

Annual Review of Economics

Informality: Causes and Consequences for Development

Gabriel Ulyssea^{1,2,3}

¹Department of Economics, University of Oxford, Oxford OX1 3UQ, United Kingdom;
email: gabriel.ulysea@economics.ox.ac.uk

²Center for Economic and Policy Research, London EC1V 0DX, United Kingdom

³IZA Institute of Labor Economics, 53113 Bonn, Germany

Annu. Rev. Econ. 2020. 12:525–46

First published as a Review in Advance on
May 1, 2020

The *Annual Review of Economics* is online at
economics.annualreviews.org

<https://doi.org/10.1146/annurev-economics-082119-121914>

Copyright © 2020 by Annual Reviews.
All rights reserved

JEL codes: O17, E24, E26, J46, O11, O12

ANNUAL
REVIEWS **CONNECT**

www.annualreviews.org

- Download figures
- Navigate cited references
- Keyword search
- Explore related articles
- Share via email or social media

Keywords

informality, development, firms, workers

Abstract

This article reviews the economic literature on informality, its causes, and its consequences for development. It covers a comprehensive body of research that ranges from well-identified experimental studies to equilibrium macro models, and which more recently includes structural models that integrate both micro and macro effects. The results available in the literature indicate that lowering the costs of formality is not an effective policy to reduce informality but may generate positive aggregate effects, such as higher output and total factor productivity (TFP). The most effective formalization policy is to increase enforcement on the extensive margin but not on the intensive margin of informality. The former generates substantial gains in aggregate TFP and output, without necessarily increasing unemployment. However, the overall welfare impacts are likely to depend on the transitional dynamics between steady states, which remains an open area for future research.

1. INTRODUCTION

The presence of a large informal sector is a defining characteristic of most developing economies. It accounts for between one-third and two-thirds of economic activity, 20–80% of the labor force, and an equally large share of firms in developed countries (e.g., Schneider & Enste 2000, Perry et al. 2007, La Porta & Shleifer 2008, Ulyssea 2018). The aim of this article is to review the extensive economic literature that studies informality, its main characteristics, causes, and consequences for firms, workers, and economic development. To do so, I cover different literature streams that include theoretical developments in conjunction with data, empirical analyses that make credible attempts to draw causal inferences from data, and a more recent but rapidly expanding body of literature that relies on structural estimation.

I start by clearly defining the object of analysis, as informality can be an elusive concept. Indeed, a substantial fraction of the earlier literature focused on the definition of informal sector. The dominant approach that has emerged from this literature has been dubbed the legalistic definition, which defines informal firms and workers as those operating at the margins of the relevant laws and regulations (e.g., Perry et al. 2007).¹ Informal firms are therefore those that do not comply with such regulations—for example, by not being registered with the tax authorities—and informal workers are those who do not have a formal labor contract. This is the definition used throughout this article.

While the definition of labor informality seems straightforward, that of firm informality is arguably less clear-cut, as firms' compliance is less likely to be a binary decision. For example, in both developed and developing countries many formally registered firms evade taxes by underreporting revenues, which implies partial compliance with tax regulations. Also crucially for the analysis of informal employment, many formal firms hire some fraction of their labor force informally to evade the costs implied by the labor regulation. I thus use the same definitions from my previous work (Ulyssea 2018) and throughout this review I distinguish between the following margins of informality: (a) extensive margin, based on whether firms register and pay entry fees to achieve a formal status; and (b) intensive margin, based on whether firms that are formal in the first sense hire workers without a formal contract. Sector membership is defined by the extensive margin, and the (in)formal sector is comprised by (in)formal firms. As will be clear throughout this article, the distinction between these two margins of informality is crucial.

Having defined what informality is, I start by reviewing some of the main stylized facts about firm and worker informality in Section 2. These facts help rationalize some of the main modeling choices found in the literature, and they also provide an important background for the discussion of the causes and consequences of informality. Section 3 thus proceeds to discussing the causes. Informality is an endogenous outcome that emerges as the result of firms' and workers' optimal behavior given their characteristics (e.g., ability) and the environment they face, which includes: institutions, in particular laws and regulations; government policies, such as enforcement and welfare programs; and the economic cycle. Therefore, in order to understand the determinants of informality one must build from the micro level—i.e., the behavior of firms and workers—to the macro level. I do so by separately discussing the determinants of firms' and workers' choices in Sections 3.1 and 3.2, respectively. In both cases, I combine the reduced-form evidence with results from structural models.

¹An alternative approach is the so-called productive definition, which defines informal firms and workers according to their productive characteristics, such as being low productivity or unskilled, being small in size, and performing low-complexity tasks in low-quality jobs (for discussions, see Fields 1990, Perry et al. 2007, Gasparini & Tornarolli 2009).

Section 4 then proceeds to discuss what we have learned from the literature about the consequences of informality. Since informality is the result of individual agents' decisions, it will clearly have direct implications for (micro) outcomes at the firm, worker, and household levels. For example, it can distort firms' size and investment decisions or individuals' investments in formal schooling and on-the-job human capital accumulation. Moreover, given the magnitude of informality and the scale of formalization policies, there are important general equilibrium effects that go beyond the simple aggregation of individual effects. Importantly, the aggregate effects of having lower or higher levels of informality are intrinsically connected to the means used to achieve them, which can be broadly classified into two groups: (a) those that increase the benefits (or decrease the costs) of formality, and (b) those that increase the costs (or decrease the benefits) of informality. The discussion in Section 4 thus follows this classification. Finally, Section 5 concludes.

2. INFORMALITY FACTS

The challenges of studying informality start with its measurement, since by definition informal activities are not officially registered. The literature has used a combination of direct and indirect approaches to measuring informality (Perry et al. 2007). The direct approach relies on micro data—typically from surveys and, to a lesser extent, tax audits—which allow both measuring the size of the informal sector and characterizing informal firms and workers.² I focus on this approach throughout this review.

2.1. Firms

Using data from a broad set of countries, the literature has established a number of facts about informal firms. It has been systematically shown that informal firms are on average smaller (in terms of both employees and revenues), pay lower wages, are run by less educated individuals, hire less educated workers, and earn lower profits than formal firms (e.g., Perry et al. 2007; La Porta & Shleifer 2008, 2014).

2.1.1. Duality (or lack thereof). The above differences between formal and informal firms have been often interpreted as evidence of a dualistic view, according to which these firms operate in entirely separate economic spaces, using completely different technologies and producing distinct goods. However, in previous work (Ulyssea 2018) I show that formal and informal firms in Brazil coexist even within narrowly defined industries, which contradicts this dualistic view. In a much earlier paper, Maloney (1999) makes the same point using worker-level data and examining transitions between sectors in the Mexican labor market.

More interestingly, not only do formality and informality coexist within the same industries producing similar products, but there is a substantial overlap in formal and informal firms' productivity distributions, even within industries (Meghir et al. 2015, Allen et al. 2018, Ulyssea 2018). Thus, one can find firms with the same level of productivity in either the formal or informal sectors, which is also inconsistent with the dual view of informality. Finally, Hsieh & Olken (2014) find no evidence of a "missing middle" in firm size distributions in the manufacturing sectors of India, Indonesia, or Mexico. The same is true in Brazil even if one incorporates informal firms and data from all industries (Ulyssea 2018). This is important, as the missing middle is an important

²The indirect approach is more aggregate in nature and attempts to measure the size of the informal economy by exploiting discrepancies between aggregate variables or by combining these variables with a model (e.g., Loayza 1996, Schneider & Enste 2000).

presumed fact behind development theories that rely on some variation of the dual-economy view (for an in-depth discussion, see Hsieh & Olken 2014).

2.1.2. The margins of informality. A well-documented fact in the literature is that the extensive margin of informality declines with firm size, that is, the share of informal firms rapidly declines as firms grow larger (e.g., Perry et al. 2007, De Paula & Scheinkman 2011). This is an important fact, as it indicates that the costs of operating in the informal sector are increasing in firm size. This will be the case, for example, if the probability of detection increases with size (as larger firms are more visible to the government). More broadly, the opportunity costs of operating in the informal sector are likely to be increasing in firm size (and productivity), as larger firms might have greater need of accessing formal credit markets or issuing tax receipts to buyers, among other factors.

Albeit less emphasized, the intensive margin represents a substantial fraction of informal employment in developing countries: 56% in Mexico (de la Parra 2016), at least 40% in Brazil (Ulyssea 2018), and 32% in Peru (Cisneros-Acevedo 2019). As observed for the extensive margin, the intensive margin also declines as firms grow larger: The average share of informal employees within formal firms declines with firm size.³ Again, this can be directly rationalized by the fact that larger firms are more visible and therefore more likely to be inspected.

2.1.3. Firm dynamics. There is growing evidence that firms in developing countries grow less as they age and that stagnant firms tend to survive longer compared to firms in developed economies (e.g., Hsieh & Klenow 2014, Akcigit et al. 2018, Eslava et al. 2019). However, most of the stylized facts about firm dynamics in developing countries rely heavily on information about formal firms.⁴ Indeed, perhaps one of the largest gaps in our knowledge about informal firms is the lack of longitudinal information about their life cycle behavior. To the best of my knowledge, the paper by Ulyssea (2019) is the only one that analyzes formal versus informal firm dynamics, combining firm-level data from Brazil with a structural model. The paper documents the following facts: (a) In both sectors, the size distribution of older cohorts first-order stochastically dominates that of younger cohorts, and firms display an increasing and concave age-size profile; (b) formal firms grow substantially more than their informal counterparts, but still much less than firms in the United States; and (c) both the extensive and intensive margins of informality decline with firms' age. These facts suggest that dynamic selection—the process whereby less productive firms exit as they age and more productive firms survive and expand—takes place in both sectors but is substantially weaker in the informal sector. Ex ante heterogeneity rather than ex post shocks seems to be the key determinant of firms' life cycle employment processes in both sectors.

