Informal employment or informal firms? Regulatory enforcement and the transformation of the informal sector

Sanjay Jain*

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: While there is general agreement that regulatory avoidance is an important part of firms' decisions to produce in the informal sector, there is much less agreement on how regulation and enforcement affect firms' decisions on, inter alia, which sector they locate in, their employment decisions, and whether to transition from one sector to another. In this paper, we focus on this set of questions: how does the regulatory regime affect these sectoral location decisions by firms? In particular, how are these decisions affected in environments where there are regulatory spillovers, so that each firm's decision, on whether to comply with applicable regulations, also carries implications for other firms? We construct a theoretical model that incorporates firms' decisions on their mode of production, encompassing not just the sector and level of production, but also the level of employment, and consider how these might be affected by varying degrees of regulatory spillovers in their operating environment. The main contribution of this research is to provide a clearer understanding of the interplay of regulation and its enforcement on the one hand, and firm decision-making about employment and output on the other, in a modelling environment where these issues are not dealt with in separate black boxes.

Key words: informal sector, regulatory spillovers, informal employment

JEL classification: O17, D04, J46, L51

* Department of Economics, University of Oxford, United Kingdom; sanjay.jain@economics.ox.ac.uk

1   Introduction

Why do firms choose to produce informally? While there is general agreement that regulatory avoidance—usually, having to do with minimum wage or health and safety laws or tax liabilities—is an important part of the picture, there is much less agreement on the specifics.[1] How do regulation and enforcement affect firms' decisions on, inter alia, which sector they locate in, their employment decisions, and whether to transition from one sector to another? In this paper, we focus on this set of questions: how does the regulatory regime affect these sectoral location decisions by firms? In particular, how are these decisions affected in environments where there are *regulatory spillovers*, so that each firm's decision, on whether to comply with applicable regulations, also carries implications for other firms?

Our starting point is the conceptual recognition that firms choose their *mode* of production, taking into account the existing regulations that they face and the intensity of regulatory enforcement. Further, this decision may be more complex than simply choosing whether to be formal or informal and can entail the firm choosing along two different margins.

The previous literature has largely focused on what might be described as 'the extensive margin', in which the firm chooses its *level* of activity in order to be compliant, or alternatively, to take itself out of the ambit of the relevant regulation.[2] More recently, however, there has been increasing recognition that even formal firms can choose along an 'intensive margin', by hiring workers 'off the books'.[3] In making these decisions, firms must take into account their legal and regulatory environment, and the *differential enforcement* of regulations relating to employment and labour conditions, versus those relating to the quality and quantity of output. Further, these decisions also depend on the extent of the externalities in the enforcement of regulations.[4] This spillover can be 'positive' or 'negative', in the following sense: if an increase in the number (or proportion) of firms that choose to locate in the informal sector also leads to an increase in the likelihood that a firm will be 'inspected' and apprehended if it is in violation of the applicable regulations, then there is a positive regulatory overspill across firms. Alternatively, the regulatory spillover may be negative if there is 'safety in numbers'—i.e. if the decision by firm A to be non-compliant makes it less likely that firm B will be apprehended and penalized.

If regulations on size (of output, or employment) are an important constraint on evasion of regulations, then, as Kanbur (2009) recognizes, firms might choose to 'relocate' away from being subject to the applicable regulation towards becoming a firm to which the pertinent regulation does not apply, by adjusting their activity to become of a smaller size. Banerji and Jain (2007) suggest that this may be due to the different factor-price ratios prevailing in the two sectors—hence, firms, in choosing to locate in either the formal or informal sector, are in effect choosing which set of factor prices they will incorporate in their profit-maximizing decisions. However, these papers do not consider regulatory spillovers in the firms' decision-making.

We construct a theoretical model that incorporates firms' choice of economic activity, encompassing not just the sector and level of production, but also the level of employment, and use the model to consider environments with varying degrees of regulatory overspill. This allows

---

[1] For an overview, see the surveys by Ulyssea (2020) and Gërxhani (2004).

[2] See Kanbur (2009) for a discussion.

[3] See Ulyssea (2018) for evidence from Brazil, and Ulyssea (2020) for a recent survey.

