Wages and Informality in Developing Countries*

Costas Meghir, Renata Narita and Jean-Marc Robin

1st September 2012

Abstract

It is often argued that informal labor markets in developing countries promote growth by reducing the impact of regulation. On the other hand informality may reduce the amount of social protection offered to workers. We extend the wage-posting framework of Burdett and Mortensen (1998) to allow heterogeneous firms to decide whether to locate in the formal or the informal sector, as well as set wages. Workers engage in both off the job and on the job search. We estimate the model using Brazilian micro data and evaluate the labor market and welfare effects of policies towards informality.

*Costas Meghir: Yale University and IFS; Renata Narita: World Bank; Jean Marc Robin: Sciences Po and UCL. We thank Joe Altonji, Corina Mommaerts and seminar participants at Yale University for comments. Costas Meghir thanks the ESRC for funding under the Professorial Fellowship RES-051-27-0204. Robin gratefully acknowledges financial support from the Economic and Social Research Council through the ESRC Centre for Microdata Methods and Practice grant RES-589-28-0001, and from the European Research Council (ERC) grant ERC-2010-AdG-269693-WASP. We also thank the ESRC Centre for Microeconomic Analysis of Public Policy at the Institute for Fiscal Studies. Responsibility for any errors is ours.

1 Introduction

Informal labour markets are a standard feature of labour markets in developing countries. These labour markets are generally seen as operating outside the tax and regulatory framework of the country, not paying taxes or social security contributions of any sort, violating minimum wage laws and not complying with employment protection regulation. It is often argued that as a result they are the engine of growth because their existence allows firms to operate in an environment where wage and regulatory costs are lower. On the other hand, informality implies that the amount of insurance offered to workers is lower. Moreover, informal markets are also subject to regulatory costs: while formal firms pay income taxes and severance, informal firms are subject to being caught and fined by the labour authorities. An interesting policy question is to which degree stricter regulatory codes affect output, sector of employment and the distribution of wages in the formal and the informal sector.

The most traditional view associates informality with a subsistence sector in a segmented labor market market, restricted by the minimum wage and tax laws. Heterogeneous workers sort themselves out of heterogeneous sectors according to the classical representation of a competitive, segmented economy à la Roy. To date, a large empirical literature has shown evidence against the segmented market view. They usually find significant job mobility across sectors or workers reporting being better off by taking up an informal job.[1] In accordance with the data and this literature, our model allows for transitions between formal and informal sectors, subject to informational frictions and choice. Workers can be exogenously laid off or can take up a job opportunity in an alternative firm either in the same sector or in

[1] For example, Maloney (1999) shows no evidence of segmented markets for Mexico, where transitions between formal and informal sector seem to be equally probable in both directions. Barros, Sedlacek and Varandas (1990), Neri (2002) and Curi and Menezes-Filho (2006) analyze Brazil and also point to the significant mobility between sectors. Furthermore, Maloney et al (2007) shows for Colombia that informal workers are more satisfied than formal workers in terms of job flexibility. For Argentina, Pratap and Quintin (2006) findings suggest that informal workers can be as well off as similar formal workers.

the other, and all the directions of mobility are potentially observable because jobs (firms) are heterogeneous both between and within sectors. In order to account for worker heterogeneity, we segment the market across observed characteristics, such as completed education and gender. Finally, the policy environment is described by corporate and labour taxes, severance payment, unemployment insurance, a legal minimum wage and an intensity of monitoring of compliance by firms.

Our paper relates to that of Albrecht, Navarro and Vroman (2009) who use the matching framework of Mortensen-Pissarides (1994) to model the informal sector as unregulated self-employment with fixed productivity, while allowing for heterogeneity in the formal sector. Bosch (2006) uses a similar framework and adds heterogeneous productivity in the informal sector. The author assumes the two markets are subject to the same frictions and direct job flows only take place from the informal to the formal sector, with the assumption that formal workers never accept an offer from the informal one.[2]

The model in this paper presents one major theoretical innovation with respect to the existing literature. Specifically, our framework adds to the literature of equilibrium search models with heterogeneous firms and on-the-job search[3] by allowing endogenous choice of sector by firms. We thus allow firms to differ in their productivity regardless of the sector in which they operate, implying that any type of firm could act in a sector, with no ex-ante restriction on whether a sector is more productive than the other.

