VoxDevLits are wiki-inspired literature reviews that aim to summarise the evidence base on narrowly defined topics related to development economics. Each Lit is written by a community of scholars working on the specific topic addressed in the review. They are intended for both policymakers and researchers. We aim to describe what we have learned from research and to highlight the important questions for which evidence is lacking. The Lits are living documents that will be updated several times each year. All published versions will be available on the VoxDev website so that scholars can cite the reviews with confidence that the version cited will be accessible in the future.

Please contact us at engage@voxdev.org if you have any questions or feedback.
Most low- and middle-income countries are characterised by a large informal sector, which implies that a substantial fraction of economic activity in these countries is completely unregulated. This has important implications for the behaviour of firms, workers, families, and consumers, with these micro-level consequences of informality potentially combining to have important effects on aggregate outcomes such as productivity, output, and growth. This VoxDevLit summarises a wide range of studies ranging from well-identified empirical analyses to macro and structural equilibrium models of informality. The final goal is to provide an accessible, integrated understanding of the causes and consequences of informality for economic development.

Citation: Gabriel Ulyssea, Matteo Bobba and Lucie Gadenne, “Informality” VoxDevLit, 6.1, January 2023
Informality

Low- and middle-income countries are typically characterised by a high level of informality, which has significant implications for firms, workers, families, and consumers that, in turn, can result in potentially large aggregate effects. This VoxDevLit summarises a wide range of studies to understand what we have learned about the causes and consequences of informality for economic development.

We start by defining what informality is. We define informal firms and workers as those who do not comply with the relevant laws and regulations. For workers, this boils down to having a formal labour contract or not. For firms, we distinguish between (i) the extensive margin of informality, whether firms register and pay entry fees to achieve a formal status; and (ii) the intensive margin, whether firms that are formal in the first sense hire workers without a formal contract. With these definitions in hand, we summarise the main informality facts established in the literature regarding firms and workers, respectively.

The survey then moves on to discuss the main determinants of informality. We start with empirical studies that use both experimental and non-experimental designs to investigate the causal effects of different determinants of firms’ decisions to formalise. These can be classified into two broad groups: (i) those associated with the costs of entering the formal sector; and (ii) those related to the costs of remaining formal (such as tax payments). On the determinants of workers’ informality, we emphasise the role of public policies – such as conditional cash transfer programs – and the role of human capital.

In the final part of the review, we summarise the literature that investigates the consequences of informality for firms, as well as the aggregate effects on taxation, redistribution, human capital accumulation, productivity, output and growth. Most of the literature discussed in this part of the survey combines the use of rich micro data with partial or general equilibrium models. Hence, most of them build from micro behaviour to aggregate outcomes to assess the potential effects of different formalisation policies.

This survey shows that the literature has made substantial progress in understanding the main determinants of firms’ choices regarding informality, both theoretically and empirically. However, some important questions remain unanswered. In particular, the literature has only started to explore the dynamics of firms’ decisions regarding the different margins of informality and how they interact with other frictions that firms face. For example, does informality work as a stepping-stone for entrepreneurs with high-growth potential but who might be constrained by, say, credit constraints?

On the workers’ side, the gaps are arguably larger. We need a deeper understanding of the determinants of workers’ allocation between formal and informal jobs, and the main trade-offs they face. A particularly important gap refers to how much workers value the greater job security provided by formal employment relative to informal jobs. As in the case of firms, we also do not know how much informal jobs represent a stepping-stone for younger workers versus the extent to which there is an “informality trap” that makes future transitions into formal employment very unlikely. More broadly, there are still very few studies that investigate the life-cycle dimensions of informality.
1 Introduction

Most low- and middle-income countries are characterised by a large informal sector, which accounts for 30-70% of GDP, 20-80% of the labour force and an equally large share of firms (Ulyssea 2020). This means that a substantial fraction of economic activity in these countries is completely unregulated, taking place at the margin of tax, labour and other relevant regulatory frameworks. This has deep implications for individual behaviour – not only firms and workers, but also families and consumers – which build all the way up to aggregate outcomes, generating important effects on outcomes such as productivity, output and growth. Thus, a complete understanding of the causes and consequences of informality requires integrating both its micro and macro dimensions. This VoxDevLit aims at doing that, covering a wide range of studies that go from well-identified empirical analyses to macro and structural equilibrium models of informality.

We start by defining what informality is. We define informal firms and workers as those who do not comply with the relevant laws and regulations. For firms, this can correspond to, for example, not being registered with the tax authorities, while informal workers are those who do not have a formal labour contract. This is the definition used throughout this review. While this criterion might be straightforward for workers, it is less so for firms, as their compliance decision is unlikely to be a binary choice. In both developed and developing countries, many formally registered firms underreport revenue to evade taxes and are therefore in partial compliance with tax regulations. Similarly, many formal firms hire informal workers to evade labour regulation costs. We thus use the same definitions from Ulyssea (2018) and distinguish between the following margins of informality: (i) extensive margin, whether firms register and pay entry fees to achieve a formal status; and (ii) intensive margin, whether firms that are formal in the first sense hire workers without a formal contract. We restrict the definition of the intensive margin to compliance or not with labour regulations because of data limitations, as information on tax evasion at the firm level – another important intensive margin of informality – is typically unavailable.

Using these definitions as a starting point, in Sections 2 and 3 we summarise the main informality facts established in the literature regarding firms and workers, respectively. These facts set the scene for the discussion that follows and highlight many of the key dimensions of informality. Section 4 discusses the main determinants of informality investigated in the existing literature. We start in Section 4.1 by covering the main empirical studies that use both experimental and non-experimental designs to investigate the causal effects of different determinants of firms’ decisions to formalise. These determinants can be classified into two broad groups: (i) those associated to the costs of entering the formal sector, such as the costs of formally registering a business; and (ii) those related to the costs of remaining formal, such as tax payments and administrative costs associated to being formal (e.g. tax compliance). In Section 4.2 we focus on the evidence on the roles of tax structure (and not only the tax burden) and trade in determining the levels of informality. Sections 4.3 and 4.4 move to discuss the determinants of workers’ informality, emphasising the role of public policies – such as means-tested conditional cash transfer programs – and the role of human capital, respectively.