[4] See Heyes and Kapur (2009) for a discussion.

us to also shed light on some important policy questions relating to the informal sector and to dualistic economies more generally: under what circumstances should we expect to find geographical or spatial clustering of informal activity? Again, spillovers are important—is there in fact 'safety in numbers'? What implications does this have for the structure of the labour market? What are the implications of regulatory spillovers on the size distribution of firms?

The rest of the paper is organized as follows. In the next section, we briefly discuss the three main strands of the related literature on informality, on which the model in this paper is built. In Section 3, we construct a simple framework to model the firm's sector-choice decision, and the effect of different regulatory regimes, together with the spillovers generated across firms that in turn affect the firm's sector-choice decision. Section 3.1 lays out the baseline model, and Section 3.2 describes the spillovers across firms created by the enforcement mandate given to the regulator(s). Section 4 presents a discussion of the model in the context of related themes in the literature on informality, and also of possible extensions to other areas not considered in detail here, such as the operation of labour markets and the inter-sectoral mobility of different kinds of labour. Section 5 concludes.

2   Background and related literature

While there is a large literature on the 'informal sector', there is little consensus on even the most basic questions—for example, what is the informal sector exactly? Why do firms choose to produce informally? Perhaps the best that can be said is that there is loose agreement that regulatory avoidance is an important part of the picture.[5] However, the specifics of the interplay between different regulatory regimes on the one hand, and the extent to which they differentially affect firms in different sectors on the other, are much less well developed.

This paper builds on three main strands of recent research. We briefly discuss each of these in turn below.

2.1   Adjusting output and employment: extensive versus intensive margins

As discussed briefly earlier, there has been a recent recognition that firms can adjust their status and, by extension, the degree to which they fall within the ambit of regulatory oversight, in different ways. While it has long been recognized that firms could choose to operate 'in the informal sector' or 'in the formal sector' (Rauch 1991; Kanbur 2017), there is increasing recognition that firms can adjust output by exploiting '…two margins of informality: (i) not register their business, the extensive margin; and (ii) hire workers "off the books", the intensive margin' (Ulyssea 2018: 2015).

The firm's ability to adjust these two margins is at the centre of the analysis in Ulyssea (2018), which also assumes that there is heterogeneity in firms' productivity—*à la* Melitz (2003) and Yeaple (2005)—and among workers, who can be of high or low skill. The firm's decision on whether to register with the authorities and pay the fixed cost of registration determines whether it is formal or informal, and this sector-membership decision constitutes the extensive margin on which it can operate. Alternatively, even a formal firm can hire (some) labour informally, and underreport output, thereby reducing its variable costs—this is acting at the intensive margin. In equilibrium,

---

[5] See, for example, Basu et al. (2010) on minimum wage legislation and noncompliance; Fajnzylber et al. (2011) on the role of taxes and tax regime simplification; de Mel et al. (2013) on administrative costs related to entry into the formal sector; and de Andrade et al. (2016) on the impact of greater auditing by governments.

more productive firms (in expected terms) self-select into the formal sector, while less productive firms choose to operate informally.

With this basic framework in place, Ulyssea (2018) shows that the model can be used to tease out the distinctions between three contrasting views of the informal sector. The first view (which might be attributed to de Soto (1989) or dubbed the *Doing Business* view[6]) suggests that informality is a consequence of the costs imposed by high regulatory burdens—for example, associated with entry into the formal sector. A second ('parasitic') view refers to those informal firms that might be productive enough to survive in the formal sector but find it more profitable to stay informal and save on the costs of compliance and taxes.[7] Finally, the third ('survivalist') view suggests that some low skill individuals/firms are too unproductive to compete as formal enterprises, and hence informality is a survival strategy for them. Using Brazilian data, Ulyssea finds support for all three views—at the bottom end of the skill distribution are small microenterprises ('survivalists'), followed by the 'parasite firms', and it is only further up the firm-productivity range that de Soto's view holds. Of course, if the firm's productivity is sufficiently high, it ends up in the formal sector.