The model is designed for analyzing economies with substantial informal and formal sectors, found across a wide range of developing economies. Here we estimate it using data from Brazil where informality of labour is about 40 percent of

[2]Other related papers are for example Gabriel Ulyssea (2010), El-Badaoui, Strobl and Walsh (2010), Boeri and Garibaldi (2005), and Fugazza and Jacques (2003). They use a more simplified structure for dual economies than that of Albrecht et al. (2009) and Bosch (2006).

[3]See e.g. Burdett and Mortensen (1998), Van den Berg and Ridder (1998), Van den Berg (2003) and Bontemps, Robin and Van den Berg (2000).

the salaried labour force.[4] Our main source is the Brazilian Labour Force Survey, *Pesquisa Mensal de Emprego*, which provides a rotating panel of individuals sampled from the six main metropolitan regions of Brazil. Finally, the model allows us to discuss the relative merits of alternative policies towards informality.

We draw the following conclusions. First, increasing the cost of informality by 10% actually improves welfare. The resulting increased competition among firms in the formal sector is the main cause, which pushes up wages and thus welfare for individuals in all states. Moreover, firms that remain informal are more than compensated by the increase in profit margins, following the move of marginal firms to the formal sector. Second, if we go as far as abolishing informality the results are more complex. First, in all cases workers' welfare (including those unemployed) increases substantially. This is both because they obtain formal jobs that are more valuable and because in most cases formal sector wages go up due to increased competition for workers among firms. Average firm profits can either increase or decrease, depending on the specific market. The extent to which they decrease determines whether welfare will increase or not. Although the model does not predict just one direction of welfare, in most markets we consider overall welfare goes up with the abolition of informality.

In the next section, we present the model. In Section 3, we describe the data and the details of estimation of the model. In Section 4, we present and comment on the main results. In Section 5, we examine the effects of changes in the compliance costs and other policies such as changes in severance and unemployment compensation. Conclusions are in Section 6.

[4]Estimate based on recent cross sectional data (PNAD) and the entire salaried workforce.

2 The Equilibrium Search Model

We have in mind a pool of low skilled homogeneous workers that will typically engage in jobs requiring low training input. This explains why we have decided to construct a wage-posting model, instead of assuming a bargaining mechanism for wage setting. Monopsonistic models are indeed usually thought as best suited to labour markets with an abundant workforce. Productivity differences will arise in this model because of firm level heterogeneity.

There are two sectors in the economy, the formal and the informal one. The two sectors arise because of the existence of taxes and regulations governing the employment of workers. Imperfect monitoring of compliance with the legal framework creates profitable opportunities for lower productivity firms to ignore the regulations and operate in the informal sector. The policy environment is described by the corporation tax on profits, income tax, social security contributions, severance pay upon laying off a worker and unemployment insurance, which is implicitly funded by taxes. All these features can be avoided when the worker is employed informally. However, firms are monitored and if caught not complying they pay a fine. Firms have a given productivity level, maximize profits and have to decide whether to comply with the regulations or employ in the informal sector, risking a fine. So the choice of sector is endogenous, which greatly complicates the determination of the equilibrium *vis-à-vis* standard wage-posting models *à la* Burdett and Mortensen (1998).

Workers seek to maximize their expected lifetime income. The flow utility of workers depends linearly on the wage they receive plus the value of the social security contributions made by the firm on their behalf, which we include in the wage measure and are net of any taxes due. Workers also value severance pay and unemployment insurance as will be evident in the value function. The economy is subject to search frictions and workers search both when unemployed and when they are

employed. They also receive competing offers from both sectors. Subscripts with value 0 denote the unemployed, with value 1 denote the formal sector and with value 2 the informal one.

2.1 Workers

Workers maximize the expected lifetime income discounted at a rate of r. At any instant, unemployed workers receive an income stream b, taken to be constant across individuals, regardless of their history. Let $W_1(w)$ and $W_2(w)$ denote the values of a wage contract w in the formal (sector 1) and the informal sectors (sector 2), and let U be the value of unemployment.

