms with and without informal competition.

Formal firms that face informal competition are, on average, 24 percent less productive than those that do not (figure B4.2.3; table 4.5). After controlling for informal competition, formal firms in the manufacturing and retail industries have higher productivity than those in other services. Older, exporting, and foreign-owned formal firms also have higher productivity even if they face competition from informal firms.

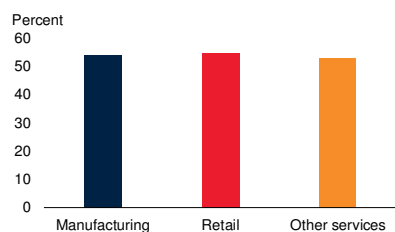
Role of the business climate and development. Economic development and the business climate may substantially shape the productivity gap between formal firms that face informal competition and those that do not. This is captured in interaction terms between the share of similar formal firms reporting informal competition and indicators of development (the logarithm of per capita GDP), the quality of the business climate (proxied by the distance to the frontier in the World Bank Doing Business Dataset), the control of corruption as measured by the World Governance Indicators, and the Business Freedom Index of the Economic Freedom indicators (table 4.5). Higher GDP per capita, better control of corruption, and a business environment that is freer and closer to best practices were all found to dampen the detrimental impact of informal competition on formal firm productivity (figure B4.2.3).

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

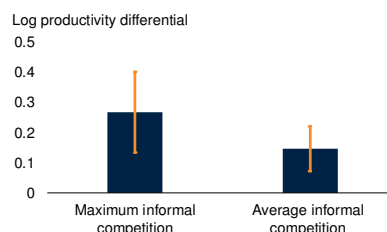
FIGURE B4.2.3 Labor productivity differential of formal firms with and without informal competition

On average, labor productivity in formal firms that face informal competition is only three-quarters of that of firms that do not face informal competition, after controlling for firm characteristics. Better business climates and governance and more advanced economic development can narrow this productivity differential.

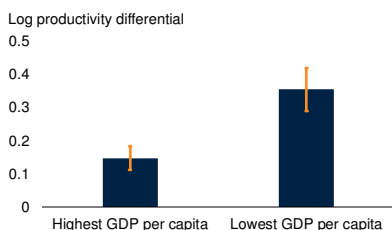
A. Formal firms reporting competition from informal firms: By firm sector



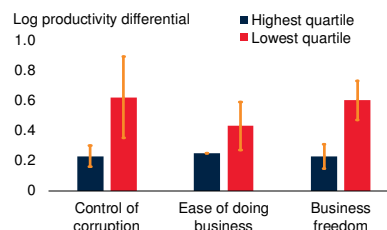
B. Labor productivity differential of formal firms with and without informal competition: By intensity



C. Labor productivity differential of formal firms with and without informal competition: By level of development



D. Labor productivity differential of formal firms with and without informal competition: By business climate indicator



Sources: World Bank; World Bank (Enterprise Surveys).

A. Figures show the shares of formal firms.

B.-D. Based on coefficient estimates from table 4.5. It shows the results from an OLS regression with labor productivity as dependent variable, proxied by annual sales (in 2009 U.S. dollars) per worker. Sample of World Bank's Enterprise Surveys collected during 2007-14, including 4,036 informal firms and 7,558 formal firms in 18 countries. Whiskers show the corresponding ± 2 standard errors.

B. Figure shows log productivity differential between formal firms facing informal competition and formal firms not facing informal competition. "Maximum informal competition" assumes that all firms in a cell face informal competition. "Average informal competition" assumes that 55 percent of firms in a cell face informal competition.

C.D. Figure shows log productivity differential between formal firms facing informal competition and formal firms not facing informal competition (conditional on development and institution). Assumes that 55 percent of firms in a cell face informal competition. Each bar shows average labor productivity of the median country in the top ("highest quartile") or bottom ("lowest quartile") quarter of countries in terms of GDP per capita (C), control of corruption (D), ease of doing business (D) or business freedom index (D).

BOX 4.2 Casting a shadow: Productivity in formal and informal firms (continued)

- *Development.* Two groups of economies in the sample were examined: those economies with per capita income in the highest quartile in the sample and those in the lowest quartile. In formal firms that face informal competition in the average economy in the highest quartile of per capita incomes, labor productivity is only 14 percentage points less than in formal firms that do not face such competition. In contrast, on average in economies in the lowest quartile of per capita incomes, labor productivity in formal firms facing informal competition is 30 percent less than in those firms that do not face such competition.
- *Control of corruption.* Again, two groups of economies were examined: those in the quartile of economies with the strongest control of corruption and those in the quartile with the weakest control of corruption. In economies with the strongest control of corruption, on average, labor productivity in formal firms that face informal competition is only 22 percentage points less than formal firms that do not face such competition, whereas in the economies with the weakest control of corruption, this differential grows to 35 percent.
- *Ease of doing business.* Similarly, the labor productivity differential between formal firms that face informal competition and those that do not might halve (to 21 percent) if an economy like Angola (in the quartile of economies with the most difficult business climates) were to improve its business climate to the level of an economy like North Macedonia (among the economies with the most conducive business climates).

Robustness tests. The impact of informal competition on formal firm productivity is robust to alternative specifications. In particular, it is robust to controlling for characteristics that make the underreporting of output difficult (Amin and Okou 2020).

Conclusion

The productivity gap between informal and formal firms is substantial in EMDEs, averaging 75 percent in a sample of 18 EMDEs between 2007-14. Competition from informal firms also appears to weigh on the productivity of exposed formal firms: the productivity of formal firms that compete with informal firms is only three-quarters that of formal firms that do not compete with informal firms, after controlling for other firm characteristics. Improvements in the business climate, and economic development more broadly, can mitigate some of these negative productivity spillovers from informal to formal firms.

Enterprise Surveys for 2010-18, more than one-third of firms in the average EMDE with above-median informality identified access to finance as a major constraint—9 percentage points higher than in the average EMDEs with below-median informality—and a need to resort to internal finance to fund investment (figure 4.3). On average, three out of four firms in EMDEs with above-median informality depend on internal finance for investment, and only one out of five firms can utilize bank funds to finance their investment needs. Poor access to public services and markets also discourages entrepreneurs from entering the formal sector (Oviedo, Thomas, Karakurum-Özdemir 2009). In EMDEs with above-median employment informality, support for small and medium-sized enterprises (SMEs), such as access to business infrastructure and physical infrastructure, is significantly poorer than in EMDEs with below-median employment informality.⁸

Social safety net. In EMDEs, access to social security programs is often limited for informal workers (Medina and Schneider 2018; WIEGO 2019). As a result, in EMDEs with above-median output informality, only about 4 percent of the population are covered by social security programs such as unemployment insurance—about two-thirds of the level in EMDEs with below-median informality (figure 4.4). Without the ability to fall back on social safety nets and with limited personal savings, informal workers are vulnerable to adverse shocks, which can tip them into poverty. For example, in the one-third of EMDEs with the most pervasive informality, large health care expenses could impoverish more than one-third of households, which is 15 percentage points more than in the one-third of EMDEs with the least informality (box 2.1).

Distorted incentives. In weak institutional and regulatory environments, informal activity can perpetuate and deepen institutional weaknesses.⁹ For example, high taxation and heavy-handed regulation will increase firms' incentives to avoid taxation and regulatory compliance by remaining informal.¹⁰ Excessive labor regulations will encourage informal employment by increasing the cost of formal employment.¹¹ Corruption and rent-seeking bureaucracies will increase firms' incentives to avoid interaction with the state by remaining informal (Choi and Thum 2005; Friedman et al. 2000; Sarte 2000).¹² Governments may strategically design and implement systems of poor governance to promote informality for the poor as an alternative redistributive strategies (Marjit, Mukherjee, and Kolmar 2006). But poor governance stymies development.

⁸ Business infrastructure refers to the presence of property rights, commercial, accounting, and other legal and assessment services and institutions that support or promote SMEs.

⁹ Conversely, access to productivity-enhancing public goods, such as electricity and the legal system, can raise the share of formal production (Mendicino and Prado 2014).

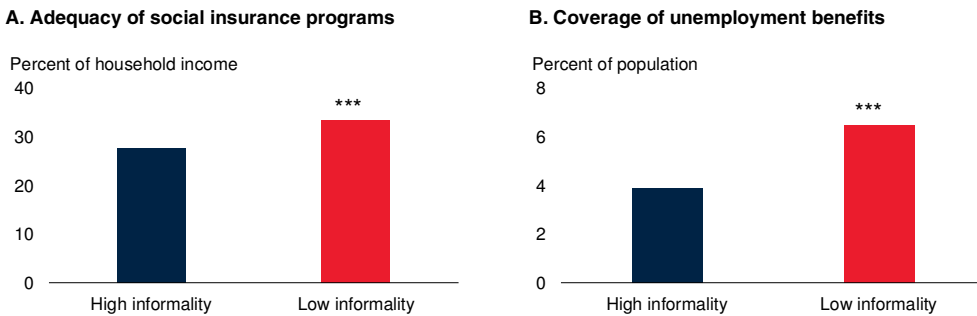
¹⁰ See Amaral and Quintin (2006); Auriol and Warlters (2005); D'Erasmus and Moscoso Boedo (2012); Dabla-Norris et al. (2018); Kanbur (2017); Ihrig and Moe (2004); Prado (2011); and Ulyseas (2018).

¹¹ See, for instance, Loayza (2016); Oviedo, Thomas, and Karakurum-Özdemir (2009); and Rauch (1991).

¹² In turn, widespread informality incentivizes government officials to impose excessive regulations that confers on them the power to collect bribes in return for providing permits (Shleifer and Vishny 1993).

FIGURE 4.4 Access to social benefits

Limited access to social protection programs makes informal workers more vulnerable during economic downturns.



Source: World Bank (World Development Indicators).

A.B. Bars are group means calculated for emerging market and developing economies with “high informality” (above-median DGE-based informal output in percent of GDP) and those with “low informality” (with below-median DGE-based informal output measure) over the period 1990-2018. DGE-based informal output are dynamic general equilibrium model-based estimates. Adequacy of social insurance programs is measured in percent of total welfare of beneficiary households. *** indicates group differences are not zero at 10 percent significance level.

Lack of resilience against the COVID-19 pandemic. The COVID-19 pandemic has exacerbated these development challenges (box 2.1). The global recession caused by the pandemic hit firms and workers in the informal sector particularly hard. Lockdowns have had a particularly disruptive effect on services activities involving human interaction, where informal firms are common, and have thus hit informal employment particularly hard. Large-scale fiscal support implemented in 2020 primarily targeted formal workers and formal firms, with limited support for informal workers or firms (chapter 2; World Bank 2020a). The unprecedented surge in unemployment caused by the global lockdown after the pandemic disproportionately affected jobs in low-value-added services with a large presence of informal jobs (Al Masri, Flamini, and Toscani 2021). A portion of job losses in the service sector may be permanent (Autor and Reynolds 2020; Zenker and Kock 2020).

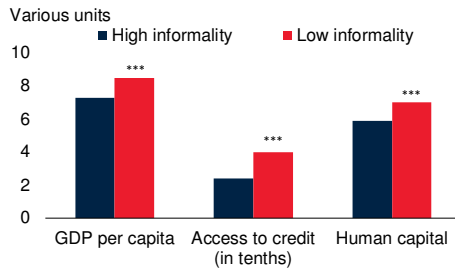
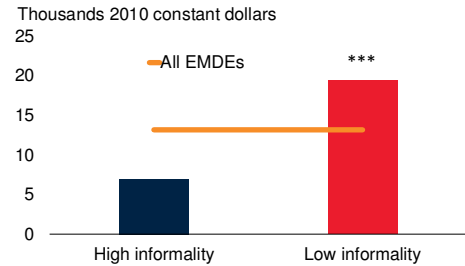
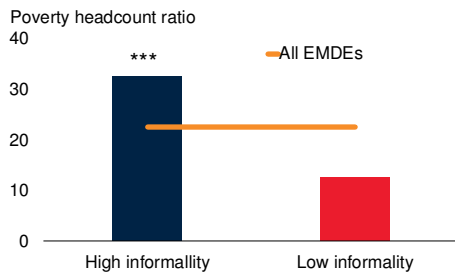
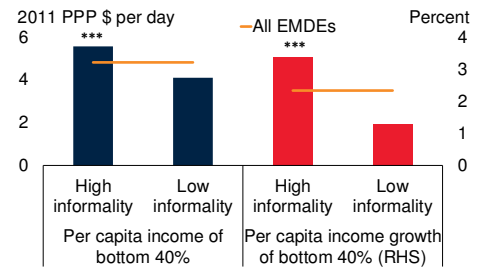
Informality and economic correlates

A large empirical literature has documented the links between informality and poor economic conditions. In particular, a large informal economy is associated with lower per capita incomes, greater poverty, less financial development, limited trade openness, and weaker output growth. These indicators differ significantly between EMDEs with high and low informality (figure 4.5).

Methodology. The next sections rely on a comprehensive literature review as well as several empirical approaches to identify and illustrate the main correlates of informality. The data is drawn from the database detailed in chapter 2 and includes data for up to 160 EMDEs and 36 advanced economies for 1990-2018. First, a descriptive statistical approach is used. The sample of more than 122 EMDEs for 1990-2018 is split into

FIGURE 4.5 Informality and economic correlates

Lower GDP per capita, lower labor productivity, and greater poverty, but also better-shared, prosperity tend to be found in countries with higher informality.

A. Economic correlates and informality**B. Labor productivity and informality****C. Informality and extreme poverty****D. Informality and shared prosperity**

Sources: Barro and Lee (2013); Dieppe (2020); World Bank (World Development Indicators).

Note: EMDEs = emerging market and developing economies; DGE = dynamic general equilibrium. “High-informality” (“Low-informality”) are EMDEs with above-median (below-median) DGE-based estimates of informal output (in percent of official GDP). *** indicates group differences are not zero at 10 percent significance level.

A. “GDP per capita” is in thousands of 2010 U.S. dollars, “Access to credit” is domestic credit to the private sector in percent of GDP, and “Human capital” is measured as average years of schooling. Data are for EMDEs over the period 1990-2018

B-D. Simple group averages over the period 1990-2018 for countries with higher informality (above median) and those with lower informality (below median) are shown in bars. EMDE averages over the period 1990-2018 are shown in orange lines.

B. Labor productivity is proxied by output per employment in thousands of 2010 U.S. dollar.

C-D. “Per capita income of bottom 40 percent” measures the annualized average growth of per capita real survey mean consumption or income for the bottom 40 percent of population.

those with above-median and below-median shares of informality by output (DGE-based estimates) and by employment (proxied by self-employment shares, see table 4.12; for other measures of informality, see table 4.13). Average development outcomes for these two groups are then compared provided that differences between the two groups are statistically significant. In the following sections, results are obtained using output informality, unless otherwise specified. The findings are robust to using employment informality, to the use of alternative definitions of informality or to using a regression that differentiates between quartiles of economies by informality (tables 4.14-4.15). These comparisons examine each correlate of informality individually. There is no presumption of causality going either from development to informality or vice versa.

BOX 4.3 Informality, poverty, and income inequality

Economies with greater informality also feature more poverty and greater income inequality. The coincidence of informality, poverty, and income inequality partly reflects wage differentials between formal and informal employment and the presence in those EMDEs where informality is more prevalent of weaker fiscal positions, social safety nets, and growth. Declines in informality have been associated with poverty reduction but not systematically with declines in income equality.

Introduction

Prevalence of informality is associated with persistent poverty and income inequality.^a This may reflect higher wages for formal than informal workers and more limited income redistribution in EMDEs with widespread informality. Limited redistribution, in turn, in part reflects a lack of fiscal resources; inefficient or inadequate tax and social security systems; burdensome regulations and taxation; and relatively slow economic growth.^b The association with inequality has been noted especially in poorer countries (Elgin and Elveren 2019; Dell’Anno 2016); other researchers have found an association with increases in income inequality (for instance, Mishra and Ray 2010; Rosser, Rosser, and Ahmed 2000).^c

The COVID-19 pandemic has led to an increase in global poverty for the first time in decades (World Bank 2020b, 2021). It has hit informal firms and informally employed workers particularly hard, in part because they struggled to adjust to digital operations and to access government support schemes (box 2.1; Yoshida, Narayan, and Wu 2020). Meanwhile, because of severe economic contractions in the formal sector, the share of informal activity is likely to have increased as a result of the pandemic.

Against this backdrop, this box explores and quantifies the relationships between informality and poverty and income inequality by addressing the following questions:

- How strong is the association between informality and poverty or income inequality?

Note: This box was prepared by Sergiy Kasyanenko.

a. For cross-country studies, see Devicienti, Groisman, and Poggi (2009); Fields (2019); Fields and Pieters (2018); Gasparini and Tornaroli (2007); ILO (2018b); and World Bank (2019, 2020b).

b. See World Bank (2019, 2020b); Perry et al. (2007); Besley and Persson (2014); and Ordóñez (2014).

c. Country level studies often find that higher informality is associated with greater income inequality (Amarante, Arim, and Yapor 2016; Ariza and Montes-Rojas 2017; Docquier and Ifikhar 2019).

BOX 4.3 Informality, poverty, and income inequality (continued)

- Which factors could account for the correlation between informality and poverty or income inequality?
- Is declining informality associated with reductions in poverty and inequality?

This box finds that more pervasive informality is associated with higher poverty and income inequality in EMDEs and that high pre-existing informality may substantially slow subsequent poverty reduction and improvements in shared prosperity. Economies with high informality have slower output and labor productivity growth, holding back poverty reduction and income growth for low-income workers. Declines in informality turn out to be significantly associated with poverty reduction—but not with reductions in income inequality. The link between declining informality and poverty reduction is weaker in regions that started with above-average poverty levels, such as South Asia (SAR) and Sub-Saharan Africa (SSA).

This box contributes several new findings to the literature. It uses a large sample of EMDEs to quantify the correlation between changes in informality and changes in poverty and income inequality over time. The existing literature predominately focuses on country-level studies of poverty and informality.^d Berdiev, Saunoris, and Schneider (2020) find a positive association between the *levels* of poverty and informality in a sample of over 100 countries. This box explores the links between *changes* in informality and poverty or inequality over time, as well as the associations between the levels of these variables. It shows that declines in poverty tend to follow declines in informality, while income inequality remains largely stable.

Cross-country patterns in informality, poverty, and income inequality

A large empirical and theoretical literature has documented the associations between informality and poverty and income inequality. This box updates and extends previous studies.

Methodology. The sample of 122 EMDEs for 1990-2018 is split into those with above-median and below-median shares of output informality. Output informality is measured by DGE-based estimates of informal output in percent of official GDP. Average indicators of poverty and income inequality for these two groups are tested for statistically significant differences. Cross-economy regressions are estimated for the associations between informality and poverty or

d. See Amuedo-Dorantes (2004); Canelas (2019); Devicienti, Groisman, and Poggi (2009); Kim (2005); and Nazier and Ramadan (2015b).

BOX 4.3 Informality, poverty, and income inequality (continued)

income inequality, controlling for the differences in overall development (proxied by real GDP per capita; table 4.6). Similar regressions are also conducted for the relationships between the changes in poverty and income inequality over time and pre-existing levels of informality (annex 4B).

Poverty and informality. On average between 1990 and 2018, more than one in four people in EMDEs with above-median informality lived below the international extreme poverty line of \$1.90 per day (in 2011 U.S. dollars)—about five times as many as in EMDEs with below-median informality (figure B4.3.1). Workers in the informal sector face higher risks of impoverishment than their formal counterparts (Chen, Vanek, and Heintz 2006).

Income inequality and informality. Between 2009 and 2018, average consumption or income of the poorest 40 percent of the population in EMDEs with above-median informality amounted to about \$4 per day (measured in 2011 PPP terms)—statistically significantly below the \$5-6 per day average in EMDEs with below-median informality. Over the same period, income growth for the bottom 40 percent of the population outstripped average income growth by significantly more in EMDEs with below-median informality than in EMDEs with above-median informality. This suggests that progress in improving shared prosperity was more rapid in EMDEs with below-median informality than in those with above-median prosperity.