The key contribution of Ulyssea (2018) and related papers (e.g. Dix-Carneiro et al. 2021), is to recognize two aspects that have perhaps not received adequate attention in the literature on informality. One is the idea, now widely accepted in the trade literature, of firm heterogeneity in productivity. The second, more closely related to this paper, is the recognition that firms can exploit two different margins—intensive and extensive—to adjust labour employment and output, rather than just the 'traditionally modelled' extensive margin, by which the firm decides whether to operate in the formal sector or informally. However, this particular strand of the literature does not consider the regulatory spillovers created by different enforcement regimes, which is the main focus of this paper.

2.2   Interactions between (firms in) the formal and informal sector

Mishra (2021) models the links between formal and informal sector firms and considers the implications of two contrasting cases. In one case, formal and informal firms provide products that are vertically differentiated in quality—in this case, the products are 'final' products, in the sense that they are purchased directly by consumers.[8] In the other case, informal sector firms act as suppliers of intermediate goods to formal sector firms—in this case, the output of informal sector firms is not provided directly to consumers, but instead to other firms who produce the final output (and are located in the formal sector).

The common interest—in Mishra (2021) as in this paper—is the recognition of the heterogeneity in the composition of the firms that constitute the two sectors. However, the focus in Mishra (2021) is on formally modelling the interactions between firms in the two sectors, both in production and in the provision of products. Further, Mishra (2021) applies the model to some data from India, in order to relate the predictions from his model to the pattern of income distribution in various India states. By contrast, the focus in this paper is on the spillover across firms caused by the regulatory regime. From a modelling point of view, this is a subtle but important distinction—by definition, these spillovers are unaccounted for in the decision-making of the individual firm, so that there is no strategic interaction per se among firms.

---

[6] See World Bank (2019).

[7] See, for example, Farrell (2004) and Levy (2008).

[8] Banerji and Jain (2007) offer a model of 'quality dualism', in the same spirit.

2.3   The 'missing middle' and the 'quality gap'

It is a commonly noted stylized fact that informal sector activity differs from that of firms in the formal in the manner of production—in particular, in the levels of labour employed by firms in the respective sectors. Rauch (1991) refers to this as the 'missing middle' in the size distribution of firms, with the informal sector being populated by small firms, and the formal sector being populated by large firms.[9] As Kanbur (2017) points out, in many countries, this is the consequence of 'size-dependent regulation'.[10] However, Rauch (1991) presents a theoretical model to argue that this pattern in the size distribution of firms (as measured by employment) might arise endogenously, rather than by administrative classification.

Banerji and Jain (2007) extend this idea to argue that there is often a systematic pattern in the distribution of the quality of output produced by firms in the informal and formal sectors. They model the provision of outputs of heterogenous quality by firms, and argue that a 'quality gap', similar to the missing middle in the size distribution, can be identified in the goods and services produced in the respective sectors. Specifically, firms located in the informal sector are associated with the production of low-quality output, while those in the formal sector tend to produce output at the higher end of the quality spectrum. Banerji and Jain (2007) suggest that this may be due to two different features of the 'quality production function'—one, that firms in the two sectors face different factor-price ratios, and two, that the quality 'isoquals' are non-homothetic.[11] Firms can choose to locate in either the formal or informal sector, thereby effectively choosing which set of factor prices they will incorporate in their profit-maximizing decisions, and this sector-choice decision depends on their quality choice decision. Hence, the sector locational choice is made jointly with the quality choice decision. By linking this sectoral-choice decision to the underlying technology for producing different types of quality, their model suggests that there may be a 'quality gap' between the outputs produced by firms which choose to locate in the formal and informal sectors. However, they do not consider the implications of regulatory spillovers, and how those considerations might affect firms' decision-making regarding the sector in which to locate.[12]

In the next section, we draw on themes from these different strands of the literature to construct a model that explicitly considers the effect of regulation on firms' decision-making on their formal/informal status and on their level of labour employment, in environments where the effect of regulatory overspill varies.

---

[9] See also Sleuwaegen and Goedhuys (2002) and Gall (2010).

[10] For example, Kanbur (2017) cites the specific case of the Indian Factories Act of 1948, which requires the registration of firms in manufacturing if the enterprise employs ten or more workers and uses electricity. Similarly, French labour regulations increase labour costs discontinuously once firm employment reaches 50 workers (Garicano et al. 2016).