Individuals receive job offers according to a Poisson process with arrival rate λ_{ij}, where $i = 0,1,2$ denotes the current state (unemployed, or working in the formal/informal sectors) and j denotes the source of the offer. An offer is an employment contract promising a *fixed* wage and, implicitly, specific outside options. In particular, a worker can receive offers from either sector – indeed we also allow offers from the informal sector to the formal one and some of these offers may be worth accepting – and can be laid off at sector-specific rates $\lambda_{i0}, i = 1,2$. Lastly, $F_j,\ j = 1,2$, defined on $[\underline{W}_j, \overline{W}_j]$, denotes the (equilibrium) distribution of (present values of) contracts from which workers sample their offers. These distributions are endogenous and the rest of the paper will explain how they are determined.

The wage in the formal sector represents the entire monetary compensation for the worker: thus it is after tax but *before* social security deductions, which are effectively part of their compensation as it entitles them to a pension and to health benefits. Pay also includes contributions to pensions made by the employer on behalf of the worker; in the informal sector no taxes or contributions are made so the wage is just the gross wage.

The value functions for each state, namely employed in the informal sector, em-

ployed in the formal sector and unemployed describe the optimal behavior of workers. As usual these values combine the immediate gains of being in the sector (e.g. the wage) together with the resulting option values, such as the possibility of moving to better jobs within or between sector or the impact of exogenous shocks, such as the possibility of job destruction leading to unemployment. Thus the value of working in the informal sector, is

$$rW_2(w) = w + \lambda_{20}\left[U - W_2(w)\right] + \lambda_{21}\,\mathbb{E}_{F_1}\max\{W - W_2(w), 0\} + \lambda_{22}\,\mathbb{E}_{F_2}\max\{W - W_2(w), 0\},$$

where $\mathbb{E}_{F_j}, j = 1,2$, takes expectations over a generic contract value W distributed as F_j in sector j. Later in solving for equilibrium it is useful to rewrite this expression after integrating by parts,[5]

$$rW_2(w) = w + \lambda_{20}\left[U - W_2(w)\right] + \lambda_{21}\int_{W_2(w)}^{\overline{W}_1}\overline{F}_1(x)\,dx + \lambda_{22}\int_{W_2(w)}^{\overline{W}_2}\overline{F}_2(x)\,dx, \quad (1)$$

where overlines on distribution functions denote survival functions: $\overline{F} = 1 - F$. Thus the flow utility in the informal sector is the wage rate (w) plus the loss that the individual may incur if laid off, which happens at rate λ_{20}, as well as the "capital gain" of obtaining a better offer either from the formal or the informal sector with rates λ_{21} and λ_{22} respectively.

A similar expression can be derived for the value of working in the formal sector. The key difference here will be in the definition of the wage, which we discussed

[5]We make use of the following property. For any random variable X with distribution (CDF) F on $[\underline{x}, \overline{x}]$, and for all $u \in \mathbb{R}$,

$$\mathbb{E}_F \max\{X - u, 0\} = \int_{\underline{x}}^{\overline{x}} \max\{x - u, 0\}\,dF(x) = \int_u^{\overline{x}} \overline{F}(x)\,dx.$$

before and in the expression for the loss incurred when moving to unemployment. We write the value of employment in the formal sector (using the second expression derived above) as

$$rW_1(w) = w + \lambda_{10}\left[U + UI + s \cdot w - W_1(w)\right] + \lambda_{11}\int_{W_1(w)}^{\overline{W}_1} \overline{F}_1(x)\,dx + \lambda_{12}\int_{W_1(w)}^{\overline{W}_2} \overline{F}_2(x)\,dx. \quad (2)$$

The cost of becoming unemployed is mitigated by two factors. The first is unemployment insurance (UI) which we assume for simplicity[6] is paid upfront. The second term is severance pay $s \cdot w$, s being the compensation rate in the case of termination of employment. As we show below, we determine the level of UI endogenously based on the tax rate used to fund it and on the overall number of unemployed. Both UI and severance pay increase the value of employment in the formal sector – and in the informal sector since a transition between the two is possible – and both affect the equilibrium distribution of wages. The only difference of UI from severance pay is that the firm directly pays the latter, whereas UI is funded by general taxation. This distinction will be of importance when we define the firm's problem. Finally, since there are no shocks to productivity, jobs are only closed down because of exogenous job destruction, which may differ depending on the sector, λ_{10} and λ_{20}.