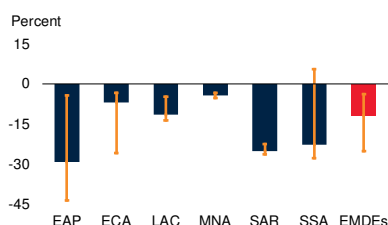
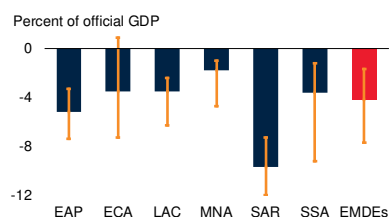
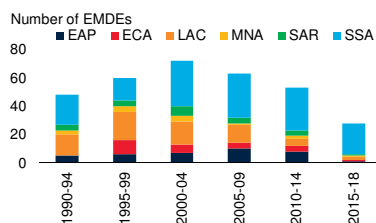
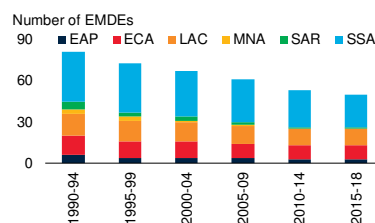
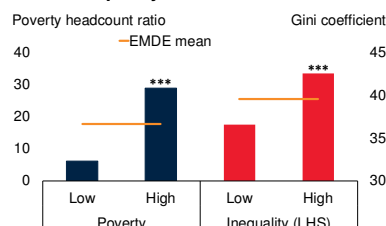
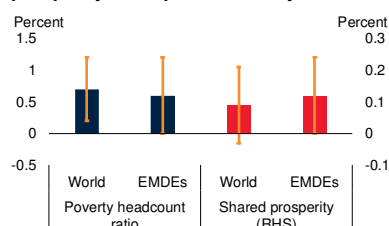
Channels of interaction

Several channels have been suggested for links between informality, poverty, and inequality: wage differentials between formal and informal workers; less redistributive and effective fiscal policies, including poorer coverage of social safety nets, in more informal economies; and slower economic growth in more informal economies (Besley and Persson 2014; Perry et al. 2007; World Bank 2019, 2020a).

Informality and fiscal indicators. Pervasive informality erodes the tax base and constrains governments' ability to provide public services, conduct countercyclical policies, service debt, and implement inequality-reducing redistributive measures (Besley and Persson 2014; Ordóñez 2014). Government revenues in EMDEs with above-median informality were 5-12 percentage points of GDP below those in EMDEs with below-median informality and were tilted towards trade taxes and away from income taxes (chapter 6). Higher informality was also associated with statistically significantly lower public spending on education and health, contributing to the slower accumulation of human capital and income convergence.

BOX 4.3 Informality, poverty, and income inequality (continued)**FIGURE B4.3.1 Informality and poverty and income inequality**

More pervasive informality is associated with significantly more prevalent poverty and higher income inequality. A reduction in informality is positively associated with poverty alleviation, albeit the strength of this link varies across EMDE regions. The link between formalization and improvement in income inequality is largely missing.

A. Change in poverty by EMDE region**B. Change in output informality by EMDE region****C. Extreme poverty above 5 percent across EMDE regions****D. Output informality above 35 percent across EMDE regions****E. Output informality and poverty and income inequality****F. Elasticities of poverty and shared prosperity to output informality**

Source: World Bank (World Development Indicators).

Note: Data for the period 1990-2018. Informality is proxied by dynamic general equilibrium model-based estimates on informal output in percent of official GDP. Low/high informality indicates informality below/above median output informality. "Poverty" refers to income below the \$1.90 per day poverty line. "Shared prosperity" refers to the income share of the bottom 40 percent of population. EAP = East Asia and Pacific; ECA = Europe and Central Asia; EMDEs = emerging market and developing economies; LAC = Latin America and the Caribbean; MNA = Middle East and North Africa; SAR = South Asia; SSA = Sub-Saharan Africa.

A.B. Median change from 1990-99 to 2010-18. Whiskers show 25-75 percentile range.

C.D. Based on average poverty/output informality during the period. Sample includes 122 EMDEs.

E. Averages over the period 1990-2018 for countries with high and low informality. The world averages are shown in orange lines. *** indicates group differences are significant (at least at 10 percent level).

F. Bars show coefficient estimates from regressions of changes in poverty (table 4.7; and changes in shared prosperity in table 4.8) on changes in output informality. A positive coefficient indicates that a decline in output informality reduces (increases) poverty (income inequality). Whiskers show 90 percent confidence intervals.

BOX 4.3 Informality, poverty, and income inequality (continued)

Greater reliance on indirect taxation in more informal economies made the tax system less progressive than systems based more on direct taxation. The reduction in income inequality due to taxes and transfers (i.e., the difference between Gini coefficients before and after taxes and transfers) was about 6 Gini points on average in EMDEs with above-median output informality—statistically significantly less, by 13 Gini points, than in EMDEs with below-median output informality.

Informality and social safety nets. Limited government spending envelopes and poor spending efficiency also constrain the coverage of social protection programs (Dabla-Norris, Gradstein, and Inchauste 2008; Joshi, Prichard, and Heady 2014; Ordóñez 2014). During 1998-2018, only one-third of people in the poorest quintile of the population in EMDEs with above-median employment informality were covered by social protection and labor market programs—considerably less than the more than half in EMDEs with below-median employment informality. In EMDEs with above-median informality, the average daily transfer per capita (in PPP terms) was one third less than in EMDEs with below-median informality.

In EMDEs with above-median informality, social protection and labor programs lowered the Gini coefficient by 2-4 points, significantly less than the 8-11-point reduction in EMDEs with below-median informality. Social protection and labor programs accounted for only about one-tenth of poverty reduction in EMDEs with above-median informality during 2009-2018, again significantly less than the one-quarter of poverty reduction in EMDEs with below-median informality.

Worker earnings differentials. A formal wage premium—systematically higher wages in the formal than informal sector—is a long established finding in the literature. In a meta-analysis of empirical studies on formal wage premiums, wages in the formal sector were found to be 19 percent above informal wages on average (box 4.1).

Over-time patterns in informality, poverty, and inequality

While informality is typically accompanied by poverty, the evidence that informality reduction is associated with lower poverty or income inequality is less conclusive. In some cases, the informal sector provides a critical source of income for the poorest and helps ward off poverty during adverse events (Rogan and Cichello 2020). While policies to reduce informality in EMDEs have often been complemented by poverty-alleviation initiatives such as income-support or unemployment protection programs, these initiatives can fail to reach the very poorest groups and can create their own disincentives for formal employment (Bloeck, Galiani, and Weinschelbaum 2019; Gasparini, Haimovich, and Olivieri

BOX 4.3 Informality, poverty, and income inequality (continued)

2009). Declines in informality and poverty often coincide with rapid economic growth, which can raise income inequality when structural factors that widen income disparities (such as urbanization and demographic change) are not offset by other policies (for example, tax and expenditure reforms; Jain-Chandra et al. 2018).

Methodology. A similar methodology to that used by Dollar and Kraay (2002) and Dollar, Kleineberg, and Kraay (2013) is applied here to investigate the link between within-country changes in informality over time and changes in poverty and inequality reduction (annex 4B). Income inequality is proxied by the Gini coefficient or the share of income of the poorest 40 percent of the population (“bottom 40 percent”) within each country. The sample of country-year observations is assembled by starting with the first available observation for each country and selecting all available consecutive observations that are at least five years apart. This approach yields about 428 economy-year pairs over the period 1990-2018 for 32 advanced economies and up to 119 EMDEs. A median distance between observations is 5.5 years for EMDEs and 5 years for advanced economies. A cross-country fixed-effects regression estimates the changes in poverty and inequality associated with changes in informality.

Evolution of informality. Output and employment informality have declined steadily over the past three decades, but with considerable heterogeneity across and within EMDE regions (figure B4.3.1). In the median EMDE, output informality contracted by 4 percentage points of GDP between 1990-99 and 2010-18. It declined most in SAR (by about 10 percent of official GDP) and East Asia and the Pacific (EAP, by about 6 percent of official GDP). In Europe and Central Asia (ECA) and Latin America and the Caribbean (LAC), output informality fell by about 3.5 and 2 percentage points of official GDP, respectively.

Evolution of poverty and income inequality. Most EMDEs saw a substantial reduction in the share of population living below the \$1.90 a day poverty line, although the pace of poverty reduction was slow in SSA (World Bank 2020b). Progress in improving shared prosperity was more heterogeneous across regions. While the share of income captured by the bottom 40 percent of population increased by about 2 percentage points from 1990-99 to 2010-18 in a median EMDE, it did not change much in EAP and SAR despite a notable reduction in poverty.

Decline in informality and poverty reduction. The regression results suggest that a 1-percentage-point reduction in output informality in EMDEs has been associated with a 0.6-percentage-point decrease in the share of population living in extreme poverty (on less than US\$1.90 a day in 2011 U.S. dollars) and a

BOX 4.3 Informality, poverty, and income inequality (continued)

significant decrease in the share of the population living on less than US\$3.20 a day (figure B4.3.1; table 4.7).

Decline in informality and improvements in income inequality. Consistent with the literature on income inequality and economic development, no strong or statistically significant relationship was found here between declines in informality and changes in inequality indicators (Adams 2003; Banerjee and Duflo 2003). A reduction in informality may play only a small role in inequality reduction partly because workers who transition from informal to formal jobs are already in the upper tail of the wage distribution, such as informal workers in formal firms (Messina and Silva 2021). Unlike in poverty regressions, where GDP growth and reductions in informality are strongly associated with poverty reduction, neither economic growth nor declines in informality appear to be statistically significantly correlated with changes in shared prosperity indicators (table 4.8).

Conclusion

Economies with higher informality also tend to have higher poverty and greater income inequality. Several forces contribute to this pattern. More constrained public spending envelopes in EMDEs with widespread informality limit governments' ability to provide public services, resulting in more limited coverage and adequacy of social security programs. Workers in the formal economy earn, on average, about a fifth more than workers in the informal economy; this earnings gap contributes to the higher levels of poverty and income inequality in EMDEs with more pervasive informality. Without establishing the causal relationship between changes in informality and changes in poverty or income inequality, the box finds that reductions in informality were typically accompanied by poverty reduction, but not by reductions in income inequality. The confirmation of causal relationship between informality and poverty (and income inequality) presents an area for future research.

To identify the most robust correlates of informality among this large menu, this chapter conducts a Bayesian Model Averaging (BMA) estimation. The large number of (often inter-correlated) correlates of informality gives rise to concerns about model uncertainty (annex 4C). The BMA approach addresses model uncertainty formally—by recognizing that the identity of the true model is unknown and by combining various sets of potential correlates of informality, ranging from economic development to infrastructure quality. This approach also addresses concerns about omitted variables in the bilateral correlations depicted in the stylized facts.

Lack of development. Higher informality is associated with lower levels of per capita income and other measures of economic development.¹³ For both output and employment informality, GDP per capita in EMDEs with below-median (“low”) informality is about four times that in EMDEs with above-median (“high”) informality (figure 4.5).¹⁴ Since informal activity is concentrated in services and agriculture, more manufacturing-based and more urban economies feature less informality. On average, about three fifths of the population in EMDEs with below-median informality, but only two fifths of the population in EMDEs with above-median informality, resides in urban areas. Past studies also show that more widespread informality has been associated with slower growth of output and investment and—in some studies—labor productivity (Perry et al. 2007; World Bank 2019).¹⁵

Poverty. Greater prevalence of informality is associated with greater prevalence of extreme poverty (figure 4.5; World Bank 2019). In 2000, the share of the population living on less than \$1.90 a day was 44 percent in EMDEs with above-median informality—more than twice the share in EMDEs with below-median informality. In 2018, one in four people remained in extreme poverty in EMDEs with above-median informality, while fewer than one in ten people were in extreme poverty in EMDEs with below-median informality.

Financial development. Financial development, which reduces external financing costs, can incentivize entrepreneurs to join the formal sector and comply with tax obligations.¹⁶ EMDEs with above-median informality significantly lag in financial development behind EMDEs with below-median informality (figure 4.3; chapter 6). People in EMDEs with below-median informality have access to significantly more commercial bank branches, automated teller machines (ATMs), and credit facilities than those in EMDEs with above-median informality. About half of the population in EMDEs with below-median informality hold an account at a financial institution or use a mobile money service, which is significantly higher than the share in EMDEs with above-median informality, by about 18 percentage points. In financially less developed EMDEs, a higher share of entrepreneurs rely more on internal funding for starting, operating, or expanding their firms (figure 4.3; Epstein and Shapiro 2017; Farazi 2014).

International trade and financial openness. While greater financial openness has been associated with lesser informality, the link between trade openness and informality is unclear. Higher capital account openness was associated with less output and

¹³ See La Porta and Shleifer (2014) and Loayza, Servén, and Sugawara (2010). However, there is also empirical evidence that the relationship between the extent of informality and the level of development has a U shape, with the informal economy tending to expand again (or at least to not shrink further) when economic development surpasses a certain threshold (Wu and Schneider 2019).

¹⁴ Median informality amounts to about 36 percent of GDP for DGE-based informal output and 43 percent of total employment for self-employment.

¹⁵ Also see D’Erasmus and Moscoso Boedo (2012); Docquier, Müller, and Naval (2017); and Medina and Schneider (2019).

¹⁶ See Blackburn, Bose, and Capasso (2012); Capasso and Jappelli (2013); D’Erasmus and Moscoso Boedo (2012); and chapter 6.

employment informality (World Bank 2019). That said, the impact of major trade liberalizations on informality has varied across countries, differed between the short- and the long-term, and between workers and firms (Dix-Carneiro and Kovak 2017; Fugazza and Fiess 2010; Goldberg and Pavcnik 2003). Several studies show informality declining after trade liberalization.¹⁷ In the presence of labor market rigidities, however, informal employment may rise in the short term after trade liberalization but not necessarily in the long-term.¹⁸ In Brazil and Peru, trade liberalization and increased import competition were associated with increases in informality as informal firms exited and formal firms increasingly hired informal workers (Cisneros-Acevedo, forthcoming) or workers increasingly worked informally (Dix-Carneiro and Kovak 2019).

Informality and institutions

More informality is also associated with lower government revenues and expenditures, less effective public institutions, more burdensome tax and regulatory regimes, and weaker governance (e.g., Dabla-Norris, Gradstein, and Inchauste 2008; Enste and Schneider 1998; World Bank 2019).¹⁹

Government revenues and expenditures. On the basis of the various measures of informality, government revenues in EMDEs with above-median informality were, on average, 5-12 percentage points of GDP below those with below-median informality during 2000-2018 (figure 4.6). The composition of tax revenues is also tilted toward trade taxes in economies with more pronounced informality, making the tax system less progressive but facilitating tax collection when income underreporting is widespread. Similarly, in EMDEs with more pervasive informality, government expenditures were 5-10 percentage points of GDP lower than in those with less informality.

Such constrained government spending is reflected in more limited provision of government services, contributing to poorer human development outcomes (Gaspar et al. 2019). During 2000-18, EMDEs with above-median informality spent about 2 percent of GDP on health, which was 1 percentage point of GDP lower than in EMDEs with below-median informality (figure 4.6). The average number of pupils per teacher in primary schools was about 35 in EMDEs with above-median informality—significantly higher, by eight student per teacher, than in EMDEs with below-median informality. Access to medical resources, such as physicians and nurses, was also significantly more limited in EMDEs with a more pervasive informal sector (World Bank 2020a).

Regulatory burdens. Both empirical and theoretical studies suggest that heavier regulatory (or administrative) burdens are associated with greater informality (figure

¹⁷ See Boly (2018); Goldberg and Pavcnik (2004, 2007); and McCaig and Pavcnik (2018).

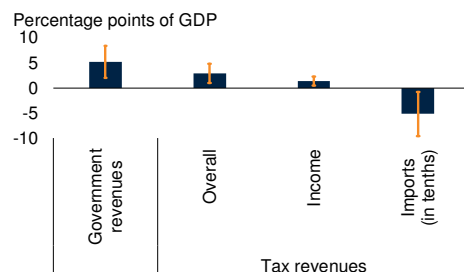
¹⁸ See Attanasio, Goldberg, and Pavcnik (2004); Bosch, Gofii-Pacchioni, and Maloney (2012); Goldberg and Pavcnik (2003); Ponczek and Ulyssea (2018); and World Bank (2019).

¹⁹ Access to the court system can also encourage formal production (Mendicino and Prado 2014; Schneider, Buehn, and Montenegro 2010).

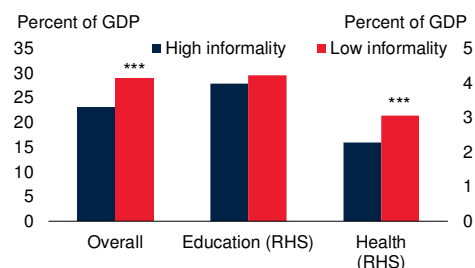
FIGURE 4.6 Informality, fiscal indicators, and institutional quality

Informality is associated with lower government revenues and expenditure capacity, hindering economies' abilities to provide health and education services. EMDEs with above-median informality tend to have more burdensome regulations and weaker governance.

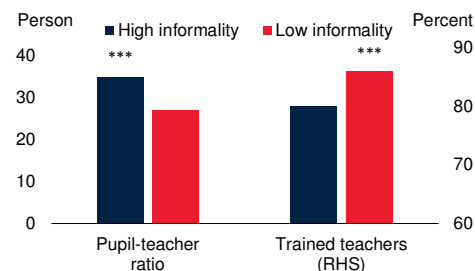
A. Differential in government revenues between EMDEs with above- and below-median informality



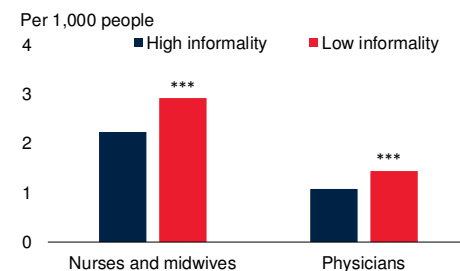
B. Government spending and informality



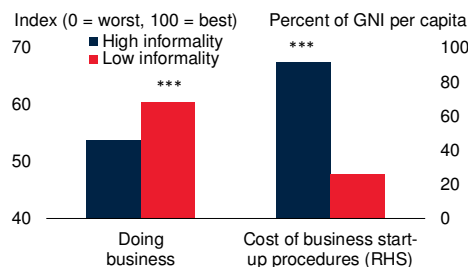
C. Access to education resources



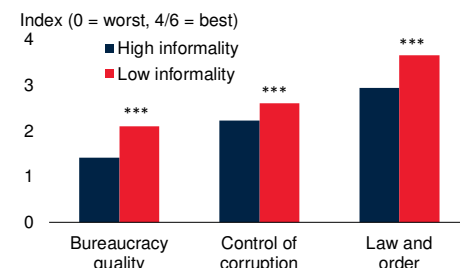
D. Access to medical resources



E. Regulatory burdens



F. Governance



Sources: *International Country Risk Guide (ICRG)*; International Monetary Fund (Government Finance Statistics, World Revenue Longitudinal Data); World Bank (*Doing Business*, World Development Indicators).

Note: Informality is proxied by dynamic general equilibrium (DGE)-based estimates of informal output in percent of GDP. Data for emerging market and developing economies (EMDEs). "High-informality" ("Low-informality") are EMDEs with above-median (below-median) DGE-based informal output measures. *** indicates group differences are not zero at 10 percent significance level.

A. Differences in percentage points of GDP between the average fiscal indicators among EMDEs with above-median and below-median informality are in bars, with corresponding 90 percent confidence intervals shown by whiskers. All fiscal indicators and informality measures are 2000-18 averages for EMDEs with populations above 3.5 million (several oil-exporting outliers are dropped).