2.2. Workers

It has been extensively documented that the incidence of informal employment displays a U-shape pattern with respect to age (it is larger among younger and older workers) and that it is higher among women and decreases with schooling, even though it remains large among high-skill workers (e.g., Funkhouser 1996, Perry et al. 2007, Gasparini & Tornarolli 2009, Galiani &

³ Perry et al. (2007) provide evidence for different Latin American countries, while Ulyssea (2018) documents this regularity using data on Brazilian firms.

⁴ Mexico has partial coverage of informal firms, as the Mexican Economic Census substantially underreports urban informal workers and firms (Busso et al. 2012). In other countries, coverage is typically limited to formal firms and workers (e.g., Indonesia) (Peters 2019).

Weinschelbaum 2012). Beyond this basic characterization, the two main sets of moments documented in the literature refer to the formal-informal wage gap and the transitions in and out of informality.

2.2.1. Formal-informal wage gaps. There is an extensive literature that has documented, for a vast array of countries, the existence of a substantial formal-informal wage gap, which persists even after controlling for several observable characteristics (e.g., Funkhouser 1996, Maloney 1999, Gong & Van Soest 2002, Prata & Quintin 2006, Perry et al. 2007, Gasparini & Tornarolli 2009, Botelho & Ponczek 2011).⁵ The analysis of formal-informal wage gaps is often linked to the goal of testing for the existence of labor market segmentation. However, different papers argue that one cannot rely on wage gaps to test for the existence of segmentation and, more broadly, that the segmentation hypothesis seems to have little empirical content (e.g., Paes de Barros 1988, Magnac 1991, Maloney 1999, Ulyssea 2010a).

More recently, I use matched employer-employee data on both formal and informal firms in Brazil to estimate the same log-wage regression estimated in the literature, but adding firm fixed effects (Ulyssea 2018). When doing that, the wage gap between formal and informal workers completely vanishes.⁶ If there is positive assortative matching between firms and workers, firm fixed effects also capture (at least partially) workers' unobserved quality, and therefore these results suggest that (a) self-selection is one of the main drivers of the wage gap between observably equivalent workers, and (b) conditional on skill, formal and informal workers perform the same tasks within the firm.⁷

2.2.2. Ins and outs of informality. Unsurprisingly, transitions in and out of informality follow a similar pattern to that observed for the stock of informal employment. In particular, at any point in time the young, women, and low-skill workers have a higher probability of transiting from unemployment and formal jobs into informal employment (e.g., Funkhouser 1996, Gong & Van Soest 2002, Pagés & Stampini 2009, Bosch & Maloney 2010). Perhaps more interestingly, most of the existing literature has focused on how the ins and outs of informality behave over the business cycle. A well-established stylized fact is that informal employment (like unemployment) is counter-cyclical, expanding during contractions and decreasing during booms as a fraction of employment (Loayza 1996, Perry et al. 2007, Bosch & Maloney 2010, Bosch & Esteban-Pretel 2012). This behavior is explained by three important flows. First, the job finding rate in the formal sector is strongly pro-cyclical, while it remains relatively stable in the informal sector. Second, informal-to-formal transitions are pro-cyclical. Third, separation rates are counter-cyclical in both sectors but more volatile in the informal sector. These features combined explain the cyclical behavior of informality (see Perry et al. 2007, Bosch & Esteban-Pretel 2012).

3. CAUSES

Part of the literature has relied on cross-country data to investigate the correlation between informal sector's size and a series of potential determinants, such as formal sector's entry costs (Djankov

⁵The results are more nuanced when one examines wage gaps across educational levels or for different quantiles of the earnings distribution (e.g., Prata & Quintin 2006, Bargain & Kwenda 2014).

⁶The within-firm, formal-informal wage gap is identified from formal firms that hire both types of workers. However, the result is not driven by the sample used: If I remove the firm fixed effects, I obtain a wage gap of similar magnitude to that typically found in the literature.

⁷Using semiparametric estimation similar to the one used by Prata & Quintin (2006) and data from Ecuador, El Badaoui et al. (2010) show that the formal-informal wage gap disappears when controlling for firm size.

et al. 2002, Auriol & Warlters 2005), labor and tax regulations, corruption, and institutional quality (e.g., Loayza 1996, Johnson et al. 1998, Friedman et al. 2000, Botero et al. 2004, Dabla-Norris et al. 2008). These results are quite informative, but they do not provide causal evidence on the determinants of informality. This section therefore focuses on the empirical literature that exploits within-country experimental and quasi-experimental variation to identify the causal effects of potential determinants. When relevant, I combine the reduced-form evidence with results from structural models.

3.1. Firms

This section focuses on the determinants of firms' decisions regarding both margins of informality. Before discussing the empirical results, I lay out a very simple framework to provide some structure to the discussion that follows. Since the empirical literature has mostly focused on the extensive margin, the framework abstracts from the intensive margin of informality and also assumes no uncertainty. The profit functions of formal and informal firms are thus given by

$$\pi_f(\theta) = (1 - \tau_y) \underbrace{\theta F(k, \ell)}_{\equiv y(\theta)} - (1 + \tau_w)w_f \ell - r_f k - \bar{c}_f$$

and

$$\pi_i(\theta) = (1 - p(y(\theta))) \theta F(k, \ell) - w_i \ell - r_i k,$$

where θ denotes firms' productivity; τ_y is the revenue tax [which corresponds to a value-added tax (VAT)]; $F(\cdot)$ is the production function, which is increasing and concave in both labor, ℓ , and capital, k ; w_f , w_i and r_f , r_i denote labor and capital prices in the formal and informal sector; τ_w denotes the payroll taxes; \bar{c}_f denotes the per-period fixed cost of operation that formal firms must pay, which aims at capturing the administrative burden of paying taxes and contributions, as emphasized by Djankov et al. (2010); and $p(y(\theta))$ denotes the cost of informality, which is assumed to be increasing and convex in firms' output.⁸

Firms do not face uncertainty about the evolution of their productivity, but they are subject to an exogenous death shock, denoted by δ_s , $s = i, f$. In this very simple setting, firms' value functions in either sector are given by the present value of the infinite stream of profits, discounted by the exogenous exit probability, $V_s(\theta) = \frac{\pi_s(\theta)}{\delta_s}$, $s = i, f$, where to simplify notation I normalize the discount rate to one. The formal and informal value functions are thus given by

$$V_f(\theta) = \frac{\theta F(k, \ell)}{\delta_f} - \underbrace{\left(\frac{w_f \ell + r_f k}{\delta_f} - \frac{\tau_y \theta F(k, \ell) + \tau_w w_f \ell + \bar{c}_f}{\delta_f} \right)}_{\text{Costs of remaining formal (A)}}$$

and

$$V_i(\theta) = \frac{\theta F(k, \ell)}{\delta_i} - \underbrace{\left(\frac{w_i \ell + r_i k}{\delta_i} - \frac{p(y(\theta)) \theta F(k, \ell)}{\delta_i} \right)}_{\text{Costs of remaining informal (B)}},$$

⁸This is a standard formulation in the literature (e.g., Rauch 1991, Fortin et al. 1997, De Paula & Scheinkman 2011, Meghir et al. 2015, Ulyssea 2018).

and firms decide to formalize if

$$V_f(\theta) - c_f^e \geq V_i(\theta),$$

where c_f^e represents the fixed registration costs (see Djankov et al. 2002).

Hence, the costs of being formal can be broadly divided into the costs of entering the formal sector, summarized by c_f^e , and the costs of remaining formal, given by the term A above. The latter corresponds to tax payments and all the recurrent administrative costs associated to being formal (captured by \bar{c}_f). If policy makers want to incentivize firms to formalize, the first alternative is to reduce the costs of formality, which can be achieved by reducing the costs of entering the formal sector, the costs of remaining formal, or both. The second alternative is to increase the benefits of formality, which in this simple framework could be achieved by, for example, increasing access to capital (i.e., a lower r_f). The third option is to increase the costs of informality (the term B), which could be done by increasing enforcement of the existing laws and regulations (e.g., increasing the intensity of government inspections). In the equations above, this could represent a higher exit rate, δ_i , or a steeper cost function, $p(y(\theta))$.

3.1.1. Evidence from experimental and quasi-experimental studies. The papers that seek to identify the causal effect of potential determinants of firms' formalization decision estimate variations of the general model

$$y_{it} = \alpha + \beta \text{Treatment}_{it} + \gamma \mathbf{X}_{it} + \epsilon_{it},$$

where t denotes time and i indexes the unit of analysis, which is typically the firm or entrepreneur but can also be a more aggregate unit such as the local economy (e.g., municipality or commuting zone) or industry; y_{it} is a dummy for whether the firm is formal, or, when more aggregate units are being used, it can be the share of formal firms in a given municipality or industry; Treatment_{it} can denote a dummy for whether the firm was offered a formalization treatment or whether a given municipality was included in a formalization program; \mathbf{X}_{it} is a vector of i -level controls; and ϵ_{it} is the error term. In the experimental benchmark at the individual level, the β typically corresponds to the intention-to-treat (ITT) parameter of a given intervention.

