EconomiX

A Search Model with Self-Employment and Heterogeneity in Managerial Ability

Eliane Badaoui
Olivier Bargain
Prudence Magejo
Eric Strobl
Frank Walsh

2023-9 Document de Travail/ Working Paper

EconomiX - UMR 7235 Bâtiment Maurice Allais
Université Paris Nanterre 200, Avenue de la République
92001 Nanterre Cedex

Site Web : economix.fr
Contact : secreteriat@economix.fr
Twitter : @EconomixU

cnrs Université Paris Nanterre

A Search Model with Self-Employment and Heterogeneity in Managerial Ability

Eliane Badaoui* Olivier Bargain† Prudence Magejo‡
Eric Strobl§ Frank Walsh¶

Abstract

The view of informal employment as a last resort in the labour market has recently been challenged by numerous studies documenting the existence of a high degree of heterogeneity within the formal and informal sectors - in particular the presence of high-tier informal work corresponding to voluntary self-employment. There is currently not much theoretical support for these observations. We develop a formal model to explain this growing empirical evidence about substantial heterogeneity within formal/informal labour markets. In our model, workers may enter self-employment or search for jobs as employees, while allowing for heterogeneity across workers' managerial ability. While workers with higher managerial ability will manage larger firms, workers with lower managerial ability will manage smaller firms and be in self-employment only when they cannot find a salaried formal/informal job. For the latter, self-employment in the informal sector is the outside employment option.

Keywords: Self-employment; Managerial ability; Informal sector.

JEL Classification: J31; O17.

*EconomiX-CNRS, University of Paris Nanterre, France.
†Bordeaux School of Economics, France, and IZA
‡University of Witwatersrand, South Africa.
§University of Bern, Switzerland.
¶*Corresponding author*. University College Dublin, Ireland. E-mail: frank.walsh@ucd.ie.

1 Introduction

Many of the earlier empirical studies comparing the wages of formal and informal sector workers suggest that informal sector workers, even if equally productive, are paid less than their formal sector counterparts.[1] A number of explanations have been offered in this regard, mostly based on a segmented view of the labour market. For instance, using a search matching model calibrated to Mexican data, Satchi and Temple (2009) conclude that workers who do not find a formal job are likely to join the informal sector.[2] In practice, the evidence has been much more mixed. For example, Cunningham and Maloney (2001) suggest that there may be an upper and lower tier of informal enterprises. In Maloney (2004), evidence from Latin American countries shows that the informal sector is mostly voluntary and that the informal labour market is heterogeneous and composed of both self-employed and salaried workers. Marcouiller et al. (1997) find a wage premium associated with informal work in Mexico but not in other countries, while Günther and Launov (2012) find evidence of both competitive and segmented labour markets in Côte d'Ivoire. According to Bosch and Maloney (2006), the performance of labour markets is asymmetric during business cycles, i.e., decisions are voluntary during expansionary periods but segmentation prevails during recessions. Similarly, Perry et al. (2007) argue that voluntary entry of workers in one of the two sectors may not rule out some degree of market segmentation.

In other instances, the existence of a wage penalty may depend on gender (Tansel, 1999), the level of education (Gong and van Soest, 2002), unobserved ability (Falco et al., 2011), firm size (El Badaoui et al., 2010; Falco et al., 2011), or on the point in the wage distribution (Bargain and Kwenda, 2011; Tannuri-Pianto and Pianto, 2002). In some studies, any payment differential can actually be entirely explained by unobserved heterogeneity (Pratap and Quintin, 2006). Considering the wide range of evidence overall, it is perhaps best to conclude that, if anything, there is a good deal of heterogeneity in the size of the informal premium/penalty and that for different groups of workers, as suggested by Fields (2009), the reality may be very diverse, i.e. the desirability of informal sector employment in terms of remuneration depends on the segment of the labour market in question. Gollin (2014), in a retrospective article on the seminal paper by Lewis (1954) which outlined a model of a dual labour market, notes: "*Perhaps in the end, Lewis's dualism is too stark. The dichotomy between capitalist and subsistence sectors appears*

[1] For further support of the segmented view, see Gong and van Soest (2002); Heckman and Hotz (1986); Mazumdar (1981); Pradhan and van Soest (1995); Roberts (1989); Tansel (1999).

[2] For instance, the presence of barriers to entry into the formal sector could pose a possible cause, so that working in the informal sector is associated with a negative wage premium even for equally productive workers; see Fields (1975) and Mazumdar (1981).

on closer examination to be more of a continuum."

In this paper we develop a model that attempts to capture this heterogeneity using a framework with search frictions. Our model applies mainly to labour markets in developing countries and considers three different employment states (wage employment, self-employment, and unemployment). It combines heterogeneity in managerial ability with search frictions in the labour market, where self-employment may be either desirable or an undesirable but an unavoidable state for different groups of workers. To clarify our contribution, we will discuss the existing body of literature that uses search models to analyse the informal sector.

Albrecht et al. (2009) extend the Mortensen and Pissarides (1994) matching model to incorporate a self-employed informal sector where there is heterogeneity in workers' productivity in that more productive workers may opt to wait for a formal sector job, while others may select into the informal sector. Ulyssea (2010) develops a matching model incorporating key institutional features of the Brazilian economy and conducts policy simulations. Also, Boeri and Garibaldi (2006) develop a matching model with supervision where workers in the informal sector cannot avail of unemployment benefits and heterogeneity in worker skill. Their model suggests that policies aimed at reducing the size of the informal sector may increase unemployment. Alternatively, Fugazza and Jacques (2004) incorporate psychic costs as part of the cost of being in the informal economy in a matching model where workers direct their search at informal sector firms.[3] Poschke (2019) develops a matching model and examines the co-existence of high rates of unemployment and self-employment in developing countries. A key difference between this model and our framework is that the equilibrium search model we suggest here allows for on-the-job search where workers do not lose the opportunity to seek employment when they enter self-employment.

Meghir et al. (2015) outline an equilibrium search model with a formal and informal sector. They simulate the model for the Brazilian economy and find that tighter enforcement can increase wages without increasing employment by improving the allocation of workers across sectors. Narita (2020) incorporates self-employment into a two-sector model similar to Meghir et al. (2015) and simulates the model using Brazilian data and analyses the impact of payroll taxes. The model is most similar to the model we develop here in that it is an equilibrium search framework which builds on the Burdett and Mortensen (1998) model and allows for self-employment and heterogeneity in managerial ability. As with Meghir et al. (2015), Narita (2020)

[3] Bosch (2006) and Bosch and Maloney (2006) look at the effect of an informal sector on patterns of job creation and job destruction.

allows for separate wage distributions for the formal and informal sectors in a model with on the job search. In contrast, our model has a single wage distribution, which greatly simplifies the analysis and has a number of advantages. Our framework allows us to explicitly trace out a hierarchy of firms with the following properties. Firms with higher ability managers offer higher wages. The smaller low wage firms run by managers with lower managerial ability will be in the informal sector. Self-employment will be the residual sector for workers with poor managerial skills who would prefer any job to self-employment but may be stuck in self-employment because search frictions make it harder to find a job. For the highest skill managers, self-employment will be more attractive than any salaried job. We solve the model by initially assuming an exogenous wage distribution in a general framework and then solve two simple examples using the Burdett and Mortensen (1998) model. When we solve the model explicitly, we can conduct comparative static analysis with interesting results[4]. The first of the two examples we solve extends the model in El Badaoui et al. (2010) (EB from now on) to incorporate self-employment. In this model managers default to avoid paying taxes but risk being caught and having to pay penalties. The second example incorporates differences in capital costs between formal and informal firms as the benefit of being formal, a benefit that is increasing in managerial ability. For the first example we will see that the tax rate is independent of the share of workers in defaulting firms. The intuition typically is that the penalties for tax defaulters are a function of the amount of tax owed. This means that a higher tax rate will increase the expected penalties facing defaulters as well as the costs of complying. In our first example, these two effects cancel out in a manner that seems consistent with empirical evidence. For example, Rocha et al. (2018) present evidence from Brazil that lowering taxes had limited effects on the likelihood of a firm being formal. In the second example, where the tax can be thought of as the fixed costs of formalising when formal firms have lower capital costs to invest in worker specific capital, the share of employees in the formal sector will rise as the costs of compliance falls or as the gap between capital costs in the formal and informal sectors increases. The analytic comparative static results demonstrate that managerial productivity has an important role in determining the size of these effects. That is, the extent to which the share of workers in the formal sector increases as the costs of becoming formal fall or as the benefits of being formal rise will be

[4]Our main theoretical results are established in a general model where large firms are assumed to pay higher wages. As in the Lucas (1978) span of control framework, workers differ in their relative productivity in employment and self-employment (managing a firm). Amaral and Quintin (2006), Blau (1985) and Rauch (1991) develop models where workers with higher managerial ability select into self-employment in a developing country context. There are no search frictions in these models.

higher the higher is managerial productivity in the firm managed by the marginal manager who is indifferent between being in the formal or informal sector. These results are consistent with the empirical literature discussed later which suggests that the degree to which firms respond to reductions in the cost of formalising or other incentives varies greatly with firm/managerial productivity.

The remainder of our paper is organized as follows. In the next section, we present a review of the literature on firm size, access to the credit market, and informality. Our theoretical model is outlined in Section 3. We use the Burdett and Mortensen (1998) framework to solve for the wage distribution explicitly in Section 4. Concluding remarks are given in the final section.

2 Firm size, access to capital markets and informality

There are many possible reasons for being in the informal sector (see Dabla-Norris et al. (2006) for a survey of the literature on informality). The model we adopt captures two of the most important stylised facts characterising informal employment. Firstly, we argue that large firms are more visible and will typically be formal so that, as most models of informality assume, more productive and larger firms will be formal (El Badaoui et al., 2010). La Porta and Shleifer (2014) outline a range of stylised facts about informal firms drawing on their more detailed analysis outlined in La Porta and Shleifer (2008) which is based on World Bank surveys of both formal and informal firms across a range of poor countries. Many of the key features of informal firms from this evidence are in keeping with our model. In particular, La Porta and Shleifer (2014) show that informal firms are small, unproductive and stagnant[5]. They note that "*Many informal entrepreneurs would gladly close their businesses to work as employees in the formal sector if offered the chance...*". They also note that "*In production function estimates, managerial human capital emerges as a quantitatively large and statistically significant determinant of productivity*".