[11] Analogous to isoquants, isoquals are defined as the combinations of factors (e.g. capital and labour) that can be used to produce a given level of *quality* of one unit of output. See also Copeland and Kotwal (1996, 1997).

[12] In contrast to Banerji and Jain (2007), where heterogeneity in tastes among consumers was the driver for different qualities being produced, in the model presented here consumers are assumed to be identical in their preferences, and the differences in firm outputs come solely from differences in the extent to which they are affected by regulatory pressures.

3   A simple model of regulatory spillovers and the informal sector

In this section, we construct a simple model of the (in)formality decision of the firm and consider the impact of different regulatory mandates on the spillover effects across firms, which in turn affects the (in)formality decision of the firm.

We take as a starting point that firms can choose to be formal or informal. Firms' choice of sector depends on the factor prices prevailing in those sectors (as in Banerji and Jain 2007), but in addition, the firm's choice of sector to operate in also implies a choice of regulatory regime to live under. More specifically, for informal firms, there are additional costs of evasion and possible penalties if detected/inspected. We focus on the latter aspect.

3.1   The firm's sector-choice decision

We start with the canonical firm profit-maximization problem: the firm chooses inputs $\{K, L\}$ to maximize profits, but with a potential additional cost if the firm were to choose to operate informally. The firm's decision on whether to operate informally, or to instead join the formal sector, will also depend on the vector of factor prices $\{w, \rho\}$ that it would face in each sector. (For ease of notation, in what follows, we dispense with different subscripts for the different sectors).

Hence, we can write the (informal) firm's profit-maximization problem as:

Choose K, L to

$$max\ \pi(q)\ =\ p.q(K, L) - (wL + \rho K) - \gamma(q) * f(K)$$

where the notation follows the usual conventions, except that $\gamma(q)$ is the probability of being 'caught' by the regulator, and $f(K)$ is the penalty or 'fine' (or harassment or bribe) if caught. We assume that:

$$\gamma'(q)\ >\ 0, \gamma''(q)\ >\ 0$$

and $f'(K)\ >\ 0$

In other words, both the probability of being caught and the fine to be paid if caught are assumed to be increasing in the (physical) capital stock employed by the firm. Hence, the firm's 'effective' cost of capital, r, can be calculated as:

$$\frac{\Delta(Costs)}{\Delta K} = \frac{\Delta[wL + \rho K + \gamma(q) * f(K)]}{\Delta K}$$

$$\Rightarrow\ r(q)\ =\ \rho\ +\ \gamma(q) * f'(K)$$

In other words, for a given level of output, even for a price-taking firm operating in a perfectly competitive factor market, its effective cost of capital, r, depends on its production choice q.

Further,

$$r'(q)\ =\ \gamma'(q) * f'(K)\ >\ 0$$

and

$$r''(q) = \gamma''(q) * f'(K) > 0$$

so that the firm's effective cost of capital is rising in output at an increasing rate. Note that although this assumption—that it is the effective cost of capital which is rising in output—may seem intuitive, the model can be modified to consider the case where it would be the effective cost of labour that would be rising in output.[13]

We turn next to deriving the firm's cost function. For simplicity, in what follows, we assume that the firm's production function is of the Cobb-Douglas form, so that we can simply write:

$$q = A * K^{\alpha} L^{\beta}$$

Hence, using standard optimization methods and skipping algebraic details, the firm's cost function can be derived as:

$$c(q) = [w * \Phi(q)^{-\left(\frac{\alpha}{\alpha+\beta}\right)} + r(q) * \Phi(q)^{\left(\frac{\beta}{\alpha+\beta}\right)}] * \left(\frac{q}{A}\right)^{\left(\frac{1}{\alpha+\beta}\right)} \tag{1}$$

where we define: $\Phi(q) = \frac{\alpha * w}{\beta * r(q)}$.