To write the value of unemployment note that in equilibrium, firms will only offer acceptable wages so that the value at the minimum offered wage, $\underline{W}_1$ and $\underline{W}_2$, are greater than U, otherwise no production would take place. So the equilibrium value of unemployment is such that

$$rU = b + \lambda_{01}(\mu_1 - U) + \lambda_{02}(\mu_2 - U), \quad (3)$$

[6]Specifically it avoids making the duration of unemployment a state variable if UI is time limited for example.

where $\mu_1 = \int_{\underline{W}_1}^{\overline{W}_1} x dF_1(x), \mu_2 = \int_{\underline{W}_2}^{\overline{W}_2} x dF_2(x)$ denote the mean contract values offered in the formal and the informal sector respectively, and b is the flow-value of leisure..

Contract values reflect the benefits, opportunities and costs of working in each sector. They are increasing functions of wages, yet the wage rate alone is not sufficient for ranking jobs across sectors, because each sector comes with different future opportunities. It is thus possible that a move across sectors is accompanied by a pay-cut. However in this model mobility within the sector can only take place when accompanied by a wage rise, which has to imply a move to a higher productivity firm. This is because there is no other source of heterogeneity (such as productivity shocks) and because firms do not respond to outside offers.

2.2 Steady-State Worker Flows

The value functions discussed above describe the optimal choices of workers and are conditional on the wage offer distributions in the formal and informal sectors. These are equilibrium objects. To derive them we need to define the steady state flow of workers between the three states (unemployment, formal and informal employment) as well as the behavior of the firms. In steady state, the stocks of workers and firms in each sector and in each part of the contract value distribution remains stable. We now define these flows and use them to solve for the steady state stocks and for the relationship between the equilibrium contract offer distribution and accepted offers.

Define the fraction of the labour force in each sector to be $m_i,\ i = 1,2$, and the unemployment rate to be $u = 1 - m_1 - m_2$. Let $G_1(W)$ and $G_2(W)$ be the distributions of accepted contract values in the formal and informal sectors, respectively: they denote the proportion of the stock of individuals with a contract value lower than or equal to W, respectively. First we define flows in and out of the formal sector along

the cumulative distribution of accepted contracts. Thus, for any $W \in [\underline{W}_1, \overline{W}_1]$,

$$\left[\lambda_{10} + \lambda_{11}\overline{F}_1(W)\right] m_1 G_1(W) + \lambda_{12} m_1 \int_{\underline{W}_1}^{W} \overline{F}_2(x)\, dG_1(x)$$
$$= \lambda_{01} u F_1(W) + \lambda_{21} m_2 \int_{\underline{W}_2}^{W} [F_1(W) - F_1(x)]\, dG_2(x). \quad (4)$$

The mass of workers in the formal sector at or below contract value $G_1(W)$ is $m_1 G_1(W)$. Some of these are destroyed because of exogenous layoffs (λ_{10}), receipt of offers valued more than W from other formal firms, and receipt of acceptable offers from the informal sector.[7] On the right hand side is the balancing job creation. Jobs are created with contract values below W in the formal sector, in when the unemployed accept offers less than W or workers in the informal sector receive and accept offers whose value is lower than W.[8]

Similarly we can also define the flow equation for the informal sector. For $W \in [\underline{W}_2, \overline{W}_2]$,

$$\left[\lambda_{20} + \lambda_{22}\overline{F}_2(W)\right] m_2 G_2(W) + \lambda_{21} m_2 \int_{\underline{W}_2}^{W} \overline{F}_1(x)\, dG_2(x)$$
$$= \lambda_{02} u F_2(W) + \lambda_{12} m_1 \int_{\underline{W}_1}^{W} [F_2(W) - F_2(x)]\, dG_1(x). \quad (5)$$

In Appendix A we show how to (uniquely) solve equations (4) and (5) for the distributions of accepted contracts G_1 and G_2 given the distribution of offered contracts F_1 and F_2. There exists an equilibrium relationship between the distribution of

[7] This is reflected in the integral since departures from all parts of $G_1(W)$ need to be recorded and not only the ones leading to higher contract values than W.