B. Simple group averages of public expenditure over the period 2000-18. Overall expenditures are for EMDEs with population above 3.5 million.

C. Simple group averages are in bars over the period 1990-2018. "Pupil-teacher ratio" measures the average number of pupils per teacher in primary school. "Trained teachers" captures the percentage of primary school teachers who have received the minimum organized teacher training (pre-service or in-service) required for teaching in a given country.

D. Simple group averages over the period 2010-18. Two outlier countries, Belize and Belarus, are excluded from the sample.

E.F. Simple group averages are in bars over the period 1990-2018. Doing business score ranges from 0 (lowest performance) to 100 (best performance). Data for business registration costs are for 2003-18 (2010-18 for doing business scores). "Bureaucracy quality," "Control of corruption," and "Law and order" are from ICRG, with higher values corresponding to better outcomes.

4.6).²⁰ Over the period 2010-18, the average ease of doing business score for EMDEs with below-median informality (by DGE estimates) was higher by 7 points—two-thirds of a standard deviation—than for EMDEs with above-median output informality.²¹ Similarly, the average cost of business start-up procedures amounted to 92 percent of gross national income (GNI) per capita in EMDEs with above-median output informality, significantly higher than in EMDEs with below-median output informality, by about 65 percentage points.

Governance. A large literature has documented the coincidence of poor governance with pervasive informality in many EMDEs, especially in Latin America and the Caribbean (LAC) and Europe and Central Asia (ECA).²² On average, EMDEs with above-median informality have had significantly poorer-quality bureaucracies, by about one standard deviation in terms of the ICRG's rating, than EMDEs with below-median informality.²³ Similar differences pertain to the control of corruption and law and order. There are also country cases suggesting that informality declined faster in the presence of greater improvements in governance and better initial governance (World Bank 2019). For example, in Georgia, during 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 (World Bank 2019).

Informality and human development-related SDGs

Greater informality is associated with weaker human development outcomes. People living in EMDEs with more widespread informality suffer from a greater prevalence of hunger, poorer health and education, greater gender inequality, and lower human capital (figure 4.7; Docquier, Müller and Naval 2017; Maloney 2004).

Hunger. EMDEs with more pervasive informality fared far worse during 1990-2018 in terms of the hunger-related SDGs than those with less pervasive informality. The share of the population suffering from stunting and wasting was significantly higher in EMDEs with above-median informality: more than a quarter of children under five years of age in EMDEs with above-median informality suffered from stunting, with

²⁰ See Bruhn (2011); de Mel, McKenzie, and Woodruff (2013); Perry et al. (2007); Rocha, Ulyssea, and Rachter (2018); and Ulyssea (2010).

²¹ Similarly, the Heritage Foundation's business freedom index was about three-quarters of a standard deviation higher in countries with below-median output informality than in countries with above-median informality.

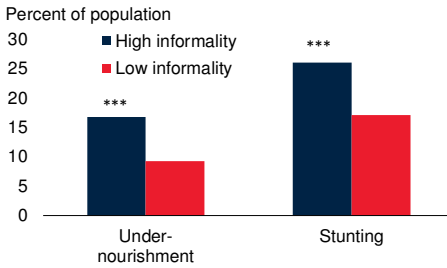
²² Sarte (2000) suggests that firms operate in the informal sector to avoid rent-seeking bureaucrats. Dreher and Schneider (2010) and Choi and Thum (2005) report an association between higher informality and weaker law and order and control of corruption. Dabla-Norris, Gradstein, and Inchauste (2008) show that the quality of the legal framework is important in determining the size of the informal sector.

²³ The measures of institutional quality used here are taken from the International Country Risk Guide (ICRG 2014). A higher value indicates better institutional quality. In the case of bureaucracy quality, high points are given to countries where the bureaucracy has the strength and expertise to govern without drastic changes in policy or interruptions in government services.

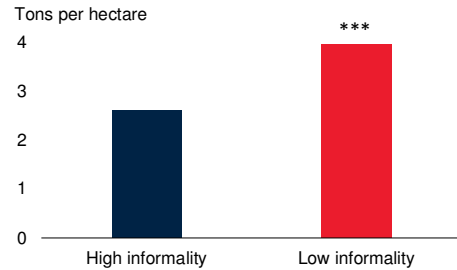
FIGURE 4.7 Informality and human development-related SDGs

EMDEs with more pervasive informality face more severe challenges in progressing toward human development-related SDGs. People in EMDEs with more informality tend to suffer more from hunger, poor health and education, and gender inequality.

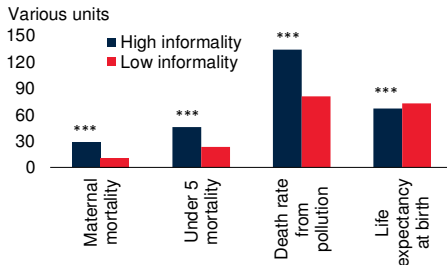
A. Hunger and informality



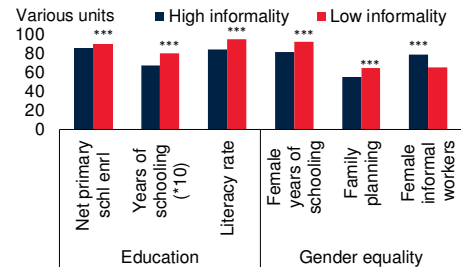
B. Agricultural productivity and informality



C. Health outcome indicators and informality



D. Educational attainment, gender equality, and informality



Sources: Sachs et al. (2018); Sachs et al. (2020); World Bank (World Development Indicators).

Note: "High informality" ("Low informality") are emerging market and developing economies (EMDEs) with above-median (below-median) dynamic general equilibrium model-based informal output measure over the period 1990-2018. Figures show the latest available data. *** indicates that group differences are significant at 10 percent level.

A. Average shares of population (children under 5 years of age) that suffer from undernourishment (stunting) for corresponding country groups. Based on 144 EMDEs.

B. Agricultural productivity is measured as cereal yield (tons per hectare of harvested land). Based on 141 EMDEs.

C. Bars for group averages for the latest year available. "Maternal mortality" is measured in per 10,000 live births. "Under 5 mortality" is measured in per 1,000 live births. "Death rate from pollution" are age-standardized death rates (per 100,000 persons) due to household and ambient (outdoor) pollution. "Life expectancy" at birth is measured in years. Based on up to 153 EMDEs.

D. Bars for group averages for the latest year available. "Net primary school enrollment" rate ("net primary schI enrI") is measured in percentage points. Mean years of schooling are measured in years (in multiplier of 10; taken from Sachs et al. 2018). "Literacy rate" of 15-24 years old is measured in percent of population. "Female years of schooling" is measured in percent of male schooling years. "Family planning" measures the percentage of women reporting having their family planning needs attended. "Female informal workers" captures the percent of female workers informally employed among all female workers.

more than 15 percent of the population being undernourished (Sachs et al. 2020). Both shares are significantly higher than in EMDEs with below-median informality, by about 10 percentage points. The higher prevalence of hunger in EMDEs with above-median informality is partly explained by lower agricultural productivity. On average, the cereal yield in EMDEs with above-median informality is less than two-thirds of the yield in EMDEs with below-median informality (figure 4.7).

Health. People in EMDEs with more widespread informality face poorer health outcomes. According to the latest available data, in 2016, the average life expectancy at birth in EMDEs with above-median informality was around 67 years, which is about 6 years shorter than in EMDEs with below-median informality. Both the maternal mortality rate and the mortality rate of children under 5 years old in EMDEs with above-median informality were nearly twice the rates in EMDEs with below-median informality. In EMDEs with above-median informality, on average, about 133 deaths per 100,000 persons were caused by household air pollution and ambient pollution, which is significantly higher than in EMDEs with below-median informality by more than 50 deaths per 100,000 persons.

Education. Access to education is less available in EMDEs with more pervasive informality (Docquier, Müller, and Naval 2017). On average in EMDEs with above-median informality, people spent less than 7 years in schooling, compared with 8-9 years in EMDEs with below-median informality. Less than 85 percent of the population aged 15-24 years in EMDEs with above-median informality is literate—more than 10 percentage points less than the population in EMDEs with below-median informality.

Gender equality. Female workers make up a disproportionate share of workers in the informal sector (Bonnet, Vanek, and Chen 2019; ILO 2018b). In EMDEs with above-median employment informality, 87 percent of employed women work in the informal sector, which is about three-quarters higher than in EMDEs with below-median employment informality (table 4.12). In South Asia (SAR) and Sub-Saharan Africa (SSA), the regions where informality is most pervasive, about 80 percent of female workers in the non-agricultural sector are informally employed (UN Women 2016). In LICs, up to 92 percent of all employed women work in the informal sector (ILO 2018b; OECD/ILO 2019). Working in the informal sector exposes female workers to low remuneration, poor working conditions, and lack of or limited access to social protection and rights at work (ILO 2019; Otobe 2017).

EMDEs with more widespread informality are also associated with greater gender inequality in broader terms. The average years of schooling received by women in EMDEs with above-median informality are, on average, 20 percent less than the average years received by men, in stark contrast to EMDEs with below-median informality where no significant gender gap in schooling prevails. Only 55 percent of women in EMDEs with above-median informality have their family planning needs attended to, which is 10 percentage points lower than in EMDEs with below-median informality. In addition to factors such as traditional gender roles, lack of access to education and insufficient coverage of family planning needs constrain women's ability to participate in the formal sector (Malta et al. 2019).

Informality and infrastructure-related SDGs

More widespread informality is associated with poorer access to, and lower overall quality of, infrastructure, with causality running in both directions. Thus widespread informality tends to limit government revenue and hence public expenditure on

infrastructure but, conversely, poor access to infrastructure can discourage firms or workers from joining the formal sector and engaging with the government (Perry et al. 2007). Infrastructure weaknesses create an additional obstacle to human capital accumulation and job creation, entrenching informality (Vorisek and Yu 2020). Lack of energy and communications infrastructure limits access to more and better job opportunities and slows productivity gains from digital technologies (Zaballos, Iglesias, and Adamowicz 2019). Poor transportation networks restrict factor mobility and market access, thus slowing productivity growth.²⁴ Weak within-city transportation networks prevent workers from accessing formal employment opportunities (Zarate 2019).

Sanitation and clean water. Greater informality is associated with more limited access to sanitation facilities and clean water (figure 4.8). In EMDEs with above-median informality, only around 55 percent of the population have access to basic sanitation services—about 30 percentage points fewer than in EMDEs with below-median informality. While almost all people in EMDEs with below-median informality have access to clean drinking water, this essential infrastructure service is only available to less than 80 percent of the population in EMDEs with above-median informality.

Access to energy and ICT infrastructure. Access to infrastructure services such as electricity, clean fuels, the internet, and mobile broadband is more limited in EMDEs with above-median informality. As shown by the latest available data, about one-third of people living in EMDEs with above-median informality lacked access to electricity, while almost all in EMDEs with below-median informality had such access. The share of the population with access to clean fuels in EMDEs with above-median informality is about half of the share in EMDEs with below-median informality. Access to the internet and mobile broadband was available to 30-40 percent of the population in EMDEs with above-median informality—less than three-fifths of the share in EMDEs with below-median informality.

Road access. Despite progress made in recent years, road access remains more limited in EMDEs with above-median informality. The latest data show that only around 15 percent of roads were paved in EMDEs with above-median informality—one-third of the share in EMDEs with below-median informality. This could in part reflect higher costs of road construction in EMDEs with more pervasive informal sectors: on average, road construction costs amounted to US\$ 0.6 million per km in EMDEs with above-median informality—significantly higher than the cost in EMDEs with below-median informality, by US\$ 0.2 million per km.

Finding the needle in the haystack: The most robust correlates

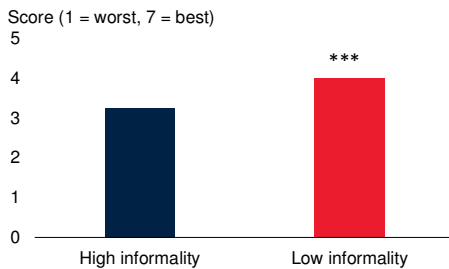
The analysis so far has considered individual correlates of informality in isolation. This section aims to identify the most robust correlates of informality via a Bayesian Model

²⁴ See Banerjee, Duflo, and Qian (2020); de Soyres, Mulabdic, and Ruta (2020); and Francois and Manchin (2013).

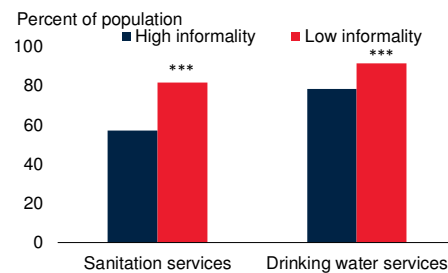
FIGURE 4.8 Informality and infrastructure-related SDGs

Higher infrastructure costs and lower quality of overall infrastructure are associated with more pervasive informality. EMDEs with above-median informality tend to have significantly less access to various types of infrastructure than those with below-median informality.

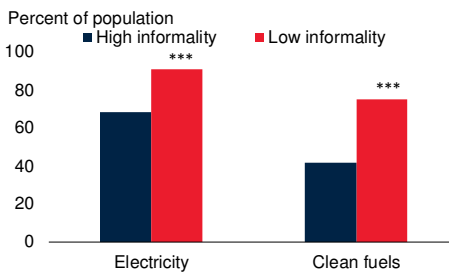
A. Quality of overall infrastructure



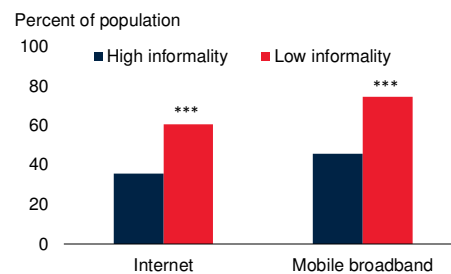
B. Access to clean water and sanitation



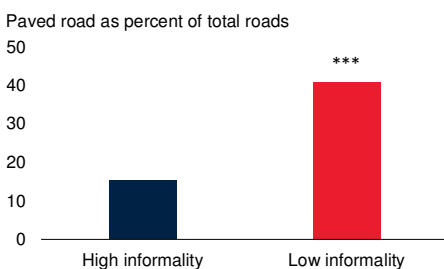
C. Access to energy



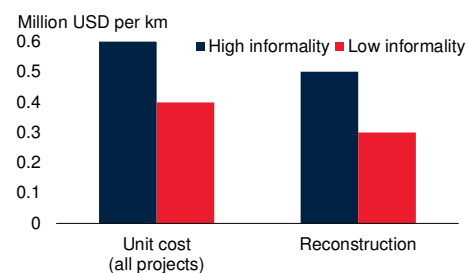
D. Access to ICT-related infrastructure



E. Access to roads



F. Road project costs



Sources: Sachs et al. (2018); Sachs et al. (2020); World Bank (Road Costs Knowledge System [ROCKS] database; World Development Indicators).

Note: "High informality" ("Low informality") are emerging market and developing economies (EMDEs) with above-median (below-median) dynamic general equilibrium (DGE)-based informal output, measured over the period 1990-2018. *** indicates group differences are significant at 10 percent level.

A. Bars show the average overall infrastructure quality for the latest year available. The index is taken from Sachs et al. (2018) and ranges from 1 to 7, with higher values representing better overall infrastructure quality. Based on 115 EMDEs.

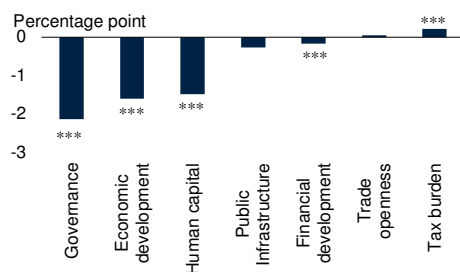
B.-E. Bars show the group means for the latest year available. Based on up to 153 EMDEs.

F. Bars show group averages for the latest year available. Data are for 51 EMDEs (outliers, GIN, MWI, and PER, are dropped). "Unit cost (all projects)" excludes new 4-lane and 6-lane expressway projects as they cost much more than other projects and do not exist in EMDEs with above-median informality. Reconstruction is the most common road project by project number. Data are from the latest year available. Based on 77 EMDEs.

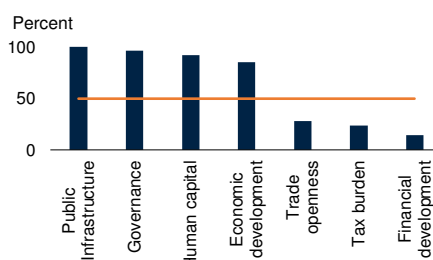
FIGURE 4.9 Results from Bayesian Model Averaging approach

A wide range of economic and institutional conditions, such as better governance, a more developed and diversified economy, and better human capital, are associated with less pervasive informality in EMDEs.

A. Effect of group correlates on the evolution of output informality



B. Probability of inclusion



Source: World Bank.

Note: Based on the panel regression result using Bayesian Model Averaging technique. In A, *** denotes that the signs of the estimated coefficients are the same for 90 percent of the model, conditional on inclusion of the variable. The dependent variable is the share of informal economy in GDP using dynamic-general-equilibrium-based estimates on informal output as a share of official GDP. From an unbalanced panel from 67 EMDEs and 5-year periods over 1998-2018. See annex 4C for details.

A. Predicted differences in the share of output informality for countries whose correlates differ by one standard deviation.

B. Probability of including at least one variable from the group to the regression (Posterior Inclusion Probability). The groups whose posterior inclusion probabilities exceeded the prior of 50 percent could be regarded as relevant.

Averaging (BMA) approach. The approach utilizes a large number of possible correlates that have been identified in the literature and reviewed in the previous sections. It recognizes the unknown nature of the true relationship between informality and various conditions and combines evidence from different models.

Methodology. Following Dieppe et al. (2020) and Durlauf, Kourtellos, and Tan (2008), independent variables that represent common concepts are grouped together to estimate their group marginal effects. The correlates of informality are sorted into seven groups, covering economic development, human capital, financial development, and governance (table 4.9). Analyses are carried out at the level of output informality, proxied by the share of informal output in GDP, while using the levels of correlates as independent variables (table 4.10). The results from a similar analysis using employment informality are broadly consistent (annex 4C). The sample comprises an unbalanced panel of 68 EMDEs, observed from 1998 to 2018 in five-year windows.

Results. The results are broadly consistent with existing studies about individual correlates. First, the posterior inclusion probabilities for several variable groups under consideration exceed 0.50 (figure 4.9). This finding suggests that most of these factors are related to the level of output informality. The group posterior inclusion probabilities (PIPs) are highest—near 1—for economic development, human capital, and governance, indicating these factors' particularly strong association with informality. The results indicate their stronger association with the level of output informality than other groups of variables.

Second, all the correlates are associated with the level of informality in the expected manner (figure 4.9). EMDEs with better governance, a more developed and less agriculture-oriented economy, and stronger human capital tend to have lower informality. In particular, informal output as a share of official GDP is significantly lower, by around 1-2 percentage points of GDP, if the EMDE has one-standard-deviation better governance, greater economic development, or larger human capital. In addition, tax burdens, measured as the share of government consumption or tax revenue in GDP, are significantly and negatively correlated with the size of output informality.

Conclusion

Pervasive informality is associated with a wide range of development challenges, from extreme poverty to lack of access to basic sanitation services. This chapter documented and quantified the wider gaps, relative to the SDGs, among EMDEs with greater informality. These shortcomings also make EMDEs with widespread informality particularly vulnerable to adverse global shocks, including the COVID-19 pandemic.

The wide range of correlates of informality suggests that informality is a phenomenon that reflects broad-based underdevelopment rather than a challenge that can be considered in isolation. Consequently, policy measures to address informality need to be equally broad-based. They include measures to enhance human capital and lift productivity, streamline regulations, and improve governance and the provision of public services and social safety nets. Policy options are discussed in detail in chapter 6.