As an inspection of the expression in Equation (1) suggests, this is exactly analogous to the standard cost function corresponding to the Cobb-Douglas production function—the only distinction being that r is written here as a function of q. It is straightforward to check, from Equation (1) above, that $c'(q) > 0$, as one would expect. What about the sign of $c''(q)$? Again, skipping algebraic details, and omitting a positive constant,[14] we can derive the expressions in Equation (2) below, which distinguish between three different expressions ('terms') as follows:

$$c''(q) = \tag{2}$$

$$\left(\frac{1}{\alpha+\beta}\right) * \left(\frac{1-(\alpha+\beta)}{\alpha+\beta}\right) * q^{\left(\frac{1}{\alpha+\beta}\right)-2} * r(q)^{\left(\frac{\alpha}{\alpha+\beta}\right)} \qquad \text{('term 1')}$$

$$+ q^{\left(\frac{1}{\alpha+\beta}\right)} \left(\frac{\alpha}{\alpha+\beta}\right) r(q)^{\left(\frac{\alpha}{\alpha+\beta}\right)-1} r'(q) * \left[\left(\frac{\alpha}{\alpha+\beta} - 1\right) * \frac{r'(q)}{r(q)} + \frac{2}{q(\alpha+\beta)}\right] \qquad \text{('term 2')}$$

$$+ r''(q) * q^{\left(\frac{1}{\alpha+\beta}\right)} \left(\frac{\alpha}{\alpha+\beta}\right) * r(q)^{\left(\frac{\alpha}{\alpha+\beta}\right)-1} \qquad \text{('term 3')}$$

We are now able to establish our first observation.

---

[13] This could happen, for example, if there was a greater danger of 'whistle-blowing' about employment levels by employees. Thanks to Ajit Mishra for this observation. The model can also easily be extended to allow for the possibility that the effective costs of *both* capital and labour might rise with output—the key assumption is that they do not rise exactly proportionately.

[14] The constant term that multiplies all the expressions on the right-hand side of Equation 2 is given by:

$$[\left(\frac{\alpha}{\beta}\right)^{\left(\frac{-\alpha}{\alpha+\beta}\right)} + \left(\frac{\alpha}{\beta}\right)^{\left(\frac{\beta}{\alpha+\beta}\right)}] * w^{\left(\frac{\beta}{\alpha+\beta}\right)} * A^{\left(\frac{-1}{\alpha+\beta}\right)}$$

It is straightforward to see that this is always positive.

Observation 1. In order for the cost function to be convex in output, it is sufficient to establish the following two conditions on the expressions in Equation (2):

(i) 'term 3' > 0

and

(ii) $r'(q) * \frac{q}{r(q)} < 1$ (i.e. the 'output elasticity' of $r < 1$)

$\Rightarrow$ 'term 2' > 0

It is straightforward to see that 'term 1' is positive. Hence, by inspection of Equation (2), we can also establish the other two conditions above as being sufficient conditions for $c''(q) > 0$.

It is worth highlighting the main point of the observation above: it suggests that the cost function can be convex, *even for* $\alpha + \beta = 1$, and indeed, even for some range of $\alpha + \beta$ in which the production function has *increasing* returns. Obviously, the source of these scale diseconomies is the modification of the cost function, by adding the additional 'cost of evasion', but it is worth noting that these diseconomies are not accruing from the production process itself. Rather, these can be viewed more accurately as 'pecuniary diseconomies of scale', coming from the additional regulatory costs.

3.2 Regulatory spillovers

So far, we have modelled the firm as optimizing 'atomistically', without considering any spillovers across firms in the implications of their individual decisions. We now turn to modifying the 'inspection function' to incorporate externalities in regulatory pressure.

To do so, we modify the 'likelihood-of-detection' function, $\gamma(.)$, and write it instead as:

$$\gamma(q_i, q_{-i})$$

In other words, the likelihood of an informal firm i being caught depends also on the activities of *other* firms, where q_{-i} denotes the vector of outputs of all other firms (in that industry or location).

The key question: is the relationship between the likelihood of firm i being detected and the levels of other firms' outputs positive or negative? More prosaically, what is the sign of the cross-partial derivative of the $\gamma(q_i, q_{-i})$ function? Here we can define the following terminology.