Secondly, in one of our examples, we allow for the possibility that capital is more expensive for informal firms, capturing the possibility that access to capital markets is an important reason for informality as the literature suggests. For instance, Amaral and Quintin (2006) outline a model where enforcement of contracts is weaker in the informal sector. High ability managers who need to borrow more find it more profitable to be formal while low ability managers stay in the informal sector. Low wages for self-employed managers of informal firms reflect returns to ability in this model. At the same time, it may be that an inefficient bureaucracy imposes

[5]Stagnant in the sense that they tend not to increase in productivity and rarely progress to being formal firms.

costs on the formal sector so that some firms will go informal. In the case of entrepreneurship, the equilibrium size for each firm depends on the abilities of the entrepreneur (Lucas, 1978). It is, however, important to state that the structure of financial institutions can determine the efficient firm size (Stein, 2002). At the firm level, access to capital may depend on the size of the firm as well as on its belonging to the formal or informal sector.[6] Straub (2005) builds a model where firms have to choose between formality and informality, and complying firms benefit from public goods, enforcement of property rights and contracts, and participate in the formal credit market. He finds that under the hypothesis of credit rationing, their decision is shaped by the interaction between the cost of entry into formality and the relative efficiency of formal versus informal credit. Stiglitz and Weiss (1981) emphasise the importance of access to credit for firm investments. Evans and Jovanovic (1989) show that if the liquidity constraint were removed the average probability of becoming an entrepreneur would increase. Moreover, according to Evans and Jovanovic (1989), the difference between optimal and actual capital is lower for people with low ability. Blanchflower and Oswald (1998) state that raising capital is the principal problem of potential entrepreneurs. Similarly, Cabral and Mata (2003) argue that the expansion of small firms is hampered by firms' financial constraints. Buyinza and Bbaale (2013) find that, within the East African Community, the majority of firms are credit constrained and that high performing firms, exporters, medium and large firms have a higher probability to access to formal credit.[7]

One obvious reason for being formal which is not captured by our model is that official certification can be important for a self-employed professional or tradesman so that particular types of small firms such as doctors and lawyers etc. will be in the formal sector for this reason. Others, such as artists or writers, may be very visible and the possibility of being caught not paying taxes may have high reputation costs. For these reasons we do not expect that all small firms will be informal and that some formal small firms will have high wage rates[8].

[6]For Germany, Audretscha and Elston (2002) find that medium sized firms are more liquidity constrained in their investment behavior than either the smallest or largest firms.

[7]Informal credit markets are strongly present in many developing countries. See, for instance, Bell (1990). Giné (2011) shows that formal and informal credit markets coexist. Banerjee and Duflo (2007) show that formal credit markets benefit to wealthy borrowers (or large farms in rural areas) while informal credit markets are mainly attributed to poor borrowers or small farms. Despite the presence of some zero-interest loans, borrowers on the informal credit market pay usually high interest rates of up to 150% per year (Banerjee, 2003).

[8]Workers with high managerial ability can be constrained, for some exogenous reasons, by lack of access to financial capital due to risk aversion, poor infrastructure or safety from crime, etc. (Gindling and Newhouse, 2012). These cases are not captured by our model.

3 The Model

We will assume that worker types (indexed by i) are ordered according to their productivity in self-employment. That is, type-0 has the lowest productivity parameter (p_0), type-1 the second lowest (p_1) and so on up to type-g the highest productivity workers. There is a mass L_i of type-i workers in the economy where the total mass of workers is $L = \sum_{i=0}^{g} L_i$. Workers can be unemployed, self-employed or salaried. Since we shall focus on a stationary equilibrium, we suppress the time subscripts to avoid clutter. The self-employed shall choose to manage a formal or an informal firm. While each worker type differs in his/her productivity as a potential manager, all workers are equally productive as employees. In reality, see below, there will be differences in workers' productivity in employment and this may differ across top skills and low skills managers and across formal and informal sectors. We can think of this assumption as a simple way of ensuring that some workers have a comparative advantage in self-employment. Albrecht et al. (2009) analyse a search model where productivity in self-employment is fixed and there is variation in skill levels in salaried employment, showing that employees with low skills may prefer to stay in self-employment rather than incur the search costs associated with finding a job. In their model, informal self-employment is a low wage sector and since low wages reflect low skills, workers are not involuntarily stuck in self-employment.[9] We do not view the results from Albrecht et al. (2009) as contradictory to the results in this paper. Rather they are another part of the heterogeneity that exists in different segments of informal/formal labour markets. For example, we might guess that, if self-employment ability were to vary in the Albrecht et al. (2009) model, some workers with poor managerial skills may be involuntarily self-employed but unable to find a job as in our model.

The model we present is one where, at any point in time, workers can choose to enter own-account self-employment or to manage a firm with some positive level of employment. The relative benefits of these different self-employment options will depend on the workers managerial ability. These workers will also receive wage offers which sometimes may attract them away from self-employment into salaried employment.

Formal sector firms are different from informal firms in that they pay tax while informal ones do not. Informal firms risk incurring the penalty associated with non-compliance. Smaller firms who are more costly to monitor are less likely to be caught and thus more likely to be in the

[9] Boeri and Garibaldi (2006) adopt the same view according to which low ability workers decide to work mainly with the status of self-employed in the informal sector.

informal sector. Here we follow EB[10] and introduce a tax rate t on wage income that is paid by firms. We outline a general model where there is a positive and continuous relationship between a firm's employment n and the wage w, $n(w)$, in a stationary equilibrium, but do not specify why this positive relationship exists. Firms have the production function $p_i q(n)$ where p_i is the productivity of a type i worker as a manager and $q(n)$ is the output. We assume that $q_n(n) > 0$ and $q_{nn}(n) \leq 0$. Explicitly solving an equilibrium search model typically involves assuming constant productivity because of the complexity of these models. Indeed, in the example we solve later in this paper, we will need to assume constant productivity to solve the model. But initially we assume an exogenously given wage distribution and under this assumption the results below hold whether marginal productivity is constant or diminishing. There is a tax rate t on wages and a Poisson arrival rate of negative shocks, δ, which will destroy the firm. The Poisson arrival rate of tax inspectors is a positive function of the number of employees at the firm, $\Theta[n(w)]$. If firms are caught not paying their taxes they are punished and must pay a fine according to the function $\Omega[wtn(w)]$, which is increasing in the per period tax bill $wtn(w)$. In a stationary equilibrium, a firm's labour supply times the separation rate equals the hiring rate: $n_i(w) = \frac{h_i(w)}{d_i(w)}$, where $h_i(w)$ is the hiring rate for a firm of type i and $d_i(w)$ is the separation rate. We assume for now that the hiring rate is increasing in the wage, noting that we will explicitly solve for the hiring rate in an example later in the paper.

The flow of profits in self-employment for a manager of type-i managing a defaulting (d) or compliant (c) firm in a stationary equilibrium at any wage w is:

$$\begin{aligned} \pi_i^d &= p_i q[n(w)] - wn(w) - \Theta[n(w)]\Omega[wtn(w)] \\ \pi_i^c &= p_i q[n(w)] - wn(w) - wtn(w) \end{aligned} \tag{1}$$

Defaulting firms pay the wage and no tax but incur a cost if they are caught defaulting, the probability of which depends on the number of workers, while compliant firms do not incur this cost but pay the tax.

The flow value of managing a firm for a defaulting (d) and a compliant (c) self-employed type-i manager is given below where w^*, which is fixed in equations (2), is the profit maximising

[10]Much of the following passage is taken straight from EB who discuss the large literature that equates informality with small firms in more detail. Note that a fundamental difference with EB is that self-employment was not modelled explicitly in this earlier contribution.

wage that the manager chooses.

$$rV_i^d(w^*) = \pi_i^d(w^*) + \lambda \int_{\underline{w}}^{\overline{w}} \left\{ \max \left[V_i^d(w^*), E_i(x) \right] - V_i^d(w^*) \right\} dF(x) - \delta V_i^d(w^*) \tag{2}$$

$$rV_i^c(w^*) = \pi_i^d(w^*) + \lambda \int_{\underline{w}}^{\overline{w}} \left\{ \max \left[V_i^c(w^*), E_i(x) \right] - V_i^c(w^*) \right\} dF(x) - \delta V_i^c(w^*)$$

The flow value of the firm, where r is the discount rate, is the dividend stream (flow of profits) plus any capital gain/loss in the value of the firm which comes from employment opportunities and from the arrival rate of negative shocks that destroy the firm. The exogenously given job offer distribution is $F(w)$ where $\underline{w}$ is the lowest wage offered and $\overline{w}$ the highest wage offered. This job offer distribution gives us the probability that any offer received has a wage of w or less. We make the simplifying assumption that the stream of job offers is the same for an employee, a self-employed worker, or an unemployed worker. We will see below that the value of these job offers will differ across workers depending on their productivity in self-employment, which varies across workers. All workers receive job offers at a Poisson arrival rate *lambda*. The gain associated with an employment offer with value $E_i(w)$ is $[E_i(w) - V_i^j(w^*)]$ where $j \in (d, c)$, or zero for offers worse than the current state. We get the expected value of offers by integrating over the wage offer distribution. Unemployment has the flow value:

$$rU = b + \lambda \int_{\underline{w}}^{\overline{w}} \left\{ \max \left[U, E_i(x) \right] - U \right\} dF(x) \tag{3}$$

The stream of unemployment benefits is b. If we think of the value of unemployment in equation (3) as an asset, the worker has the possibility of a capital gain in the asset value if a job offer that is better than unemployment arrives. The gain associated with an employment offer with value $E_i(w)$ is $[E_i(w) - U]$. If the worker receives an offer no better than the current state, he/she stays in self-employment and the capital gain term is zero. Self-employment is preferred to unemployment if the expected value of self-employment net of the setup cost of a firm is better than the expected value of unemployment:

$$\max[V_i^d, V_i^c] > U \tag{4}$$

The flow value of employment in any job is:

$$rE_i(w) = w + \lambda \int_{\underline{w}}^{\overline{w}} \{\max[E_i(w), E_i(x)] - E_i(w)\} dF(x) - \delta \left\{ E_i(w) - \max[V_i^d, V_i^c, U_i] \right\} \quad (5)$$

This is the flow value of the wage plus the value of higher wage job offers times the arrival rate of such offers integrated over the job offer distribution, plus the arrival rate of negative shocks that lead to job loss times the capital loss from losing a job. The latter is the difference between the value of the current state and the maximum of the value of unemployment and the value of self-employment (net of the cost of setting up a firm).

The manager of a firm will choose to comply with its tax obligations if:

$$V_i^c > V_i^d \quad (6)$$

We note that in a stationary equilibrium:

$$V_i^c(w^*) - V_i^d(w^*) = \frac{\pi_i^c(w^*) - \pi_i^d(w^*)}{r + \delta} = \frac{\Theta[n(w)]\Omega[wtn(w)] - wtn(w)}{r + \delta} > 0 \quad (7)$$

Initially, rather than fully specifying an equilibrium search model we adopt a simpler approach. We assume a well behaved equilibrium exists and show below that if all agents are optimising and there are no barriers to moving across states other than the costs of setting up a firm and the search frictions associated with finding a job that are outlined in the model, then optimising behaviour implies the four propositions outlined below. Explicitly solving an equilibrium search model as we do in the next section will involve putting a lot more structure on the model. In particular, we will assume constant productivity and use an extension of the Burdett and Mortensen (1998) framework, outlined in Mortensen (2003) to solve for the wage distribution explicitly and show that the propositions hold.