In Section 5 we review the literature that investigates the consequences of informality for firms, taxation and redistribution, and the aggregate effects on human capital accumulation, productivity, output and growth. On the latter, most of the literature covered in this Section combines extensive use of rich micro data, with partial or general equilibrium models that aim at rationalising a broad set of empirical regularities. In that sense, most of them build from micro behaviour to aggregate outcomes to assess the potential effects of different formalisation policies on the levels of informality and other key aggregate outcomes of interest.

This is the first release of the VoxDevLit on informality, which largely builds on a previous and quite comprehensive review by Ulyssea (2020). This VoxDevLit piece does not cover all topics reviewed in Ulyssea (2020), but we expand it to cover some new developments in the literature. As this is a dynamic literature review, this document will be updated regularly to include new and exciting studies as this body
of academic work continues to evolve. We look forward to readers’ feedback on the review, and to ongoing discussions on this fascinating topic.

2 Facts about firms

The existing literature has systematically shown that informal firms are on average smaller (both in terms of employees and revenues), pay lower wages, are run by less educated individuals, hire less educated workers and earn lower profits than formal firms (Ulyssea 2020). This is true for different countries and different data sets used. These differences have been often interpreted as evidence in favour of a dualistic view of informality, in which formal and informal firms are not integrated at all and operate in completely separate economic spaces, using different technologies and producing distinct goods. However, the data does not seem to support this view. Not only do they coexist within the same industries and produce similar products (e.g. Ulyssea 2018), but there is a substantial overlap in formal and informal firms’ productivity distributions, even within industries (Meghir et al. 2015, Ulyssea 2018, Allen et al. 2018).

Another important empirical regularity is that the share of informal firms (i.e. the extensive margin of informality) rapidly declines as firms grow larger (e.g. Perry et al. 2007, De Paula and Scheinkman 2011), as shown in Figure 1, Panel A. This fact indicates that the costs of operating in the informal sector are increasing in firm size (or the benefits decreasing). This is intuitive, as one would expect that larger informal firms have a harder time remaining undetected by the government. More broadly, the opportunity costs of operating in the informal sector are likely to be increasing in firm size. For example, larger firms might have greater need of accessing formal credit lines or issuing invoices to buyers, which is not possible if they remain informal.

Figure 1: Informality margins and firm size

Turning to the intensive margin of informality, it represents a substantial fraction of informal employment in developing countries: 56% in Mexico, at least 40% in Brazil and 32% in Peru (Bujanda and de la Parra 2020, Ulyssea 2018 and Cisneros-Acevedo 2022). The intensive margin also declines as firms grow larger, as the average share of informal employees within formal firms declines with firm size (see Figure 1, Panel B). This fact can also be rationalised by the fact that larger firms are more visible and therefore more likely...
to be inspected. Conversely, it is costly to inspect small firms and therefore government’s inspectors tend to focus on larger firms (e.g. Almeida and Carneiro 2012).

3 Facts about workers

The literature has extensively shown that informality among workers displays a U-shape pattern with respect to age (larger among younger and older workers), it is higher among women and decreases with schooling (e.g. Perry et al. 2007, Gasparini and Tornaroli 2009). Transitions in and out of informality follow a similar pattern: the young, women and low-skill workers have a higher probability of transiting from unemployment and formal jobs into informal employment (see, for example, Bosch and Maloney 2010).

A second set of facts refer to the ins and outs of informality over the business cycle. Informal employment (like unemployment) has been shown to be strongly counter-cyclical, expanding during recessions and decreasing during economic booms (as a fraction of employment). This can be explained by the combination of three important facts (see Perry et al. 2007, Bosch and Esteban-Preter 2012). First, the job finding rate in the formal sector is strongly pro-cyclical, but it is fairly stable in the informal sector. Second, informal to formal transitions are pro-cyclical. Third, separation rates are countercyclical in both sectors but more volatile in the informal sector.

A third well-established fact is the existence of a substantial formal-informal wage gap, which persists even after controlling for several observable characteristics (see Ulyssea 2020). However, Ulyssea (2018) uses 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. The estimated within-firm wage gap is statistically and economically zero, which suggests that self-selection is one of the main drivers of the wage gap between observably equivalent workers. Moreover, it also indicates that, conditional on workers’ skill, formal and informal workers perform the same tasks within the firm.

4 Determinants of informality

4.1 Firms: costs and benefits of (in)formality

This section discusses the determinants of firms’ decisions regarding the extensive and intensive margins of informality. Broadly speaking, the costs of informality can be categorised into two large groups: (i) the costs of entering the formal sector, such as the costs of formally registering a business; and (ii) the costs of remaining formal, such as tax payments and other ongoing administrative costs associated to being formal (costs associated to tax compliance beyond direct tax payments, for example). As discussed in Ulyssea (2020), if policymakers want to increase formalisation rates among firms, they can do so by reducing these costs of formality. A second approach would be to increase the benefits of formality, via greater access to capital, for example. Finally, it is possible to increase the costs of informality, which can be achieved by increasing enforcement of the existing laws and regulations (increasing the number of inspections, for example).

The policies or interventions analysed in papers that empirically investigate the potential determinants of firms’ formalisation decision can naturally be grouped into these three broad categories. The policies and interventions analysed in the literature have been highly concentrated in the first group – reducing the costs of formality – and in particular reducing the costs of entering the formal sector (Bruhn and McKenzie 2014). However, Figure 2 shows that the available results in the literature are not very encouraging. It summarises the results from both experimental and quasi-experimental studies that have a credible
research design to identify the causal effect of a given policy.¹

Figure 2: Formalisation effects on firms – experimental and quasi-experimental results

<table>
<thead>
<tr>
<th>Study</th>
<th>Intervention Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>De Giorgi &amp; Rahman (2013)</td>
<td>Information about the process of registration and its potential benefits</td>
</tr>
<tr>
<td>Andrade et al. (2016)</td>
<td>Information and reimbursing all registration costs</td>
</tr>
<tr>
<td>de Mel et al. (2013)</td>
<td>Creating national formalisation programs that substantially reduce registration costs</td>
</tr>
<tr>
<td>Bruhn (2011)</td>
<td>Registration and assistance in registering</td>
</tr>
<tr>
<td>Kaplan et al. (2011)</td>
<td>Registration and assistance in registering</td>
</tr>
<tr>
<td>Piza (2018)</td>
<td>Registration and assistance in registering</td>
</tr>
<tr>
<td>Rocha et al. (2018)</td>
<td>Registration and assistance in registering</td>
</tr>
</tbody>
</table>

Notes: Figure from Ulyssea (2020). Green circles indicate results from experimental studies and grey diamonds the results from the non-experimental literature. We only report ITT estimates from the experimental papers.