Most of the formalization policies implemented and analyzed in the literature have been in the direction of reducing the costs of formality. In particular, great emphasis has been placed on reducing registration costs, which are often regarded as a major constraint to firm creation and formalization (e.g., De Soto 1989, Djankov et al. 2002). Indeed, substantial efforts have been made to reduce registration costs around the world, but the results of these reforms have been quite limited in terms of inducing firms to formalize (Bruhn & McKenzie 2014). Taken as a whole, the results from both experimental and quasi-experimental studies indicate that reducing the costs of entering the formal sector have no (or a very limited) effect on formalization.

Figure 1 summarizes the results from different studies and provides a visual comparison of the effects from different interventions.⁹ As the figure shows, providing information about the process and potential benefits of registration (De Giorgi & Rahman 2013), providing information

⁹This figure focuses on papers that have a credible research design to identify the causal effect of a given policy. I focus on authors' preferred specifications and standardize the reported point estimate so that all results reflect effects on formalization rates. Piza (2018) revisits the results obtained by Monteiro & Assunção (2012) and Fajnzylber et al. (2011), highlighting the threats to identification and confounding factors that were not accounted for in these studies. Notwithstanding the importance of these two earlier papers, I choose to report only the latest and revised estimate provided by Piza (2018).

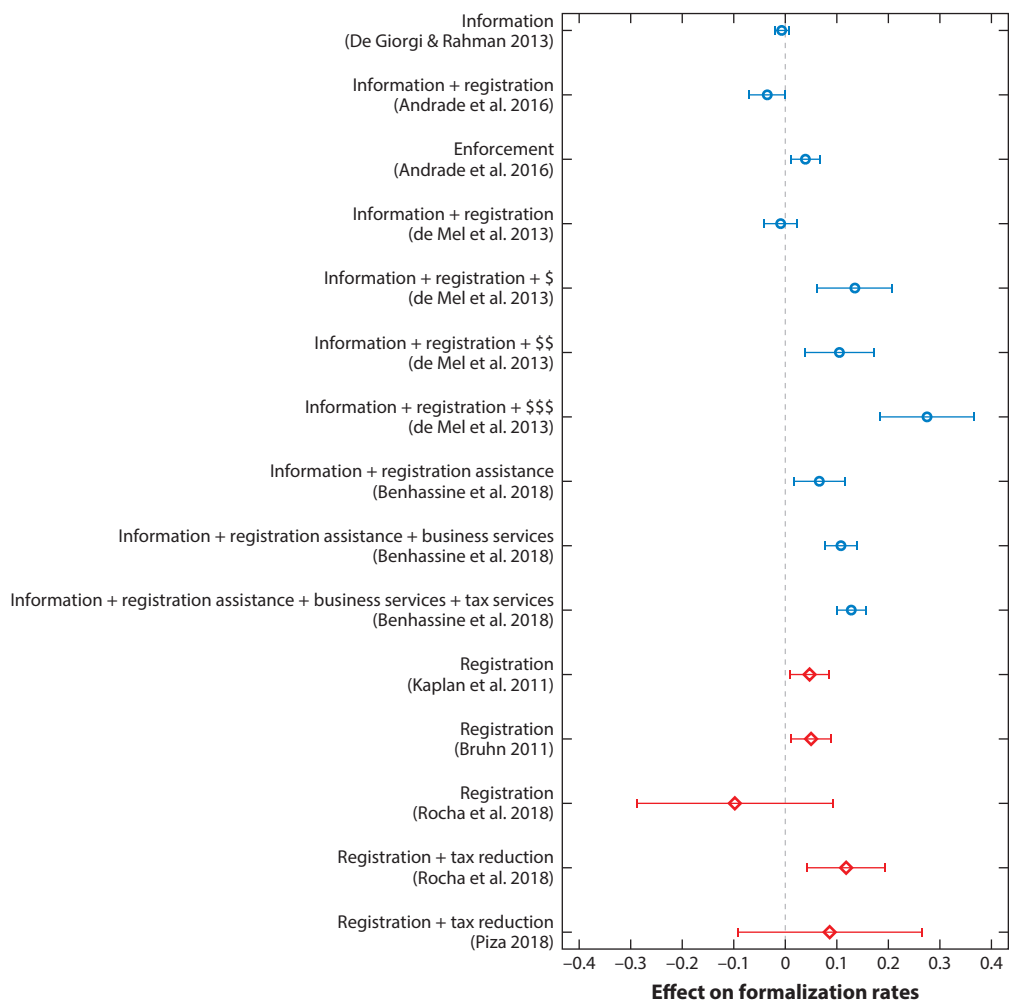


Figure 1

Estimates of formalization effects from the literature. Blue circles represent the results from experimental studies, while the red diamonds represent the results from the nonexperimental literature. In de Mel et al. (2013), \$ indicates half a month's profits, \$\$ indicates one month's profits, and \$\$\$ indicates two months' profits. Only the intention-to-treat estimates are shown for the experimental papers.

and reimbursing all registration costs (De Mel et al. 2013, de Andrade et al. 2016), and creating economy-wide programs to simplify and reduce the monetary costs of registering a firm (Bruhn 2011, Kaplan et al. 2011, Piza 2018, Rocha et al. 2018) have very limited effects on registration. The results reported by Benhassine et al. (2018) constitute an important exception, as they find a positive and significant effect (6.6 to 9.6 percentage points) even in the treatment that only provides information and assistance in registering.¹⁰ In order to separate the effect of information per se from the effect of the high-quality staff who provided the information, the authors design

¹⁰Their empirical context is the introduction of a simplified formalization regime being offered to small informal businesses in Benin.

and implement an additional experiment that provides the same information in a more standard way (without the qualified staff). In this case the authors find no effect.

As **Figure 1** shows, the largest formalization effects come from interventions that reduce the ongoing costs of formality or increase its benefits. This is very clear from the results obtained by De Mel et al. (2013). When firms were offered information and had all registration costs covered, the authors find no formalization effect. When, in addition to that, firms were offered a substantial compensation for formalizing (the equivalent of two months' profits for the median firm), 47% of firms registered [this is the treatment-on-the-treated (TOT) estimate, while **Figure 1** depicts the ITT of 27.5%]. With a few more assumptions, De Mel and colleagues are able to back out the demand curve for formalization, which is downward sloping with respect to the present discount cost of formalizing. They find a steep slope when net costs are negative and a much flatter slope when net costs are positive, which indicates that few firms perceive substantial positive benefits from being formal. Hence, to induce substantial formalization it is necessary to either substantially reduce the costs of remaining formal (term A in Equation 1) or increase its benefits.

Rocha et al. (2018) find similar results when estimating the effect of a formalization policy in Brazil targeted at entrepreneurs with at most one employee. In its first phase, the policy eliminated entry costs for eligible entrepreneurs, and in the second phase it substantially reduced the tax burden. The first phase had no effect on formalization, while the second led to an increase of around 11%. This result is entirely driven by the formalization of existing informal firms and not by the creation of new formal businesses, nor by greater survival among formal firms. Even in the case discussed by Benhassine et al. (2018)—who find strong formalization effects from a simplified formalization regime—the authors attribute this effect to the fact that their treatments ultimately enhanced the potential benefits of formalization.¹¹ However, the authors argue that these interventions are not cost effective, which is the same conclusion reached by Rocha et al. (2018) for Brazil.

Finally, a far less emphasized formalization policy is to increase enforcement of existing laws and regulations. To the best of my knowledge, the only study focusing on this aspect is the one by de Andrade et al. (2016), who randomly assign municipal inspectors to firms in order to assess whether higher enforcement can induce firms to formalize in the state of Minas Gerais, Brazil. Their results show that being assigned to the enforcement treatment increases registration rates with the municipality by 2–4 percentage points. As discussed by the authors, the ITT is not necessarily the most informative parameter in their setting, as many firms assigned to treatment were closed, could not be found, or were already formal. Indeed, the instrumental variable (IV) estimate of the impact of actually receiving an additional inspector visit is much higher, indicating an effect of 21–27 percentage points in firms' registration. The authors find no evidence of spillovers on neighboring firms, which they attribute to the relatively low increase in inspections and to the fact that many firms indicate that they do not communicate with their neighbors.

3.1.2. Interpreting the evidence. The results discussed so far indicate that reducing the costs of entering the formal sector has very limited or no formalization effects, while reducing the ongoing costs of formality (or increasing its benefits) is more effective, but the effects are not large and the policies do not seem to be cost effective. Overall, the largest formalization effects come from greater enforcement. The natural question then is, Why does reducing the costs of formality have such limited effects?

¹¹In addition to the first treatment discussed above, the second treatment offered business services—facilitating access to training and commercial banks—and the third treatment added tax services aimed at reducing the costs and uncertainty associated with tax compliance.