Proposition One:

Higher productivity firms offer higher wages than lower productivity firms.[11] *That is, if there are two firms managed by type-i and type-j managers, respectively, where* $p_i > p_j$*, then* $w_i \geq w_j$ *for all wages paid by type-i and type-j firms.*

Proof in Appendix A.1

[11] Burdett and Mortensen (1998) page 268 also show this to be true in their model.

We assume that there is a stationary equilibrium where there is a continuous positive relationship between the number of employees and the wage rate $n(w)$. We will denote the tax bill as $B = wtn(w)$ for shorthand.

Proposition Two:

If the expected penalty from non-compliance is increasing at least as fast as the tax bill in the wage (number of workers) and there are some compliant and some non-compliant firms of a given type in equilibrium, these are sufficient conditions such that there will be a cut-off point in firm size below which all firms will default on their taxes, and above which firms will be compliant. In other words, there will be a wage distribution with small low-wage firms in the informal sector and large high-wage firms in the formal sector.

Proof in Appendix A.2

Propositions One and Two establish that higher productivity firms pay higher wages and that, under some conditions, there will be a cut-off point for firm size and the wage above which firms will choose to be compliant. Large high-wage firms will be compliant and small low-wage firms will be non-compliant. This was also the case in EB where there was no difference in productivity across firms and where self-employment was not modelled explicitly[12]. We will denote the lowest and highest wage paid by each firm type in equilibrium, respectively, as $\underline{w}_i$ and $\overline{w}_i$.

Proposition Three:

Workers with the lowest managerial ability would prefer any job to self-employment.

Proof in Appendix A.3

Proposition Three shows that workers with the lowest managerial ability who are self-employed will be in a secondary informal sector in the sense that any job would be better than their current state. If unemployment benefits are sufficiently low, the condition in equation (4) will hold for the lowest ability group and this group will prefer self-employment to unemployment. But if benefits are sufficiently high, the condition in equation (4) will not hold and this group will choose unemployment over self-employment.

[12]The conditions under which the proposition held in EB are slightly different because EB did not include the difference in capital costs between formal and informal firms.

We will refer to firms managed by a type-i worker as "type-i firms" and firms with only the manager employed as "own-account firms" from now on. We assume that an own-account worker has labour supply n_0 so that output for an own-account worker of type-i is $p_i q(n_0)$. Since the only tax is a payroll tax and own-account workers do not hire others, their profit will be:

$$\pi_i = p_i q(n_0) \tag{8}$$

Proposition Four:

As long as some firms are larger than own-account firms, the highest ability managers will prefer self-employment to any employment offer.

Proof in Appendix A.4

While Proposition Three established that workers with the lowest managerial ability would prefer any job to self-employment, Proposition Four establishes that there will be a group of workers (those with the highest managerial ability) who prefer self-employment to any job offer. We note that this implies that if there are more than two ability groups, there will be a lot of heterogeneity across the pool of self-employed workers. In particular, there will be a low ability group who would prefer any job to self-employment, a high ability group who would not accept any job, and possibly an intermediate ability group who (given that there are good and bad job offers) would accept some jobs but not others.

We follow EB and assume the Poisson arrival rate of tax inspectors is a constant μ times employment to the power of a constant β so that large firms are more likely to be caught defaulting: $\Theta[n(w)] = \mu n(w)^{\beta}$. We specify the penalty for defaulting as x times the firm's per period tax bill: $\Omega[wtn(w)] = xwtn(w)$. To save on notation, we define $z = x\mu$ as the parameter that when multiplied by employment to the power of $\sigma = \beta + 1$ determines the expected punishment for defaulters at any point in time. We will see later that when we come to writing the solution for the equilibrium profit and wage distribution of non-compliant workers, using $\sigma = \beta + 1$ will be neater. In this case, from equation (7), compliance is optimal if:

$$n(w) > z^{\frac{1}{1-\sigma}} \tag{9}$$

4 The Burdett Mortensen model as an example

In the example we solve for below we will allow for two cases. The simple case is with no capital and an example where we use a version of the Burdett and Mortensen (1998) model with match specific capital outlined in Mortensen (2003). As noted earlier the possibility that capital is more costly in non-compliant firms is an important part of the difference between formal and informal firms.

Below we summarise labour market flows and then derive the equilibrium labour supply curve and wage offer distribution using the equilibrium search model outlined in Burdett and Mortensen (1998) [BM from now on] and extended to incorporate investments in firm specific capital in Mortensen (2003). This model serves as an example where a positive relationship between firm size and wages emerges endogenously in equilibrium and where we can solve the model explicitly. While equilibrium search models have been solved with heterogeneity across workers outside options and firm productivity [see Bontemps et al. (1999) or Postel-Vinay and Robin (2002), for example], the model we develop here is complicated by the fact that a worker's outside option is defined by the profit (s)he can make in self-employment, that is the distribution of outside options are not exogenous. We adopt the simplest approach possible to solve the model in a tractable way that illustrates some of its key features. In particular, we will assume that there are only two ability groups. In equilibrium the high (H) managerial ability will be employers while the low ability group (L) will be either employees or own-account workers in self-employment.

We will assume that unemployment benefits are low enough so that the low ability group choose to be in self-employment and thus there will be no unemployed workers in equilibrium.[13] As in the BM model, on the job search will ensure that wage dispersion will emerge in equilibrium, even though productivity in employment is the same for all workers. The assumption that workers transition directly between self-employment and being an employee may not fit every country. Narita (2020) shows that most transitions in and out of self-employment in Brazil are via a spell of unemployment between 2002-2007 when unemployment rates were between 8% and 9%. We report in Appendix B the rates of unemployment and self-employment across a range

[13]This assumption fits well in most developing countries where many poor people, often a majority, are self-employed in both urban and rural areas (Fields, 2019; Gollin, 2008). Fields (2019) argues that while some workers are voluntarily self-employed, others are self-employed because they do not have the possibility of being wage employees and because they are too poor to remain unemployed and earn nothing. According to Poschke (2019), a plausible reason is that unemployment insurance benefits, b, are lower than wages in salaried employment, w, since developing countries do not provide unemployment insurance benefits. He gives evidence from Ethiopia where the ratio b/w is equal to 0.4.

of developing countries. We show that, for many developing countries, unemployment rates are negligible, making the assumption that workers move directly between self-employment and employment more plausible. Fields (2019) suggests that most workers in developing countries are self-employed. The data we report in Appendix B shows that almost 54% of workers in low- and middle-income countries are self-employed, among whom the majority are own-account workers.[14] This share reaches 80% in low-income countries. Self-employment differs considerably by region, where Sub-Saharan Africa and South Asia have self-employment rates exceeding 70%. The highest share of self-employed of 95% is recorded for Niger with no unemployment. This suggests some variability among countries. Interestingly, the unemployment rate is overall low. The data indicate that unemployed constitute 5.5% of the labour force in developing countries and 4.6% among low-income countries. In this regard, unemployment is limited and unlikely the biggest challenge in most developing countries (Fields, 2019). Poschke (2019) reveals that self-employment is particularly high in the presence of labor market frictions. More precisely, he argues that "*while labor market frictions always reduce wage employment, they do so via higher unemployment when firm entry costs are high, as in rich economies, but via higher self-employment when firm entry is cheap, as in poor economies*". Narita (2020) shows that self-employment entry opportunities "*increase as the unemployed get older regardless of their ability since the option of self-employment is always preferred to remaining unemployed*".

We will proceed by deriving the labour supply curve in a model where there are search frictions and workers receive on the job offers.[15] We define m_i as the mass of type-i self-employed worker (managers) who employ others and s_i as the mass of type-i self-employed own-account workers. This means that a mass of $m_i + s_i = S_i L_i$ type-i workers are in self-employment, where S_i is the self-employment rate for type-i workers and L_i the population of type-i workers. We note from Proposition Four that $S_H = 1$, i.e., all of the high ability workers will be self-employed managers $m_H = S_H L_H = L_H$ (we will assume that the productivity of these workers is high enough such that their profit from hiring workers is greater than from being an own-account worker). This condition will be determined in the equilibrium outlined below. Low ability workers who fail to obtain a job are self-employed own-account workers and thus $s_L = S_L L_L$ with $S_L < 1$, and $(1 - S_L)$ is the fraction of employees. There is random matching so that any job offer is equally likely to come from any firm irrespective of the firm's size.[16] The distribution

[14] See, for instance, Fields (2019), Narita (2020) and Poschke (2019).
[15] See Mortensen (2003) and Burdett and Mortensen (1998) for a detailed derivation of the labour supply curve.
[16] See Manning (2003) pp. 284-286 for a discussion on the matching technology.

of wage offers which we will solve for is $F(w)$. BM assume $r = 0$ in their derivation of the labour supply curve and we follow this assumption. The other assumptions and parameter definitions from the previous section continue to hold. In a stationary equilibrium, inflows and outflows to self-employment of each worker type are equal, implying the following relationship between inflows and outflows from self-employment and employment for low ability workers (who are the only group who will be in employment in our model with two skill groups) where the share of wage offers which fall below the self-employment profit a low skill worker can earn is $F(\pi_L)$:

$$\lambda S_L[1 - F(\pi_L)] = \delta(1 - S_L) \tag{10}$$

The outflows, on the left-hand side, are the job offer arrival rate λ times the fraction of offers where the job offer is better than the profit worker type-L expects in self-employment. The inflows to self-employment, on the right hand side, are the fraction of type-L workers (we recall that type-H workers are all managers) who are employees times the job destruction rate. It follows that, since all high ability workers are managers and since wages exceed self-employment profits for all low ability workers, $F(\pi_L) = 0$ in any equilibrium where there are some employees from type-L workers:

$$S_L = \frac{\delta}{\delta + \lambda} \tag{11}$$

For high ability workers, we know from Proposition Four that $S_H = 1$. Firms managed by high ability managers that go out of business after being subjected to a negative shock are replaced by new firms which will be created by the same managers in equilibrium, implying that there are no inflows or outflows from self-employment for this group.

In a steady state equilibrium, the outflows from low skill employment (the separation rate plus the flow of job offers received from higher wage firms, times the stock of employment N_L) and inflows from low skill employment (the number of offers less than w accepted by self-employed workers of type-L) are equal, i.e.:

$$\dot{N}_L = \{\delta + \lambda[1 - F(w)]\} N_L(w) - \lambda F(w) S_L L_L = 0 \tag{12}$$

We assume that labour supply of low ability workers: $n_0 = 1$.