Figure 2 shows that providing information about the process of registration and its potential benefits (De Giorgi and Rahman 2013), information and reimbursing all registration costs (De Mel et al. 2013, de Andrade et al. 2014), or creating national formalisation programs that substantially reduce registration costs (Bruhn 2011, Kaplan et al. 2011, Piza 2018, Rocha et al. 2018) have very limited effects on firm formalisation. The exception to the rule is Benhassine et al. (2018), who find a positive and significant effect of providing information and assistance in registering in Benin. However, it seems that this positive result comes from the high-quality staff used in the experiment (who were responsible for providing the information), rather than from the informational content itself. Indeed, in a follow-up intervention, the authors provided the same information without the qualified staff used in the first intervention and found no effects.

Figure 2 shows that the largest formalisation effects come from interventions that reduce the costs of staying in the formal sector or that increase the benefits of formality. The results from De Mel et al. (2013) are particularly illustrative: the authors find no effects from the treatment arm that essentially eliminated registration costs; however, when firms are offered a substantial compensation for formalising (the equivalent of two months’ profits for the median firm), in addition to removing registration costs, there is a substantial formalisation effect of 47%. Rocha et al. (2018) find similar results in the context of a national formalisation policy in Brazil targeted at entrepreneurs with at most one employee. In its first phase, this

¹ We focus on authors’ preferred specifications and standardise the reported point estimates so that all results reflect effects on formalisation rates. For the experimental papers, we only report ITT estimates. See Ulyssea (2020) for more details.
formalisation program eliminated entry costs for eligible entrepreneurs and in the second phase it also substantially reduced the tax burden faced by firms. The first phase had a null effect on formalisation rates, but the second one generated an increase of around 11%. This result is driven by the formalisation of existing informal firms, not by the creation of new formal businesses, nor greater survival among formal incumbents.

The third group of policies – which seek to increase the costs of informality – has received far less attention by policy makers and empirical studies. The first exception is the work of Andrade et al. (2014), who randomly assign municipal inspectors to firms in order to assess whether higher enforcement could induce firms to formalise in the state of Minas Gerais, Brazil. Using an IV approach, the authors show that the impact of receiving an additional inspector visit is quite high, with an effect of 21-27 percentage points in firms’ registration. They find no spillovers to neighbouring firms.

A second important exception is the work of Bujanda and de la Parra (2020). The authors use nearly half a million random work-site inspections by the Ministry of Labour in Mexico to analyse the effects of increasing the cost of the intensive margin of informality, that is, the cost for formal firms of hiring informal workers. The authors find a 7 percentage point increase in the quarterly probability of being formalised by their employer for informal workers in establishments that were inspected (relative to similar workers in non-inspected establishments). They find no effects on wages and quarterly transitions from inspected establishments to unemployment increase from 2.9% to 4.1%. At the firm level, they document a negative and significant effect on total formal employment. In particular, a decline in formal worker poaching rate and an increase in the destruction rate of formal jobs.

**Interpreting the evidence**

The results discussed above show that reducing the formal sector’s entry costs has very limited effects on formalisation. Reducing the ongoing costs of formality (or increasing its benefits) is more effective, but the effects are not large enough to make these policies cost effective, as they typically lead to substantial forgone tax revenues (Ulyssea, 2020). Overall, the largest formalisation effects come from greater enforcement.

A useful way to organise these results is through the lenses of the three leading views about informality in the literature (La Porta and Shleifer 2014). The first view – the “De Soto view” – argues that the informal sector is composed of potentially productive entrepreneurs who are kept out of formality by high regulatory costs, in particular entry regulation. This view dates back to De Soto (1989) and has motivated numerous efforts to reduce fixed registration costs around the world. The second – termed the “Parasite view” – argues that informal firms are “parasite firms”: they could survive in the formal sector but choose to remain informal to earn higher profits by not complying with the relevant taxes and regulations. The third – the “Survival view” – sees informality as a survival strategy for low-skill individuals, who are too unproductive to ever become formal. The policy implications – and their expected results – are clear. According to the first view, reducing entry costs would lead to higher formality, productivity and growth, as it would “unleash” constrained informal entrepreneurs. The second view implies that the government should increase enforcement, which would allow the reallocation of resources from less productive informal firms to more productive formal ones. Finally, the third view implies that reducing entry costs would have very limited effects, while increasing enforcement would not lead to formalisation of informal firms, as they are not able to survive in the formal sector. Moreover, higher enforcement could have negative social consequences by eliminating the livelihoods of the most vulnerable individuals.