Several areas for further development are worth exploring. First, the chapter does not demonstrate a causal link between informality and the various development outcomes. The exploration of causal relationships between informality and these outcomes, in either direction, is left for future studies. Second, due to data limitations, some variables, such as access to paved roads and bank account ownership, that are relevant to informality are not included in the current BMA analysis. Future studies can improve upon the work reported here by incorporating those variables. Third, future research could explore asymmetries in the challenges posed by informality. There may be interactions between country circumstances and worker or firm characteristics that can mitigate some of the challenges posed by informality. For firms, some of these interactions were explored in box 4.2 but other important interactions may yet come to light in future research.

ANNEX 4A Meta regression analysis

A random-effects model assumes that there is a distribution of true effects rather than a common fixed effect across studies (DerSimonian and Laird 1986). In particular, a study-specific estimate of the informal-formal wage gap has a sampling distribution $\hat{\theta}_i \sim N(\theta_i, \sigma^2)$, where σ^2 is the within study variance of the estimate due to a sampling error; while the true effect has the following distribution $\theta_i \sim N(\mu, \tau^2)$. Meta-analysis pools information across many studies to estimate μ and τ^2 , where τ^2 measure the degree of across-study variations.²⁵ The proportion of total variation in study estimates is equal to $I^2 = \tau^2 / (\tau^2 + \sigma^2)$ and reflects the impact of across-study heterogeneity (Higgins and Thompson, 2002). The meta-regression analysis (MRA) can be performed to associate this variation with any characteristics of the study or sample.

The MRA of estimated wage differentials between formal and informal jobs uses estimates of the wage gap drawn from each study as the dependent variable. The set of regressors, or moderator variables, includes study characteristics that are deemed consequential for the reported results, for example, identification and estimation methods, study design and data sources. This, in particular, helps clarify the diversity of research outcomes on the size of the informal-formal wage gap and identify the sensitivity of reported wage gaps to study-specific methods and data. A random-effects MRA is performed by estimating the following regression:

$$\hat{\theta}_i = \mu + \sum_j^k \alpha_j X_{ij} + \epsilon_i + \vartheta_i,$$

where $\hat{\theta}_i$ is a study-specific estimate of the informal-formal wage gap, ϵ_i is a sampling error with a standard deviation that may vary across studies, ϑ_i is an error term reflecting across-study variation of true effects with a constant across-study variance τ^2 ; finally, the set of moderator variables, \mathbf{X} , includes:

- A dummy variable accounts for differences in methodology: FE_i is 1 if fixed effects were used to correct for unobserved workers' characteristics and 0 otherwise.

Two dummy variables reflect the gender composition of the sample: $FEMALE_i$ is 1 if estimates were obtained for female workers only and 0 otherwise, $-MALE_i$ 1 if estimates were obtained for male workers only and 0 otherwise. The reference categories for this set of dummy variables are estimates obtained with samples containing both female and male workers.

- Regional dummy variables are included to account for regional heterogeneity.
- $Self-employed_i$ is a dummy variable indicating that a study measured the wage gap between self-employed and formal employees.

²⁵ The random-effects meta-analysis estimate is a special case of a generalized method of moments estimator, where each estimate is weighted proportionally to its sampling error. Thus, it can only be applied to studies that reported standard errors of their inform-formal wage gap estimates.

ANNEX 4B Regression analysis

Correlation between informality, poverty, and income inequality. The following cross-country ordinary least squares regression model is estimated to show the association between informality and levels of extreme poverty and income inequality:

$$\bar{y}_i = \alpha_0 + \theta_1 \bar{x}_i + \theta_2 \overline{\text{LnGDPpc}_i} + \epsilon_i$$

The results are reported in table 4.6.

The dependent variable (\bar{y}_i) includes a range of measures for levels of poverty and income inequality averaged over 1990-2018 in country i . The level of poverty is proxied by the poverty headcount ratio at \$1.90 a day (2011 PPP) in percent of the total population. Measures for income inequality include the Gini coefficient (range from 0-100, with 0 being perfect equality and 100 being extreme inequality), survey mean consumption or income per capita of the lowest-income 40 percent of population, and the difference in consumption or income per capita levels between the bottom 40 percent of population and the total population (World Bank 2018). Lastly, the progress in shared prosperity, measured as the difference in the average annual growth in income or consumption of the poorest 40 percent of population and that of total population, is used as the dependent variable in column 6.

The variable of interest, \bar{x}_i , is the average level of informality in country i over the period 1990-2018, including the share of DGE/MIMIC-based estimates of informal output in official GDP, the share of self-employed in employed. All regressions control for income per capita, measured as the logged real GDP per capita in 2010 U.S. dollars averaged between 1990 and 2018 ($\overline{\text{LnGDPpc}_i}$). The proxies for poverty, income inequality, and shared prosperity are taken from World Development Indicators (WDI).

Declines in informality, poverty reduction and income equalization. The association between within-country changes in informality and poverty and inequality reduction is explored using a similar sample setup and methodology as in Dollar and Kraay (2002) and Dollar, Kleineberg, and Kraay (2013). In particular, the sample of country-year observations is assembled by starting with the first available observation for each country and selecting all available consecutive observations with at least a five-year distance between them (sampling window). This approach yields 428 country-year pairs for 32 advanced economies 119 EMDEs with at least 2 observation per country and a median of 4 observations per country. A median distance between observations is 5.5 years for EMDEs and 5 years for advanced economies. The sample excludes fragile and conflict states.

In table 4.7, the dependent variables are changes in poverty rates at US\$1.90 and US\$3.20 per day (in PPP terms) poverty lines at the end of the sampling window. In table 4.8, the dependent variables are changes in the Gini coefficient and shared prosperity at the end of the sampling window, where shared prosperity refers to the income share of the bottom 40 percent of the population.

The main variable of interest is the cumulative change of output (or employment) informality during the sampling windows. Employment informality is proxied by self-employment in percent of total employment, while output informality is measured by DGE-based estimates on informal output in percent of official GDP.

Additional control variables include initial poverty/inequality levels, which are measured at the start of the sampling windows, to capture persistence in poverty/inequality outcomes; initial levels of informality, cumulative GDP per capita growth during the sampling window; a constant; country and time fixed effects; and squared initial informality to control for the possible non-linear relationship between informality and poverty.

ANNEX 4C Bayesian Model Averaging approach

Model uncertainty is a common issue in regressions that investigate the correlates of informality. Past theoretical models and empirical studies have identified many potential drivers and implications of informality, ranging from social and economic factors underlying underdevelopment to institutional conditions (Schneider, Buehn, and Montenegro 2010; World Bank 2019). Bayesian Model Averaging (BMA) can address model uncertainty formally—by recognizing that the identity of the true model is unknown and that it may be preferable to combine evidence from many different models. Here the BMA model is used to show the potential correlates of output informality in EMDEs. A hyper-g prior is used for each coefficient, following Feldkircher and Zeugner (2012), which may achieve greater robustness than the priors used in the earlier literature. Priors on the inclusion probabilities are discussed below.

Grouping variables. Multiple variables can represent the same broad concepts. For example, both the share of population with primary schooling and above and the share of population with secondary schooling and above can proxy for the quality of human capital in that country. Bayesian approaches should be designed to take this into account (Durlauf, Kourtellos, and Tan 2008; Ghosh and Ghattas 2015). In the analysis underlying this chapter, variables that represent common concepts are grouped together following Dieppe (2020) and Durlauf, Kourtellos, and Tan (2008). As in their work, a group is deemed relevant if the posterior probability of including at least one variable from the group exceeds the prior inclusion probability. To account for the dependency within groups, the prior inclusion probability of each variable is defined as:

$$m_j^i = 1 - (1 - p_j)^{\frac{1}{k_j}}$$

where m_j^i , p_j , and k_j are the prior inclusion probability of variable i in the group j , the probability of including at least one variable from the group j , and the number of variables in group j , respectively. m_j^i is set so that the prior probability of including at least one variable out of the k_j variables in the group is equal to p_j . The quantity p_j is set

to 0.5 for all j , so there is no specific prior knowledge on the probability of a group's inclusion. Posterior distributions of the coefficients of the variables obtained from BMA are aggregated to the group level. The marginal impact of a group is defined as follows:

$$\beta_j^G = \sum_{\{i \in \text{Group } j\}} \beta_i PIP_i \delta_{j,i}$$

where β_j^G is the marginal impact of the group j , β_i is a posterior mean of variable given inclusion of the variable, PIP_i is a posterior inclusion probability of a variable, and $\delta_{j,i}$ is the factor loading of variable i in group j . A factor of group j is defined as the variable within a group whose coefficient posterior mean multiplied by the posterior inclusion probability is the highest. $\delta_{j,i}$ is the coefficient from the linear regression of a variable on the factor. β_j^G can be interpreted as the marginal impact of the factor, accounting for the correlations of the variables within groups. It can also be interpreted as the hypothetical posterior mean when including only one variable per group. In a linear regression, the factor-loading weighted sum of the coefficients is identical to the coefficient obtained by another regression (one that includes one variable per group).

Empirical specification. Following Loayza, Servén, and Sugawara (2010), analyses are carried out at the country level as the dependent variable. The explanatory variables ($\mathbf{X}_{j,t}$) are those identified as potential drivers or implications of informality by former theoretical or empirical studies.

Former studies show that the relationships between informality and its correlates in EMDEs may differ from those in advanced economies (Wu and Schneider 2019). To mitigate this issue, the analysis here includes only EMDEs. Based on data availability, the final sample covers an unbalanced panel for about 55 EMDEs over the period 1989-2018. Details about the list of group variables under investigation and their expected signs are summarized in table 4.9.

The level of output informality in country j in the beginning of a five-year period is modelled as follows:

$$y_{j,t} = \mathbf{X}_{j,t} \boldsymbol{\beta} + \theta_t + \varepsilon_{j,t}$$

where $y_{j,t}$ captures levels of informality at time period t in country j . Two measures of informality are used: informality as DGE-based informal output in percent of official GDP and share of self-employment. $\mathbf{X}_{j,t}$ is a vector of variables of interest, taking from the beginning year of period t and country j . θ_t controls for time fixed effects and $\varepsilon_{j,t}$ is a time varying unobserved idiosyncratic factor. To understand the cross-sectional variation of informality, as well as the time-series variation, country fixed effects are not included. The results are summarized in table 4.10 (output informality) and 4.11 (employment informality).

Limitations. The empirical strategy shown above tries to reduce country heterogeneity issues by restricting the sample to a relatively homogenous set. At the same time, it uses the BMA approach to overcome ad hoc variable selection and the arbitrary omission of

variables. However, the approach taken here cannot be used to draw any conclusions about causal effects. Since many explanatory variables (e.g., tax rates, government effectiveness, and financial development) could be viewed as equilibrium outcomes, which are jointly determined with informality, it is hard to draw conclusions about causal effects. Meanwhile, strong and valid instrumental variables that deal with this issue are hard to find. The analysis summarized here should be interpreted as correlates of informality levels. Another limitation of the model here is the omission of cyclical determinants, such as unemployment and macro-economic shocks. Due to the medium (to long)-run focus of the analysis here, cyclical factors that could be associated with short-run fluctuations in informality are not considered.

TABLE 4.1 Data: Meta regression analysis

Study	Countries/ estimates	Sample period	Methodology	Mean wage gap*
Aydin et al. (2010)	1/4	1998-2007	OLS, ML logit	57.75
Baskaya and Hulagu (2011)	1/2	2005-2009	OLS, PSM	15.45
Bargain and Kwenda (2014)	3/6	2001, 2005	OLS, FE	19.19
Botelho and Ponczek (2011)	1/2	1995-2001	OLS, FE	11.76
Earle and Sakova (2000)	6/6	1993, 1994	ML Logit	-13.33
El Badaoui, Strobl, and Walsh (2008)	1/17	2001-2003	OLS, DID, PSM	28.48
El Badaoui, Strobl, and Walsh (2010)	1/6	1994	OLS, PSM	25.65
Funkhouser (1997)	1/4	1991-1992	OLS	23.82
Gindling (1991)	1/1	1982	OLS	28.50
Huber and Rahimov (2014)	1/2	2007	OLS	-34.98
Lehmann and Pignatti (2007)	1/2	2004	OLS	-6.80
Lehmann and Zaiцева (2013)	1/5	2003-2011	OLS, QR, FE	6.90
Magnac (1991)	1/1	1980	OLS	30.30
Marcouiller, de Castilla, and Woodruff (1997)	3/6	1990	OLS	16.50
Nguyen, Nordman, Roubaud (2013)	1/4	2002-2006	OLS, FE	4.83
Nordman, Rakotomana, and Roubaud (2016)	1/6	2000-2004	OLS, FE	15.33
Paratap and Quintin (2006)	1/3	1993-1995	OLS, FE	28.49
Tansel and Kan (2012)	1/6	2006-2009	OLS, FE	11.56

Source: World Bank.

Note: DID = difference-in-difference estimators; OLS = pooled ordinary least squares; FE = fixed effects regression; ML logit = multinomial logit regression; PSM = propensity score matching; QR = quantile regression. The sample covers these EMDE countries: Argentina, Brazil, Colombia, Costa Rica, Czech Republic, Ecuador, El Salvador, Hungary, Madagascar, Mexico, Peru, Poland, Russian Federation, Slovak Republic, South Africa, Tajikistan, Turkey, Ukraine, and Vietnam.

*Average formal sector premium across all estimates, in percent; a negative number indicates a wage penalty for formal sector workers.

TABLE 4.2 Regression: Meta regression analysis

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
μ	0.195*** (0.03)	0.11** (0.04)	0.23*** (0.03)	0.21*** (0.03)	0.14*** (0.04)	0.24*** (0.04)	0.17*** (0.05)	0.18*** (0.06)
<i>Female</i>		0.16* (0.08)			0.15* (0.08)		0.12 (0.08)	0.12 (0.08)
<i>Male</i>		0.14** (0.06)			0.13** (0.06)		0.11* (0.06)	0.10 (0.06)
<i>Fixed Effects</i>			-0.15** (0.07)		-0.13** (0.06)	-0.14** (0.06)	-0.13** (0.06)	-0.13** (0.07)
<i>Self-employed</i>				-0.34* (0.14)		-0.32** (0.13)	-0.25* (0.14)	-0.26* (0.14)
Latin America and the Caribbean								0.00 (0.07)
Europe and Central Asia								-0.03 (0.07)
Adjusted R ²		7.8	5.8	6.4	12.0	11.4	14.8	12.4
Number of obs.	83	83	83	83	83	83	83	83
τ^2	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05
I^2	99.6	99.5	99.4	99.5	99.4	99.4	99.4	99.1

Source: World Bank.

Note: *** p<0.01, ** p<0.05, * p<0.1; standard errors are in parenthesis. Within study standard errors of the estimates are used as weights to correct for the heteroscedasticity. The dependent variable is the informal-formal wage gap estimates by former studies (listed in annex table B4.1.A.1). τ^2 = estimates of across-study variance. I^2 = residual variance due to study heterogeneity.

TABLE 4.3 Labor productivity differential between types of firms (percent)

	Informal firms						Informal versus formal firms
	Manager has higher education	Main owner is male	Services sector	Firm has bank loan	Single-employee firm	Young firm (<=5 years)	
Angola	45.8	70.0	44.9	-60.0	225.0	20.0	-75.5***
Argentina	25.0	200***	0.0	0.0	11.1	-16.7	-92.5***
Burkina Faso	-6.2	-6.2	28.6	6.7	66.7	-10.0	-79.8***
Botswana	89.4*	72.7**	-29.1	100.0	-35.0	-18.2	-89.8***
Côte d'Ivoire	0.0	25.0	66.7**	-40.0	50.0	40.0	-47.5*
Cameroon	-41.7*	36.4	77.8**	-24.0	140.0***	56.2**	-55.8***
Congo, Dem. Rep.	33.3	0.0	36.0**	50.0	50.0***	0.0	10.7
Cabo Verde	133.3	-25.0	185.7	1585**	566.7*	100.0	0.89
Ghana	0.0	12.5	0.0	25.0	66.7***	0.0	-51.8***
Guatemala	25.0	46.7***	33.3**	50.0	57.1***	-20.0	-86.0***
Kenya	50.0***	6.7	-40***	44.0**	12.0	-20.0**	-81.6***
Madagascar	40.0	-33.3	100***	33.3	60.0*	8.3	-88.1***
Mali	13.2	14.3	-19.4	31.4	57.1	-46.2**	-71.3***
Myanmar	80.0*	-11.1	63.6***	11.3	31.2	0.0	-89.1***
Mauritius	66.7*	6.7	114.3***	25.0	6.7	25.0	-82.9***
Nepal	11.1	0.0	0.0	33.3	150.0***	-16.7	-56.5***
Peru	28.6*	12.5	-50***	-11.1	2.9	-7.4	-74.2***
Rwanda	50.0***	28.6**	25.0*	-25.9	50.0***	-11.1	-91.4***
All countries	48.1***	10.2	8.2	20.0**	41.2***	-6.7	-79.4***

Source: World Bank.

Note: Productivity differential between the median informal and the median formal firm (last column) or between median informal firms among different groups of firms (all other columns). For example, "Manager has higher education" shows the difference in the median productivity among informal firms with managers with higher education and those without higher education. Other firm characteristics are not controlled for, hence results are similar but not identical to column (1) in table 4.4. Productivity is defined as annual sales (in 2009 U.S. dollars) relative to the number of workers. "All countries" is the unweighted average across each column. Significance is denoted by *** (1 percent), ** (5 percent), * (10 percent).

TABLE 4.4 Regression: Labor productivity of formal and informal firms

	(1)	(2)	(3)	(4)	(5)
Informal firm:	-1.400***	-0.648***	-1.131***	-1.200***	-1.008***
Yes 1 No 0	(0.091)	(0.184)	(0.131)	(0.121)	(0.160)
Firm age (logs)	0.120***	0.285***	0.118***	0.116**	0.137***
	(0.045)	(0.053)	(0.045)	(0.045)	(0.045)
Firm size (logs, workers)	-0.102***	-0.119***	-0.056*	-0.104***	-0.108***
	(0.027)	(0.027)	(0.032)	(0.028)	(0.028)
Manufacturing:	-0.402***	-0.407***	-0.401***	-0.401***	-0.399***
Yes 1 No 0	(0.056)	(0.056)	(0.056)	(0.056)	(0.056)
Capital city:	0.201***	0.190***	0.187***	0.394***	0.201***
Yes 1 No 0	(0.061)	(0.061)	(0.061)	(0.087)	(0.061)
Manager experience (logs, years)	0.094**	0.141***	0.107***	0.091**	0.190***
	(0.040)	(0.041)	(0.040)	(0.040)	(0.055)
Informal firm * Firm age (logs)		-0.353***			
		(0.069)			
Informal firm * Firm size (logs, workers)			-0.208***		
			(0.066)		
Informal firm * Capital city:				-0.360***	
Yes 1 No 0				(0.114)	
Informal firm * Manager experience (logs, years)					-0.176***
					(0.060)
Country fixed effects	Yes	Yes	Yes	Yes	Yes
Constant	9.013***	8.552***	8.859***	8.909***	8.748***
	(0.131)	(0.164)	(0.149)	(0.139)	(0.162)
Number of observations	10,527	10,527	10,527	10,527	10,527
R ²	0.291	0.296	0.293	0.293	0.292

Source: World Bank.

Note: Standard errors in parenthesis. Significance is denoted by *** (1 percent), ** (5 percent), * (10 percent). OLS regression with labor productivity as dependent variable, as proxied by annual sales (in 2009 U.S. dollars) per worker, based on a sample using World Bank's Enterprise Surveys data collected during 2007-14 for 4,036 informal firms and 7,558 formal firms in 18 countries. "Informal firm" is a dummy variable taking the value of 1 if a firm is unregistered and 0 otherwise. "Manufacturing" is a dummy variable taking the value of 1 if a firm operates in the manufacturing sector and 0 otherwise. "Capital city" is a dummy variable taking the value of 1 if a firm is located in the capital city and 0 otherwise.