We can solve for employment of type-L workers earning w or less from equation (12) as

$N_L(w) = \frac{\lambda F(w) S_L L_L}{\delta + \lambda[1-F(w)]}$. Low skill employment at a wage less than w can also be defined as $(1 - S_L)$, times the wage distribution of this group, $G_L(w)$, times the population of workers of this type L_L:

$$N_L(w) = (1 - S_L)G_L(w)L_L = \frac{\lambda F(w) S_L L_L}{\delta + \lambda[1 - F(w)]} \tag{13}$$

Using the fact that $\frac{S_L}{1-S_L} = \frac{\delta}{\lambda}$ and rearranging equation (13) we can solve for $G_L(w)$:

$$G_L(w) = \frac{\delta F(w)}{\delta + \lambda[1 - F(w)]} \tag{14}$$

The probability that any worker contacted by the firm will be hired times the number of workers contacted is the firms hiring rate:

$$h_L(w) = \lambda \frac{L_L}{L_H} [S_L + (1 - S_L)G_L(w)] = \frac{L_L}{L_H} \frac{\delta\lambda}{\delta + \lambda[1 - F(w)]} \tag{15}$$

S_L is the probability any worker contacted will be self-employed. Self-employed workers will accept all wage offers, and all wage offers will be greater than the value of self-employment, since a lower offer would attract no workers and would never be made. $(1 - S_L)$ is the probability any worker contacted is in employment, $G_i(w)$ is the probability they earn no more than w and will accept the firms wage offer and $\lambda \frac{L_L}{L_H}$ is the number of workers each firm expects to contact in each period given random matching.

The separation rate $d(w)$ at any firm is the sum of the job destruction rate δ plus the arrival rate of offers to each worker times the probability the offer comes from a higher wage firm $\lambda[1 - F(w)]$:

$$d(w) = \delta + \lambda[1 - F(w)] \tag{16}$$

The separation rate times employment equals inflows per firm in a stationary equilibrium, so that the labour supply of type-L workers is:

$$n_L(w) = \frac{h_L(w)}{d(w)} = \frac{L_L}{L_H} \frac{\delta\lambda}{\{\delta + \lambda[1 - F(w)]\}^2} \tag{17}$$

Equation (17) gives the labour supply of type-L workers to a firm paying a wage w in terms of the equilibrium wage offer distribution.

We can assume all managers are high skill when writing the profit function from Proposition Four. The profit functions are the same as presented earlier in (1) except for the additional assumption that the marginal productivity of workers is constant. Since in equilibrium in our example with two ability types there can be no job offers that a manager will receive which will be more attractive than being a manager, the value of self-employment given in equation (2) can be written as follows:

$$\begin{aligned} V_H^d(w^*) &= \frac{\pi_i^d(w^*)}{\delta} = \frac{(p_H - w)n(w) - n(w)^\sigma zwt - \gamma^d k_H^d h_H(w)}{r+\delta} \\ V_H^c(w^*) &= \frac{\pi_i^c(w^*)}{\delta} = \frac{[p_H - w(1+t)]n(w) - \gamma^c k_H^c h_H(w)}{\delta} \end{aligned} \tag{18}$$

Since we assume $n_0 = 1$, the production function for an own-account worker with low managerial ability is:

$$q_L[p_L, n_0] = p_L \tag{19}$$

Equation (19) is also the profit function of own-account low ability workers since the only tax is a payroll tax and type-L self-employed non-compliant workers do not hire others. From equation (3) the value of unemployment is increasing in benefits b. We assume that $p_L n_0 > b$. Since unemployed workers and own-account workers draw from the same wage offer distribution and have the same arrival rate of job offers, if the flow of income in unemployment b is less than the flow of profit of a low ability worker $p_L n_0$, which we assume it is, then the value of own-account self-employment is better than unemployment for low ability workers. We also see immediately that it could never be profitable for a low ability worker to hire another worker since any other worker will be at least as productive as an own-account worker.

4.1 The equilibrium wage distribution

The lowest wage a low ability worker will consider working for is π_L since this is the wage that they can earn in self-employment. We can solve for the profit of the lowest informal firm by solving for the level of profit of a firm that pays this wage, where we know from Proposition Two that this will be a defaulting firm. We also know that in equilibrium all wage strategies must have the same level of profit since any firm can choose any wage. We note that the labour supply of the lowest wage firm (where any offer is accepted by the workers so that $F(w) = 0$) can be written as $n_L \pi_L = \frac{L_L}{L_H} \frac{\delta\lambda}{(\delta+\lambda)^2}$.

The profit function of any non-compliant firm is:

$$\pi_H^d = (p_H - w)n(w) - n(w)^\sigma zwt \tag{20}$$

We solve for the level of profit π_H^d the manager of the firm paying the lowest wage earns, where the lowest wage equals the profit (π_L) a worker with low managerial ability would earn as an own-account worker, i.e., their outside option. By plugging this wage into equation (20) above we can solve for the equilibrium level of profit, which must be the same for firms offering any wage:

$$\pi_H^d = (p_H - \pi_L)n(\pi_L) - n(\pi_L)^\sigma z\pi_L t = (p_H - \pi_L)\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda\}^2} - \left[\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda\}^2}\right]^\sigma z\pi_L t \tag{21}$$

We see that equation (21) gives the level of profit of a high skill manager in terms of the exogenous parameters. Plugging this value for the level of profit into the profit function in (20) we can solve for the wage in terms of the equilibrium wage offer distribution facing any non-compliant firm:

$$w^d = \frac{p_H n(w) - \pi_H}{n(w) + n(w)^\sigma zt} = \frac{p_H\left\{\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda[1-F(w)]\}^2}\right\} - \pi_H}{\left\{\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda[1-F(w)]\}^2}\right\} + \left\{\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda[1-F(w)]\}^2}\right\}^\sigma zt} \tag{22}$$

Similarly, compliant firms will make the same profit as defaulting firms which is given in equation (21). By equalising this with the profit function of a compliant firm with a higher wage:

$$\pi_H^c = [p_H - w(1+t)]n(w) \tag{23}$$

Next, we solve for the relationship between the wage and the wage offer distribution for compliant firms by substituting the labour supply function into equation (23):

$$w^c = \frac{1}{(1+t)}\left[p_H - \frac{\pi_H}{n(w)}\right] = \frac{1}{(1+t)}\left[p_H - \frac{\pi_H}{\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda[1-F(w)]\}^2}}\right] \tag{24}$$

We note that for $n(w) > \left(\frac{1}{z}\right)^{\frac{1}{\sigma-1}}$ the firm will be compliant and for any employment level below this the firm will maximise the stream of profit by defaulting. Using the labour supply curve in

equation (17) this inequality can be written as:

$$\frac{L_L}{L_H}\frac{\delta\lambda}{\{\delta+\lambda[1-F(w)]\}^2} > \left(\frac{1}{z}\right)^{\frac{1}{1-\sigma}} \tag{25}$$

This inequality can be used to solve for the value of the wage distribution below which the firm will default:

$$F(w^*) > \frac{\delta+\lambda}{\lambda} - \left[z^{\frac{1}{\sigma-1}}\frac{L_L}{L_H}\frac{\delta}{\lambda}\right]^{0.5} \tag{26}$$

BM show some interesting comparative static results from this inequality which continue to hold in the framework with self-employment.

Proposition Five:

The share of employees in non-compliance is independent of the tax rate t, and decreasing in the expected penalty if caught z.

Proof: From equation (14), the employee wage distribution is a monotonically increasing function of the wage offer distribution. We see from inequality (26) that the share of wage offers made to employees in non-compliant firms is independent of the tax rate and decreasing in the expected penalty if caught z.

While initial intuition might tell us that an increase in the tax rate would increase the incentive to default, Proposition Five reminds us that we generally expect the tax rate to enter the expected penalty function as well as the expected cost of compliance if penalties for non-compliance are a function of the amount of tax owed. In this case, because it enters both the expected cost of compliance and the expected penalty in a linear fashion, the tax does not affect the decision on compliance/non-compliance. An important qualification to this is that the analysis in this paper is focused on an interior solution with both compliant and non-compliant firms. For example, we will show later that there is a threshold for the tax rate above which high ability managers would opt out of being an employer and revert to own-account employment. Rocha et al. (2018) present evidence from Brazil suggesting that reducing the tax rate when registration costs have been eliminated will reduce informality, but the effects are small. We can see that equation (26) is decreasing in the enforcement parameter z, indicating that an increase in enforcement decreases informality. This is consistent with empirical evidence from the literature. Brockmeyer et al. (2019) present evidence from Costa Rica that enforcement

reduces informality.

We note that $F(w)$ will be continuous since, if there was any mass point in the equilibrium wage offer distribution, a manager of a firm offering the wage at that mass point could gain a first order increase in labour supply and profits by increasing their wage infinitesimally and attracting a group of workers from the firms who are bunched at the mass point[17]. We also note that, as in BM, the lowest wage firm will attract no workers from other firms and all workers in these firms will accept any offer from any other firm. For these reasons the lowest wage firm which hopes to attract type-L workers as employees has the incentive to reduce the wage they offer to the reservation wage of a self-employed worker.

While the analysis so far has concentrated on the case where there are some compliant and non-compliant firms, we must allow for the possibility that, if the tax rate were too high, high ability managers may be indifferent between own-account work and being a manager. That is, if the tax rate is high enough, the profit of being an own-account worker will dominate that of being a worker in a compliant firm. To check this we calculate the profit of the highest wage compliant firm (where $F(w) = 1$ and so from (17) the labour supplied to the highest wage firm is $n_L \frac{L_L}{L_H} \frac{\lambda}{\delta}$). We need to ensure that own-account profit p_H is higher than the profit of the highest wage compliant firm π_H^c (which is the same as the profit of all other firms in equilibrium):

$$p_H > \pi_H^c = [p_H^c - w(1+t)]n(w) = [p_H - w(1+t)]\frac{L_L}{L_H}\frac{\lambda}{\delta} \tag{27}$$

Solving for the wage:

$$\overline{w} > \frac{p_H[\frac{L_L}{L_H}\frac{\lambda}{\delta} - 1]}{(1+t)\frac{L_L}{L_H}\frac{\lambda}{\delta}} \tag{28}$$

This implies that if the wage of the highest wage compliant firm $\overline{w}$ exceeds the right hand side of inequality (28), then the manager would earn more profit in own-account self-employment than as manager of the highest wage compliant firm. If we replace the wage in inequality (28) with the wage the lowest wage firm could ever pay (which is p_L since a firm offering a lower wage would attract no employees), the inequality becomes $\frac{L_L}{L_H} > \frac{1}{1+t}\frac{\frac{L_L}{L_H}\frac{\lambda}{\delta}-1}{\frac{L_L}{L_H}\frac{\lambda}{\delta}}$. Solving for the tax rate, we establish a sufficient condition that if the tax rate is above this threshold there is no wage at which high skilled managers could attract workers and make more profit than they would as an

[17]Showing this formally was one of the important contributions of BM and the argument is spelt out in detail there.

own-account worker:

$$t > \frac{p_H}{p_L}\left(1 - \frac{L_L}{L_H}\frac{\lambda}{\delta}\right) - 1 \tag{29}$$

Figure 1 simulates the wage offer distribution for the parameter values given in Table 1.