As shown in Ulyssea (2018), these views are not competing but complementary frameworks for understanding informality, as they simply reflect the underlying heterogeneity in the informal sector. Thus, the main task becomes determining their relative importance in the data. Ulyssea (2018) proposes a simple taxonomy of informal firms based on these views and uses a structural model to infer their relative importance in the Brazilian context. The results are reproduced in Figure 3.
Figure 3: Taxonomy of informal firms in Brazil

![Graph showing the taxonomy of informal firms in Brazil](image)


Figure 3 shows that the "De Soto view" accounts for a small fraction of informal firms, 9.3%. The "Parasite view" corresponds to 41.9%, while the remaining 48.8% correspond to the "Survival view". To the extent that these results from Brazil are informative for other contexts, Figure 3 provides a rationale for the results found in the empirical literature summarised in Figure 2. It is only a small fraction of informal firms that are constrained by registration costs and therefore reducing these costs does not lead to substantial formalisation. Increasing enforcement can have a substantial impact, since the parasite view encompasses a large fraction of informal firms. However, the survival view accounts for nearly half of all informal firms. Since enforcement policies cannot really distinguish between the two types of firms, they could lead to potentially large adverse effects by displacing a large number of very low productivity informal entrepreneurs and their employees. The extent to which these adverse effects would be observed in equilibrium crucially depends on how much reallocation of labour away from informal firms to more productive formal firms actually happens. Even if the reallocation does occur in the longer run, the transition to the new steady state can be costly to displaced individuals.

4.2 Firms: Taxes and trade

**Taxes**

The tax structure (and not only the tax burden) that firms face is a key determinant of their decisions to formalise, but one that has received much less attention in the literature. An early exception is the work of de Paula and Scheinkman (2010), who analyse the role of value-added tax (VAT) in transmitting informality via its credit scheme. In this type of scheme, establishments receive a credit for the amount of tax paid upstream in the production chain, which is used against future tax liabilities. By definition, purchases from informal suppliers do not generate tax credits and informal firms cannot generate tax credits from their own suppliers, even if those are formal. Thus, formality/informality could be transmitted throughout the production chain, as more (in)formal supply chains produce greater incentives for firms to be (in)formal.
The empirical analysis using micro data on formal and informal firms in Brazil confirms the predictions of their model. Formality of a firm’s suppliers and buyers is correlated with its own formal status, while greater enforcement upstream or downstream implies a higher probability of being formal. In related work, Naritomi (2018) uses a unique administrative data from Brazil to show that introducing incentives similar to the VAT for final sales leads to substantial increase in firms’ reported revenue.

**Trade**

Using different research designs, Paz (2014), Dix-Carneiro and Kovak (2019), Cruces et al. (2018), and Ponczek and Ulyssea (2022) find substantial effects of trade opening (tariff reduction) on informal employment. Moreover, Ponczek and Ulyssea (2022) show that these informality effects are concentrated on low-skill workers in Brazil. These results thus confirm a long-standing concern that trade reforms could lead to a reallocation of firms and workers from the formal to the informal sector due to the greater competitive pressure faced by domestic firms (Goldberg and Pavcnik 2003).

The results in Cisneros-Acevedo (2022) provide an interesting nuance to the results above and illustrate the importance of distinguishing between the two margins of informality. The author analyses the impact of import tariff reduction in Peru to show that greater trade opening had two opposing effects. On the one hand, tariff reduction leads the least productive (informal) firms to exit, which causes *informality on the extensive margin* to fall. On the other hand, the same competitive pressure induces formal firms to cut costs by hiring informal workers, causing an *increase in the intensive margin of informality*. The latter effect dominates, and overall informality increases as a result of tariff reduction.

Importantly, the results in Dix-Carneiro and Kovak (2019) and Ponczek and Ulyssea (2022) indicate that the increase in informality *per se* is not clearly a negative result, as informality can help reduce employment losses after a negative shock. In particular, Ponczek and Ulyssea (2022) show that regions with weaker enforcement observed higher informality effects but no unemployment effects. Not only that, in these regions formal plants had a larger probability of surviving, most likely due to the intensive margin (which is consistent with the results of Cisneros-Acevedo 2022).

The flipside of these more negative results can be found in the work of McCaig and Pavcnik (2018). They show that a positive export shock in Vietnam (from the U.S.-Vietnam Bilateral Trade Agreement) led to a substantial decline of informality due to the reallocation of workers away from informal firms to more productive formal firms.

Despite their importance, these empirical studies are not able to account for general equilibrium effects, which can be quite important in the case of trade opening. Dix-Carneiro et al. (2021) develop an equilibrium trade model with firm dynamics and firm heterogeneity, formal and informal sectors, labour market frictions and a rich institutional setting, which is estimated using several data sources from Brazil. Their results are broadly consistent with those from the empirical literature discussed thus far, whereby the decline in informality within the tradable sector is a consequence of trade liberalisation. However, the authors show that even though informality can be an “unemployment buffer” – as shown by Ponczek and Ulyssea (2022) – it is not a “welfare buffer”, as welfare is actually higher when enforcement is stronger, even if it comes at the cost of greater unemployment.

### 4.3 Workers: The role of public policies

One important concern in developing countries with large informal sectors is the potential effect of welfare policies shifting labour supply from the formal to the informal sector. This is clearly a concern in cash transfer programs that are means-tested, which could incentivise individuals to work informally to remain eligible for the benefit. More broadly, universal programmes that increase the benefits of informality (or reduce its costs) could produce similar effects, as extensively discussed in Levy (2010).
Bosch and Campos-Vazquez (2014) analyse one such programme, the Seguro Popular in Mexico. It was created in 2002 and introduced universal health coverage, including all informal and previously uninsured informal workers. Before that, health coverage was tied to payroll contributions, which represented a large cost of being informally employed. Hence, the programme substantially decreased the costs of informality. The authors show that Seguro Popular had a negative effect on the formality trend (measured by social security contribution) in small and medium firms. In the absence of the programme, around 4.6 and 4% additional employers and employees would have formally registered, respectively. Camacho et al. (2013) analyse a similar programme in Colombia (the Subsidised Health Regime) and find that the programme leads to an increase in informal employment of around 4 percentage points. The reported effects are not very large, which could indicate that the value of these programmes to workers is not very high. Indeed, Conti et al. (2018) develop and estimate a household search model with formal and informal sectors and show that the utility value of Seguro Popular represents 4% and 9% of the mean household income for high and low education households, respectively.

As discussed above, the availability of cash transfer programmes could also be an important determinant of informal labour supply and could discourage work more broadly (Banerjee et al. 2017). Evidence from the Bolsa Familia in Brazil (De Brauw et al. 2015), Plan de Atención Nacional a la Emergencia Social in Uruguay (Bergolo and Cruces 2021), and the Universal Child Allowance in Argentina (Garganta and Gasparini 2015) suggest that both of these effects – higher informality and non-employment – are present in the data. For example, Bergolo and Cruces (2021) find a reduction of formal employment in eligible households of around 8 percentage points, which are equally distributed between informal employment and non-employment.