To a large extent, these different policies can be directly linked to the three leading views about informality in the literature (La Porta & Shleifer 2014). The first argues that the informal sector is a reservoir of potentially productive entrepreneurs who are kept out of formality by high regulatory costs, most notably entry regulation. This view dates back to work by De Soto (1989) and provides the rationale for the numerous efforts to reduce fixed costs of registration around the world. The second view sees informal firms as parasite firms, which are productive enough to survive in the formal sector but choose to remain informal to earn higher profits from the cost advantages of not complying with taxes and regulations (e.g., Levy 2008). This view therefore justifies the use of increased enforcement as a way to effectively reduce informality. The third view posits that informality is a survival strategy for low-skill individuals, who are too unproductive to ever become formal. In this case, reducing the costs of formality is not effective, and increasing enforcement can generate substantial adverse effects, as these informal businesses would not formalize but simply exit.

Until recently these views were seen as competing frameworks for understanding informality, but in my work (Ulyssea 2018) I show that in fact they are not. They just reflect heterogeneous firms optimally choosing to comply or not with the relevant laws and regulations. These views are thus complementary and not competing frameworks for understanding informality, in which case the relevant question becomes determining the relative importance of each view in the data. In that paper I propose a simple taxonomy of informal firms based on these views and use the estimated model to infer the relative importance of these types in the Brazilian context.

As **Figure 2** shows, the mass of firms that correspond to De Soto's definition is quite small, only 9.3% of all informal firms. The informal firms that could formalize once formal sector's entry costs

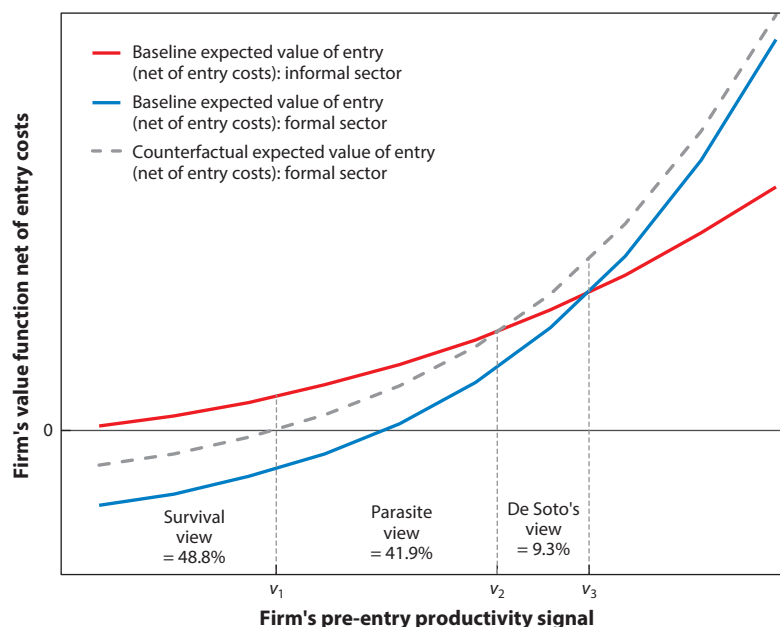


Figure 2

Distribution of informal firm types in Brazil. The figure shows, for each productivity level, firms' expected value function net of entry costs in the formal and informal sectors (*solid lines*). The dashed line represents the net expected value of entering the formal sector in a scenario in which its entry costs are equalized to the informal sector's. Figure adapted from Ulyssea (2018).

are removed, but choose to remain informal to enjoy the cost advantages of informality, correspond to 41.9% of all informal firms. Finally, the remaining 48.8% correspond to the survival view, i.e., they are too unproductive to ever become formal. If the results from Brazil are informative for other contexts, then **Figure 2** provides a rationale for the results discussed in the previous section. A very small fraction of firms are actually constrained by high registration costs, and therefore reducing them will have a limited effect on formalization. In contrast, increasing enforcement can have a substantial impact, as a large fraction of firms could be formal but optimally choose to remain informal. However, this policy can have potentially large negative effects, as these types are not directly observable, and nearly half of all informal firms are not productive enough to ever survive in the formal sector (see discussion about the consequences in Section 4).

3.1.3. Other potential channels. Access to the formal financial system and credit markets is often cited as a potential benefit of formality. Thus, the more developed the financial sector and credit markets are in a given country, the greater is the opportunity cost of being informal. The credit channel has been theoretically analyzed in early papers, such as those by Straub (2005) and Quintin (2008), which emphasize the role of weaker (or inexistent) contract and debt enforcement in the informal sector. However, the empirical evidence in support of this channel is very scarce. To the best of my knowledge, the only paper that empirically tests the relevance of this channel is by Catão et al. (2009), who find a positive association between formalization rates at the two-digit industry level and financial deepening.¹² D’Erasmus & Boedo (2012), D’Erasmus (2016), and Lopez-Martin (2019) confirm these results using general equilibrium models of firm dynamics with different forms of financial frictions, in which firms can operate formally or informally. In these models, operating in the formal sector is costly (due to taxes and regulations), but it allows firms to access better-functioning credit markets and cheaper capital. These models are calibrated to different developing economies, and the results show that reducing financial frictions in the formal sector leads to a substantial decrease in firm informality.

A less emphasized but important determinant of firm informality is the tax structure (and not only the tax burden) of the economy. In order to analyze the role of VAT in transmitting informality, De Paula & Scheinkman (2010) develop a model where the VAT collected according to a credit scheme creates a mechanism for the transmission of informality. In the credit scheme, the VAT applies to each sale, and establishments receive a credit for the amount of tax paid upstream in the production chain, which is used against future tax liabilities. Since informal transactions are not recorded, informal suppliers do not generate tax credits. Similarly, informal firms cannot use payments from formal suppliers to generate tax credits. Hence, there is a potential mechanism for the transmission of informality (or formality) depending on how informal (or formal) the production chain is. Evidence from firm-level data from Brazil confirms the predictions of the model: The formality of a firm’s suppliers and buyers is correlated with its own formal status, and greater enforcement of regulations upstream or downstream implies a higher probability of being formal.¹³

Finally, a long-standing concern is that the competitive pressure introduced by trade reforms could induce reallocation of firms and workers from the formal to the informal sector (Goldberg

¹²The authors explore how variation in financial dependence across sectors (à la Rajan & Zingales 1998) interacts with time-varying economy-wide measures of financial deepening.

¹³In a related paper, Pomeranz (2015) shows that the paper trail generated by a VAT system has a preventive deterrence effect on tax evasion, which corresponds to the intensive margin of informality discussed above but applied to tax compliance (instead of labor regulations). Similarly, Naritomi (2019) uses unique administrative data from Brazil to show, among other results, that introducing incentives similar to the VAT for final sales leads to a substantial increase in firms’ reported revenue.

& Pavcnik 2003), especially in the case of a unilateral trade-opening episode. The basic argument is that negative demand shocks would push formal firms to downsize, hire a larger share of informal workers (the intensive margin), or exit altogether. Workers laid off from the formal sector would subsequently seek employment in the informal sector, and potential entrants could be discouraged to enter the formal sector and choose informality instead (the extensive margin). Using different research designs, Paz (2014), Cruces et al. (2018), Dix-Carneiro & Kovak (2019), and Ponczek & Ulyssea (2019) find positive and substantial effects of trade opening (tariff reduction) on informal employment in Argentina and Brazil.¹⁴ These results mirror those found by McCaig & Pavcnik (2018), who show that a positive export shock in Vietnam (due to the United States–Vietnam Bilateral Trade Agreement) caused a substantial reallocation of workers from informal microenterprises to formal sector firms. In a recent article (Dix-Carneiro et al. 2019) we develop an equilibrium trade model with firm dynamics and firm heterogeneity, formal and informal sectors, labor market frictions, and a rich institutional setting, which is estimated using several data sources from Brazil. Our counterfactual results are broadly consistent with this recent literature, as we find that reductions in (bilateral) trade costs lead to a sizeable reduction in informality within manufacturing. However, the total effect on informality—which cannot be assessed in these reduced-form studies—is very small.

3.2. Workers

Similarly to firms, workers choose between formal and informal jobs comparing costs and benefits. On the one hand, formal employment implies coverage by labor regulations and social security, there are typically higher job security and lower volatility in earnings (Gomes et al. 2018), and there can be higher on-the-job human capital accumulation (Bobba et al. 2019). On the other hand, individuals can evade income taxes if working informally and can have greater flexibility in hours worked in the face of a restrictive labor market regulation. Additionally, part of the cost of a formal job might be the expected waiting time if formal jobs are in short supply due to, say, the presence of a binding minimum wage. In this case, informal jobs might represent a desirable alternative for unemployed individuals who would otherwise have to face long queues for a formal job. Some of these dimensions are more static in nature, while others have a direct connection to individuals' life cycle. In what follows I use this separation to discuss the existing literature.

3.2.1. Static determinants. The literature has largely focused on the effect of welfare policies on labor supply and the allocation between formal and informal employment. Bosch & Campos-Vazquez (2014) analyze the case of Seguro Popular (SP) in Mexico, which was created in 2002 and aimed at creating universal health coverage, in particular including all informal (previously uninsured) workers. Since SP provided free health coverage—which was before tied to payroll contributions—the program decreased the costs of informality and was thus expected to cause an increase in informal employment (Levy 2008). The authors show that the program indeed had a negative effect on the formality trend (i.e., affiliation to social security) in small and medium firms: Had the program not been in place, approximately 4.6% and 4% additional employers and employees, respectively, would have formally registered.¹⁵ This effect is not very large, which

¹⁴Goldberg & Pavcnik (2003) and Bosch et al. (2012) find no effect of trade opening on informal employment in Brazil. As discussed by Dix-Carneiro & Kovak (2019), these contrasting results can be explained by differences in research design.