TABLE 4.5 Regression: Labor productivity of formal firms facing informal competition

	(1)	(2)	(3)	(4)	(5)
Informal competition	-0.268***	-1.642***	-1.919***	-0.574***	-1.657***
<i>(Proportion of firms in the cell that report competing with informal firms)</i>	(0.067)	(0.602)	(0.618)	(0.059)	(0.307)
Number of workers (logs)	-0.197***	-0.150***	-0.175***	-0.166***	-0.179***
	(0.016)	(0.017)	(0.019)	(0.019)	(0.020)
Firm's age (logs)	0.208***	0.215***	0.296***	0.286***	0.356***
	(0.023)	(0.026)	(0.032)	(0.029)	(0.032)
Firm belongs to manufacturing sector:	0.137***	0.077*	0.164***	0.157***	0.139***
Yes 1 No 0	(0.044)	(0.046)	(0.052)	(0.048)	(0.053)
Firm belongs to retail sector:	0.695***	0.747***	0.896***	0.862***	0.879***
Yes 1 No 0	(0.045)	(0.047)	(0.053)	(0.049)	(0.054)
Top manager is female:	-0.051	-0.125**	-0.128*	-0.086	-0.063
Yes 1 No 0	(0.048)	(0.058)	(0.073)	(0.067)	(0.070)
Exports (proportion of sales)	0.268**	0.403***	0.431***	0.385***	0.397***
	(0.114)	(0.117)	(0.145)	(0.133)	(0.148)
Firm has foreign owners:	0.638***	0.836***	0.821***	0.658***	0.781***
Yes 1 No 0	(0.063)	(0.062)	(0.070)	(0.066)	(0.074)
Log GDP per capita (PPP, 2009 Int'l Dollars)		0.631***			
		(0.043)			
Informal competition * Log GDP per capita		0.138**			
		(0.067)			
Distance to Frontier (<i>Doing Business</i>)			0.031***		
<i>(Higher values imply better regulatory practices)</i>			(0.006)		
Informal competition * DTF			0.022**		
			(0.010)		
Corruption (Governance Indicators)				0.574***	
<i>(Higher values imply less corruption)</i>				(0.048)	
Informal competition * corruption				0.177**	
				(0.085)	
Business Freedom index (Economic Freedom of the World)					0.015***
<i>(Higher values imply less regulation and more freedom for businesses)</i>					(0.003)
Informal competition * Business Freedom index					0.016***
					(0.005)
Constant	8.771***	3.818***	7.469***	9.410***	8.163***
	(0.178)	(0.390)	(0.381)	(0.088)	(0.224)
Country fixed effects	YES	NO	NO	NO	NO
Number of observations	45,996	45,996	44,770	45,996	43,760
R-squared	0.404	0.259	0.184	0.191	0.154

Source: World Bank.

Note: Standard errors in parenthesis. Significance is denoted by *** (1 percent), ** (5 percent), * (10 percent). OLS regression with labor productivity as dependent variable, as proxied by annual sales (in 2009 U.S. dollars) per worker, based on a sample of formal firms only using World Bank's Enterprise Surveys data collected during 2007-14 for 4,036 informal firms and 7,558 formal firms in 18 countries. "Informal competition" is the share of firms in a cell (a group of firms of similar size in the same region and sector) that report competition from informal firms. It is worth mentioning that one could use a firm-level dummy rather than the proportion of formal firms in a cell to proxy informal competition. However, endogeneity concerns may arise because the informal competition faced by a specific firm may also be driven by its productivity. Therefore, the proportion of formal firms facing informal competition in a cell, which would be uncorrelated with the productivity of a specific firm, should be more robust to endogeneity concerns. "Manufacturing" is a dummy variable taking the value of 1 if a firm operates in the manufacturing sector and 0 otherwise. "Capital city" is a dummy variable taking the value of 1 if a firm is located in the capital city and 0 otherwise.

TABLE 4.6 Regression: Informality, poverty, and income inequality

	Extreme poverty headcount ratio	Gini coefficient	Income per capita for bottom 40 percent	Income per capita growth for bottom 40 percent	Difference in income per capita between bottom 40 percent and total population	Difference in income per capita growth between bottom 40 percent and total population
Output informality	0.08	0.16**	-0.01	-0.07**	-0.02	0.01
(DGE estimates)	(0.13)	(0.07)	(0.02)	(0.03)	(0.04)	(0.02)
GDP per capita	-16.12***	1.25*	2.38***	0.70**	-3.72***	0.50***
	(1.45)	(0.67)	(0.25)	(0.32)	(0.37)	(0.18)
Observations	110	110	58	58	58	58
R-squared	0.59	0.05	0.60	0.16	0.68	0.09
Output informality	0.02	0.13**	-0.01	-0.07**	-0.02	0.00
(MIMIC estimates)	(0.13)	(0.06)	(0.02)	(0.03)	(0.03)	(0.02)
GDP per capita	-16.42***	1.09	2.39***	0.71**	-3.71***	0.48***
	(1.44)	(0.69)	(0.25)	(0.32)	(0.37)	(0.18)
Observations	112	112	58	58	58	58
R-squared	0.59	0.04	0.60	0.17	0.68	0.09
Employment informality	0.30***	0.09**	-0.04**	-0.04	-0.02	-0.03**
(Self-employment)	(0.08)	(0.04)	(0.02)	(0.03)	(0.02)	(0.02)
GDP per capita	-10.44***	2.54**	1.62***	0.28	-3.90***	-0.03
	(2.10)	(0.99)	(0.37)	(0.53)	(0.46)	(0.25)
Observations	118	118	60	60	60	60
R-squared	0.64	0.05	0.65	0.14	0.69	0.19

Source: World Bank.

Note: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, *p<0.1. Self-employment is in percent of total employment. Output informality measures are in percent of official GDP.

TABLE 4.7 Regression: Changes in informality and poverty reduction

Poverty at US\$1.90 per day								
	Output informality				Employment informality			
	World (1)	World (2)	EMDEs (1)	EMDEs (2)	World (1)	World (2)	EMDEs (1)	EMDEs (2)
Lagged poverty	-0.73*	-0.73*	-0.76*	-0.75*	-0.42*	-0.42*	-0.44*	-0.44*
	(0.07)	(0.07)	(0.07)	(0.07)	(0.09)	(0.09)	(0.11)	(0.11)
Change in informality	0.72**	0.69*	0.63	0.60**	0.10	0.08	0.112	0.004
	(0.38)	(0.33)	(0.40)	(0.36)	(0.06)	(0.08)	(0.08)	(0.07)
x lagged poverty		0.001		0.001		0.001		0.01
		(0.02)		(0.01)		(0.03)		(0.008)
Lagged informality	2.53*	2.53*	2.61*	2.62*	-0.09	-0.11	-0.34	-0.47
	(0.82)	(0.84)	(0.88)	(0.91)	(0.23)	(0.25)	(0.40)	(0.38)
Squared lagged informality	-0.016**	-0.016**	-0.017*	-0.017*	0.001	0.001	0.004	0.005
	(0.008)	(0.008)	(0.008)	(0.009)	(0.003)	(0.003)	(0.005)	(0.005)
GDP growth	-0.08*	-0.08*	-0.07	-0.07	-0.04	-0.04	-0.02	-0.02
	(0.04)	(0.04)	(0.05)	(0.05)	(0.03)	(0.03)	(0.04)	(0.04)
R ²	0.66	0.66	0.69	0.69	0.65	0.65	0.74	0.75
Observations	366	366	262	262	266	266	164	164
Number of countries	117	117	85	85	92	92	60	60

Poverty at US\$3.20 per day								
	Output informality				Employment informality			
	World (1)	World (2)	EMDEs (1)	EMDEs (2)	World (1)	World (2)	EMDEs (1)	EMDEs (2)
Lagged poverty	-0.59*	-0.59*	-0.65*	-0.65*	-0.42*	-0.41*	-0.46*	-0.49*
	(0.07)	(0.07)	(0.08)	(0.08)	(0.1)	(0.1)	(0.13)	(0.135)
Change in informality	1.87*	1.87*	1.74**	1.74**	-0.23	0.20**	-0.57*	0.18
	(0.58)	(0.58)	(0.60)	(0.60)	(0.19)	(0.10)	(0.26)	(0.13)
x lagged poverty	-0.02**		-0.02*		0.014*		0.03*	
	(0.01)		(0.01)		(0.006)		(0.009)	
Lagged informality	2.88*	2.88*	2.88*	2.88*	-0.88**	-0.49	-1.53*	-0.98
	(0.98)	(0.98)	(1.04)	(1.04)	(0.45)	(0.42)	(0.65)	(0.67)
Squared lagged informality	-0.026*	-0.026*	-0.03*	-0.03*	0.01*	0.005	0.017*	0.011
	(0.01)	(0.01)	(0.01)	(0.01)	(0.005)	(0.005)	(0.008)	(0.007)
GDP growth	-0.15*	-0.15*	-0.12*	-0.12*	-0.14*	-0.14*	-0.15*	-0.17*
	(0.04)	(0.04)	(0.05)	(0.05)	(0.05)	(0.05)	(0.06)	(0.07)
R ²	0.59	0.59	0.64	0.64	0.64	0.61	0.72	0.68
Observations	366	366	262	262	266	266	164	164
Number of countries	117	117	85	85	92	92	60	60

Source: World Bank.

Note: Robust standard errors in parentheses; *p<0.05, ** p<0.1. The sample of country-year observations starts with the first available observation for each country and all consecutive observations with at least five-year minimum window between them. Sample excludes fragile and conflict states. "Employment informality" is self-employment in percent of total employment. "Output informality" is proxied by DGE-based estimates on informal output in percent of official GDP. Dependent variables are changes in poverty rates at US\$1.90 and US\$3.20 per day. "GDP growth" denotes cumulative change in GDP per capita during the time window between observations. "Change in informality" denotes change in output (employment) informality during the time window between observations in percent of official GDP (in percent of official employment). "Lagged informality" denotes informality levels at t-1. Control variables include a constant, country and time fixed effects, and squared lagged informality to control for a possible non-linear relationship between informality and poverty. A positive coefficient on change in informality indicates that formalization reduces poverty.

TABLE 4.8 Regression: Changes in informality and improvement in income inequality

	Gini coefficient				Shared prosperity			
	Output informality		Employment informality		Output informality		Employment informality	
	World	EMDEs	World	EMDEs	World	EMDEs	World	EMDEs
Lagged inequality	-0.90*	-0.84*	-0.86*	-0.80*	-0.90*	-0.85*	-0.84*	-0.77*
	(0.06)	(0.07)	(0.09)	(0.12)	(0.06)	(0.07)	(0.09)	(0.12)
Change in informality	-0.18	-0.22	0.02	0.006	0.10	0.12	-0.006	-0.01
	(0.17)	(0.18)	(0.07)	(0.08)	(0.07)	(0.07)	(0.03)	(0.04)
Lagged informality	-0.28	-0.36	0.29	0.28	0.16	0.21	-0.12	-0.14
	(0.41)	(0.46)	(0.18)	(0.29)	(0.20)	(0.22)	(0.08)	(0.13)
Squared lagged informality	0.006	0.007	-0.004*	-0.004	-0.003	-0.003	0.002*	0.002
	(0.004)	(0.005)	(0.002)	(0.003)	(0.002)	(0.002)	(0.001)	(0.001)
GDP growth	0.01	0.02	0.012	0.04	-0.007	-0.009	-0.003	-0.02
	(0.02)	(0.03)	(0.030)	(0.05)	(0.014)	(0.014)	(0.014)	(0.023)
R ²	0.70	0.72	0.65	0.70	0.68	0.70	0.65	0.71
Observations	366	262	266	164	366	262	366	164
Number of countries	117	85	92	60	117	85	117	60

Source: World Bank.

Note: Robust standard errors in parentheses; *p<0.05, ** p<0.1. "Shared prosperity" refers to the income share of the bottom 40 percent of population. The sample of country-year observations starts with the first available observation for each country and all consecutive observations with at least five-year minimum window between them. Sample excludes fragile and conflict states. "Employment informality" is self-employment in percent of total employment and "output informality" is proxied by DGE-based estimates on informal output in percent of official GDP. Dependent variables are two measures of inequality: the change in Gini coefficient and shared prosperity. "GDP growth" denotes cumulative change in GDP per capita during the time window between observations. "Change in informality" denotes change in output (employment) informality during the time window between observations in percent of official GDP (in percent of official employment). "Lagged informality" denotes informality levels at t-1. Control variables include a constant, country and time fixed effects, and squared lagged informality to control for a possible non-linear relationship between informality and inequality. A positive coefficient on change in informality indicates that formalization reduces shared prosperity (the income share of the bottom 40 percent of population).

TABLE 4.9 Data: Bayesian Model averaging approach

Group	Expected sign (group)	Variable	Source
Economic development	(-)	Share of population in urban area	WDI
		Labor productivity (in logs)	WDI
		Share of manufacturing (in percent of value added)	WDI
		Share of manufacturing (in percent of employment)	WDI
Human capital	(-)	Human capital	UNDP
		Share of population with primary schooling and above	Wittgenstein Center for Demography and Global Human Capital
		Share of population with secondary schooling and above	
Financial development	(-)	Domestic credit to private sector (in percent of GDP)	WDI
Trade openness	(-/+)	GVC position index	UNCTAD
		GVC participation index	UNCTAD
Governance	(-)	Political rights index	Freedom house
		Civil rights index	Freedom house
		Accountability	WGI
		Regulatory Quality	WGI
		Control of Corruption	WGI
		Political stability	WGI
Tax burdens	(+)	Government expenditure (in percent of GDP)	WDI
		Tax revenues (in percent of GDP)	IMF GFS
Access to public infrastructure	(-)	Electricity consumption per capita	WDI
		Access to Basic Water, rural/Urban	WDI
		Air goods transported	WDI
		Road length per capita	World Road Statistics

Source: World Bank

Note: The expected signs are for each group of variables. They summarize the relationship between each group of variables and the level of informality suggested by past studies. “(-)” (“+”) suggests that an increase in the corresponding group variable would be associated with a fall (increase) in output informality. Among all groups, only the group of variables on tax burdens are expected to be positively associated with output informality, suggesting that higher tax burdens are associated with a higher level of informal output. In the case of trade openness, its relationship with informality could be either positive or negative, as suggested by former studies (chapter 6).

TABLE 4.10 PIP and posterior means (output informality)

Variable	PIP	Posterior Mean	Group	Group PIP	Group coef	Group coef positive prob.
Share of manufacturing (percent of value added)	0.22	-0.30				
Labor productivity (log)	0.05	0.00				
Share of population in urban area	0.06	-0.03	Economic development	0.88	-1.66	0.00
Share of manufacturing (percent of employment)	0.67	-1.08				
Agriculture export share	0.30	-0.25				
Share of population with primary schooling and above	0.34	0.15				
Share of population with secondary schooling and above	0.08	0.00	Human capital	0.92	-1.49	0.01
Human capital	0.88	-1.65				
Domestic credit to private sector (percent of GDP)	0.13	-0.14	Financial development	0.13	-0.14	0.01
Civil rights index	0.18	-0.01				
Political rights index	0.20	-0.31				
Accountability	0.21	-0.03	Governance	0.96	-2.18	0.00
Control of corruption	0.70	-3.31				
Regulatory quality	0.58	2.18				
Political stability	0.42	-0.70				
Electricity consumption per capita	0.28	0.15				
Access to basic water, rural	0.29	0.32				
Access to basic water, urban	0.81	1.40	Access to public infrastructure	1.00	-0.36	0.26
Road length per area (log)	0.28	0.03				
Air goods transported (log)	0.97	-2.26				
Government expenditure (percent of GDP)	0.12	0.09	Tax burdens	0.25	0.21	0.99
Tax revenues (percent of GDP)	0.14	0.12				
GVC participation index	0.16	0.12	Trade openness	0.25	0.08	0.64
GVC position index	0.10	-0.05				

Source: World Bank.

Note: Based on the panel regression result using Bayesian Model Averaging technique. The dependent variable is the share of informal economy in GDP using dynamic general equilibrium-based measures. From unbalanced panel from 67 emerging market and developing economies (EMDEs), using five-year panel from 1998-2018.

TABLE 4.11 PIP and posterior means (employment informality)

Variable	PIP	Posterior mean	Group	Group PIP	Group coef	Group coef positive prob.
Share of manufacturing (percent of value added)	0.29	-1.04				
Labor productivity (log)	0.22	-0.01				
Share of population in urban area	0.03	-0.04	Economic development	0.99	-7.39	0.00
Share of manufacturing (percent of employment)	0.48	-1.71				
Agriculture export (percent exports)	0.98	-4.59				
Share of population with primary schooling and above	0.24	0.25				
Share of population with secondary schooling and above	0.12	0.34	Human capital	1.00	-8.02	0.00
Human capital	1.00	-8.60				
Domestic credit to private sector (percent of GDP)	0.09	0.23	Financial development	0.09	0.23	0.95
Civil rights index	0.32	1.36				
Political rights index	0.07	0.18				
Accountability	0.37	2.07				
Control of Corruption	0.74	-5.76	Governance	1.00	-7.74	0.00
Regulatory quality	0.04	0.03				
Political stability	0.94	-5.62				
Electricity consumption per capita	0.59	-6.62				
Access to basic water, rural	0.58	2.87				
Access to basic water, urban	0.55	2.38				
Road length per area (log)	0.96	-4.81	Access to public infrastructure	1.00	-6.59	0.16
Air goods transported (log)	0.16	-0.48				
Government expenditure (percent of GDP)	0.06	-0.10				
Tax revenues (percent of GDP)	0.07	0.16	Tax burdens	0.11	0.06	0.59
GVC participation index	0.06	0.05				
GVC position index	0.11	0.23	Trade openness	0.17	0.28	0.93

Source: World Bank.

Note: Based on the panel regression result using Bayesian Model Averaging technique. The dependent variable is the share self employment. From unbalanced panel from 64 emerging market and developing economies, using five-year panel from 1998-2018.

TABLE 4.12 Correlates of employment informality in EMDEs

	High informality	Low informality	P-value for t-test
SDG global index rank	113.6	77.0	0.00
Extreme poverty headcount (2000, percent of population)	39.7	11.6	0.01
Extreme poverty headcount (latest, percent of population)	24.0	5.4	0.00
Agriculture sector (value added, percent of GDP)	19.6	5.4	0.00
Agriculture sector (employment, percent of employment)	45.5	10.8	0.00
Finance constraint (percent of firms)	32.3	28.9	0.24
Bank finance (percent of firms)	20.7	30.9	0.00
Internal finance (percent of investment)	73.6	66.3	0.00
Basic training	1.9	2.0	0.31
Business infrastructure	2.8	3.0	0.09
Physical infrastructure	3.6	3.4	0.03
Adequacy of social insurance programs (percent of household income)	27.4	34.1	0.04
Coverage of unemployment benefits (percent of population)	5.0	5.2	0.91
GDP per capita (in thousands of 2010 U.S. dollars)	2.0	9.6	0.00
Access to credit (percent of GDP)	24.3	43.1	0.00
Human capital (years of schooling)	5.2	8.1	0.00
Trade openness (percent of GDP)	71.9	92.9	0.00
Labor productivity (in thousands of 2010 U.S. dollars)	5.5	21.3	0.00
Per capita income of bottom 40 percent (2011 PPP\$ per day)	3.2	6.6	0.00
Per capita income growth of bottom 40 percent (2011 PPP\$ per day)	1.5	3.1	0.02
Government revenues (percent of GDP)	20.4	31.9	0.00
Tax revenues (overall, percent of GDP)	13.4	19.0	0.00
Tax revenues (income, percent of GDP)	4.4	6.0	0.00
Tax revenues (imports, percent of GDP)	1.3	0.3	0.00
Government expenditures (overall, percent of GDP)	23.2	32.8	0.00
Government expenditures (education, percent of GDP)	4.0	4.6	0.00
Government expenditures (health, percent of GDP)	2.2	3.6	0.00
Undernourishment (percent of population)	17.7	6.6	0.00
Stunting (percent of population)	28.1	12.1	0.00
Agricultural productivity (tons per hectare)	2.3	4.3	0.00
Maternal mortality (per 10,000 births)	267.4	65.1	0.00
Under 5 mortality (per 1,000 births)	43.8	17.4	0.00

TABLE 4.12 Correlates of employment informality in EMDEs (continued)

	High informality	Low informality	P-value for t-test
Death rate from pollution (per 100,000 persons)	136.7	67.2	0.00
Life expectancy at birth (years)	67.4	73.4	0.00
Net primary school enrollment	88.6	90.6	0.24
Literacy rate (percent of people aged 15-24)	84.6	97.9	0.00
Female years of schooling (percent of male schooling)	77.8	97.8	0.00
Family planning	55.4	65.3	0.00
Female informal workers	86.6	50.3	0.00
Doing business score	53.8	61.8	0.00
Cost of business start-up procedures (percent of GNI per capita)	79.9	22.8	0.00
Bureaucracy quality	1.49	2.03	0.00
Control of corruption	2.27	2.57	0.01
Law and order	3.02	3.59	0.00
Quality of overall infrastructure	4.0	3.3	0.00
Access to sanitation services (percent of population)	55.6	88.8	0.00
Access to drinking water services (percent of population)	78.3	95.1	0.00
Access to electricity (percent of population)	69.4	96.3	0.00
Access to clean fuels (percent of population)	36.7	83.2	0.00
Access to Internet (percent of population)	35.2	65.3	0.00
Access to mobile broadband (percent of population)	46.2	73.4	0.00
Paved road (percent of total roads)	14.0	47.8	0.00
Road project unit costs (all, \$U.S. million per km)	0.5	0.4	0.22
Road project unit costs (reconstruction, \$U.S. million per km)	0.5	0.3	0.22

Source: World Bank.