Figure 1 graphs the wage offer distribution. The lowest wage defaulting firm pays workers their outside option (the own account profit of $p_L = 1$). There is a vertical line at F^* where any firm offering a wage that puts them at a higher value than this in the wage offer distribution will find it more profitable to be compliant.

4.2 The case with higher capital costs in the informal sector

The per period cost of the match specific capital is $\gamma^d k$ and $\gamma^c k$ for a non-compliant and compliant firm, respectively. As we discussed earlier, a prominent explanation for informality in the literature is that informal firms may be restricted to informal capital markets where poorer enforcement implies higher risk and higher interest rates, as in the models of Amaral and Quintin (2006) and Straub (2005), for example. We capture this feature by assuming that the price of firm specific investments in new hires is higher for an informal manager than that for a formal manager: $\gamma^d > \gamma^c$[18]. In what follows we change the tax from a percentage of the wage bill to a fixed cost of formalising T which the firm pays in each period. In the interests of tractability we ignore any potential penalties from not paying tax other than that the firm faces higher capital costs. We can think of this example as a framework to analyse a firm's decision on whether to formalise, where there is a fixed cost involved in registering etc., but where registered firms have other benefits such as access to cheaper capital. To incorporate the match specific capital, we note that firms have the production function $p_H k^\eta n(w)$. The productivity parameter of the manager is multiplied by the number of workers $n(w)$ and the amount of capital per worker k that the firm chooses when a worker is hired to the power of $\eta < 1$. As in the previous example we assume output is fixed at p_J where $J \in (H, L)$ indicates ability. We assume that own account workers do not invest in job specific training and so have no reason to pay t implying that output equals their profit. We will proceed taking it as given that only high ability managers find it

[18]While we could assume interest rates are higher by assuming a higher discount rate for informal firms in equation (2), this would assume the informal firm has higher costs of borrowing and returns to saving. The assumption we make implicitly is that when the manager incurs the sunk cost of setting up the firm he/she needs some credit and this is more expensive for informal firms relative to formal firms. This is a way of capturing the possibility that the cost of accessing credit for firm start-up is higher for informal firms who rely on more expensive informal lenders.

profitable to hire workers and will outline the parameter restrictions that ensure that this is true after we have solved for the equilibrium level of profit. The value of a firm for a high ability manager who is defaulting or compliant respectively is:

$$\begin{aligned} V_H^d(w) &= \frac{\pi_H^d(w)}{\delta} = \frac{[p_H(k_H^d)^\eta - w]n(w) - \gamma^d k_H^d h_H(w)}{\delta} \\ V_H^c(w) &= \frac{\pi_H^c(w)}{\delta} = \frac{[p_H(k_H^c)^\eta - w]n(w) - \gamma^c k_H^c h_H(w) - T}{\delta} \end{aligned} \tag{30}$$

Using the fact that $n_H(w) = \frac{h_H(w)}{d(w)}$, this can be written as:

$$\begin{aligned} V_H^d(w) &= \frac{\pi_H^d(w)}{\delta} = \frac{\left[\frac{p_H k_H^{d\eta} - w}{d(w)} - \gamma^d k_H^d\right] h_H(w)}{\delta} \\ V_H^c(w) &= \frac{\pi_H^c(w)}{\delta} = \frac{\left[\frac{p_H k_H^{c\eta} - w}{d(w)} - \gamma^c k_H^c\right] h_H(w) - \gamma^c T}{\delta} \end{aligned} \tag{31}$$

The firm's optimal choice of capital implies $\frac{\eta p_H (k_H^j)^{\eta-1}}{d(w)} = \gamma^j$, so that:

$$k_H^j = \left[\frac{\eta p_H}{\gamma^j d(w)}\right]^{\frac{1}{1-\eta}} = \left[\frac{\eta p_H}{\gamma^j\{\delta + \lambda[1 - F(w)]\}}\right]^{\frac{1}{1-\eta}} \quad \text{where} \quad j \in (c, d) \tag{32}$$

We see that in equilibrium the optimal level of investment depends on the firm's choice of wage and on whether the firm is compliant. We can substitute the optimal value of capital back into the profit function for defaulting and compliant firms respectively:

$$\begin{aligned} \pi_H^d(w) &= \left[p_H(k_H^d)^\eta - w\right] n(w) - \gamma^d k_H^d h_H(w) \\ &= \left\{\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right)(\gamma^d)^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} \left\{\delta + \lambda\left[1 - F(w)\right]\right\}^{-\frac{\eta}{1-\eta}} - w\right\} \frac{L_L}{L_H} \frac{\delta\lambda}{\left\{\delta + \lambda\left[1 - F(w)\right]\right\}^2} \end{aligned} \tag{33}$$

$$\begin{aligned} \pi_H^c(w) &= \left[p_H(k_H^c)^\eta - w\right] n(w) - \gamma^c k_H^c h_H(w) \\ &= \left\{\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right)(\gamma^c)^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} \left\{\delta + \lambda\left[1 - F(w)\right]\right\}^{-\frac{\eta}{1-\eta}} - w\right\} \frac{L_L}{L_H} \frac{\delta\lambda}{\left\{\delta + \lambda\left[1 - F(w)\right]\right\}^2} - T \end{aligned} \tag{34}$$

A firm will default if:

$$\pi_H^d(w) = p_H(k_H^d)^\eta n(w) - \gamma^d k_H^d h(w) > \pi_H^c(w) = p_H(k_H^c)^\eta n(w) - \gamma^c k_H^c h(w) - T \tag{35}$$

This condition can be reorganized as follows:

$$T > \Delta\pi^G = \left[p_H(k_H^c)^\eta - p_H(k_H^d)^\eta\right] n(w) - \left[\gamma^c k_H^c - \gamma^d k_H^d\right] h(w) \tag{36}$$

This just says that the tax liability of the firm is greater than the difference in gross (that is before tax) profits between compliant and defaulting firms, so that the benefit of the lower capital costs associated with compliance are outweighed by the tax. The difference in gross profit on the right-hand side can be decomposed into the difference in revenue between compliant and defaulting firms (which is in red) and the difference in capital costs in blue (there is no difference in labour costs at a given wage). Using the solution for capital in equation (32), we see that revenue for either defaulting or compliant, i.e. $j \in (c, d)$, firms can be written as follows:

$$\begin{aligned} p_H(k_H^j)^\eta n(w) &= p_H^{\frac{1}{1-\eta}} \left(\frac{\eta}{\gamma^j}\right)^{\frac{\eta}{1-\eta}} \frac{L_L}{L_H} \frac{\delta\lambda}{\{\delta + \lambda[1 - F(w)]\}^{\frac{2-\eta}{1-\eta}}} \\ k_H^j &= \left[\frac{\eta p_H}{\gamma^j\{\delta + \lambda[1 - F(w)]\}}\right]^{\frac{1}{1-\eta}} \\ n_L(w) &= \frac{h_L(w)}{d(w)} = \frac{L_L}{L_H} \frac{\delta\lambda}{\{\delta + \lambda[1 - F(w)]\}^2} \end{aligned}$$

We also see that capital costs can be written:

$$\gamma^j k_H^j h(w) = (\gamma^j)^{-\frac{\eta}{1-\eta}} (\eta p_H)^{\frac{1}{1-\eta}} \frac{L_L}{L_H} \frac{\delta\lambda}{\{\delta + \lambda[1 - F(w)]\}^{\frac{2-\eta}{1-\eta}}}$$

Using these equations where $j \in (c, d)$, inequality (35) can be rewritten as follows where, as above, the difference in capital costs is in blue and the difference in revenue is in red:

$$\begin{aligned} T > \Delta\pi^G &= \left[p_H(k_H^c)^\eta - p_H(k_H^d)^\eta\right] n(w) - \left[\gamma^c k_H^c - \gamma^d k_H^d\right] h(w) \\ &= \left((\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right) \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) p_H^{\frac{1}{1-\eta}} \frac{L_L}{L_H} \frac{\delta\lambda}{\{\delta + \lambda[1 - F(w)]\}^{\frac{2-\eta}{1-\eta}}} \end{aligned} \tag{37}$$

Since $\eta < 1$ by assumption then $\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} > 0$. Also, since $\gamma^c < \gamma^d$, we can see that $\left((\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right) > 0$, so the right-hand side of inequality (36) is positive and is also increasing in the difference in the cost of capital between defaulting and compliant firms. We also note that a higher separation rate lowers $\Delta\pi^G$. Since the separation rate $d(w) = \delta + \lambda[1 - F(w)]$ is decreasing in the wage as we expect, this implies that other things equal a higher wage firm will have a lower separation rate and increase the profit of compliance relative to defaulting. It

follows from this that Proposition Two will continue to hold for this example[19].

The equilibrium wage distribution can be solved as before by noting that the lowest wage firm will offer a wage of p_L (that corresponds to the reservation wage of a low ability worker) and solving for the profit of this firm. In equilibrium, all firms must make this level of profit:

$$\pi_H = \left\{ \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right) \left[\gamma^d (\delta + \lambda) \right]^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} - p_L \right\} \frac{L_L}{L_H} \frac{\delta \lambda}{(\delta + \lambda)^2} > p_H \tag{38}$$

We must ensure that this level of profit is higher than high-ability workers outside option to ensure that high ability workers choose to be managers in equilibrium. Similarly, we can use this condition to check that a low productivity worker could never make more profit as a manager than as an own account worker by replacing p_H with p_L:

$$\pi_H = \left\{ \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right) \left[\gamma^d (\delta + \lambda) \right]^{-\frac{\eta}{1-\eta}} p_L^{\frac{1}{1-\eta}} - p_L \right\} \frac{L_L}{L_H} \frac{\delta \lambda}{(\delta + \lambda)^2} < p_L \tag{39}$$

This condition can be re-arranged as:

$$\left\{ \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right) \left[\frac{p_L}{\gamma^d (\delta + \lambda)} \right]^{\frac{\eta}{1-\eta}} - 1 \right\} \frac{L_L}{L_H} \frac{\delta \lambda}{(\delta + \lambda)^2} < 1 \tag{40}$$

We can check that this condition holds in equilibrium. We note that, since all firms must make equal profit, we can set this level of profit equal to that of any higher wage defaulting firm:

$$\pi_H = \left\{ \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right) \left(\gamma^d \right)^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} \left\{ \delta + \lambda [1 - F(w)] \right\}^{-\frac{\eta}{1-\eta}} - w \right\} \frac{L_L}{L_H} \frac{\delta \lambda}{\left\{ \delta + \lambda [1 - F(w)] \right\}^2} \tag{41}$$