¹⁵Camacho et al. (2013) find similar effects from a related program in Colombia (the Subsidized Health Regime), which led to an increase in informal employment of around 4 percentage points. In contrast, Azuara & Marinescu (2013) find a very small effect of SP on informality among unskilled workers

suggests that the value of the health insurance provided by the program was not substantial. Conti et al. (2018) develop and estimate a household search model with formal and informal sectors and show that this is indeed the case: They estimate the utility value of SP to be around 4% and 9% of the mean household income for high- and low-education households, respectively.

Another potentially important determinant is the availability of cash transfer programs. Despite the gains from these programs in terms of reducing poverty, improving educational outcomes, and increasing access to health services, a long-standing concern is that they might discourage work (Banerjee et al. 2017). Additionally, since these programs are often means-tested and governments can only verify earnings from registered employment, they could generate strong incentives for workers to reallocate their labor supply to the informal sector. Evidence from the Bolsa Família in Brazil (De Brauw et al. 2015), the Plan de Atención Nacional a la Emergencia Social in Uruguay (Bergolo & Cruces 2018), and the Asignación Universal por Hijo in Argentina (Garganta & Gasparini 2015) systematically suggest that this is the case. In particular, Bergolo & Cruces (2018) find a reduction of formal employment in eligible households of around 8 percentage points, which are equally distributed between informal employment and nonemployment. Garganta & Gasparini (2015) find an even stronger effect: The probability of formalization declines 28–43% with respect to the counterfactual with no program.

Unemployment insurance (UI) programs have also been investigated as potential determinants of formalization decisions. On the one hand, a more generous UI system could increase the attractiveness of formal jobs. On the other hand, once workers qualify for UI there is an incentive to accumulate the benefits with nontaxed earnings from an informal job [Gerard & Gonzaga (2018) provide evidence that this is indeed the case in Brazil]. Bosch & Esteban-Pretel (2015) develop a search-and-matching model that is one of the few that accounts for both directions of UI effects on informality. Their results indicate that introducing a UI system in Mexico with a relatively low replacement rate (30%) decreases the share of formal employment with no effect on unemployment. Increasing the generosity of UI to a replacement rate of 70% generates different effects, increasing unemployment and formality. Their results therefore suggest that a more generous UI system induces more people to queue for a formal job, and this effect dominates the incentives to accumulate both the benefit and informal earnings. However, the literature as a whole shows mixed results regarding the implications of UI for unemployment and informality, as the results largely depend on the models' details and the country/data used to calibrate/estimate them (see, for example, Fugazza & Jacques 2004, Boeri & Garibaldi 2005, Albrecht et al. 2009, Ulyssea 2010b, Margolis et al. 2012, Meghir et al. 2015).

3.2.2. Dynamics. Some important determinants of informal labor supply can have life cycle effects, such as the design of social security systems. In particular, it is often the case that pension systems include some form of noncontributory, government-provided safety net that aims to guarantee a minimum income level for the elderly (Joubert 2015). These noncontributory benefits are typically decreasing in workers' contributions over their life cycle. Thus, they implicitly represent a tax on pension contributions, which reduces incentives to work formally, especially for low-skill individuals (for whom these benefits represent a greater share of lifetime income). Joubert (2015) develops and estimates a dynamic discrete choice model in the spirit of work by Keane & Wolpin (1997), which captures households' labor supply choice between formal and informal employment as well as saving decisions under the rules of Chile's pension system. His results indicate

(0.9 percentage points for a baseline informality rate of 60%) and no effect on the overall sample. The main difference between the two studies is that Bosch & Campos-Vazquez (2014) use administrative data, which substantially reduces measurement error.

that mandatory pension contributions can induce substantial shifts to informality: Increasing the contribution rate by 5 percentage points increases the size of the informal sector by 12.5% and 9.3% for men and women, respectively.

Finally, as discussed in Section 2.2, education and informality are strongly and negatively correlated in the data. Thus, to study individuals' decisions regarding human capital investments and sectoral choice in a unified framework seems arguably important. However, very few papers have used life cycle models to jointly analyze human capital investments and formal/informal labor supply decisions. Two notable exceptions are the studies by Garcia (2015) and Joubert (2015), who use dynamic life cycle models (in both cases estimated using Chilean micro data) to analyze this joint decision. Their results show that differences in perceived returns (pecuniary and nonpecuniary) explain most of the sorting between sectors (as opposed to potential barriers to workers' mobility), and that returns are higher in the formal sector for high-skill individuals. Moreover, similarly to work by Bobba et al. (2017), their results indicate that the interactions between schooling and formal/informal labor supply decisions have important implications for the analysis of formalization policies, as I discuss in the next section.

4. CONSEQUENCES

This section starts by reviewing the literature on the firm-level effects of formalization and then proceeds to a discussion of more aggregate effects. The latter relies on an extensive macro literature but also on a fairly recent but growing structural literature that integrates both micro and macro effects (Meghir et al. 2015; Bobba et al. 2017, 2019; Ulyssea 2018). The aggregate consequences are organized according to the two broad sets of determinants mentioned in the introduction: those that increase the costs (decrease the benefits) of informality and those that decrease the costs (increase the benefits) of formality.

4.1. Firm-Level Effects of Formalization

Since the seminal paper by Rauch (1991), the theoretical literature has modeled informality as the outcome of firms self-selecting into the formal and informal sectors based on their productivity and relative payoffs. In that sense, informality does not cause low productivity, but rather lower-quality firms self-select into informality (e.g., Arias et al. 2010). Thus, in a static world the causal effect of formality on firm productivity is not well defined, although one can still observe differences in firms' outcomes such as sales and size, since newly formalized firms no longer have to stay small to hide from the government. Dynamically, however, informality can have a negative effect on firms' productivity, as it can affect firms' investment decisions, technology adoption, and access to capital, among other factors.

The papers that seek to identify the effect of formality on firms' performance typically estimate some version of the general model

$$y_{it} = \alpha + \beta \text{Formal}_{it} + \gamma \mathbf{X}_{it} + \epsilon_{it},$$

where i indexes firms and t indexes time; y_{it} is a measure of firm performance, such as sales per worker, profits, number of employees, and so on; Formal_{it} is a dummy variable for being formal (i.e., registered); and \mathbf{X}_{it} is a vector of firm-level controls.

The identification problem is that the decision to be formal is clearly not exogenous and can be affected by elements that are unobservable to the econometrician (but not to the entrepreneurs), such as firm-level demand or productivity shocks and unobserved, time-varying entrepreneurial quality. The basic approach to circumvent this problem has been to use experimental or

quasi-experimental variation in access to policies or interventions that change the costs of formalizing, such as those discussed in Section 3.1, to instrument the formalization dummy $Formal_{it}$.

The results in the literature are somewhat mixed, but they generally indicate that formalizing has no statistically significant effects on different measures of firm performance, such as sales, profits, and number of employees (e.g., Benhassine et al. 2018, Rocha et al. 2018). For example, De Mel et al. (2013) find a positive average effect of formalization on profits, but this impact seems to be driven by few firms experiencing substantial growth. Looking at the distribution of profits, their results show that they are almost identical for treatment and control firms over most of the support. When examining the channels through which formalization could benefit firms, they find that firms that formalized increased advertising and use of receipt books but saw no increases in government contracts or participation in government programs, nor greater use of bank accounts or loans.¹⁶ This body of evidence is thus consistent with the argument that the perceived benefits of formalization are very low for most small-scale entrepreneurs (Bruhn & McKenzie 2014). Importantly, however, these firm-level outcomes are typically measured in relatively short time horizons (up to 2–3 years postformalization), while it might be the case that the positive effects of formality take longer to materialize.

4.2. Increasing the Costs of Informality

As discussed in Section 2.1, data from different countries show that both the extensive and intensive margins of informality decline with firm size, which suggests that the costs of informality are increasing with firm size. Accordingly, a common feature of most models in the literature is to assume some form of an increasing cost of informality with respect to firm size.¹⁷ Thus, a prototypical counterfactual experiment in equilibrium models is to simulate higher enforcement on informal firms by making this cost function steeper. This would correspond to a government crackdown on all informal firms, leading to a substantial reduction or complete eradication of informal firms.

All results available in the literature indicate that such policy could lead to substantial gains in aggregate productivity. These positive effects can come from a number of different channels that have been documented in the literature. First, the positive effects may operate via mechanical composition effects and reallocation of resources, as greater enforcement eliminates many low-productivity (informal) firms and resources are reallocated to more productive formal firms (e.g., Ulyssea 2010b, Bosch & Esteban-Pretel 2012, Charlot et al. 2015, Meghir et al. 2015, Ulyssea 2018). Second, and closely related to the previous point, in the presence of labor market frictions, reducing the availability of low-quality informal jobs can make it easier for workers to find higher-quality formal jobs (e.g., Meghir et al. 2015). Third, higher enforcement may lead to greater overall capital accumulation, as informal firms face higher financial frictions and are more credit constrained than formal sector firms (e.g., D’Erasmus & Boedo 2012, Ordonez 2014). Fourth, such policy may change occupational choices and discourage low-skill individuals to self-select into informal entrepreneurship, therefore increasing labor supply (e.g., Ordonez 2014, López 2017). Fifth, the policy may prompt higher investments in human capital, both before entering the labor market (i.e., formal schooling) and through on-the-job accumulation (Bobba et al. 2017, 2019).