(i) Negative spillover: an increase in q_{-i} leads to **decreased** probability of inspection (this parallels the notion that firms' output levels are 'strategic complements'). In other words, there is 'safety in numbers' (Heyes and Kapur 2009).

(ii) Positive spillover: an increase in q_{-i} leads to **increased** probability of inspection (paralleling the idea that firms' output levels are 'strategic substitutes').[15]

---

[15] Or, in other words: 'There is no safety in numbers, or in anything else' (James Thurber in 'The Fairly Intelligent Fly', *The New Yorker*, February 4, 1939).

With this structure in place, we can now turn to the question: how does the regulatory authority affect the firm's decision-making? Clearly, as Heyes and Kapur (2009) argue in the context of environmental regulation, mandates matter: the direction of the spillover depends on the task with which the regulator is charged. Consider two cases:

- Case 1: the regulator is given a fixed budget B, and is asked to minimize, for example, the incidence of informality.
- Case 2: the regulator is set a target, which could be, for example, to restrict the number of informal firms to a certain number or proportion of firms; to restrict informal output to a certain level or proportion of output; etc.

These mandates generate contrasting effects: when the mandate results in a constraint on regulatory attention, this creates negative spillovers across firms, i.e. as informal output grows, the probability of an individual firm being inspected/caught is decreasing. By contrast, if there are 'neighbourhood' or industry effects, these create positive spillovers, i.e. as informal output grows, the probability of an individual firm being inspected/caught is increasing. The key point here is the spillovers created across firms, depending on the regulatory mandate.

We can summarize this discussion in the following remark.

> **Observation 2.** The incidence of informality, driven by the different spillovers under different regulatory regimes, will be higher under the former regime, i.e. when the regulatory mandate results in a constraint on regulatory attention. This could come, for example, if the regulator operates with a fixed budget.

It is worth briefly discussing the implications of this observation, before we turn to discussing further extensions in the next section. By explicitly incorporating regulatory *spillovers* in firms' decision-making vis-à-vis formality or informality, we are able to see more clearly that (regulators') mandates matter: by shaping regulators' incentives, they determine whether these regulatory spillovers across firms are positive or negative. That, in turn, is an element in firms' decisions to operate in the formal or informal sector—the 'extensive margin', to use Ulyssea's (2018) terminology—or to preserve its status as either formal or informal, and use variations in its labour hiring to adjust for increases or decreases in its output—the 'intensive margin', in Ulyssea's (2018) terminology.

What the simple analysis above does is to take a small first step in explicitly incorporating regulatory pressures in firms' decision-making vis-à-vis formality or informality. The large and growing literature on firms and informality is a testament that this decision 'matters': for factor usage decisions, for the cost structure that firms face, for the market structure that emerges, and indeed for the structure of the larger economy as a whole.

4   Discussion and extensions

The baseline model laid out above can be used to provide some insight on a number of related questions. So far, we have largely constructed the model as applying to a generic industry or activity. However, the model can be applied to examine other questions of interest: Why are some industries more likely to have a greater share of their activity in the informal sector than others? One explanation may have to do with technology. For example, if a large amount of physical capital is needed to start production, then informality is less likely to be feasible. Other explanations may have to do with 'salience'—some activities, such as retailing, must be carried out

in a 'visible' context, while, for instance, small-scale production in a 'sweatshop' or home-based care services are harder to spot by regulators. The model presented in this paper suggests an additional explanation based on spillovers across firms, in the extent to which regulatory attention is increased or decreased by the activity of competing firms in the same industry.

To take another implication of our model: under what circumstances should we expect to find geographical or spatial clustering of informal activity? Again, this is a commonly cited stylized fact, most notably in the context of informal markets[16] but also more widely. The model suggests an explanation based on regulatory spillovers: if there is in fact 'safety in numbers', then it is straightforward to see geographic clusters as arising out of the regulatory environment and, in particular, the incidence of the regulatory spillovers across firms.[17]

In further research, one could extend the model to consider the implications for the structure of the labour market. It is sometimes argued that the simple classification of firms as 'belonging' to either the informal or formal sector, while ignoring the possibility of inter-sectoral transition, may be unnecessarily reductive—see Ulyssea (2020) for a discussion. This criticism applies especially strongly when speaking of workers. As Natarajan et al. (2020) point out, it is by no means the case that workers are 'locked into' working in one sector or another.[18] The model presented here does not speak directly to this phenomenon, but an extension might incorporate a measure of the extent to which different sectors are populated by self-employed workers vs wage workers (for example, in the provision of food preparation and delivery) to draw the implications of regulatory changes on the inter-sectoral mobility of labour.