[8] Note that if needed G_1 is extended outside its support by making it nil to the left of $\underline{W}_1$ and equal to 1 to the right of $\overline{W}_1$.

accepted (G) and offered (F) contract values:

$$m_1 G_1(W) = \frac{\lambda_{01} F_1(W) - \Phi(W)}{d_1(W)} u, \tag{6}$$

$$m_2 G_2(W) = \frac{\lambda_{02} F_2(W) + \Phi(W)}{d_2(W)} u, \tag{7}$$

where $\Phi(W) \equiv \Phi[F_1, F_2](W)$ is an operator on F_1 and F_2 that is derived in Appendix A, and that is nil for all $W \leq \max\{\underline{W}_1, \underline{W}_2\}$, and where, in the denominator,

$$d_1(W) = \lambda_{10} + \lambda_{11}\overline{F}_1(W) + \lambda_{12}\overline{F}_2(W), \tag{8}$$

$$d_2(W) = \lambda_{20} + \lambda_{21}\overline{F}_1(W) + \lambda_{22}\overline{F}_2(W), \tag{9}$$

are the total job destruction rates in sectors 1 and 2.

Straightforwardly, we can also derive expressions for the proportion of workers in each sector and in unemployment, by setting W (in equations (6) and (7) equal to its largest value and making use of the fact that $m_1 + m_2 + u = 1$:

$$\frac{m_1}{u} = \frac{\lambda_{01} - \Phi(\overline{W}_1)}{\lambda_{10} + \lambda_{12}\overline{F}_2(\overline{W}_1)}, \tag{10}$$

$$\frac{m_2}{u} = \frac{\lambda_{02} + \Phi(\overline{W}_2)}{\lambda_{20} + \lambda_{21}\overline{F}_1(\overline{W}_2)}, \tag{11}$$

$$\frac{1}{u} = 1 + \frac{m_1}{u} + \frac{m_2}{u}. \tag{12}$$

Hence, knowledge of the distribution of wage offers by the formal sector, F_1, and the informal sector F_2, allows us to infer the steady state stocks of employment (m_1 and m_2) and unemployment (u) as well as the equilibrium distribution of accepted wages G_1 and G_2 that are observable. This is not a full characterization of equilibrium; we now need to show how the offer distributions F_1 and F_2 and the decision to post offers in one or the other sector are determined. This depends on firm behavior

to which we now turn.

2.3 Firms

Firms maximize profits by choosing in which sector to operate and the wage they will post, which determines the size of their labour force, given their specific productivity p. In the formal sector there are a number of costs associated with hiring a worker at a wage rate w. These include payroll taxes (τ), corporate taxes on profits (t) and severance payments ($s \cdot w$) to workers who are laid off. Finally, these firms may be subject to minimum wage laws $w_{\min}$, which imply that firms cannot necessarily adjust pay to offset the effects of severance pay (Lazear, 1990). Informal labour markets are monitored randomly by the government authorities whose role is to enforce tax and labour laws. When caught a firm has to pay a fine depending on its size. We denote as $C(\ell_2)$ the expected cost of informality, assumed convex in firm size ℓ_2. This function will have to be estimated from the data, based on firm behavior.

The strategy of the firm is to choose a contract value (or wage) to offer any worker it contacts. The strategy will determine the attractiveness and hence the size of the firm and hence in equilibrium there is a tradeoff between low labour costs and size. There are no adjustment costs and, conditional on the wage they pay workers, no dynamics in the decision problem of the firms. They just choose a wage and thus implicitly a contract value W to maximize profit flows.

Specifically, firms will offer optimal contracts $K_1(p)$ and $K_2(p)$ that solve the following profit maximization problems given productivity p:

$$\pi_1(p) = \max_{W \geq \max\{U, W_1(w_{\min})\}} (1-t)\left[p - (1+\tau+\lambda_{10}s)w_1(W)\right]\ell_1(W), \tag{13}$$

$$\pi_2(p) = \max_{W \geq \max\{U, W_2(w_{\min})\}} \left[p - w_2(W)\right]\ell_2(W) - C(\ell_2(W)), \tag{14}$$

where $\ell_1(W)$ and $\ell_2(W)$ are the size of informal and formal firms respectively, offering a wage contract worth W, and $w_j(W)$ denotes the wage to be paid to a worker in sector j corresponding to a contract value W.