¹⁶It is worth noting that the point estimates in these studies are often large but very imprecisely estimated. Thus, these null results are not very precisely estimated zeros, but rather they are often economically meaningful point estimates with large standard errors.

¹⁷Following work by Rauch (1991), the cost of informality assumes the form of either a perfect enforcement (penalty) after a size threshold—typically measured by number of employees, revenues, or capital—or a smoother and increasing cost function with respect to size (e.g., Fortin et al. 1997, De Paula & Scheinkman 2011, Ordonez 2014, Meghir et al. 2015, Ulyssea 2018).

The relevant trade-off therefore comes from the potential adverse consequences of greater enforcement. Indeed, in an earlier paper Boeri & Garibaldi (2005) argue that large informal sectors are widely tolerated, despite substantial improvements in detection technologies, because increasing enforcement would lead to higher unemployment. Using different equilibrium matching models calibrated to the Brazilian economy, Ulyssea (2010a) and Charlot et al. (2015) find results consistent with this trade-off: Even though greater enforcement substantially reduces informality, it also increases unemployment and reduces welfare. More recently, however, Meghir et al. (2015), Haanwinckel & Soares (2016), and Dix-Carneiro et al. (2019) find no unemployment effects from higher enforcement, and Meghir et al. (2015) even show that higher enforcement could lead to higher wages and lower inequality among their sample of low-skill males in the state of São Paulo, Brazil.¹⁸ Importantly, these papers move away from the one-to-one matching structure, so their models allow for the presence of large and productive formal firms that can absorb workers displaced from the informal sector.

Similarly, D’Erasmus & Boedo (2012), Ordóñez (2014), and Ulyssea (2018, 2019) show in their counterfactual analyses that output increases with higher enforcement. Thus, even though their models do not include labor market frictions (and therefore cannot speak to unemployment effects), their results indicate that the reallocation of resources to the formal sector more than compensates for the potential losses of output due to many informal firms shutting down entirely (i.e., never formalizing). The welfare implications in all of these exercises, however, must be considered with care for at least two reasons. First, these counterfactuals refer to steady state equilibria and cannot speak to the adjustment costs between steady states. Second, there are no good estimates of implementation costs. Given that informal firms are numerous, small in scale, and often geographically spread, these costs are likely to be very high and therefore can have a first-order effect on any welfare measure.

So far, the discussion has focused on enforcement on the extensive margin of informality. However, greater enforcement on the intensive margin generates very different results (Ulyssea 2018): It reduces the share of informal employment but can actually lead to a higher share of informal firms. The reason is very simple: By increasing enforcement of costly labor regulations, the government increases the de facto costs of formality for smaller formal firms. Thus, in the new steady state equilibrium many firms self-select out of formality and into the informal sector. This policy also has redistributive effects across formal firms: Low-productivity firms incur losses in their lifetime profits, while high-productivity firms benefit from the policy. The effects on aggregate total factor productivity (TFP) are very small (1.7% increase), and output actually decreases (by 1.6%) because the net displacement of firms dominates the gains in aggregate TFP.

The reduced-form evidence on the effects of enforcement is restricted to the intensive margin. Using micro data from Brazil, Almeida & Carneiro (2009) exploit the fact that enforcement of labor regulation in the country is decentralized and shows substantial geographic variation (their unit of analysis is the municipality). They use a credible IV strategy to show that an increase in the number of inspections per hundred formal firms leads to modest reductions in average firm size, output, and sales at the municipality level. Hence, their results are very much in line with those found by Ulyssea (2018). In a follow-up paper (Almeida & Carneiro 2012), the authors show that more inspections cause modest reductions in the proportion of formal and self-employed workers and an increase in nonemployment. Using the same data on enforcement, but with a different empirical strategy, Ponczek & Ulyssea (2019) assess whether greater enforcement of labor regulations can amplify the adverse labor market effects of a negative economic shock. For that, we

¹⁸Prado (2011) shows a more nuanced view by exploring differences across a broad set of countries. His results indicate that the welfare effects of greater enforcement largely depend on the baseline regulatory costs of each country.

exploit the strong, largely negative shock to local labor markets brought about by the unilateral trade liberalization in Brazil. We show that for a given trade-induced (negative) shock, regions within Brazil that had a weaker enforcement of labor regulations experienced higher informality effects but no unemployment effects. Symmetrically, regions with stricter enforcement did not experience higher informality but had substantial increases in unemployment.¹⁹ Hence, these results suggest that the de facto labor market flexibility introduced by informality leads to lower employment losses in face of an adverse economic shock.

Finally, using unique firm-level data from Mexico, de la Parra (2016) investigates the consequences of increasing enforcement on the intensive margin of informality. She uses data on inspections at randomly selected formal establishments in Mexico to analyze the effect of enforcement of labor regulation on informal and formal employment, turnover, and wages. Her results show that inspections increase the transition probability from informal to formal jobs within the same establishment from 14% to 20% within the first three months after the inspection occurs. She also finds disemployment effects, as informal workers at inspected establishments are more likely to transit into unemployment. Hence, de la Parra's results confirm the potential disemployment effects of greater enforcement on the intensive margin.

4.3. Reducing the Costs or Increasing the Benefits of Formality

As mentioned in Section 3.1, perhaps the most studied formalization policy is the reduction of formal sector's entry costs. Even though these reforms have not been effective in inducing firms to formalize, they might still produce important aggregate effects.

Indeed, different papers in the literature show that reducing entry costs into the formal sector can produce positive and sizeable aggregate effects (Ulyssea 2010a, D'Erasmus & Boedo 2012, Charlot et al. 2015, Ulyssea 2018). In a recent article (Ulyssea 2018) I show that reducing entry costs into the formal sector eliminates deadweight losses from wasteful barriers to entry, which leads to increases in competition, aggregate production in the formal sector, and high-skill wages (the formal sector is more intensive in high-skill labor). Since the formalization process is concentrated among low-productivity firms, there is a negative composition effect that leads to a decrease in aggregate TFP. Nevertheless, total output increases because there is a substantial increase in the mass of active firms in the economy. Ulyssea (2010a) and Charlot et al. (2015) find similar positive aggregate effects, which also translate into substantially lower unemployment and a smaller formal-informal wage gap. D'Erasmus & Boedo (2012) argue that differences in formal sector's entry costs can explain roughly three-quarters of the gap in TFP between the United States and their calibration of two prototypical low- and middle-income countries.

In contrast, Lopez-Martin (2019) finds limited aggregate effects from reducing entry costs in Mexico and Egypt: In Mexico the gains in aggregate TFP and output per capita would be of 0.5 and 0.7 percentage points, respectively, with even more limited gains in Egypt. The main reason for this difference seems to be the fact that his model strongly emphasizes the role played by financial frictions in the formal sector (i.e., the ability of firms to collateralize their assets and access credit), and these are the binding constraints that, once relaxed, generate substantial gains in aggregate productivity, output, and welfare [D'Erasmus (2016) finds similar results].

As argued in Section 3.1, another important dimension of the costs of formality relates to tax payments. Indeed, the reduced-form evidence suggests that reducing the tax burden can induce some formalization, albeit the elasticity seems to be low. The counterfactual results from both

¹⁹ Almeida & Poole (2017) find related results when analyzing industry-specific exchange rate shocks in the late 1990s in Brazil and their effects on formal employment.

macro and structural models seem to corroborate that, because reductions in payroll tax seem to generate some formalization but with a low elasticity (e.g., D’Erasmus & Boedo 2012, Haanwinckel & Soares 2016). If one differentiates the effects on the intensive and extensive margins of informality, the effects are stronger on labor informality (via the intensive margin) and weaker on firm informality (Ulyssea 2018). Consistently with the limited effects on informality, the results from different papers indicate that reducing payroll taxes has very limited (but overall positive) aggregate effects (Ulyssea 2010a, D’Erasmus & Boedo 2012, Haanwinckel & Soares 2016, Ulyssea 2018).

5. CONCLUSION

This article reviews and synthesizes an extensive economic literature that investigates informality, its causes, and its consequences for development. As this review makes clear, this body of research is not only extensive but also quite diverse in terms of the approaches and tools used, ranging from well-identified experimental studies to equilibrium macro models and, more recently, structural models. This variety of approaches is a natural consequence of the very nature of the object of interest. On the one hand, informality is a micro phenomenon, as it is the result of individual agents maximizing their payoffs in the economic environment they face. On the other hand, the sheer magnitude of informality and the scale of the policies used to address it make informality a macro phenomenon as well, and one that has deep implications for the economy as a whole. Thus, a complete understanding of the consequences of informality must encompass both dimensions.