These results refer to the direct effects of the transfer on individuals’ labour supply decisions. However, some of these programs are quite large in scale and could have sizable general equilibrium effects in local economies. Indeed, Gerard et al. (2021) show in recent work that Bolsa Familia (PBF) led to an increase in local formal employment in Brazil. They exploit a large increase in the number of PBF beneficiaries in 2009, as well as a change in the methodology used to allocate slots across municipalities. Their main result indicates that municipalities more positively affected by the increase in total PBF payments observed an increase in the number of formal jobs of up to 2% by 2011 (two years after the reform). The effects are concentrated in low-skill, private sector jobs with no effects on public sector jobs. Their results suggest that larger PBF benefits had an important local multiplier effect, as they also find a similar increase in formal employment for workers who were never part of the program. Importantly, they show that these positive local aggregate effects are observed despite the negative direct effect on beneficiaries’ formal labour supply. Their results at the individual level – contrasting beneficiaries and non-beneficiaries in a regression discontinuity design – are consistent with the evidence mentioned above, as beneficiaries seem to reduce their formal labour supply to retain eligibility.

The discussion thus far has focused on the static relationship between public policies and individuals’ informality decisions. However, there can be important dynamic implications as well, especially when one considers the effects of social security systems that have some form of non-contributory safety net that provides a minimum benefit to the elderly. These non-contributory benefits can represent a tax on individuals who contribute, as they are typically decreasing in one's contributions, and could hence discourage formal work, especially for low skill workers. Joubert (2015) investigates the importance of these forces in a life-cycle, discrete choice model that captures household’s labour supply choice between formal and informal employment, and saving decisions under the rules of Chile’s pension system. The results from counterfactual analysis using the estimated model show increasing the mandatory contribution rate by 5 percentage points increases informality by 12.5% and 9.3% for men and women, respectively.

2 Azuara and Marinescu (2013) also study the Seguro Popular programme and find a very small effect of SP on informality among unskilled workers (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 and Campos-Vazquez (2014) use administrative data, which substantially reduces measurement error.
4.4 Workers: The role of human capital

As discussed in Section 3, a well-established fact about informality is that it is (sometimes sharply) decreasing with individuals’ schooling levels. This fact could thus suggest that changes in the composition of the labour force toward a higher share of more-educated workers could be an important force to reduce informality rates. This seems consistent with broad trends observed in many Latin American countries, which have recently shown sustained reductions in informality among employees, without experiencing major changes in labour regulations, minimum wages, or enforcement.

A recent study by Haanwinkel and Soares (2021) investigates to what extent the changes in composition of the labour force could explain the observed strong reduction in informality levels observed in Brazil in the 2000s. For this, the authors develop a model with two types of workers – skilled and unskilled – and a large number of firms, which differ in productivity but are not intrinsically connected to any sector. Decisions related to formality status are the result of a labour market equilibrium where each agent is choosing optimally. The model incorporates many features of Brazilian labour regulation: payroll taxes, mandated benefits, and the minimum wage. Finally, it also includes an informality penalty that increases with firm size to account for the risk of being caught by labour inspectors and for the eventual punishment. The model is able to reproduce several patterns in the Brazilian labour market, especially those related to labour informality, even though it does not impose structural differences across sectors.

Their key result is that increased schooling is the most important factor in explaining the decline in informality observed in Brazil between 2003 and 2012. If the schooling composition of the labour force had remained the same as in 2003, but all other factors had changed according to what was observed during the period, there would have been an increase in the informality rate instead of a large decrease. In the model, increased schooling alone is able to generate a large decline in informality rates, suggesting that education is a key determinant of formalisation. Two main channels explain this result. First, an expansion in schooling levels leads to higher low-skill wages due to scarcity and increased productivity of these workers, which results in higher wages in the informal sector. Second, it stimulates increases in firm size. As a result of the increased incentive to grow, formal firms hire more workers and, simultaneously, firms operating at the margin of informality find it profitable to move into the formal sector (since it is difficult to hide in the informal sector if a firm becomes too large).

5. Consequences of informality

5.1 The consequences for firms

The papers that seek to identify the effect of formality on firms’ performance typically estimate a regression of the outcome of choice (profits, for example) onto a dummy for whether the firm is formal or not, and additional covariates that capture firms’ (and owners’) observable characteristics. The identification problem then comes from the self-section of firms into formality, which is expected to be positive – i.e. better firms/entrepreneurs self-select into being formal. This decision is likely to be influenced by elements that are unobservable to the econometrician, such as firm-level demand or productivity shocks and unobserved, time-varying entrepreneurial quality. To overcome this identification problem, the literature has largely resorted to experimental or quasi-experimental variation in access to policies and interventions that change the costs and benefits of formalisation, such as those discussed in Section 4.1, to construct instruments for the endogenous regressor of interest, i.e. the formalisation dummy.

The results in the literature generally indicate that formalisation has no statistically significant effects on different measures of firm performance, such as sales, profits and number of employees (e.g. Rocha et al. 2018, Benhassine et al. 2018). Even when the study does find a positive average effect of formalisation
on profits, as in De Mel et al. (2013), this seems to be driven by few firms experiencing substantial growth. This lack of effect is consistent with the argument that the perceived benefits of formalisation are very low for most small-scale entrepreneurs (Bruhn and McKenzie 2014). It might be the case that the positive effects of formality take a long time to appear, while most of these studies follow these firms for up to three years. Even if this is the case, these are not very encouraging results, as the costs of formality kick in immediately upon formalisation (such as tax payments), while the benefits would take much longer, if at all.

Perhaps more revealing, even when firms do formalise as a result of the incentives provided, they do not seem to change any meaningful behaviour. De Mel et al. (2013) show that formalisation does not increase tax payments, the likelihood of holding a business bank account nor of applying for a business or personal loan. The only dimension of such intermediary outcomes affected by formalisation is the probability of keeping a receipt book. These results therefore challenge the idea that formalisation per se could have an important causal effect on firms’ performance.