Note: See the notes to figure 4.1-4.8 for detailed definitions and data sources. The column "high (low) informality" shows simple averages for emerging market and developing economies (EMDEs) with above (below)-median employment informality (proxied by self-employment as a share of total employment).

TABLE 4.13 Correlates of informality in EMDEs: MIMIC-based informal activity and WEF index

	MIMIC-based output informality			Perceived informality (WEF)		
	High	Low	P-val	High	Low	P-val
GDP per capita	7.3	8.4	0.00	7.4	8.4	0.00
Access to credit	23.0	39.1	0.00	22.7	42.3	0.00
Human capital	5.8	7.1	0.01	5.7	7.5	0.00
Trade openness	73.7	82.3	0.15	7.0	8.7	0.00
Doing business score	53.3	59.8	0.00	53.3	62.1	0.00
Cost of business start-up procedures	96.2	26.8	0.00	86.7	22.5	0.00
Bureaucracy quality	1.4	2.1	0.04	1.5	2.0	0.00
Control of corruption	2.2	2.6	0.00	2.3	2.6	0.00
Law and order	2.9	3.6	0.00	2.9	3.8	0.00

Sources: Barro and Lee (2013); *International Country Risk Guide (ICRG)*; World Bank (*Doing Business*, World Development Indicators).

Note: Data are from emerging market and development economies (EMDEs) and the period 1990-2018. "High" are EMDEs with above median Multiple Indicators and Multiple Causes model (MIMIC)-based informal output measures (perceived informality measured by the World Economic Forum's index, reversed order), while "Low" are EMDEs with below median MIMIC-based informal output measures (perceived informality). "P-val" shows the p-values of the t-tests conducted for the group comparisons. The correlates include GDP per capita (in logs; 2010 U.S. dollars); access to credit (domestic credit to private sector in percent of GDP); human capital (average years of schooling); trade openness (the sum of imports and exports in percent of GDP); the ease of doing business score (the score is reflected on a scale from 0 to 100, where 0 represents the lowest and 100 represents the best performance); bureaucracy quality; control of corruption; and law and order (ICRG). Higher values corresponding to better outcomes.

TABLE 4.14 Regression: Developmental challenges and DGE-based output informality in EMDEs

	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
Dependent variable =	[1]	[2]	[3]	[4]	[5]	[6]
SDG global index rank	84.500*** (6.380)	3.638 (9.165)	28.017*** (9.973)	20.362** (9.837)	117	0.088
Extreme poverty headcount (2000, percent of population)	13.140** (5.439)	13.760 (17.815)	29.560** (13.628)	40.310** (18.508)	17	0.298
Extreme poverty headcount (latest, percent of population)	7.336*** (1.908)	1.418 (3.204)	23.579*** (5.318)	12.435** (4.925)	111	0.199
Agriculture sector (value added, percent of GDP)	4.846*** (1.074)	3.589** (1.705)	14.358*** (2.281)	15.348*** (2.288)	157	0.298
Agriculture sector (employment, percent of employment)	10.366*** (2.372)	10.350*** (3.907)	32.082*** (4.433)	34.126*** (3.939)	157	0.360
Finance constraint (percent of firms)	24.295*** (2.325)	5.775* (3.371)	14.365*** (3.958)	8.257* (4.201)	109	0.108
Bank finance (percent of firms)	28.650*** (2.406)	-0.276 (3.482)	-8.426** (3.324)	-10.278*** (3.367)	109	0.126
Internal finance (percent of investment)	68.843*** (2.015)	-1.446 (3.152)	6.834** (2.775)	5.311* (3.165)	109	0.086
Basic training	1.937*** (0.102)	0.128 (0.127)	-0.076 (0.149)	-0.028 (0.123)	68	0.041
Business infrastructure	2.941*** (0.075)	-0.001 (0.085)	-0.053 (0.104)	-0.081 (0.094)	68	0.019
Physical infrastructure	3.675*** (0.088)	-0.201 (0.129)	-0.296** (0.121)	-0.230 (0.157)	68	0.070
Adeq. of social insurance program (percent household income)	36.281*** (3.530)	-6.164 (4.372)	-10.095* (5.335)	-6.798 (4.811)	93	0.050
Coverage of unemployment benefits (percent of population)	6.519*** (1.823)	0.326 (2.399)	-2.290 (2.069)	-3.653* (2.025)	59	0.092
GDP per capita (logs)	8.769*** (0.220)	-0.697** (0.293)	-1.645*** (0.296)	-1.317*** (0.274)	121	0.263
Access to credit (percent of GDP)	44.818*** (4.988)	-9.466 (6.533)	-23.791*** (5.731)	-18.603*** (6.649)	122	0.139
Human capital (years of schooling)	7.253*** (0.363)	-0.524 (0.573)	-1.432* (0.747)	-1.270** (0.594)	99	0.055
Trade openness (percent of GDP)	83.269*** (6.215)	-8.021 (8.878)	-7.312 (8.266)	-9.606 (8.375)	119	0.013
Labor productivity (in thousands of 2010 U.S. dollars)	27.342*** (5.092)	-15.96*** (5.338)	-20.522*** (5.270)	-20.254*** (5.328)	116	0.235
Per capita income of bottom 40 percent (2011 PPP\$ per day)	6.344*** (0.790)	-1.911* (1.054)	-2.426** (1.206)	-2.044* (1.107)	58	0.095
Per capita income growth of bottom 40 percent (2011 PPP\$ per day)	3.215*** (0.702)	0.307 (1.178)	-1.581* (0.841)	-2.270** (1.071)	58	0.132

TABLE 4.14 Regression: Developmental challenges and DGE-based output informality in EMDEs (continued)

	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
Dependent variable =	[1]	[2]	[3]	[4]	[5]	[6]
Government revenues (percent of GDP)	26.745*** (1.678)	-1.311 (2.704)	-4.068 (2.459)	-4.735* (2.578)	83	0.050
Tax revenues (overall, percent of GDP)	15.655*** (1.297)	-0.020 (1.705)	-0.730 (1.892)	-1.706 (1.605)	83	0.016
Tax revenues (income, percent of GDP)	5.517*** (0.677)	-0.929 (0.825)	-1.075 (0.906)	-1.212 (0.815)	83	0.035
Tax revenues (imports, percent of GDP)	0.968* (0.492)	0.152 (0.593)	0.317 (0.546)	0.023 (0.549)	69	0.008
Government expenditures (overall, percent of GDP)	28.156*** (1.755)	0.071 (2.685)	-3.477 (2.357)	-5.270** (2.520)	83	0.075
Government expenditures (education, percent of GDP)	4.332*** (0.247)	-0.233 (0.391)	-0.085 (0.373)	-0.610* (0.365)	118	0.025
Government expenditures (health, percent of GDP)	3.240*** (0.287)	-0.372 (0.417)	-1.090*** (0.345)	-0.838** (0.407)	122	0.076
Pupil - teacher ratio (primary education)	21.369*** (1.489)	4.821* (2.561)	14.413*** (3.008)	12.430*** (3.584)	120	0.162
Trained teachers in primary edu (percent of teachers)	86.299*** (3.034)	0.438 (3.891)	-12.399** (4.954)	-0.440 (4.589)	103	0.100
Undernourishment (percent of population)	8.738*** (1.667)	0.830 (2.137)	6.340** (2.797)	10.087*** (3.440)	113	0.119
Stunting (percent of population)	15.486*** (2.390)	3.293 (3.288)	11.487*** (3.493)	9.455*** (3.197)	118	0.121
Agricultural productivity (tons per hectare)	4.778*** (0.812)	-1.697* (0.876)	-2.550*** (0.841)	-1.726 (1.175)	119	0.074
Maternal mortality (per 10,000 births)	82.677*** (20.641)	50.356 (38.247)	206.452*** (50.318)	198.189*** (58.552)	122	0.146
Under 5 mortality (per 1,000 births)	19.368*** (3.040)	7.919 (5.157)	28.871*** (6.358)	23.192*** (7.116)	122	0.156
Death rate from pollution (per 100,000 persons)	75.419*** (9.292)	11.181 (14.843)	70.000*** (15.546)	46.281*** (17.098)	122	0.155
Life expectancy at birth (years)	72.871*** (1.098)	-0.784 (1.491)	-6.739*** (1.593)	-5.198*** (1.767)	122	0.169
Net primary school enrollment	91.268*** (1.448)	-1.443 (2.238)	-6.683** (3.163)	-3.401 (2.312)	117	0.053
Literacy rate (percent of people aged 15-24)	96.585*** (0.942)	-3.306 (2.093)	-14.395*** (3.562)	-10.544*** (3.538)	118	0.142
Female years of schooling (percent of male schooling)	8.497*** (0.400)	-0.930 (0.587)	-2.210*** (0.735)	-1.313** (0.624)	122	0.083
Family planning (percent of women with needs)	69.000*** (2.709)	-8.123* (4.642)	-16.607*** (4.304)	-11.750** (4.902)	121	0.094

TABLE 4.14 Regression: Developmental challenges and DGE-based output informality in EMDEs (continued)

Dependent variable =	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
	[1]	[2]	[3]	[4]	[5]	[6]
Female informal workers (percent of female workers)	60.138*** (4.865)	11.028 (7.157)	19.612** (8.026)	17.188** (6.939)	67	0.112
Doing business score	62.683*** (1.253)	-4.368* (2.258)	-9.435*** (2.099)	-8.453*** (2.481)	122	0.133
Cost of business start-up procedures (percent of GNI per capita)	17.574*** (4.245)	18.013** (8.728)	70.237*** (17.750)	78.590*** (22.164)	122	0.153
Bureaucracy quality	2.271*** (0.105)	-0.340* (0.184)	-0.927*** (0.173)	-0.824*** (0.148)	95	0.279
Control of corruption	2.808*** (0.108)	-0.394** (0.151)	-0.550*** (0.142)	-0.640*** (0.166)	95	0.183
Law and order	4.029*** (0.171)	-0.746*** (0.255)	-1.147*** (0.220)	-1.050*** (0.255)	95	0.230
Quality of overall infrastructure	4.295*** (0.152)	-0.576** (0.227)	-1.158*** (0.201)	-0.936*** (0.211)	108	0.246
Access to sanitation services (percent of population)	83.694*** (3.702)	-4.642 (5.880)	-27.935*** (6.741)	-25.588*** (6.561)	122	0.179
Access to drinking water services (percent of population)	93.046*** (1.952)	-4.307 (3.300)	-15.721*** (3.660)	-14.017*** (3.886)	122	0.157
Access to electricity (percent of population)	92.750*** (2.817)	-3.920 (4.544)	-28.945*** (6.764)	-19.955*** (6.152)	122	0.176
Access to clean fuels (percent of population)	82.268*** (4.393)	-14.314* (7.399)	-43.833*** (8.216)	-37.288*** (8.000)	120	0.225
Access to Internet (percent of population)	65.353*** (4.137)	-9.966* (5.811)	-32.517*** (6.151)	-26.623*** (5.780)	122	0.245
Access to mobile broadband (percent of population)	83.712*** (8.220)	-18.885* (9.981)	-42.583*** (9.599)	-33.104*** (10.050)	122	0.182
Paved road (percent of total roads)	40.100*** (8.154)	-1.265 (12.626)	-19.772** (9.118)	-29.360*** (8.416)	45	0.247
Road project unit costs (all, \$US million per km)	0.389*** (0.108)	-0.085 (0.124)	0.156 (0.254)	0.209 (0.186)	67	0.037
Road project unit costs (reconstruction, \$US million per km)	0.441*** (0.157)	-0.149 (0.188)	0.184 (0.412)	-0.094 (0.223)	31	0.044

Source: World Bank.

Note: See the notes to figures 4.1-4.8 for detailed definitions and data sources for various dependent variables. Development outcomes are regressed against a set of dummies that categorize emerging market and developing economies (EMDEs) into quantiles of shares of informality using dynamic general equilibrium (DGE)-based estimates on informal output in percent of official GDP. The constant from the regression results show the development outcomes for the lowest quartile; coefficients show the difference between the corresponding quartile and the lowest quartile. All regressions are cross-sectional, with variables averaged during (up to) 1990-2018 (or otherwise specified period or latest year available). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

TABLE 4.15 Regression: Developmental challenges and self-employment in EMDEs

Dependent variable =	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
	[1]	[2]	[3]	[4]	[5]	[6]
SDG global index rank	67.033*** (5.576)	22.467*** (8.264)	26.733*** (8.004)	65.863*** (7.356)	119	0.382
Extreme poverty headcount (2000, percent of population)	9.283* (5.126)	5.577 (7.034)	12.457 (9.287)	53.497*** (7.376)	21	0.752
Extreme poverty headcount (latest, percent of population)	4.163*** (1.210)	2.420 (2.432)	7.777** (3.392)	32.323*** (3.835)	119	0.439
Agriculture sector (value added, percent of GDP)	3.241*** (0.612)	4.387*** (1.109)	9.684*** (1.243)	23.135*** (1.808)	170	0.580
Agriculture sector (employment, percent of emp.)	6.184*** (1.148)	9.979*** (1.977)	25.441*** (2.190)	53.426*** (2.602)	161	0.765
Finance constraint (percent of firms)	26.429*** (2.172)	5.117 (3.743)	-3.481 (2.945)	15.571*** (3.699)	114	0.220
Bank finance (percent of firms)	34.543*** (2.214)	-7.428** (3.669)	-11.221*** (3.004)	-16.531*** (3.009)	114	0.196
Internal finance (percent of investment)	65.445*** (2.072)	1.701 (3.161)	3.691 (3.419)	12.792*** (2.680)	114	0.147
Basic training	2.000*** (0.084)	0.013 (0.114)	-0.084 (0.129)	-0.079 (0.125)	72	0.015
Business infrastructure	2.966*** (0.056)	-0.021 (0.077)	-0.140* (0.081)	-0.095 (0.100)	72	0.044
Physical infrastructure	3.643*** (0.079)	-0.093 (0.121)	-0.159 (0.126)	-0.395*** (0.148)	72	0.111
Adequacy of social insurance programs (percent of household income)	36.782*** (3.319)	-5.376 (4.421)	-8.182* (4.350)	-10.646** (5.074)	100	0.058
Coverage of unemployment benefits (percent of population)	6.644*** (1.705)	-3.069 (1.851)	-1.378 (2.385)	-1.909 (2.021)	61	0.044
GDP per capita (in logs)	9.131*** (0.153)	-0.862*** (0.190)	-1.471*** (0.221)	-2.461*** (0.185)	133	0.580
Access to credit (percent of GDP)	50.592*** (4.629)	-15.457*** (5.709)	-19.843*** (5.964)	-32.891*** (5.228)	131	0.245
Human capital (years of schooling)	8.626*** (0.258)	-1.131** (0.499)	-2.201*** (0.463)	-4.686*** (0.406)	99	0.500
Trade openness (percent of GDP)	98.526*** (5.327)	-11.076 (10.325)	-24.556*** (7.770)	-29.558*** (7.393)	129	0.095
Labor productivity (in thousands of 2010 dollars)	30.441*** (4.531)	-18.349*** (4.654)	-22.537*** (4.812)	-28.026*** (4.543)	119	0.383
Per capita income of bottom 40 percent (2011 PPP\$ per day)	8.200*** (0.766)	-3.622*** (0.839)	-3.746*** (0.899)	-6.341*** (0.812)	61	0.583
Per capita income growth of bottom 40 percent (2011 PPP\$ per day)	3.393*** (0.812)	-0.506 (1.155)	-1.152 (0.891)	-2.762** (1.100)	61	0.130

TABLE 4.15 Regression: Developmental challenges and self-employment in EMDEs (continued)

Dependent variable =	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
	[1]	[2]	[3]	[4]	[5]	[6]
Road project unit costs (all, \$US million per km)	0.536*** (0.152)	-0.247 (0.171)	0.088 (0.254)	-0.112 (0.203)	69	0.041
Road project unit costs (reconstruction, \$US million per km)	0.447*** (0.114)	-0.241* (0.138)	-0.170 (0.144)	0.282 (0.407)	34	0.112
Government revenues (percent of GDP)	33.473*** (1.829)	-9.163*** (2.608)	-11.665*** (2.409)	-13.253*** (2.227)	83	0.327
Tax revenues (overall, percent of GDP)	19.586*** (1.255)	-6.703*** (1.651)	-4.357** (1.812)	-6.630*** (1.334)	83	0.242
Tax revenues (income, percent of GDP)	5.867*** (0.683)	-1.504* (0.809)	-1.093 (0.913)	-1.853** (0.790)	83	0.076
Tax revenues (imports, percent of GDP)	0.669 (0.435)	1.016 (0.935)	0.484 (0.539)	0.755 (0.489)	70	0.033
Government expenditures (overall, percent of GDP)	33.438*** (2.074)	-5.909** (2.862)	-10.147*** (2.468)	-10.691*** (2.344)	83	0.247
Government expenditures (education, percent of GDP)	4.674*** (0.376)	-0.371 (0.449)	-0.873** (0.432)	-0.475 (0.440)	123	0.042
Government expenditures (health, percent of GDP)	4.089*** (0.371)	-0.812* (0.471)	-1.374*** (0.435)	-2.542*** (0.392)	126	0.277
Pupil - teacher ratio (primary education)	17.443*** (1.066)	5.350*** (1.499)	11.154*** (2.544)	24.265*** (2.106)	125	0.501
Trained teachers in primary edu. (percent of teachers)	85.695*** (2.789)	0.894 (3.856)	-0.401 (4.362)	-9.355* (4.779)	107	0.060
Undernourishment (percent of population)	4.607*** (0.871)	3.935** (1.541)	8.496*** (2.049)	18.253*** (2.606)	123	0.346
Stunting (percent of population)	8.969*** (1.351)	6.473*** (2.329)	14.698*** (2.631)	23.889*** (2.109)	125	0.450
Agricultural productivity (tons per hectare)	5.715*** (1.014)	-2.830*** (1.058)	-2.921*** (1.038)	-3.896*** (1.035)	124	0.188
Maternal mortality (per 10,000 births)	37.242*** (6.469)	55.851* (29.142)	98.539*** (24.735)	368.820*** (45.042)	129	0.434
Under 5 mortality (per 1,000 births)	12.850*** (1.794)	9.253** (4.052)	14.768*** (3.712)	47.802*** (4.704)	133	0.470
Death rate from pollution (per 100,000 persons)	58.061*** (4.963)	19.158* (11.268)	36.939*** (10.401)	122.033*** (11.641)	129	0.466
Life expectancy at birth (years)	74.394*** (0.617)	-1.991* (1.125)	-3.122*** (1.156)	-11.022*** (1.117)	129	0.423
Net primary school enrollment	91.533*** (1.111)	-1.935 (2.087)	0.751 (1.613)	-6.626*** (2.443)	128	0.092
Literacy rate (percent aged 15-24)	98.876*** (0.228)	-2.581** (1.106)	-5.672*** (1.587)	-22.994*** (3.375)	123	0.424