Finally, one can consider the implications of regulatory spillovers for the size distribution of firms, a criterion that is commonly used in the definition and classification of firms as formal/informal, as discussed earlier. If regulations on size (of output, or employment) are an important constraint on evasion of regulations, then, as Kanbur (2009) recognizes, firms might choose to 'relocate'— away from being subject to the applicable regulation towards becoming a firm to which the pertinent regulation does not apply, by adjusting their activity to become of a smaller size. Of course, this comes at the possible cost of loss of economies of scale. Our model suggests an additional motive for the willingness of firms to incur these costs of inefficiently small size. If there are negative regulatory spillovers across firms, then the optimal strategy might be one of 'proliferation', in which production is, in effect, split across loosely affiliated, decentralized units, so that the likelihood of apprehension is lower. In a sense, this is analogous to limited liability in imperfect credit markets—while the incentive to limit detection, and consequent punitive fines, exists in any regulatory avoidance context, these motives might be especially strong when firms take into account the regulatory spillovers across firms in their sector.

---

[16] See, for example, Thomas (1992) and Gërxhani (2004) for a discussion.

[17] It is straightforward to observe that, even within the informal sector, if there are variations (possibly geographic) in the intensity or nature of regulatory enforcement, then the costs, and therefore the optimal outputs, could vary within the informal sector.

[18] They present evidence on the inter-sectoral movement of labour, using longitudinal data from India, to suggest a more nuanced view of this pattern of movement. In particular, they find that self-employed workers are more mobile than wage workers.

5   Conclusion

While the main research strategy here has been the construction of a theoretical model, the analysis has also been informed by the evidence on employment in formal and informal firms—in Brazil (Ulyssea 2018) and some recent evidence from a case study of the Indian context (Mishra 2021). The main contribution of this research is to provide a clearer understanding of the interplay of regulation and its enforcement on the one hand, and firm decision-making about employment and output on the other, in an environment where these issues are not dealt with in separate black boxes. While we do not endogenize the particular form of regulations, the model can be generalized to also incorporate regulators' responses to firm strategies—what sorts of enforcement regimes should regulators choose? The answers to these questions should be of both research and policy interest, in thinking about the policy interventions that might enhance productive employment, especially for informal workers.

References

Banerji, A., and S. Jain (2007). 'Quality Dualism'. *Journal of Development Economics*, 84(1): 234–50. https://doi.org/10.1016/j.jdeveco.2005.09.010

Basu, A.K., N.H. Chau, and R. Kanbur (2010). 'Turning a Blind Eye: Costly Enforcement, Credible Commitment and Minimum Wage Laws'. *Economic Journal*, 120(543): 244–69. https://doi.org/10.1111/j.1468-0297.2009.02298.x

Copeland, B.R., and A. Kotwal (1996). 'Product Quality and the Theory of Comparative Advantage'. *European Economic Review*, 40(9): 745–60. https://doi.org/10.1016/0014-2921(95)00061-5

Copeland, B.R., and A. Kotwal (1997). 'Quality-Biased Technical Progress and North-South Trade'. *Journal of International Trade and Economic Development*, 6(1): 1–14. https://doi.org/10.1080/09638199700000001

de Andrade, G.H., M. Bruhn, and D. McKenzie (2016). 'A Helping Hand or the Long Arm of the Law? Experimental Evidence on What Governments Can Do to Formalize Firms'. *World Bank Economic Review*, 30(1): 24–54. https://doi.org/10.1093/wber/lhu008

de Mel, S., D. McKenzie, and C. Woodruff (2013). 'The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka'. *American Economic Journal: Applied Economics*, 5(2): 122–50. https://doi.org/10.1257/app.5.2.122

de Soto, H. (1989). *The Other Path: The Economic Answer to Terrorism*. New York: Harper & Row.