Overall, the existing studies seem to indicate that reducing the costs of entering the formal sector is not the most effective way to reduce informality, but it can nevertheless have important positive effects on the economy. The magnitude of these positive effects, however, remains a point of some controversy, with different studies finding different results. Reducing the ongoing costs of formality, in particular taxes, has stronger formalization effects but not strong enough to make these policies cost effective. On the macro side, they seem to produce some positive but very modest effects on output and TFP. Finally, the most effective policy to reduce informality seems to be intensifying enforcement. However, it is crucial to differentiate between enforcement on the extensive and the intensive margins, as higher enforcement on the latter tends to produce more disemployment effects and worse aggregate effects. As for greater enforcement on the extensive margin, it generates substantial positive effects on aggregate TFP and output, while the key point is the net effect on employment. More recent results from structural studies seem to indicate that greater enforcement does not increase long-run unemployment, and therefore this policy should generate net welfare gains. However, a key unexplored area of research is the transitional dynamics of such policies. In particular, if the adjustment of labor from the informal to the formal sector takes a lot of time and differs across skill levels, then the overall welfare cost might be substantially higher.

Finally, another interesting and largely unexplored area of research relates to the role of informality in developed economies. In particular, the importance of the intensive margin of informality is likely to increase with the rise of nonstandard (or atypical) work arrangements, also commonly referred to as gig economy jobs. Not only has the visibility of these nonstandard work arrangements substantially increased with the rise of companies such as Uber, Airbnb, and TaskRabbit, but also the actual fraction of the workforce in the gig economy seems to have reached quite substantial levels, at least in the United States (Bracha & Burke 2016, 2018). Clearly, the economic contexts are quite different. In particular, the emergence of gig jobs seems to be disproportionately benefiting large and high-productivity firms, while the intensive margin of informality in developing countries is concentrated on the left tail of the productivity distribution. Nevertheless, many of

the main economic forces at play are similar, and a better understanding of this rise of informality in developed countries is arguably an exciting area for future research.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I have deeply benefited from discussions with different colleagues, at different points in time, about the various dimensions of informality. In particular, but without implicating them, I am extremely grateful to José Márcio Camargo, Ricardo Paes de Barros, James J. Heckman, Chang-Tai Hsieh, Steven Durlauf, Penny Goldberg, Rafael Dix-Carneiro, Costas Meghir, and Claudio Ferraz.

LITERATURE CITED

- Akcigit U, Alp H, Peters M. 2018. *Lack of selection and limits to delegation: firm dynamics in developing countries*. NBER Work. Pap. 21905
- Albrecht J, Navarro L, Vroman S. 2009. The effects of labour market policies in an economy with an informal sector. *Econ. J.* 119:1105–29
- Allen J, Nataraj S, Schipper TC. 2018. Strict duality and overlapping productivity distributions between formal and informal firms. *J. Dev. Econ.* 135:534–54
- Almeida R, Carneiro P. 2009. Enforcement of labor regulation and firm size. *J. Comp. Econ.* 37:28–46
- Almeida R, Carneiro P. 2012. Enforcement of labor regulation and informality. *Am. Econ. J. Appl. Econ.* 4:64–89
- Almeida RK, Poole JP. 2017. Trade and labor reallocation with heterogeneous enforcement of labor regulations. *J. Dev. Econ.* 126:154–66
- Arias J, Azuara O, Bernal P, Heckman JJ, Villarreal C. 2010. *Policies to promote growth and economic efficiency in Mexico*. NBER Work. Pap. 16554
- Auriol E, Warlters M. 2005. Taxation base in developing countries. *J. Public Econ.* 89:625–46
- Azuara O, Marinescu I. 2013. Informality and the expansion of social protection programs: evidence from Mexico. *J. Health Econ.* 32:938–50
- Banerjee AV, Hanna R, Kreindler GE, Olken BA. 2017. Debunking the stereotype of the lazy welfare recipient: evidence from cash transfer programs. *World Bank Res. Obs.* 32:155–84
- Bargain O, Kwenda P. 2014. The informal sector wage gap: new evidence using quantile estimations on panel data. *Econ. Dev. Cult. Change* 63:117–53
- Benhassine N, McKenzie D, Pouliquen V, Santini M. 2018. Does inducing informal firms to formalize make sense? Experimental evidence from Benin. *J. Public Econ.* 157:1–14
- Bergolo M, Cruces G. 2018. *The anatomy of behavioral responses to social assistance when informal employment is high*. Unpublished manuscript, Univ. República, Montevideo, Urug.
- Bobba M, Flabbi L, Levy S. 2017. *Labor market search, informality and schooling investments*. IZA Discuss. Pap. 11170, Inst. Labor Econ., Bonn, Ger.
- Bobba M, Flabbi L, Levy S, Tejada M. 2019. Labor market search, informality, and on-the-job human capital accumulation. *J. Econom.* In press
- Boeri T, Garibaldi P. 2005. Shadow sorting. In *NBER Macroeconomics Annual*, ed. C. Pissarides, J. Frenkel, pp. 125–63. Cambridge, MA: MIT Press
- Bosch M, Campos-Vazquez RM. 2014. The trade-offs of welfare policies in labor markets with informal jobs: the case of the “Seguro Popular” program in Mexico. *Am. Econ. J. Econ. Policy* 6:71–99
- Bosch M, Esteban-Pretel J. 2012. Job creation and job destruction in the presence of informal markets. *J. Dev. Econ.* 98:270–86

- Bosch M, Esteban-Pretel J. 2015. The labor market effects of introducing unemployment benefits in an economy with high informality. *Eur. Econ. Rev.* 75:1–17
- Bosch M, Goñi-Pacchioni E, Maloney W. 2012. Trade liberalization, labor reforms and formal–informal employment dynamics. *Labour Econ.* 19:653–67
- Bosch M, Maloney WF. 2010. Comparative analysis of labor market dynamics using Markov processes: an application to informality? *Labour Econ.* 17:621–31
- Botelho F, Ponczek V. 2011. Segmentation in the Brazilian labor market. *Econ. Dev. Cult. Change* 59:437–63
- Botero JC, Djankov S, Porta RL, Lopez-de-Silanes F, Shleifer A. 2004. The regulation of labor. *Q. J. Econ.* 119:1339–82
- Bracha A, Burke MA. 2016. *Who counts as employed? Informal work, employment status, and labor market slack.* Work. Pap., Fed. Reserve Bank Boston, Boston, MA
- Bracha A, Burke MA. 2018. Wage inflation and informal work. *Econ. Lett.* 171:159–63
- Bruhn M. 2011. License to sell: the effect of business registration reform on entrepreneurial activity in Mexico. *Rev. Econ. Stat.* 93:382–86
- Bruhn M, McKenzie D. 2014. Entry regulation and the formalization of microenterprises in developing countries. *World Bank Res. Obs.* 29:186–201
- Busso M, Fazio M, Algazi S. 2012. (In)formal and (un)productive: the productivity costs of excessive informality in Mexico. IDB Work. Pap. IDB-WP-341, Inter-Am. Dev. Bank, Washington, DC
- Camacho A, Conover E, Hoyos A. 2013. Effects of Colombia's social protection system on workers' choice between formal and informal employment. *World Bank Econ. Rev.* 28:446–66
- Catão LA, Pagés C, Rosales MF. 2009. *Financial dependence, formal credit, and informal jobs: new evidence from Brazilian household data.* IZA Discuss. Pap. 4609, Inst. Labor Econ., Bonn, Ger.
- Charlot O, Malherbet F, Terra C. 2015. Informality in developing economies: regulation and fiscal policies. *J. Econ. Dyn. Control* 51:1–27
- Cisneros-Acevedo C. 2019. *Unfolding trade effect in two margins of informality: the Peruvian case.* Job Mark. Pap., Univ. Nottingham, Nottingham, UK
- Conti G, Ginja R, Narita R. 2018. *The value of health insurance: a household job search approach.* IZA Discuss. Pap. 11706, Inst. Labor Econ., Bonn, Ger.
- Cruces G, Porto G, Viollaz M. 2018. Trade liberalization and informality in Argentina: exploring the adjustment mechanisms. *Lat. Am. Econ. Rev.* 27:13
- Dabla-Norris E, Gradstein M, Inchauste G. 2008. What causes firms to hide output? The determinants of informality. *J. Dev. Econ.* 85:1–27
- de Andrade GH, Bruhn M, McKenzie D. 2016. A helping hand or the long arm of the law? Experimental evidence on what governments can do to formalize firms. *World Bank Econ. Rev.* 30:24–54
- De Brauw A, Gilligan DO, Hoddinott J, Roy S. 2015. Bolsa Familia and household labor supply. *Econ. Dev. Cult. Change* 63:423–57
- De Giorgi G, Rahman A. 2013. SME's registration: evidence from an RCT in Bangladesh. *Econ. Lett.* 120:573–78
- de la Parra B. 2016. *Formal firms, informal workers and household labor supply in Mexico.* Job Mark. Pap., Univ. Chicago, Chicago
- De Mel S, McKenzie D, Woodruff C. 2013. The demand for, and consequences of, formalization among informal firms in Sri Lanka. *Am. Econ. J. Appl. Econ.* 5:122–50
- De Paula A, Scheinkman JA. 2010. Value-added taxes, chain effects, and informality. *Am. Econ. J. Macroecon.* 2:195–221
- De Paula A, Scheinkman JA. 2011. The informal sector: an equilibrium model and some empirical evidence. *Rev. Income Wealth* 57:S8–26
- De Soto H. 1989. *The Other Path: The Economic Answer to Terrorism.* New York: Harper & Row
- D'Erasmus PN. 2016. Access to credit and the size of the formal sector. *Economia* 16:143–99
- D'Erasmus PN, Boedo H. 2012. Financial structure, informality and development. *J. Monet. Econ.* 59:286–302
- Dix-Carneiro R, Goldberg PK, Meghir C, Ulyssea G. 2019. *Trade and informality in the presence of labor market frictions and regulations.* Unpublished manuscript, Duke Univ., Durham, NC
- Dix-Carneiro R, Kovak BK. 2019. Margins of labor market adjustment to trade. *J. Int. Econ.* 117:125–42