TABLE 4.15 Regression: Developmental challenges and self-employment in EMDEs (continued)

	Lowest quartile (Q1)	2nd quartile (Q2) -Q1	3rd quartile (Q3) -Q1	Highest quartile (Q4) -Q1	Obs	R-sq
Dependent variable =	[1]	[2]	[3]	[4]	[5]	[6]
Female years of schooling (percent of male schooling)	9.788*** (0.234)	-1.300*** (0.442)	-2.239*** (0.475)	-5.222*** (0.384)	131	0.500
Family planning (percent of women with needs)	63.142*** (3.244)	3.554 (4.293)	-1.058 (4.844)	-13.977*** (4.278)	129	0.125
Female informal workers (percent of female workers)	37.304*** (3.654)	26.815*** (5.184)	42.878*** (5.013)	56.091*** (3.835)	70	0.724
Doing business score	64.187*** (1.194)	-4.942*** (1.849)	-6.484*** (1.977)	-14.312*** (1.828)	134	0.293
Cost of business start-up procedures (percent of GNI per capita)	14.904*** (3.551)	16.091** (7.362)	25.022*** (7.597)	106.249*** (17.265)	134	0.353
Bureaucracy quality (ICRG)	2.275*** (0.132)	-0.485*** (0.178)	-0.547*** (0.161)	-1.020*** (0.210)	96	0.253
Control of corruption (ICRG)	2.803*** (0.108)	-0.463*** (0.164)	-0.552*** (0.143)	-0.517*** (0.145)	96	0.161
Law and order (ICRG)	4.055*** (0.152)	-0.921*** (0.255)	-0.944*** (0.218)	-1.134*** (0.217)	96	0.229
Quality of overall infrastructure	4.352*** (0.131)	-0.788*** (0.183)	-0.826*** (0.194)	-1.227*** (0.209)	111	0.270
Access to sanitation services (percent of population)	92.238*** (1.760)	-6.793** (3.213)	-18.251*** (4.457)	-56.234*** (3.942)	133	0.599
Access to drinking water services (percent of population)	96.573*** (0.879)	-3.219** (1.495)	-8.487*** (2.380)	-28.303*** (2.739)	133	0.522
Access to electricity (percent of population)	98.099*** (1.175)	-3.622 (2.521)	-10.983*** (3.811)	-47.426*** (4.852)	133	0.517
Access to clean fuels (percent of population)	88.398*** (3.502)	-10.309* (5.779)	-30.273*** (6.279)	-73.766*** (4.645)	131	0.586
Access to Internet (percent of population)	71.024*** (3.444)	-12.388*** (4.340)	-24.094*** (4.984)	-47.731*** (4.337)	133	0.498
Access to mobile broadband (percent of population)	82.942*** (8.002)	-18.727* (9.703)	-28.295*** (9.093)	-46.673*** (9.114)	133	0.209
Paved road (percent of total roads)	57.567*** (10.160)	-23.393* (12.608)	-41.977*** (10.394)	-45.587*** (10.399)	41	0.446

Source: World Bank.

Note: See the notes to figures 4.1-4.8 for detailed definitions and data sources for various dependent variables. Development outcomes are regressed against a set of dummies that categorize emerging market and developing economies into quantiles of shares of informality using self-employment in percent of total employment. The constant from the regression results show the development outcomes for the lowest quartile and coefficients show the difference between the corresponding quartile and the lowest quartile. All regressions are cross-sectional regressions with variables averaged during (up to) 1990-2018 (or otherwise specified period or latest year available). Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

References

- Abraham, R. 2019. "Informal Employment and the Structure of Wages in India: A Review of the Trends." *Review of Income and Wealth* 65 (S1): S102-22.
- Adams, R. 2003. "Economic Growth, Inequality, and Poverty: Findings from a New Data Set." Policy Research Working Paper 2972, World Bank, Washington, DC.
- Al Masri, D., V. Flamini, and F. G. Toscani. 2021. "The Short-Term Impact of COVID-19 on Labor Markets, Poverty and Inequality in Brazil." IMF Working Paper 2166, International Monetary Fund, Washington, D.C.
- Ali, N., and B. Najman. 2017. "Informal competition, firms' productivity and policy reforms in Egypt." In *The Informal Economy: Exploring Drivers and Practices*, edited by I. A. Horodnic, P. Rodgers, C. C. Williams, and L. Momtazian. Abingdon, U.K.: Routledge (in press).
- Allen, J., and T. Schipper. 2016. "Understanding the Informal Sector: Do Formal and Informal Firms Compete?" Mimeograph.
- Amaral, P. S., and E. Quintin. 2006. "A Competitive Model of the Informal Sector." *Journal of Monetary Economics* 53 (7): 1541-53.
- Amarante V., R. Arim, and M. Yapora, 2016. "Decomposing Inequality Changes in Uruguay: The Role of Formalization in the Labor Market." *IZA Journal of Labor & Development* 5 (1): 1-20.
- Amin, M., and X. Huang. 2014. "Does Firm-size Matter in the Informal Sector?" Enterprise Note 28, World Bank, Washington, DC.
- Amin, M., and A. Islam. 2015. "Are Large Informal Firms More Productive than the Small Informal Firms? Evidence from Firm-level Surveys in Africa." *World Development* 74 (C): 374-85.
- Amin, M., F. Ohnsorge, and C. Okou. 2019. "Casting a Shadow: Productivity of Formal Firms and Informality." Policy Research Working Paper 8945, World Bank, Washington, D.C.
- Amin, M., and C. Okou. 2020. "Casting a Shadow: Productivity of Formal Firms and Informality." *Review of Development Economics* 24 (4): 1610-30.
- Amuedo-Dorantes, C. 2004. "Determinants and Poverty Implications of Informal Sector Work in Chile." *Economic Development and Cultural Change* 52 (2): 347-68.
- Arias, O., and M. Khamis. 2008. "Comparative Advantage, Segmentation and Informal Earnings: A Marginal Treatment Effects Approach." IZO Discussion Paper 3916, Institute for the Study of Labor, Bonn, Germany.
- Ariza, J., and G. Montes-Rojas. 2017. "Labour Income Inequality and the Informal Sector in Colombian Cities." *Cuadernos de Economía* 36 (72): 77-98.
- Attanasio, O., P. Goldberg, and N. Pavcnik. 2004. "Trade Reforms and Wage Inequality in Colombia." *Journal of Development Economics* 74 (2): 331-66.
- Auriol, E., and M. Walters. 2005. "Taxation Base in Developing Countries." *Journal of Public Economics* 89 (4): 625-46.
- Autor, D. and Reynolds, E., 2020. "The Nature of Work After the COVID Crisis: Too Few Low-Wage Jobs." The Hamilton Project, Brookings Institution, Washington, DC.

- Aydin, E., M. Hisarciklilar, and I. Ikkaracan. 2010. "Formal versus Informal Labor Market Segmentation in Turkey in the Course of Market Liberalization." *Topics in Middle Eastern and North African Economies* 12 (September).
- Banerjee, A., and E. Duflo. 2003. "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8: 267-99.
- Banerjee, A., E. Duflo, and N. Qian. 2020. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." *Journal of Development Economics* 145 (June): 102442.
- Bargain, O., and P. Kwenda. 2014. "The Informal Sector Wage Gap: New Evidence Using Quantile Estimations on Panel Data." *Economic Development and Cultural Change* 63 (1): 117-53.
- Barro, R. J., and L. W. Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950-2010." *Journal of Development Economics* 104 (September): 184-98.
- Barro, R. J. 2015. "Convergence and Modernisation." *Economic Journal* 125: 911-42.
- Baskaya, Y. S., and T. Hulagu. 2011. "Informal-Formal Worker Wage Gap in Turkey: Evidence from A Semi-Parametric Approach." Working Paper 11/15, Central Bank of the Republic of Turkey, Ankara.
- Benjamin, N., and A. Mbaye. 2012. "The Informal Sector, Productivity, and Enforcement in West Africa: A Firm-level Analysis." *Review of Development Economics* 16 (4): 664-80.
- Berdiev, A., J. Saunoris, and F. Schneider. 2020. "Poverty and The Shadow Economy: The Role Of Governmental Institutions." *The World Economy* 43 (4): 921-47.
- Besley, T., and T. Persson. 2014. "Why Do Developing Countries Tax so Little?" *Journal of Economic Perspectives* 28 (4): 99-120.
- Blackburn, K., N. Bose, and S. Capasso. 2012. "Tax Evasion, the Underground Economy and Financial Development." *Journal of Economic Behavior & Organization* 83 (2): 243-53.
- Blanchflower, D.G., A. Oswald, and A. Stutzer. 2001. "Latent Entrepreneurship Across Nations." *European Economic Review* 45 (4-6): 680-91.
- Bloeck, M., S. Galiani, and F. Weinschelbaum. 2019. "Poverty Alleviation Strategies under Informality: Evidence for Latin America." *Latin American Economic Review* 28 (1): 14.
- Boly, A. 2018. "On the Short- and Medium-Term Effects of Formalisation: Panel Evidence from Vietnam." *The Journal of Development Studies* 54 (4): 641-56.
- Bonnet, F., J. Vanek, and M. Chen. 2019. *Women and Men in the Informal Economy—A Statistical Brief*. Manchester, U.K.: WIEGO.
- Bosch, M., E. Goñi-Pacchioni, and W. Maloney. 2012. "Trade Liberalization, Labor Reforms and Formal-Informal Employment Dynamics." *Labour Economics* 19 (5): 653-67.
- Bose, N., S. Capasso, and M. Wurm, 2012. "The Impact of Banking Development on the Size of Shadow Economies." *Journal of Economic Studies* 39 (6): 620-38.
- Botelho, F., and V. Ponczek. 2011. "Segmentation in the Brazilian Labor Market." *Economic Development and Cultural Change* 59 (2): 437-63.
- Brandt, L., J. Van Biesebroeck, and Y. Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics* 97 (2): 339-51.

- 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): 382-86.
- Canelas, C. 2019. "Informality and Poverty in Ecuador." *Small Business Economics* 53: 1097-115.
- Capasso, S., and T. Jappelli. 2013. "Financial Development and the Underground Economy." *Journal of Development Economics* 101 (March): 167-78.
- Capp, J., H. Elstrodt, and W. Jones Jr. 2005. "Reining in Brazil's Informal Economy." *McKinsey Quarterly*. <http://www.mckinseyquarterly.com>.
- Chen, M., J. Vanek, and J. Heintz. 2006. "Informality, Gender, and Poverty: A Global Picture." *Economic and Political Weekly* 41 (21): 2131-39.
- Cho, J., and D. Cho. 2011. "Gender Difference of the Informal Sector Wage Gap: A Longitudinal Analysis for the Korean Labor Market." *Journal of the Asia Pacific Economy* 16 (4): 612-29.
- Choi, J., and M. Thum. 2005. "Corruption and the Shadow Economy." *International Economic Review* 46 (3): 817-36.
- Chong, A., and M. Gradstein. 2007. "Inequality and Informality." *Journal of Public Economics* 91(1-2): 159-79.
- Cisneros-Acevedo, C. Forthcoming. "Unfolding Trade Effect in Two Margins of Informality: The Peruvian Case." *The World Bank Economic Review*.
- Cunningham, W. V., and W. F. Maloney. 2001. "Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis." *Economic Development and Cultural Change* 50 (1): 131-56.
- Cusolito, A. P., and W. F. Maloney. 2018. *Productivity Revisited: Shifting Paradigms in Analysis and Policy*. Washington, DC: World Bank.
- D'Erasmus, P. N., and H. J. Moscoso Boedo. 2012. "Financial Structure, Informality and Development." *Journal of Monetary Economics* 59 (3): 286-302.
- Dabla-Norris, E., M. Gradstein, and G. Inchauste. 2008. "What Causes Firms to Hide Output?" *Journal of Development Economics* 85 (1-2): 1-27.
- Dabla-Norris, E., L. J. Mayor, F. Lima, and A. Sollaci. 2018. "Size Dependent Policies, Informality, and Misallocation." IMF Working Paper 18/179, International Monetary Fund, Washington DC.
- De Mel, S., D. McKenzie, and C. Woodruff. 2011. "What is the Cost of Formality? Experimentally Estimating the Demand for Formalization." Working Paper, University of Warwick, Coventry, U.K.
- De Mel, S., D. McKenzie, and C. Woodruff. 2013. "The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka." *American Economic Journal: Applied Economics* 5 (2): 122-50.
- De Soto, Hernando. 1989. *The Other Path: The Invisible Revolution in the Third World*. New York: Harper & Row.
- De Soyres, F., A. Mulabdic, and M. Ruta. 2020. "Common Transport Infrastructure: A Quantitative Model and Estimates from the Belt and Road Initiative." *Journal of Development Economics* 143 (March): 102415.
- Deininger, K., S. Jin, and M. Sur. 2007. "Sri Lanka's Rural Non-Farm Economy: Removing Constraints to Pro-Poor Growth." *World Development* 35 (12): 2056-78.

- Dell'Anno, R. 2016. "Inequality and Informality in Transition and Emerging Countries." *IZA World of Labor*: 325. doi:10.15185/izawol.325.
- DerSimonian, R., and N. Laird. 1986. "Meta-Analysis in Clinical Trials." *Controlled Clinical Trials* 7 (3): 177-88.
- Devicienti F., F. Groisman, and A. Poggi. 2009. "Informality and poverty: Are these processes dynamically interrelated? Evidence from Argentina," Working Papers 146, Society for the Study of Economic Inequality, Palma, Spain.
- Dieppe, A. (ed.). 2020. *Global Productivity: Trends, Drivers, and Policies*. Washington, DC: World Bank.
- Dieppe, A., A. Kawamoto, Y. Okawa, C. Okou, and J. Temple. 2020. "What Explains Productivity Growth." in *Global Productivity: Trends, Drivers, and Policies*, edited by A. Dieppe, 39-94. Washington, DC: World Bank.
- Distinguin, I., C. Rugemintwari, and R. Tacneng. 2016. "Can Informal Firms Hurt Registered SMEs' Access to Credit?" *World Development* 84 (August): 18-40.
- Dix-Carneiro, R., and B. Kovak. 2017. "Trade Liberalization and Regional Dynamics." *American Economic Review* 107 (10): 2908-46.
- Dix-Carneiro, R., and B. Kovak. 2019. "Margins of Labor Market Adjustment to Trade." *Journal of International Economics* 117 (C): 125-42.
- Docquier, F., and Z. Iftikhar. 2019. "Brain Drain, Informality and Inequality: A Search-and-matching Model for Sub-Saharan Africa." *Journal of International Economics* 120 (September): 109-25.
- Docquier, F., T. Müller, and J. Naval. 2017. "Informality and Long-Run Growth." *The Scandinavian Journal of Economics* 119 (4): 1040-85.
- Dollar, D., T. Kleineberg, and A. Kraay. 2013. "Growth Still is Good for the Poor." Policy Research Working Paper 6568, World Bank, Washington DC.
- Dollar, D., and A. Kraay. 2002. "Growth is Good for the Poor." *Journal of Economic Growth* 7: 195-225.
- Dreher, A., and F. Schneider. 2010. "Corruption and The Shadow Economy: An Empirical Analysis." *Public Choice* 144 (1-2): 215-38.
- Durantón, G., and D. Puga. 2004. "Chapter 48—Micro-Foundations of Urban Agglomeration Economies." *Handbook of Regional and Urban Economics* 4: 2063-117.
- Durlauf, S. N., A. Kourtellos, and C. M. Tan. 2008. "Are Any Growth Theories Robust?" *The Economic Journal* 118 (527): 329-46.
- Earle, J. S., and Z. Sakova. 2000. "Business Start-Ups or Disguised Unemployment? Evidence on the Character of Self-Employment from Transition Economies." *Labour Economics* 7 (5): 575-601.
- El Badaoui, E., E. Strobl, and F. Walsh. 2008. "Is There an Informal Employment Wage Penalty? Evidence from South Africa." *Economic Development and Cultural Change* 56 (3): 683-710.
- El Badaoui, E., E. Strobl, and F. Walsh. 2010. "The Formal Sector Wage Premium and Firm Size." *Journal of Development Economics* 91 (1): 37-47.
- Elgin, C., and A. Elveren. 2019. "Informality, Inequality, and Feminization of Labor." Working Paper 483, Political Economy Research Institute, University of Massachusetts Amherst.

- Enste, D., and F. Schneider. 1998. "Increasing Shadow Economies all over the World—Fiction or Reality." IZA Discussion Paper 26, IZA Institute of Labor Economics, Bonn, Germany.
- Epstein, B., and A. F. Shapiro. 2017. "Employment and Firm Heterogeneity, Capital Allocation, and Countercyclical Labor Market Policies." *Journal of Development Economics* 127 (July): 25-41.
- 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): 262-76.
- Falco, P., A. Kerr, N. Rankin, J. Sandefur, and F. Teal. 2011. "The Returns to Formality and Informality in Urban Africa." *Labour Economics* 18: S23-S31.
- Farazi, S. 2014. "Informal Firms and Financial Inclusion: Status and Determinants." *Journal of International Commerce, Economics and Policy* 5 (3): 1440011.
- Farrell, D. 2004. "The Hidden Dangers of Informal Economy." *McKinsey Quarterly* 3: 27-37.
- Feldkircher, M., and S. Zeugner. 2012. "The Impact of Data Revisions on the Robustness of Growth Determinants—A Note on 'Determinants of Economic Growth: Will Data Tell?'" *Journal of Applied Econometrics* 27(4): 686-94.
- Fernandes, A. M. 2008. "Firm Productivity in Bangladesh Manufacturing Industries." *World Development* 36 (10): 1725-44.
- Fernandez, C., E. Ley, and M. Steel. 2001. "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics* 16 (5): 563-76.
- Fields, G. S. 1975. "Rural-Urban Migration, Urban Unemployment and Underemployment, and Job-Search Activity in LDCs." *Journal of Development Economics* 2 (2): 165-87.
- Fields, G. S. 2005. "A Guide to Multisector Labor Market Models." Employment Policy Primer, World Bank, Washington, DC.
- Fields, G. S. 2019. *Employment and Development: How Work Can Lead From and Into Poverty*, edited by J. Pieters. Oxford, U.K.: Oxford University Press.
- Fields, G. S., and J. Pieters. 2018. *Employment and Development*. Oxford, U.K.: Oxford University Press.
- Francois, J., and M. Manchin. 2013. "Institutions, Infrastructure, and Trade." *World Development* 46 (June): 165-75.
- Friedman, E., S. Johnson, D. Kaufmann, and P. Zoido-Lobaton. 2000. "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries." *Journal of Public Economics* 76 (3): 459-93.
- Friesen, J., and K. Wacker. 2013. "Do Financially Constrained Firms Suffer from More Intense Competition by the Informal Sector? Firm-Level Evidence from the World Bank Enterprise Surveys." Discussion Paper 139, Courant Research Centre, Göttingen, Germany.
- Fugazza, M., and N. Fiess. 2010. "Trade Liberalization and Informality: New Stylized Facts." Policy Issues in International Trade and Commodities Study Series No. 43, United Nations, New York and Geneva.
- Funkhouser, E. 1997. "Mobility and Labor Market Segmentation: The Urban Labor Market in El Salvador." *Economic Development and Cultural Change* 46 (1): 123-53.