Using equation (41), we can solve for the value of the wage in terms of the wage offer distribution for a defaulting firm:

$$w = \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}} \right) \left(\gamma^d \right)^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} \left\{ \delta + \lambda [1 - F(w)] \right\}^{-\frac{\eta}{1-\eta}} - \pi_H \frac{L_H}{L_L} \frac{\left\{ \delta + \lambda [1 - F(w)] \right\}^2}{\delta \lambda} \tag{42}$$

Similarly, since all firms must make the same profit in equilibrium we can also solve for the relationship between the wage offer distribution and the wage of a compliant firm by using the level of profit given in equation (38) in the profit function of a compliant firm provided in

[19] We note that in this simple example where all managers have the same productivity in equilibrium Proposition One does not arise.

equation (34) and solving.

$$w = \left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right)(\gamma^c)^{-\frac{\eta}{1-\eta}} p_H^{\frac{1}{1-\eta}} \{\delta + \lambda[1 - F(w)]\}^{-\frac{\eta}{1-\eta}} - (\pi_H + T)\frac{L_H}{L_L}\frac{\{\delta + \lambda[1 - F(w)]\}^2}{\delta\lambda} \tag{43}$$

Next, we need to ensure that there are some defaulting and non-compliant firms in equilibrium. We see that if inequality (37) is satisfied for the lowest wage firm, then there will be some defaulting firms, while if it is violated for the highest wage firm the highest wage firm will be compliant and there will be some defaulting and compliant firms. We set $F(w) = 0$ for the lowest wage firm in inequality (37) and $F(w) = 1$ for the highest wage firm to get the following condition for the tax rate such that it is high enough so that some firms will default, but not too high so that there will be some compliant firms:

$$\begin{aligned}\left[(\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right]\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) p_H^{\frac{1}{1-\eta}} \frac{L_L}{L_H}\frac{\delta\lambda}{(\delta+\lambda)^{\frac{2-\eta}{1-\eta}}} &< T \\ < \left[(\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right]\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) p_H^{\frac{1}{1-\eta}} \frac{L_L}{L_H}\frac{\delta\lambda}{\gamma^{\frac{2-\eta}{1-\eta}}}&\end{aligned} \tag{44}$$

Before plotting the equilibrium wage distribution we can use inequality (37) to solve for the threshold value of the wage offer distribution $F(w) = F^*$ such that for values of the wage offer distribution that are higher/lower than the threshold level, the firm prefers complying/defaulting to defaulting/complying:

$$F^* = \frac{\delta + \lambda}{\lambda} - \frac{1}{\lambda}\left[\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) p_H^{\frac{1}{1-\eta}} \frac{L_H}{L_L}\delta\lambda\right]^{\frac{2-\eta}{1-\eta}} T^{-\frac{1-\eta}{2-\eta}} \left[(\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right]^{\frac{1-\eta}{2-\eta}} \tag{45}$$

We proceed by differentiating equation (45) to calculate comparative static effects where:

$$\frac{dF^*}{dT} = \frac{1-\eta}{2-\eta}\frac{1}{\lambda}\left[\left(\eta^{\frac{\eta}{1-\eta}} - \eta^{\frac{1}{1-\eta}}\right) p_H^{\frac{1}{1-\eta}} \frac{L_H}{L_L}\delta\lambda\right]^{\frac{1-\eta}{2-\eta}} \left[(\gamma^c)^{-\frac{\eta}{1-\eta}} - (\gamma^d)^{-\frac{\eta}{1-\eta}}\right]^{\frac{1-\eta}{2-\eta}} T^{\frac{-3+2\eta}{2-\eta}} \tag{46}$$

The comparative static result in equation (46) shows that, the fraction of wage offers and workers who are in non-compliant firms will increase with the tax rate. The comparative static

result with respect to the cost of capital is given below:

$$\begin{aligned}\frac{dF^*}{d\gamma^d} &= -\frac{1-\eta}{2-\eta}\frac{1}{\lambda}\left[\left(\eta^{\frac{\eta}{1-\eta}}-\eta^{\frac{1}{1-\eta}}\right)p_H^{\frac{1}{1-\eta}}\frac{L_H}{L_L}\delta\lambda\right]^{\frac{1-\eta}{2-\eta}}T^{-\frac{1-\eta}{2-\eta}} \\ &\left[(\gamma^c)^{-\frac{\eta}{1-\eta}}-(\gamma^d)^{-\frac{\eta}{1-\eta}}\right]^{\frac{-1}{2-\eta}}\frac{\eta}{1-\eta}(\gamma^d)^{-\frac{\eta}{1-\eta}}<0\end{aligned} \tag{47}$$

As we expect, an increase in the cost of capital in the defaulting firms reduces the share of offers from and workers in defaulting firms. The empirical results from the literature are mixed in terms of efforts to increase the share of formal firms. Floridi et al. (2020) present evidence from a meta-analysis that interventions to encourage firms to become formal have had little or no impact. Benhassine et al. (2018) present evidence from Benin that while providing information on how to become formal process had no impact, a wider range of policies aimed at enhancing the benefits of becoming formal did lead to a rise in the share of formal firms, although there were no clear ex-post benefits to these firms from being formal. There are similar findings in de Mel et al. (2013) in that information on its own had little impact on the share of formal firms while other incentives did lead to an increase in firms becoming formal. The benefits of being formal were shown to be concentrated in a small number of high productivity firms. The results from de Andrade et al. (2014) also show that providing information support to firms to encourage them to become formal had little effect while increasing inspection had a modest effect. Once again suggesting that for many firms not becoming formal was a deliberate choice and that they only became formal when inducements to doing so were offered and often gained little from this. Arguably these results show the importance of accounting for heterogeneity across firm productivity as emphasised in the model we outlined earlier. When modelling the decision to formalise. This can also be seen in the comparative static results given in equations (46) and (47) where we see that managerial productivity is a key determinant of the comparative static effect. If managers have lower productivity their responsiveness to a change in the difference in capital costs between complying and defaulting firms or a change in the costs of formalisation T.

5 Conclusion

This paper provides a formal model that can make sense of the growing body of empirical evidence indicating, in contrast to the earlier literature, which tended to analyse whether formal status was an indicator of being in a secondary sector, that in fact there may be a good deal

of heterogeneity within the formal/informal sectors. An important aspect of the model is that while formality is highly correlated with high or low wage employment, the structure of wages is determined by the underlying structure of the economy, not by formality status. We outline a model which combines heterogeneity in managerial ability with search frictions in the labour market, allowing for a richer set of outcomes, where self-employment may be either desirable or an undesirable, but an unavoidable state for different groups of workers.

The model incorporates two examples. Firstly, the Burdett and Mortensen (1998) model is solved for the case where informality is chosen to avoid paying tax, but where the cost is the probability of being caught and penalised for tax evasion. The comparative static results for the model demonstrate the importance of taking equilibrium effects into account when analysing the impact of changes in the tax rate. In a model with some defaulting firms, a change in the tax rate will affect the penalties of those caught defaulting as well as the costs of complying, so that the share of the workers in defaulting firms may rise or fall. Secondly, in the example where firms pay a cost to formalise in return for lower capital costs, the comparative static results are as one would expect. Lower formalisation costs will increase the share of the formal sector, as will lower relative capital costs in the formal sector. The size of these effects depends on the productivity of the manager, which is consistent with the findings of the empirical literature which tends to find that for many low productivity firms, the costs of formalising are prohibitive. Indeed, it maybe that many of the lowest productivity firms would not be able to bear the costs of formalising and would close under its weight if forced to comply[20].

[20]For example, Badaoui and Walsh (2022) show, using World Bank informal surveys of firms across a range of countries, that a substantial share of informal firms have output per worker which is less than the statutory minimum wage in their country.

References

Albrecht, J., L. Navarro, and S. Vroman (2009). The effects of labour market policies in an economy with an informal sector. *The Economic Journal 119*, 1105–1129.

Amaral, P. and E. Quintin (2006). A competitive model of the informal sector. *Journal of Monetary Economics 53*, 1541–1553.

Audretscha, D. and J. Elston (2002). Does firm size matter? Evidence on the impact of liquidity constraints on firm investment behavior in Germany. *International Journal of Industrial Organization 20*, 1–17.

Badaoui, E. and F. Walsh (2022). Productivity, non-compliance and the minimum wage. Technical Report 102778.

Banerjee, A. (2003). Contracting constraints, credit markets, and economic development. In L. P. H. Mathias Dewatripont and S. J. Turnovsky (Eds.), *Advances in Economics and Econometrics: Theory and Applications*. Eighth World Congress, Volume III.

Banerjee, A. and E. Duflo (2007). The economic lives of the poor. *Journal of Economic Perspectives 21*(1), 141–167.

Bargain, O. and P. Kwenda (2011). Earnings structures, informal employment, and self-employment: New evidence from Brazil, Mexico, and South Africa. *Review of Income and Wealth 57*, S100–S122.

Bell, C. (1990). Interactions between institutional and informal credit agencies in rural India. *The World Bank Economic Review 4*(3), 297–327.

Benhassine, N., D. McKenzie, V. Pouliquen, and M. Santini (2018). Does inducing informal firms to formalize make sense? experimental evidence from Benin. *Journal of Public Economics 157*.

Blanchflower, D. and A. Oswald (1998). What makes an entrepreneur? *Journal of Labor Economics 16*(1), 26–60.

Blau, D. (1985). Self-employment and self-selection in developing country labor markets. *Southern Economic Journal 52*(2), 351–363.

Boeri, T. and P. Garibaldi (2006). Shadow sorting. Technical report, CEPR Discussion Paper No. 5487.

Bontemps, C., J.-M. Robin, and G. Van Den Berg (1999). An empirical equilibrium job search model with search on the job and heterogeneous workers and firms. *International Economic Review 40*(4), 1039–1074.

Bosch, M. (2006). Job creation and job destruction in the presence of informal labour markets. Technical report, CEP Discussion paper no. 761.

Bosch, M. and W. Maloney (2006). Gross worker flows in the presence of informal labor markets. the Mexican experience 1987-2002. Technical report, CEP Discussion paper no. 753.

Brockmeyer, A., S. Smith, M. Hernandez, and S. Kettle (2019). Casting a wider tax net: Experimental evidence from Costa Rica. *American Economic Journal: Economic Policy 11*(3), 55–87.

Burdett, K. and D. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review 39*(2), 257–273.

Buyinza, F. and E. Bbaale (2013). Access to credit and the effect of credit constraints on the performance of manufacturing firms in the east african region: Micro analysis. *International Journal of Economics and Finance 5*(10), 85–99.

Cabral, L. and J. Mata (2003). On the evolution of the firm size distribution: Facts and theory. *American Economic Review 93*(4), 1075–1090.