5.2 Consequences for taxation and redistribution

The existence of large informal sectors in low- and middle-income countries could potentially change the redistributive properties of taxation. This is particularly important for consumption taxes, on which these countries rely to raise a large share of their revenues. Consumption purchased from informal retailers is, by definition, untaxed. If the budget share that households spend in the informal sector varies systematically with income, the presence of informal sectors will change the incidence of consumption taxes. Informal consumption patterns could therefore also affect the desirability of different consumption tax policies. In particular, if much of consumption (especially of poorer households) is outside the formal sector, then this could, for example, substantially reduce the motivation for taxing necessities (food products in particular) at a reduced rate compared to other products.

Bachas et al. (2021) investigate the patterns of informal consumption and their implications for tax policy in 32 developing countries of varying levels of economic development (from Burundi to Chile). Informal sector purchases are by definition hard to observe and link to consumers’ incomes. To overcome this challenge, they use the places of purchase reported by households in expenditure surveys to proxy for the share of consumption in the informal sector. Building on evidence from retail censuses and existing literature, they assign each place of purchase to the informal or formal sector, based on the idea that large modern retailers are much more likely to remit taxes than smaller traditional ones (Lagakos 2016, Kleven et al. 2016). They assume, for example, that consumption from home production, markets and street stalls is informal, whilst consumption from supermarkets is formal.

The key descriptive result of their paper is the existence of a downward-sloping Informality Engel Curve (IEC): they find that, in all countries, the informal budget share declines steeply with household income. Figure 4 shows this in Rwanda and Mexico as an example. In Rwanda, the informal budget share falls from 90% for the poorest decile of households to 70% for the richest decile. In Mexico, it falls from 55% to 25%.

The shape of the informality Engel curves implies that informal consumption patterns make consumption taxes progressive. Simply setting a uniform tax rate on all goods would lead to the richest quintile paying twice as much in taxes as the poorest quintile in the average (Figure 5, red line). As a comparison, Figure 5 plots the taxed budget share obtained under the ‘naïve’ assumption that governments can tax all expenditures but choose to exempt food from taxation (green line). Comparing the green and red lines, one can see that the de facto exemption of the informal sector from taxation is clearly more progressive than this naïve scenario. Results country-by-country in the paper show that this ‘progressivity dividend’ from exempting the informal sector is largest in the poorest countries. The small size of the formal sector in these countries, together with the downward-sloping informality Engel curves, makes formal transactions a particularly good ‘tag’ for household income. Finally, once the informal sector is taken into account, we
see that exempting food items from taxation only slightly increases progressivity (orange line). This is because the formal food Engel curve actually has a small but positive slope in the poorest countries.

**Figure 4:** Informality Engel curves in Rwanda and Mexico

![Informality Engel curves](image)

**Notes:** Figures reproduced from Bachas et al (2021). These panels show the local polynomial fit of the informality Engel curves in Rwanda (Panel A) and Mexico (Panel B). Per person total expenditure on the horizontal axis is measured in log. Informal budget share is on the vertical axis. The shaded area around the polynomial fit corresponds to the 95% confidence interval. The solid grey line corresponds to the median of each country’s expenditure distribution, while the dotted lines correspond to the 5th and 95th percentiles.

The authors then use a simple model to study optimal consumption tax policy in the presence of an informal sector, and calibrate it using their data. The existence of the informal sector affects both the equity characteristics of consumption taxes (via the shape of the Informality Engel Curves) and their efficiency: informality increases the efficiency cost of taxing consumption because households can switch to informal varieties of products when taxes on formal ones increase. They find that, in some of the poorest countries, subsidising food relative to non-food is simply not optimal once the informal sector is taken into account. Since, in these countries, poor households consume most of their food from the informal sector, the subsidy ends up either redistributing very little or benefiting mostly richer households.

Overall, the evidence indicates that the presence of large informal sectors in developing countries make consumption taxes progressive. These results caution that any benefits from reducing the informal sector’s size should be weighed against potential equity costs. These findings more generally suggest that informality may affect the distributional consequences of tax policy in developing countries in subtle ways, an important avenue for future research.
5.3 Aggregate consequences

Informality is an endogenous outcome as much as the unemployment rate or wages observed in a given economy. As such, the aggregate effects of having lower or higher levels of informality are ultimately determined by the means used to achieve a lower level of informality. Both the (more extensive) macro literature and the recent structural literature have approached this question using calibrated/estimated models and relying on counterfactual exercises relative to specific policy experiments. These can be largely classified as those that increase the costs of informality or reduce the costs (or increase the benefits) of formality. In what follows we discuss the main results of the literature by different types of aggregate outcomes in counterfactual exercises that emphasise policy experiments in both of these categories.

5.3.1 Human capital

Two recent papers (Bobba et al. 2021, Bobba et al. 2022) investigate the negative relationship between informality and the stock of human capital in the economy, considering both investments prior to labour market entry and on-the-job human capital accumulation. The broad idea behind both papers is that informality can work as a “tax” on human capital in developing countries. The authors consider a class of equilibrium search and matching models of the labour market aimed at capturing two empirical regularities that are hard to explain in models where the market is competitive or where there is segmentation between the formal and the informal sector: (i) individual workers transition frequently between formal and informal jobs; and (ii) conditional on workers’ schooling level, formal and informal earnings distributions overlap.

The framework in Bobba et al. (2022) allows for workers to decide on their schooling level prior to entering the labour market. The education decision depends on the present discounted value of participating in a schooling-specific labour market with returns resulting from all the factors discussed above. A key finding is that informality depresses the returns to schooling. Since workers anticipate this, they will invest less in their education, and as a consequence the proportion of workers acquiring a given level of schooling will
fall. The mechanism is as follows: the institutions that generate informality tax high-productivity formal jobs and increase the relative profitability of informal self-employment and of low-productivity informal salaried jobs. These taxes and subsidies differ across schooling levels, hurting relatively more those workers with more schooling. Precisely because those workers are more productive, it is harder for firms to offer them informal jobs (since expected penalties are higher). In addition, because some benefits are pooled, when workers with more schooling are formally employed, they subsidise those with less schooling.