- Galiani, S., F. Weinschelbaum. 2012. "Modeling Informality Formally: Households and Firms." *Economic Inquiry* 50 (3): 821-38.
- 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.
- García, G. A., and E. R. Badillo. 2018. "Rationing of Formal Sector Jobs and Informality: The Colombian Case." *Journal of International Development* 30 (5): 760-89.
- Gaspar, V., D. Amaglobeli, M. Garcia-Escribano, D. Prady, and M. Soto. 2019. "Fiscal Policy and Development: Human, Social, and Physical Investments for the SDGs." Staff Discussion Note 19/03, International Monetary Fund, Washington, DC.
- Gasparini, L., F. Haimovich, and S. Olivieri. 2009. "Labor Informality Bias of a Poverty-Alleviation Program in Argentina." *Journal of Applied Economics* 12 (2): 181-205.
- Gasparini, L., and L. Tornarolli. 2007. "Labor Informality in Latin America and the Caribbean: Patterns and Trends from Household Survey Microdata." CEDLAS Working Paper, Centro de Estudios Distributivos, Laborales y Sociales, Universidad Nacional de La Plata, Argentina.
- Ghosh, J. and A. E. Ghattas. 2015. "Bayesian Variable Selection Under Collinearity." *The American Statistician* (69): 165-73.
- Gindling, T. 1991. "Labor Market Segmentation and the Determination of Wages in the Public, Private- Formal, and Informal Sectors in San José, Costa Rica." *Economic Development and Cultural Change* 39 (3): 585-605.
- Gindling, T., N. Mossaad, and D. Newhouse. 2016. "Earnings Premiums and Penalties for Self-Employment and Informal Employees around the World." Policy Research Working Paper 7530, World Bank, Washington, DC.
- Gindling, T., N. Mossaad, and D. Newhouse. 2020. "Self-Employment Earnings Premiums/Penalties and Regulations: Evidence from Developing Economies." *Small Business Economics* 55 (1): 507-27.
- Goldberg, P. K., and N. Pavcnik. 2003. "The Response of the Informal Sector to Trade Liberalization." *Journal of Development Economics* 72 (2): 463-96.
- Goldberg, P. K., and N. Pavcnik. 2004. "Trade, Inequality, and Poverty: What Do We Know? Evidence from Recent Trade Liberalization Episodes in Developing Countries." NBER Working Paper 10593, National Bureau of Economic Research, Cambridge, MA.
- Goldberg, P. K., and N. Pavcnik. 2007. "Distributional Effects of Globalization in Developing Countries." *Journal of Economic Literature* 45 (1): 39-82.
- Gong, X., A. van Soest, and E. Villagomez. 2004. "Mobility in the Urban Labor Market: A Panel Data Analysis for Mexico." *Economic Development and Cultural Change* 53 (1): 1-36.
- Gonzalez, A., and F. Lamanna. 2007. "Who Fears Competition from Informal Firms?" Policy Research Working Paper 4316, World Bank, Washington, DC.
- Grimm, M., P. Knorringa, and J. Lay. 2012. "Constrained Gazelles: High Potentials in West Africa's Informal Economy." *World Development* 40 (7): 1352-68.
- Günther, I., and A. Launov. 2006. "Competitive and Segmented Informal Labor Markets." IZO Discussion Paper 2349, Institute for the Study of Labor, Bonn, Germany.

- Haltiwanger, J. C., J. I. Lane, and J. R. Spletzer. 1999. "Productivity Differences Across Employers: The Roles of Employer Size, Age, and Human Capital." *American Economic Review* 89 (2): 94-8.
- Harris, J. R., and M. P. Todaro. 1970. "Migration, Unemployment, and Development: A Two Sector Analysis." *American Economic Review* 60 (1): 126-42.
- Hazans, M. 2011. "Informal Workers Across Europe: Evidence from 30 Countries." IZO Discussion Paper 5871, Institute for the Study of Labor, Bonn, Germany.
- Heckman, J. J., and X. Li. 2003. "Selection Bias, Comparative Advantage and Heterogeneous Returns to Education: Evidence from China in 2000." NBER Working Paper 9877, National Bureau of Economic Research, Cambridge, MA.
- Heredia, J., A. Flores, C. Geldes, and W. Heredia. 2017. "Effects of Informal Competition on Innovation Performance: The Case of Pacific Alliance." *Journal of Technology Management and Innovation* 12 (4): 22-8.
- Higgins, J. P. T., and S. G. Thompson. 2002. "Quantifying Heterogeneity in a Meta-Analysis." *Statistics in Medicine* 21 (11): 1539-58.
- Huber, P., and U. Rahimov. 2014. "Formal and Informal Sector Wage Differences in Transition Economies: Evidence from Tajikistan." Working Papers in Business and Economics 48/2014, Mendel University in Brno, Czech Republic.
- Husmanns, R. 2004. "Measuring the Informal Economy: From Employment in the Informal Sector to Informal Employment." ILO Working Paper 42, International Labour Office, Geneva.
- ICRG (International Country Risk Guide). 2014. "ICRG Methodology." PRS Group, Syracuse, NY. <https://www.prsgroup.com/wp-content/uploads/2014/08/icrgmethodology.pdf>.
- Ihrig, J., and K. S. Moe. 2004. "Lurking in The Shadows: The Informal Sector and Government Policy." *Journal of Development Economics* 73 (2): 541-57.
- ILO (International Labour Organization). 2013. *Measuring Informality: A Statistical Manual on The Informal Sector and Informal Employment*. International Labor Office, Geneva.
- ILO (International Labour Organization). 2018a. *World Employment and Social Outlook—Trends 2018*. International Labor Office, Geneva.
- ILO (International Labour Organization). 2018b. *Women and Men in the Informal Economy: A Statistical Picture*. International Labour Office, Geneva.
- ILO (International Labour Organization). 2019. *Global Wage Report 2018/19: What Lies behind Gender Pay Gaps*. Geneva: International Labour Office.
- Iriyama, A., R. Kishore, and D. Talukda. 2016. "Playing Dirty or Building Capability? Corruption and HR Training as Competitive Actions to Threats from Informal and Foreign Firm Rivals." *Strategic Management Journal* 51 (2): 315-34.
- Islam, A. 2018. "The Burden of Water Shortages on Informal Firms." Policy Research Working Paper 8457, World Bank, Washington, DC.
- Jain-Chandra, S., N. Khor, R. Mano, J. Schauer, J. Wingender, and J. Zhuang. 2018. "Inequality in China—Trends, Drivers and Policy Remedies." IMF Working Paper 2018/127, International Monetary Fund, Washington, DC.
- Jones, I. C., and W. Nordhaus. 2008. "Comment on" A. Shleifer and R. La Porta's "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity* 2008 (2): 353-63.

- Joshi, A., W. Prichard, and C. Heady. 2014. "Taxing the Informal Economy: The Current State of Knowledge and Agendas for Future Research." *Journal of Development Studies* 50 (10): 1325-47.
- Jovanovic, B. 1982. "Selection and Evolution of Industry." *Econometrica* 50 (3): 649-70.
- Kahyalar, N., S. Fethi, S. Katircioglu, and B. Ouattara. 2018. "Formal and Informal Sectors: Is There Any Wage Differential?" *The Service Industries Journal* 38 (11-12): 789-823.
- Kanbur, R. 2017. "Informality: Causes, Consequences and Policy Responses." *Review of Development Economics* 21 (4): 939-61.
- Kim, B. 2005. "Poverty and Informal Economy Participation: Evidence from Romania." *Economics of Transition* 13 (1): 163-85.
- La Porta, R., and A. Shleifer. 2008. "The Unofficial Economy and Economic Development." *Brookings Papers on Economic Activity* 2008 (2): 275-352.
- La Porta, R., and A. Shleifer. 2014. "Informality and development." *Journal of Economic Perspectives* 28 (3): 109-26.
- Lehmann, H., and N. Pignatti. 2007. "Informal Employment Relationships and Labor Market Segmentation in Transition Economies: Evidence from Ukraine." IZO Discussion Paper 3269, Institute for the Study of Labor, Bonn, Germany.
- Lehmann, H., and A. Zaiceva. 2013. "Informal Employment in Russia: Incidence, Determinants and Labor Market Segmentation." DSE Working Paper 903, Department of Economics, University of Bologna, Italy.
- Loayza, N. V. 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: 129-62.
- Loayza, N. V. 2016. "Informality in the Process of Development and Growth." *The World Economy* 39 (12): 1856-916.
- Loayza, N. V. 2018. "Informality: Why Is It So Widespread and How Can It Be Reduced?" Research & Policy Brief, World Bank, Kuala Lumpur, Malaysia.
- Loayza, N., and J. Rigolini. 2006. "Informality Trends and Cycles." Policy Research Working Paper 4078, World Bank, Washington, DC.
- 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.
- Luttmer, E. G. J. 2007. "Selection, Growth, and the Size Distribution of Firms." *The Quarterly Journal of Economics* 122 (3): 1103-44.
- Magnac, T. 1991. "Segmented or Competitive Labor Markets." *Econometrica* 59 (1): 165-87.
- Maloney, W. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *World Bank Economic Review* 13 (2): 275-302.
- Maloney, W. F. 2004. "Informality Revisited." *World Development* 32 (7): 1159-78.
- Malta, V., L. Kolovich, A. Marinez, and M. Tavares. 2019. "Informality and Gender Gaps Going Hand in Hand?" IMF Working Paper 19/112, International Monetary Fund, Washington, D.C.

- Marcouiller, D., V. R. de Castilla, and C. Woodruff. 1997. "Formal Measures of the Informal-Sector Wage Gap in Mexico, El Salvador, and Peru." *Economic Development and Cultural Change* 45 (2): 367-92.
- Marjit S., V. Mukherjee, and M. Kolmar. 2006. "Poverty, Taxation, and Governance." *The Journal of International Trade & Economic Development* 15 (3): 325-33.
- Matano, A., M. Obaco, and V. Royuela. 2020. "What Drives the Spatial Wage Premium for Formal and Informal Workers? The Case of Ecuador." *Journal of Regional Science* 60 (4): 823-47.
- McCaig, B., and N. Pavcnik. 2018. "Export Markets and Labor Allocation in a Low-Income Country." *American Economic Review* 108 (7): 1899-1941.
- 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, D.C.
- Medina, M., and F. Schneider. 2019. "Shedding Light on the Shadow Economy: A Global Database and the Interaction with the Official One". CESifo Working Paper 7981. CESifo Group Munich.
- Meghir, C., R. Narita, and J. M. Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review* 105(4): 1509-46.
- Melitz, M. J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71(6): 1695-725.
- Mendi, P., and Costamagna, R. 2017. "Managing Innovation under Competitive Pressure from Informal Producers." *Technological Forecasting and Social Change* 114: 192-202.
- Mendicino, C., and M. Prado. 2014. "Heterogeneous Firms and the Impact of Government Policy on Welfare and Informality." *Economics Letters* 124 (1): 151-6.
- Messina, J., and J. Silva. 2021. "Twenty Years of Wage Inequality in Latin America." *The World Bank Economic Review* 35 (1): 117-47.
- Mishra, A., and R. Ray. 2010. "Informality, Corruption, and Inequality," Department of Economics Working Papers 13/10, University of Bath, Department of Economics.
- Nazier, H., & Ramadan, R. 2015. "Informality and poverty: a causality dilemma with application to Egypt." *Advances in Management and Applied Economics* 5 (4): 1-4.
- Nguyen, H. C., C. J. Nordman, and F. Roubaud. 2013. "Who Suffers the Penalty? A Panel Data Analysis of Earnings Gaps in Vietnam." *Journal of Development Studies* 49 (12): 1694-710.
- Nordman C. J., F. Rakotomana, and F. Roubaud. 2016. "Informal versus Formal: A Panel Data Analysis of Earnings Gaps in Madagascar?" IZO Discussion Paper 9970, Institute for the Study of Labor, Bonn, Germany.
- OECD/ILO (Organisation for Economic Co-operation and Development/International Labour Organization). 2019. "Addressing the Gender Dimension of Informality." In *Tackling Vulnerability in the Informal Economy*. Paris: OECD Publishing.
- Ordóñez, J. 2014. "Tax Collection, the Informal Sector, and Productivity." *Review of Economic Dynamics* 17 (2): 262-86.

- Otobe, N. 2017. "Gender and the Informal Economy: Key Challenges and Policy Response." ILO Employment Working Paper 236, International Labour Office, Geneva.
- Oviedo, A. M. 2009. *Economic Informality: Causes, Costs, and Policies, A Literature Survey of International Experience*. Washington, DC: World Bank.
- 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.
- Ponczek, V., and G. Ulyssea. 2018. "Enforcement of Labor Regulation and the Labor Market Effects of Trade: Evidence from Brazil," IZA Discussion Papers 11783, IZA Institute of Labor Economics, Bonn, Germany.
- Prado, M. 2011. "Government Policy in the Formal and Informal Sectors." *European Economic Review* 55 (8): 1120-36.
- Pratap, S., and E. Quintin. 2006. "Are Labor Markets Segmented in Developing Countries? A Semiparametric Approach." *European Economic Review* 50 (7): 1817-41.
- Quintin, E. 2008. "Contract Enforcement and the Size of the Informal Economy." *Economic Theory* 37 (3): 395-416.
- Rauch, J. E. 1991. "Modelling the Informal Sector Formally." *Journal of Development Economics* 35 (1): 33-47.
- Rocha, R., G. Ulyssea, and R. Rachter. 2018. "Do Lower Taxes Reduce Informality? Evidence from Brazil." *Journal of Development Economics* 134 (September): 28-49.
- Rogan, M., and P. Cichello. 2020. "(Re)conceptualising Poverty and Informal Employment." In *The Informal Economy Revisited: Examining the Past, Envisioning the Future*, edited by M. Chen and F. Carré, Routledge: New York.
- Rosenthal, S. S., and W. C. Strange. 2004. "Chapter 49—Evidence on the Nature and Source of Agglomeration Economies." *Handbook of Regional and Urban Economics* 4: 2119-71.
- Rosser B., M. Rosser, and E. Ahmed. 2000. "Income Inequality and the Informal Economy in Transition Economies." *Journal of Comparative Economics* 28 (1): 156-71.
- Rozenberg, J., and M. Fay. 2019. *Beyond the Gap: How Countries Can Afford the Infrastructure They Need While Protecting the Planet*. Washington, D.C.: World Bank.
- Sachs, J. D., G. Schmidt-Traub, C. Kroll, G. Lafortune, and G. Fuller. 2018. *SDG Index and Dashboards Report 2018*. New York: Bertelsmann Stiftung and Sustainable Development Solutions Network.
- Sachs, J. D., G. Schmidt-Traub, C. Kroll, G. Lafortune, G. Fuller, and F. Woelm. 2020. *Sustainable Development Report 2020: The Sustainable Development Goals and COVID-19*. Cambridge, U.K.: Cambridge University Press.
- Sanfey, P., and U. Teksoz. 2007. "Does Transition Make You Happy?" *Economics of Transition* 15 (4): 707-31.
- Saracoğlu, D. S. 2008. "The Informal Sector and Tax on Employment: A Dynamic General Equilibrium Investigation." *Journal of Economic Dynamics and Control* 32 (2): 529-49.

- Sarte, P.-D. G. 2000. "Informality and Rent-Seeking Bureaucracies in a Model of Long-Run Growth." *Journal of Monetary Economics* 46 (1): 173-97.
- Schipper, T. 2020. "Informality, Innovation, and Aggregate Productivity Growth." *Review of Development Economics* 24 (1): 125-43.
- Schneider, F. 2005. "Shadow Economies Around the World: What do we Really Know?" *European Journal of Political Economy* 21 (3): 598-642.
- 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.
- Shleifer, A., and R. W. Vishny. 1993. "Corruption." *The Quarterly Journal of Economics* 108 (3): 599-617.
- Stanley, T. D., H. Doucouliagos, M. Giles, H. Heckemeyer, R. J. Johnston, P. Laroche, J. P. Nelson. 2013. "Meta-Analysis of Economics Research Reporting Guidelines: Reporting Guidelines for Meta-Regression Analysis in Economics." *Journal of Economic Surveys* 27 (2): 390-94.
- Tansel, A., and E. Kan. 2012. "The Formal/Informal Employment Earnings Gap: Evidence from Turkey." ERC Working Paper 1204, Economic Research Center, Middle East Technical University, Ankara, Turkey.
- Tansel, A., H. Keskin, and Z. Ozdemir. 2020. "Is there an Informal Employment Wage Penalty in Egypt? Evidence from Quantile Regression on Panel Data." *Empirical Economics* 58 (6): 2949-79.
- Ulyseea, G. 2010. "Regulation of Entry, Labor Market Institutions and the Informal Sector." *Journal of Development Economics* 91 (1): 87-99.
- Ulyseea, G. 2018. "Firms, Informality, and Development: Theory and Evidence from Brazil." *American Economic Review* 108 (8): 2015-47.
- Ulyseea, G. 2020. "Informality: Causes and Consequences for Development." *The Annual Review of Economics* 12 (1): 525-46.
- UN SDSN (Sustainable Development Solutions Network). 2019. *SDG Costing & Financing for Low-income Developing Countries*. SDSN: New York.
- UN Women. 2016. *Progress of the World's Women 2015/2016: Transforming Economies, Realizing Rights*. UN Women: New York.
- van der Sluis, J., M. van Praag, and W. Vijverberg. 2005. "Entrepreneurship Selection and Performance: A Meta-Analysis of the Impact of Education in Developing Economies." *World Bank Economic Review* 19 (2): 225-61.
- Vargas, J. P. M. 2015. "Informality in Paraguay: Macro-Micro Evidence and Policy Implications." Working Paper 15/245, International Monetary Fund, Washington, DC.
- Vorisek, D., and S. Yu. 2020. "Understanding the Cost of Achieving the Sustainable Development Goals." Policy Research Working Paper 9164, World Bank, Washington, DC.
- WHO (World Health Organization). 2017. *Tracking Universal Health Coverage: 2017 Global Monitoring Report*. Geneva: World Health Organization.
- WIEGO (Women in Informal Employment: Globalizing and Organizing). 2019. "Extending Social Protection to Informal Workers." Briefing Note March 2019. Cambridge, MA: WIEGO.

World Bank. 2019. *Global Economic Prospects: Darkening Skies*. January. Washington, DC: World Bank.

World Bank. 2020a. *Global Economic Prospects*. June. Washington, DC: World Bank.

World Bank. 2020b. *Poverty and Shared Prosperity 2020: Reversals of Fortune*. Washington, DC: World Bank.

World Bank. 2021. *Global Economic Prospects*. January. Washington, DC: World Bank.

Wu, D., and F. Schneider. 2019. “Nonlinearity Between the Shadow Economy and Level of Development.” IMF Working Paper 19/48, International Monetary Fund, Washington, DC.

Yoshida, N., A. Narayan, and H. Wu. 2020. “How COVID-19 Affects Households in Poorest Countries—Insights from Phone Surveys.” *Voices: Perspectives on Development* (blog). December 10, 2020. <https://blogs.worldbank.org/voices/how-covid-19-affects-households-poorest-countries-insights-phone-surveys>.

Zaballos, A., E. Iglesias, and A. Adamowicz. 2019. *The Impact of Digital Infrastructure on the Sustainable Development Goals: A Study for Selected Latin American and Caribbean Countries*. Washington, DC: Inter-American Development Bank.

Zarate, R. 2019. “Does Transit Infrastructure Reduce Informality in Developing Countries? Guest Post by Román David Zárate.” *Development Impact* (blog). December 6, 2019. <https://blogs.worldbank.org/impactevaluations/does-transit-infrastructure-reduce-informality-developing-countries-guest-post>.

Zenker, S., and F. Kock. 2020. “The Coronavirus Pandemic—A Critical Discussion of a Tourism Research Agenda.” *Tourism Management* 81 (December): 104-64.