Dix-Carneiro, R., P.K. Goldberg, C. Meghir, and G. Ulyssea (2021). 'Trade and Informality in the Presence of Labor Market Frictions and Regulations'. IZA Discussion Paper 14060. Bonn: Institute of Labor Economics.

Fajnzylber, P., W.F. Maloney, and G.V. Montes-Rojas (2011). 'Does Formality Improve Micro-Firm Performance? Evidence from the Brazilian SIMPLES Program'. *Journal of Development Economics*, 94(2): 262–76. https://doi.org/10.1016/j.jdeveco.2010.01.009

Farrell, D. (2004). 'The Hidden Dangers of the Informal Economy'. *McKinsey Quarterly*, 26–37.

Gall, T. (2010). 'Inequality, Incomplete Contracts, and the Size Distribution of Business Firms'. *International Economic Review*, 51(2): 335–64. https://doi.org/10.1111/j.1468-2354.2010.00582.x

Garicano, L., C. Lelarge, and J. Van Reenen (2016). 'Firm Size Distortions and the Productivity Distribution: Evidence from France'. *American Economic Review*, 106(11): 3439-79. https://doi.org/10.1257/aer.20130232

Gërxhani, K. (2004). 'The Informal Sector in Developed and Less Developed Countries: A Literature Survey'. *Public Choice*, 120: 267–300. https://doi.org/10.1023/B:PUCH.0000044287.88147.5e

Heyes, A., and S. Kapur (2009). 'Enforcement Missions: Target vs Budgets'. *Journal of Environmental Economics and Management*, 58(2): 129–40. https://doi.org/10.1016/j.jeem.2009.04.005

Kanbur, R. (2009). 'Conceptualising Informality: Regulation and Enforcement'. IZA Discussion Paper 4186. Bonn: Institute of Labor Economics.

Kanbur, R. (2017). 'Informality: Causes, Consequences and Policy Responses'. *Review of Development Economics*, 21(4): 939–61. https://doi.org/10.1111/rode.12321

Levy, S. (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Washington, DC: Brookings Institution Press.

Melitz, M.J. (2003). 'The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity'. *Econometrica*, 71(6): 1695–725. https://doi.org/10.1111/1468-0262.00467

Mishra, A. (2021). 'What Sustains Informality?: A Study Of The Interactions Between Formal- And Informal-Sector Firms'. WIDER Working Paper 2021/71. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2021/009-2

Natarajan, R.R., S. Schotte, and K. Sen (2020). 'Transitions between Informal and Formal Jobs in India: Patterns, Correlates, And Consequences'. WIDER Working Paper 2020/101. Helsinki: UNU-WIDER. https://doi.org/10.35188/UNU-WIDER/2020/858-0

Rauch, J. (1991). 'Modelling the Informal Sector Formally'. *Journal of Development Economics*, 35(1): 33–47. https://doi.org/10.1016/0304-3878(91)90065-4

Sleuwaegen, L., and M. Goedhuys (2002). 'Growth of Firms in Developing Countries, Evidence from Côte d'Ivoire'. *Journal of Development Economics*, 68(1): 117–35. https://doi.org/10.1016/S0304-3878(02)00008-1

Thomas, J.J. (1992). *Informal Economic Activity*. New York: Harvester Wheatsheaf.

Ulyssea, G. (2018). 'Firms, Informality, and Development: Theory and Evidence from Brazil'. *American Economic Review*, 108(8): 2015–47. https://doi.org/10.1257/aer.20141745

Ulyssea, G. (2020). 'Informality: Causes and Consequences for Development'. *Annual Review of Economics*, 12: 526–46. https://doi.org/10.1146/annurev-economics-082119-121914

World Bank (2019). *Doing Business 2020*. Washington, DC: The World Bank. Available at: www.doingbusiness.org/en/reports/global-reports/doing-business-2020 (accessed 31 March 2021).

Yeaple, S.R. (2005). 'A Simple Model of Firm Heterogeneity, International Trade, and Wages'. *Journal of International Economics*, 65(1): 1–20. https://doi.org/10.1016/j.jinteco.2004.01.001