Cunningham, W. and W. Maloney (2001). Heterogeneity in the Mexican micro-enterprise sector: An application of factor and cluster analysis. *Economic Development and Cultural Change 50*(1), 131–156.

Dabla-Norris, E., M. Gradstein, and G. Inchauste (2006). What causes firms to hide output? the determinants of informality. *Journal of Development Economics 85*(1-2), 1–27.

de Andrade, G., M. Bruhn, and D. McKenzie (2014). A helping hand or the long arm of the law? experimental evidence on what governments can do to formalize firms. *The World Bank Economic Review 30*(1), 24–54.

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), 122–150.

El Badaoui, E., E. Strobl, and F. Walsh (2010). The formal sector wage premium and firm size. *Journal of Development Economics 91*(1), 37–47.

Evans, D. and B. Jovanovic (1989). An estimated model of entrepreneurial choice under liquidity constraints. *Journal of Political Economy 97*(4), 808–827.

Falco, P., A. Kerr, N. Rankin, S. Sandefur, and F. Teal (2011). The returns to formality and informality in urban Africa. *Labour Economics 18*(1), s23–s31.

Fields, G. (1975). Rural-urban migration, urban unemployment and underemployment, and job-search activity in LDC's. *Journal of Development Economics 2*(2), 165–187.

Fields, G. (2009). Segmented labor market models in developing countries. In H. Kincaid and D. Ross (Eds.), *The Oxford Handbook of the Philosophy of Economic Science*. Oxford University Press.

Fields, G. (2019). Self-employment and poverty in developing countries. *IZA World of Labor 60v2*.

Floridi, A., B. Demena, and W. N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics 67*(101925).

Fugazza, M. and J. Jacques (2004). Labor market institutions, taxation and the underground economy. *Journal of Public Economics 88*(1-2), 395–418.

Gindling, T. and D. Newhouse (2012). Self-employment in the developing world. Technical report, Policy Research Working Paper 6201.

Giné, X. (2011). Access to capital in rural thailand: An estimated model of formal vs. informal credit. *Journal of Development Economics 96*(1), 16–29.

Gollin, D. (2008). Nobody's business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics 55*, 219–233.

Gollin, D. (2014). The Lewis model: A 60-year retrospective. *Journal of Economic Perspectives 28*(3), 71–88.

Gong, X. and A. van Soest (2002). Wage differentials and mobility in the urban labour market: a panel data analysis for Mexico. *Labour Economics 9*(4), 513–529.

Günther, I. and A. Launov (2012). Informal employment in developing countries: Opportunity or last resort. *Journal of Development Economics 97*(1), 88–98.

Heckman, J. and V. Hotz (1986). An investigation of the labor market earnings of Panamanian males: Evaluating the sources of inequality. *Journal of Human Resources 21*(4), 507–542.

La Porta, R. and A. Shleifer (2008). The unofficial economy and economic development. *Brookings Papers on Economic Activity 47*(1), 275–352.

La Porta, R. and A. Shleifer (2014). Informality and development. *Journal of Economic Perspectives 28*(3), 109–126.

Lewis, W. (1954). Economic development with unlimited supplies of labor. *Manchester School 22*(2), 139–191.

Lucas, R. (1978). On the size distribution of business firms. *Bell Journal of Economics 9*, 508–523.

Maloney, W. (2004). Informality revisited. *World Development 32*(7), 1159–1178.

Marcouiller, D., V. Ruiz de Castilla, and C. Woodruff (1997). Formal measures of the informal-sector wage gap in Mexico, El Salvador, and Peru. *Economic Development and Cultural Change 45*(2), 367–392.

Mazumdar, D. (1981). The urban labor market and income distribution: a study of Malaysia. Technical report, Washington, D.C.: World Bank Group.

Meghir, C., R. Narita, and J.-M. Robin (2015). Wages and informality in developing countries. *American Economic Review 105*(4), 1509–1546.

Mortensen, D. (2003). *Wage Dispersion: Why Are Similar Workers Paid Differently?* MIT Press.

Mortensen, D. and C. Pissarides (1994). Job creation and job destruction in the theory of unemployment. *Review of Economic Studies 61*, 397–415.

Narita, R. (2020). Self-employment in developing countries: A search-equilibrium approach. *Review of Economic Dynamics 35*, 1–34.

Perry, G., W. Maloney, O. Arias, P. Fajnzylber, A. Mason, and J. Saavedra-Chanduvi (2007). *Informality: Exit and exclusion.* The World Bank, Washington, D.C.

Poschke, M. (2019). Wage employment, unemployment and self-employment across countries. Technical report, IZA DP No. 12367.

Postel-Vinay, F. and J.-M. Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica 70*(6), 2295–2350.

Pradhan, M. and A. van Soest (1995). Formal and informal sector employment in urban areas of Bolivia. *Labour Economics 2*, 275–297.

Pratap, S. and E. Quintin (2006). Are labor markets segmented in developing countries? a semiparametric approach. *European Economic Review 50*, 1817–1841.

Rauch, J. (1991). Modelling the informal sector formally. *Journal of Development Economics 35*, 33–47.

Roberts, B. (1989). The other working class: Uncommitted labor in britain, spain, and mexico. In M. L. Kohn (Ed.), *Cross-National Research in Sociology*, pp. 352–372. Newbury Park, CA: Sage Publications.

Rocha, R., G. Ulyssea, and L. Rachter (2018). Do lower taxes reduce informality? evidence from Brazil. *Journal of Development Economics 134*, 28–49.

Satchi, M. and J. Temple (2009). Labor markets and productivity in developing countries. *Review of Economic Dynamics 26*, 333–345.

Stein, J. (2002). Information production and capital allocation: Decentralized versus hierarchical firms. *Journal of Finance 57*.

Stiglitz, J. and A. Weiss (1981). Credit rationing in markets with imperfect information. *American Economic Review 71*(3), 393–410.

Straub, S. (2005). Informal sector: The credit market channel. *Journal of Development Economics 78*, 299–321.

Tannuri-Pianto, M. and D. Pianto (2002). Informal employment in Brazil - a choice at the top and segmentation at the bottom: a quantile regression approach. Technical report, Mimeo.

Tansel, A. (1999). Formal versus informal sector choice of wage earners and their wages in Turkey. Technical report, Economic Research Forum Paper 9927.

Ulyssea, G. (2010). Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics 91*, 87–99.

Table 1: Parameter Assumptions for Simulations in Figures 1 and 2

$L_H = 10$	$L_L = 100$
$z = 0.2$	$t = 0.1$
$p_L = 1$	$p_H = 2$
$\delta = 0.207$	$\lambda = 0.287$
$\sigma = 2$	$\eta = 0.5$
$\gamma^c = 0.8$	$\gamma^d = 1$
$T = 5$	

Figure 1: Simulated Wage Offer Distribution, Compliant and Non-Compliant Firms

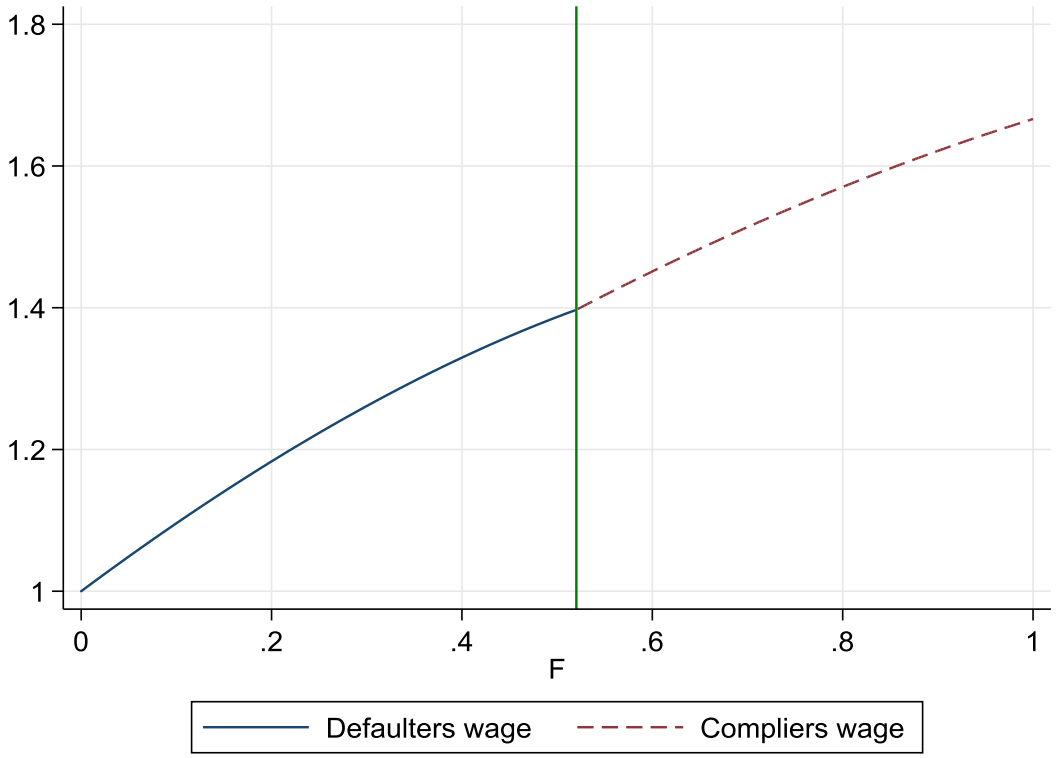
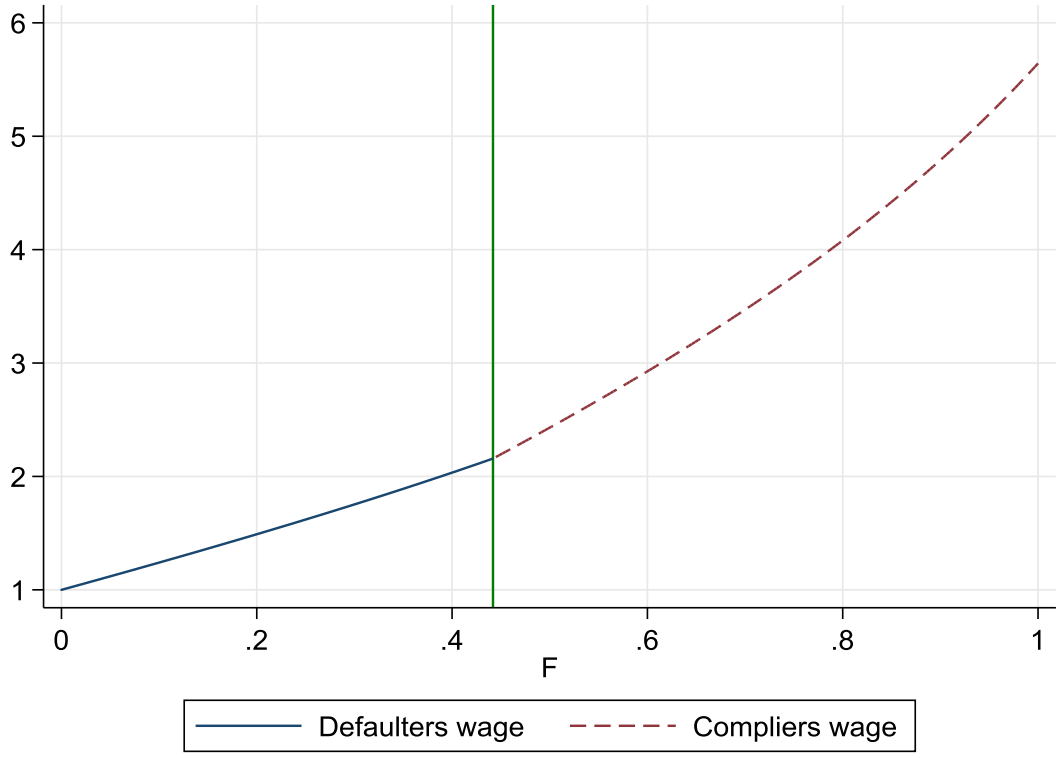