The parameters of the model are estimated using data from Mexico, a country where more than half of the labour force is informally employed. Two sets of counterfactual experiments based on the estimated model quantify the effects of informality on schooling. First, eliminating informal jobs enhances labour market returns to schooling and increases schooling investments by 10% but at the cost of decreasing welfare for both workers (5%) and firms (28%). This trade-off originates from the fact that taxes and subsidies operate through the labour market and are associated with the formality status of the job. Second, the proportion of individuals who acquire the higher schooling level decreases monotonically with the payroll tax rate – as shown in panel a of Figure 6. Interestingly, these changes in schooling are paired with an almost constant informality rate (panel b). The phenomenon is explained by major composition effects resulting from the progressive features of the contributory social security benefits: as the tax rate increases, a proportionally larger benefit is available to lower earnings individuals.

Figure 6: The effect of changes in the payroll tax rate on schooling and informality

Note: Outcomes from policy experiments that change the social security contribution rate from 0 to 0.66. The baseline contribution rate is represented by the vertical lines in the figure. Informality is the proportion of informal employees and self-employed in the population. Schooling is the proportion of individuals who complete secondary education by ability type. Source: Bobba, Flabbi and Levy (2022).

Bobba et al. (2021) consider a similar framework where workers are homogenous before entering the labour market and are allowed to accumulate human capital while on-the-job. The human capital evolution while participating in the labour market captures the additional productivity that may be acquired on the job. The authors allow this dynamic to depend on the formality status of a job. While off-the-job and searching (either as unemployed or as self-employed), individuals may even lose previously accumulated knowledge, leading to a depreciation of human capital. Estimation results show that human capital accumulation on the job occurs more rapidly when workers are formally employed. For first entrants in the labour market, it takes on average 1.4 years to start upgrading their human capital if they work formally but it takes 40% longer to do so if they work informally. Human capital upgrading is harder the higher the stock of skills already acquired on the job but, at any human capital level, the probability of upgrading remains higher if working formally. Policy experiments reveal that on-the-job human capital accumulation magnifies the negative impact on productivity of labour market institutions that give rise to informality. For example, the increase in Seguro Popular benefits (see Section 4.3) over the course of 10 years is
associated with a drop in aggregate human capital of about 5 percentage points.

5.3.2 Aggregate productivity, output and growth

Higher enforcement

A common feature of most informality models is the presence of a cost of informality that is increasing in firm size, which is typically measured as number of employees, capital or revenues (e.g. Fortin et al. 1997, De Paula and Scheinkman 2011, Ordonez 2014, Meghir et al. 2015, Ulyssea 2018). Thus, a common counterfactual experiment found in the literature is to simulate higher enforcement on informal firms by making this cost function steeper. The basic intuition is that higher enforcement – via intensified inspections, say – increases the cost of operating as an informal firm (due to a higher probability of detection, for example), which would lead to a substantial reduction of informality.

The results from different studies indicate that reducing the size of the informal sector by increasing enforcement could lead to substantial gains in aggregate productivity. As summarised in Ulyssea (2020), different mechanisms contribute to generating these positive effects. First, there are positive composition effects, as greater enforcement eliminates many low-productivity (informal) firms, which then frees up resources that are reallocated to more productive formal firms (e.g. Ulyssea 2010, Charlot et al. 2015, Bosch and Esteban-Pretel 2012, Meghir et al. 2015, Ulyssea 2018). Second, reducing the availability of low-quality informal jobs can make it easier for workers to find higher quality formal jobs, especially if there are substantial labour market frictions, (e.g. Meghir et al. 2015). Third, because informal firms face higher financial frictions and are more credit constrained, formalisation can lead to greater capital accumulation (e.g. D’Erasmo and Boedo 2012, Ordonez 2014). Fourth, it affects occupational choices by discouraging low-skill individuals to self-select into informal entrepreneurship, therefore increasing labour supply in the formal sector (e.g. Ordonez 2014, Lopez 2017). Fifth, as discussed in the previous section, there can be higher investments in formal schooling (before entering the labour market) and on-the-job human capital accumulation (Bobba et al. 2021, 2022).

Despite these positive aggregate effects, higher enforcement can have adverse effects on those directly affected and potentially on aggregate outcomes as well. In an earlier paper, Boeri and Garibaldi (2005) argued that large informal sectors are “tolerated” by governments because increasing enforcement could lead to substantial increases in unemployment. The results in Ulyssea (2010) and Charlot et al. (2015) are consistent with this conjecture. Using equilibrium matching models calibrated to the Brazilian economy, they show that greater enforcement substantially reduces informality, but at the cost of increasing unemployment. More recently, Meghir et al. (2015) and Haanwinckel and Soares (2021) find no unemployment effects from higher enforcement. Dix-Carneiro et al. (2021) show non-monotonic effects: stricter enforcement barely changes unemployment, but completely eradicating the informal sector causes the unemployment rate to increase substantially.

One of the key dimensions to determine the extent of the positive effects on productivity and the negative effects on unemployment is how much employment reallocation there can be from low-productivity informal firms to more productive formal firms. A second important, and completely overlooked, question is how lengthy the transition between steady states is, and therefore how long this reallocation process can take. This is key for both the political economy of implementing these measures, but also to determine the welfare costs in the short- and medium-run for those negatively affected by these policies.

Even though most of the literature focuses on enforcement policies on the extensive margin of informality (i.e. increasing the costs of informal firms), Ulyssea (2018) shows that increasing enforcement on the intensive margin can generate very different results. In particular, even though it reduces informality amongst workers, it can in fact increase the share of informal firms. This occurs because higher enforcement in the intensive margin effectively increases the costs of operating in the formal sector for small formal firms. Hence, many of these firms choose to enter the informal rather than the formal sector in the new equilibrium. As a consequence, this policy generates losses for low-productivity formal firms,
while high-productivity firms benefit from it in terms of higher life-time profits. The effects on aggregate productivity are small – an increase of 1.7 percent – and output decreases by 1.6 percent, as the reduction in the number of firms operating in the economy more than compensates the small gains in aggregate productivity.