- Djankov S, Ganser T, McLiesh C, Ramalho R, Shleifer A. 2010. The effect of corporate taxes on investment and entrepreneurship. *Am. Econ. J. Macroecon.* 2:31–64
- Djankov S, Porta RL, Lopez-de-Silanes F, Shleifer A. 2002. The regulation of entry. *Q. J. Econ.* 117:1–37
- El Badaoui E, Strobl E, Walsh F. 2010. The formal sector wage premium and firm size. *J. Dev. Econ.* 91:37–47
- Eslava M, Haltiwanger JC, Pinzón A. 2019. *Job creation in Colombia versus the US: “Up or out dynamics” meets the “life cycle of plants.”* NBER Work. Pap. 25550
- Fajnzylber P, Maloney WF, Montes-Rojas GV. 2011. Does formality improve micro-firm performance? Evidence from the Brazilian SIMPLES program. *J. Dev. Econ.* 94:262–76
- Fields G. 1990. Labour market modelling and the urban informal sector: theory and evidence. In *The Informal Sector Revisited*, ed. D Turnham, B Salome, A Schwarz, pp. 49–69. Paris: OECD
- Fortin B, Marceau N, Savard L. 1997. Taxation, wage controls and the informal sector. *J. Public Econ.* 66:293–312
- Friedman E, Johnson S, Kaufmann D, Zoido-Lobaton P. 2000. Dodging the grabbing hand: the determinants of unofficial activity in 69 countries. *J. Public Econ.* 76:459–93
- Fugazza M, Jacques JF. 2004. Labor market institutions, taxation and the underground economy. *J. Public Econ.* 88:395–418
- Funkhouser E. 1996. The urban informal sector in Central America: household survey evidence. *World Dev.* 24:1737–51
- Galiani S, Weinschelbaum F. 2012. Modelling informality formally: households and firms. *Econ. Inq.* 50:821–38
- Garcia IL. 2015. *Human capital and labor informality in Chile: a life-cycle approach.* RAND Work. Pap. WR-1087, RAND Corp., Santa Monica, CA. https://www.rand.org/pubs/working_papers/WR1087.html
- Garganta S, Gasparini L. 2015. The impact of a social program on labor informality: the case of AUH in Argentina. *J. Dev. Econ.* 115:99–110
- Gasparini L, Tornarolli L. 2009. Labor informality in Latin America and the Caribbean: patterns and trends from household survey microdata. *Desarrollo. Soc.* 2009:13–80
- Gerard F, Gonzaga G. 2018. *Informal labor and the efficiency cost of social programs: evidence from the Brazilian unemployment insurance program.* NBER Work. Pap. 22608
- Goldberg PK, Pavcnik N. 2003. The response of the informal sector to trade liberalization. *J. Dev. Econ.* 72:463–96
- Gomes D, Iachan F, Santos C. 2018. *Labor earnings dynamics with a large informal sector.* Unpublished manuscript, Univ. Alberta, Edmonton, Can.
- Gong X, Van Soest A. 2002. Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico. *Labour Econ.* 9:513–29
- Haanwinckel D, Soares RR. 2016. *Workforce composition, productivity, and labor regulations in a compensating differentials theory of informality.* IZA Discuss. Pap. 9951, Inst. Labor Econ., Bonn, Ger.
- Hsieh CT, Klenow PJ. 2014. The life cycle of plants in India and Mexico. *Q. J. Econ.* 129:1035–84
- Hsieh CT, Olken B. 2014. The missing “missing middle.” *J. Econ. Perspect.* 28:89–108
- Johnson S, Kaufmann D, Zoido-Lobaton P. 1998. Regulatory discretion and the unofficial economy. *Am. Econ. Rev.* 88:387–92
- Joubert C. 2015. Pension design with a large informal labor market: evidence from Chile. *Int. Econ. Rev.* 56:673–94
- Kaplan DS, Piedra E, Seira E. 2011. Entry regulation and business start-ups: evidence from Mexico. *J. Public Econ.* 95:1501–15
- Keane MP, Wolpin KI. 1997. The career decisions of young men. *J. Political Econ.* 105:473–522
- La Porta R, Shleifer A. 2008. The unofficial economy and economic development. *Brook. Pap. Econ. Act.* 105:473–522
- La Porta R, Shleifer A. 2014. Informality and development. *J. Econ. Perspect.* 28:109–26
- Levy S. 2008. *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico.* Washington, DC: Brook. Inst. Press
- Loayza NA. 1996. The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conf. Ser. Public Policy* 45:129–62
- López JJ. 2017. A quantitative theory of tax evasion. *J. Macroecon.* 53:107–26

- Lopez-Martin B. 2019. Informal sector misallocation. *Macroecon. Dyn.* 23:3065–98
- Magnac T. 1991. Segmented or competitive labour markets. *Econometrica* 59:165–87
- Maloney W. 1999. Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico. *World Bank Econ. Rev.* 13:275–302
- Margolis DN, Navarro L, Robalino DA. 2012. *Unemployment insurance, job search and informal employment*. IZA Discuss. Pap. 6660, Inst. Labor Econ., Bonn, Ger.
- McCaig B, Pavcnik N. 2018. Export markets and labor allocation in a low-income country. *Am. Econ. Rev.* 108:1899–941
- Meghir C, Narita R, Robin JM. 2015. Wages and informality in developing countries. *Am. Econ. Rev.* 105:1509–46
- Monteiro J, Assunção JJ. 2012. Coming out of the shadows? Estimating the impact of bureaucracy simplification and tax cut on formality in Brazilian microenterprises. *J. Dev. Econ.* 99:105–15
- Naritomi J. 2019. Consumers as tax auditors. *Am. Econ. Rev.* 109:3031–72
- Ordóñez JCL. 2014. Tax collection, the informal sector, and productivity. *Rev. Econ. Dyn.* 17:262–86
- Paes de Barros R. 1988. *On the empirical content of the formal-informal labor market segmentation hypothesis*. Work. Pap., Econ. Res. Cent.–NORC, Univ. Chicago, Chicago
- Pagés C, Stampini M. 2009. No education, no good jobs? Evidence on the relationship between education and labor market segmentation. *J. Comp. Econ.* 37:387–401
- Paz LS. 2014. The impacts of trade liberalization on informal labor markets: a theoretical and empirical evaluation of the Brazilian case. *J. Int. Econ.* 92:330–48
- Perry G, Maloney W, Arias O, Fajnzylber P, Mason A, Saavedra-Chanduvi J. 2007. *Informality: Exit or Exclusion*. Washington, DC: World Bank
- Peters M. 2019. *Heterogeneous mark-ups, growth and endogenous misallocation*. Unpublished manuscript, Yale Univ., New Haven, CT
- Piza C. 2018. Out of the shadows? Revisiting the impact of the Brazilian SIMPLES program on firms' formalization rates. *J. Dev. Econ.* 134:125–32
- Pomeranz D. 2015. No taxation without information: deterrence and self-enforcement in the value added tax. *Am. Econ. Rev.* 105:2539–69
- Ponczek V, Ulyssea G. 2019. *Is informality an employment buffer? Evidence from the trade liberalization in Brazil*. Unpublished manuscript, São Paulo Sch. Econ., São Paulo, Braz.
- Prado M. 2011. Government policy in the formal and informal sectors. *Eur. Econ. Rev.* 55:1120–36
- Pratap S, Quintin E. 2006. Are labor markets segmented in developing countries? A semiparametric approach. *Eur. Econ. Rev.* 50:1817–41
- Quintin E. 2008. Contract enforcement and the size of the informal economy. *Econ. Theory* 37:395–416
- Rajan RG, Zingales L. 1998. Financial dependence and growth. *Am. Econ. Rev.* 88:559–86
- Rauch JE. 1991. Modelling the informal sector formally. *J. Dev. Econ.* 35:33–47
- Rocha R, Ulyssea G, Rachter L. 2018. Do lower taxes reduce informality? Evidence from Brazil. *J. Dev. Econ.* 134:28–49
- Schneider F, Enste D. 2000. Shadow economies: size, causes and consequences. *J. Econ. Lit.* 38:77–114
- Straub S. 2005. Informal sector: the credit market channel. *J. Dev. Econ.* 78:299–321
- Ulyssea G. 2010a. The formal-informal labor market segmentation hypothesis revisited. *Braz. Rev. Econom.* 30:311–34
- Ulyssea G. 2010b. Regulation of entry, labor market institutions and the informal sector. *J. Dev. Econ.* 91:87–99
- Ulyssea G. 2018. Firms, informality, and development: theory and evidence from Brazil. *Am. Econ. Rev.* 108:2015–47
- Ulyssea G. 2019. *Formal and informal firm dynamics*. Unpublished manuscript, Univ. Oxford, Oxford, UK