Figure 2: Simulated Wage Offer Distribution, Compliant and Non-Compliant Firms



Appendix A - Proof of Propositions

A.1 Proof of Proposition One

From (2) a firm will maximize the expected value of the firm by maximizing the stream of expected profits. In equilibrium, a type-i manager pays a wage w_i and a type-j manager a wage w_j. We define $\pi_i^{si}(p_i, w_i)$ as the flow of profit for the optimal choice of wage and compliance status of a type-i firm, where $s_i \in (c, d)$ is the compliance status of an optimizing type-i firm and $s_j \in (c, d)$ the compliance status of an optimizing type-j firm. $\pi_i^{sj}(p_i, w_j)$ is the profit a type-i firm would make if they adopted the same wage and compliance status as an optimal type-j firm. $\pi_j^{sj}(p_j, w_j)$ is the flow of profit implied by the optimal wage and compliance status of a type-j firm and $\pi_j^{si}(p_j, w_i)$ is the profit a type-j firm would make if it adopted the wage and compliance status of an optimal type-i firm. It follows that:

$$\pi_i^{si}(p_i, w_i) \geq \pi_i^{sj}(p_i, w_j) > \pi_j^{sj}(p_j, w_j) \geq \pi_j^{si}(p_j, w_i) \tag{A.1}$$

The first inequality in (A.1) holds since no firm can do no better than its optimal choice of wage and compliance status. Since we have assumed that $p_i > p_j$, the definition of profit in (1) implies the second inequality will hold. Since $\pi_j^{sj}(p_j, w_j)$ is the level of profit for an optimal choice of wage for this firm type, the third inequality will hold. We note that the inequalities in (A.1) imply that:

$$\pi_i^{si}(p_i, w_i) - \pi_j^{si}(p_j, w_i) - \pi_i^{sj}(p_i, w_j) - \pi_j^{sj}(p_j, w_j) = q[n(w_i)] - q[n(w_j)] > 0 \tag{A.2}$$

This implies that $n(w_i) \geq n(w_j)$ which implies that $w_i \geq w_j$.

A.2 Proof of Proposition Two

Assume that w_i and w_j are both wages offered in equilibrium by a firm of type-i and a firm of type-j, respectively, where $p_j \geq p_i$. From Proposition One we know that $w_j > w_i$. We assume w_i is the wage paid by the lowest wage compliant firm across all firms. Since the firm chooses to

be compliant, then the value of compliance is higher than non-compliance at the optimal wage:

$$V_i^c(w_i) > V_i^d(w_i) \tag{A.3}$$

We note that for any firm of type-g the value of the firm is determined by choosing the wage that will maximise the stream of profits to the manager. We also note that for firms i and j, respectively, the difference between profit in compliance and non-compliance at the optimal wage can be written as:

$$\begin{aligned} \Delta_i &= \pi_i^c(w_i) - \pi_i^d(w_i) = -w_i tn(w_i) + \Theta[n(w_i)]\Omega[w_i tn(w_i)] > 0 \\ \Delta_j &= \pi_j^c(w_j) - \pi_j^d(w_j) = -w_j tn(w_j) + \Theta[n(w_j)]\Omega[w_j tn(w_j)] > 0 \end{aligned} \tag{A.4}$$

We note that if firm j were non-compliant it must be that $\Delta_j < 0$. This implies that if $\Delta_j - \Delta_i \geq 0$ then firm j cannot be non-compliant. The difference in the two difference terms above can be written as:

$$\Delta_j - \Delta_i = \underbrace{\Theta[n(w_j)]\Omega[w_j tn(w_j)] - \Theta[n(w_i)]\Omega[w_i tn(w_i)]}_{\text{Difference in expected penalty}} - \underbrace{[w_j tn(w_j) - w_i tn(w_i)]}_{\text{Difference in tax bill}} \tag{A.5}$$

We note that if the expected penalty is increasing at least as fast as the tax bill as the wage (number of workers) increase then firm j must also be compliant. This establishes that any firm j which has higher productivity to the lowest wage compliant firm must also be compliant. We note that firm i is the lowest wage compliant firm by definition so that any firm with lower wage is non-compliant.

A.3 Proof of Proposition Three

The value of the lowest equilibrium wage offer is $E(\underline{w})$. Say this wage is offered by a type-j firm in equilibrium where $j \geq i$ and where i is the lowest ability type. If a type-i worker preferred self-employment to this offer then:

$$V_i^h[p_i, w, n(w)] > E(\underline{w}) \qquad \text{where} \quad h \in (c, d) \tag{A.6}$$

But since for any other worker type $k \neq i$ it must be that $p_k > p_i$, in which case it follows from the definition of profits and the value of the firm (1) and (2) that $V_k^h[p_k, w, n(w)] >$

$V_i^h[p_i, w, n(w)] > E(\underline{w})$. That is, it must be that if type-i workers would prefer self-employment to the lowest wage job, the same would be true for all workers. This means that no firm could offer such a contract and attract any workers and such an offer cannot be an equilibrium contract. It must be that in equilibrium the lowest wage contract is sufficiently attractive to attract at least the lowest ability worker or else it cannot attract any workers.

A.4 Proof of Proposition Four

As long as some firms are larger than own-account firms, the highest ability managers will prefer self-employment to any employment offer. From Proposition One, the highest ability managers (type-g) will manage the largest, highest wage firm. Since any manager can choose to be an own-account firm at any point in time, the profit of being the highest wage manager must be at least as high as the profit of being an own-account worker in equilibrium:

$$\pi_z^j(\overline{w}) > p_z q(n_0) \tag{A.7}$$

where $j \in (c, d)$. Since this is the largest firm and since $q_n(n) > 0$ and $q_{nn}(n) \leq 0$ by assumption and $(\overline{w}_t) >_0$ by assumption, the additional output generated by the last $_0$ workers in the highest wage firm is less than or equal to the output a high ability manager could produce as an own-account worker:

$$p_z \left\{q[n(\overline{w}_t)] - q[n(\overline{w}_t) -_0]\right\} \leq p_z q(_0) \tag{A.8}$$

In the case where $q_{nn}(n) \leq 0$, even if the highest wage firm paid a wage to the marginal worker equal to the value of their marginal product and this worker was from the highest ability group, from (A.5) this would be less than this worker's earnings in self-employment (if the worker is from the highest ability group g). That is, it could never be profitable for even the highest wage firm to offer the highest ability worker a wage that would be attractive enough to attract this worker from self-employment. In the case of $q_{nn}(n) = 0$ offering workers their marginal product as a wage could not be an optimal strategy for any manager since they would make no surplus and could do better in self-employment. Firms must offer wages less than the marginal product of workers to make it profitable to hire them. No firm could profitably offer the highest ability worker a wage equal to their productivity in self-employment.

Appendix B - Data

Table B.1: Share of Self-employed and Unemployed in a Range of Countries

Region/Country	(1)	(2)	Region/Country	(1)	(2)
Niger	95.06	0.46	Ghana	72.25	4.12
Central African Republic	93.16	4.04	Bhutan	72.11	2.27
Chad	92.61	1.91	Gambia	72.01	8.94
Guinea	91.89	4.14	Cote d'Ivoire	70.93	3.17
Somalia	91.67	12.79	Timor-Leste	69.12	4.42
South Sudan	91.57	12.01	Zimbabwe	68.75	5.02
Sierra Leone	90.35	4.36	Bolivia	68.25	3.46
Equatorial Guinea	88.41	7.95	Rwanda	66.29	0.99
Benin	88.10	2.32	Myanmar	64.92	0.50
Madagascar	87.84	1.67	Senegal	63.67	6.47
Eritrea	86.20	6.34	Malawi	62.10	5.56
Burundi	85.80	1.42	Bangladesh	59.27	4.22
Burkina Faso	85.69	4.62	Pakistan	56.32	3.98
Mozambique	84.28	3.19	Peru	55.46	3.03
Ethiopia	84.15	2.04	Yemen	55.03	12.9
Tanzania	83.68	1.96	Sudan	54.35	16.76
Afghanistan	82.19	10.98	Vietnam	54.30	2.04
Guinea-Bissau	81.31	2.79	Indonesia	51.75	3.62
Mali	80.38	7.24	Ecuador	51.23	3.81
Nigeria	79.87	8.53	Thailand	50.28	0.72
Congo	79.48	4.13	Honduras	50.18	5.57
Liberia	78.75	2.89	Colombia	49.57	9.96
Angola	78.51	6.93	Kenya	49.27	2.60
Congo, Rep.	78.36	9.60	Morocco	48.56	9.01
Nepal	77.34	2.85	Mongolia	48.53	5.31
Uganda	77.32	1.72	Tonga	47.52	3.01
Togo	76.43	3.60	Iran	47.22	11.14
India	75.83	5.27	Lesotho	47.06	23.86
Papua New Guinea	74.80	2.37	Cambodia	47.05	0.13
Lao PDR	74.72	0.62	China	44.66	4.60
Cameroon	74.39	3.32	Venezuela	43.63	7.16
Haiti	73.51	13.48	Paraguay	43.01	6.60
Zambia	73.45	11.91	Nicaragua	42.70	5.14
East Asia & Pacific	43.51	3.86	High income	12.16	4.80
South Asia	71.60	5.04	Low & middle income	53.97	5.50
Sub-Saharan Africa	74.99	6.28	Low income	80.33	4.61
Latin America & Caribbean	37.86	7.99	OECD members	16.34	5.39
World	46.49	5.37			

Note: We report, in column (1), the share of self-employed (% of total employment) and, in column (2), the share of unemployed (% of total labour force) in 2019. Source: World Development Indicators.