Finally, Almeida and Carneiro (2009) use micro data to estimate average aggregate effects of enforcement across municipalities in Brazil. For that, they exploit the fact that enforcement of labour regulation (the intensive margin of informality) is implemented in a decentralised way and displays a lot of variation across local economies. They use an IV estimator to show that an increase in the number of inspections per hundred formal firms leads to small reductions in output and firm size. In a follow-up paper Almeida and Carneiro (2012) show that more inspections also lead to small negative effects on the share of formal workers and self-employed and an increase in non-employment.

The results by Ponczek and Ulyssea (2022) mentioned in previous sections directly speak to this discussion as well. They examine the effects of local economic shocks generated by the unilateral trade liberalisation in Brazil, and how they varied across regions with weaker and stronger enforcement levels. This unilateral trade opening essentially represented a negative demand shock to the affected industries in Brazil, which generated heterogeneous effects across regions where employment was more and less concentrated in these industries. As mentioned in Section 4.2, the authors show that regions with stricter enforcement experienced lower informality effects, but greater losses in overall employment and greater reductions in the number of formal plants. Regions with weaker enforcement had opposite effects and all the effects are concentrated among low-skill workers. Thus, similarly to Bujanda and de la Parra (2020), these results indicate that greater enforcement can lead to greater formalisation but with potentially adverse employment effects. Conversely, the greater flexibility introduced by informality might allow formal firms and low-skill workers to cope better with adverse labour market shocks (Ponczek and Ulyssea, 2022).

**Reducing the costs of formality**

As discussed in Section 4.1, the literature that uses experimental and non-experimental empirical designs to estimate the effects of reducing the costs of formality on firms’ decisions to formalise shows essentially zero or very limited effects. Despite the importance of these results, it can still be the case that these policies could have important aggregate effects. Indeed, one of the regularities that emerges from counterfactual exercises in different papers is that reducing fixed entry costs into the formal sector can produce positive and sizeable aggregate effects (Ulyssea 2010, D’Erasmo and Boedo 2012, Charlot et al. 2015, Ulyssea 2018). For example, Ulyssea (2018) shows that reducing entry costs into the formal sector leads to higher competition, aggregate production in the formal sector and high-skill wages. However, because it is mostly low-productivity firms that formalise, there is a negative composition effect that leads to a decrease in aggregate productivity. Total output still increases because there is a substantial increase in the number of firms in the economy led by an increase in the number of formal firms (Ulyssea (2010) and Charlot et al. (2015) find similar positive aggregate effects, including lower unemployment).

If, however, there are important frictions in the formal sector as well – such as financial frictions – then these positive effects can be limited. That is the case in Lopez-Martín (2018), who finds limited aggregate effects from reducing entry costs in Mexico and Egypt, of at most 0.5 and 0.7 percentage points in aggregate productivity and output per capita, respectively. It is only when financial frictions in the formal sector (modeled as the ability of firms to collateralize their assets) are relaxed that the economies observe substantial gains in aggregate productivity, output and welfare (D’Erasmo (2016) finds similar results).

As also discussed in Section 4.1, the empirical evidence suggests that reducing the tax burden can induce some formalisation of informal firms, but even in this case the effects are not very large. The counterfactual results from macro and structural models corroborate that: reductions in payroll tax seem to generate some positive but limited formalisation effects (e.g. D’Erasmo and Boedo 2012, Haanwinckel and Soares 2021). Ulyssea (2018) shows that these effects are stronger on labour informality (via the intensive margin) and weaker on firm informality. The overall effects on aggregate productivity and output
are also positive but quite limited (Ulyssea 2010, D’Erasmo and Boedo 2012, Haanwinckel and Soares 2021, Ulyssea 2018).

6. Final remarks

This VoxDevLit review attempts to provide a comprehensive coverage of the extensive economic literature that analyses the causes and consequences of informality in developing countries. As discussed in Ulyssea (2020), this literature is not only extensive but also quite broad in terms of methodologies used, ranging from well-identified empirical studies (both experimental and non-experimental) to equilibrium macro models and structural empirical models. This seems not only natural but necessary for a more complete understanding of this important economic phenomenon.

We believe that this review piece shows that the literature has made substantial progress in understanding the main determinants of firms’ choices regarding informality, both theoretically and empirically. Indeed, the literature has put a lot more emphasis on the analysis of firms’ than workers’ behaviour. However, even on firms’ side there remain many important unanswered questions. In particular, the literature has only started to explore more the dynamics of firms’ decisions regarding the different margins of informality and how they potentially interact with different frictions that firms face. For example, does informality work as a stepping-stone for entrepreneurs with high-growth potential but who might be constrained by, say, credit constraints?

On the workers’ side, the gaps in knowledge are arguably wider. We need a deeper understanding of the determinants of workers’ choice/allocation between formal and informal jobs, what determines their permanence and evolution in either, as well as the main tradeoffs they face. Regarding the tradeoffs, for example, we do not know how much workers value the greater job security provided by formal employment relative to informal jobs. As in the case of firms, we do not know how much informal jobs represent a stepping-stone for younger workers versus the extent to which there is an “informality trap” that makes future transitions into formal employment very unlikely. Related to this point, there are still very few studies that investigate the life-cycle dimensions of informality.

Finally, on the analysis of the aggregate implications of informality, a very important unexplored dimension is the transition dynamics between equilibria. This is arguably very important to understand both the political economy of informality and formalisation policies, but also to provide a more accurate assessment of the welfare implications of different formalisation policies.

References


Bergolo, M and G Cruces (2021), "The anatomy of behavioral responses to social assistance when informal employment is high", Journal of Public Economics, 193: 104313


Bujanda, L F and B S de la Parra (2020), "Increasing the Cost of Informal Workers: Evidence from Mexico", Mimeo


Informality