The wage functions $w_1(W)$ and $w_2(W)$ are the wages defined by $W_1(w) = W$ and $W_2(w) = W$, from equations (2) and (1) respectively. That is,

$$(1+\lambda_{10}s)w_1(W) = (r+\lambda_{10})W - \lambda_{10}(U+UI) - \lambda_{11}\int_W^{\overline{W}_1} \overline{F}_1(x)\,dx - \lambda_{12}\int_W^{\overline{W}_2} \overline{F}_2(x)\,dx, \tag{15}$$

and

$$w_2(W) = (r+\lambda_{20})W - \lambda_{20}U - \lambda_{21}\int_W^{\overline{W}_1} \overline{F}_1(x)\,dx - \lambda_{22}\int_W^{\overline{W}_2} \overline{F}_2(x)\,dx. \tag{16}$$

In steady-state, the flow of workers leaving the workforce of any firm ($d_1(W)\ell_1(W)$ and $d_2(W)\ell_2(W)$ for the two sectors respectively) should be equal to the inflow of new hires. Hence,

$$\ell_1(W) = \frac{1}{n_1}\frac{h_1(W)}{d_1(W)}, \tag{17}$$

$$\ell_2(W) = \frac{1}{n_2}\frac{h_2(W)}{d_2(W)}, \tag{18}$$

where n_1 and n_2 are the (endogenous) proportions of firms in the formal and informal sector respectively, $h_1(W)$ and $h_2(W)$ denote the share of contacts between firms and workers willing to accept a job paid W, i.e.

$$h_1(W) = \lambda_{01}u + \lambda_{11}m_1G_1(W) + \lambda_{21}m_2G_2(W), \tag{19}$$

$$h_2(W) = \lambda_{02}u + \lambda_{12}m_1G_1(W) + \lambda_{22}m_2G_2(W), \tag{20}$$

and $d_1(W)$ and $d_2(W)$ are the total job destruction rates (equation (8) and (9))).

2.4 Equilibrium Productivity Distributions

A key element of our model is that firms decide whether to post vacancies in the formal or the informal sector as well as what wage to post. In equilibrium all strategies (given productivity) will yield equal profits, a property we now use to determine how firms locate. Because of the various costs of employing workers in the formal sector, we can expect firms with lower productivity to locate in the informal sector, at least in the presence of minimum wages, if expected fines for informality are not too high. However, there may be a range of productivities over which, in equilibrium, firms are indifferent between the two sectors; indeed this turns out to be the case. This is a particularly important feature of the model with key implications for the welfare effects of policies towards informality. Of course, the fact that firms of both types coexist over a productivity range does not mean they will have the same size or pay the same rates; quite the contrary and we will discuss this later.

We assume that there exists a number of potential entrants, normalized to one, with a distribution of productivity $\Gamma_0(p)$ on $[\underline{p}, \overline{p}]$. In equilibrium we will obtain a measure of productivities in each sector. We denote the equilibrium measure of productivity in each sector by $\Gamma_j(p)$ $(j = 1, 2)$. At the equilibrium, each firm maximizes profit flows given the equilibrium contract distributions. Hence,

$$\Gamma_1(p) = n_1 F_1(K_1(p)), \tag{21}$$

$$\Gamma_2(p) = n_2 F_2(K_2(p)). \tag{22}$$

Denote the support of the measure for informal firms to be $[\underline{p}_2, \overline{p}_2]$ and for formal firms $[\underline{p}_1, \overline{p}]$, where it is possible that to have overlap in the supports, i.e. $\overline{p}_2 > \underline{p}_1$. As discussed above, we expect that the equilibrium is such that the initial interval of productivity will be occupied by informal firms only, i.e. $\underline{p}_2 < \underline{p}_1$, and wage offers may be below the formal minimum wage. For $\underline{p}_1 \leq p \leq \overline{p}_2$, firms operate in

both sectors. We also allow for the possibility that there is a range of productivities ($p > \overline{p}_2$) where firms operate only in the formal sector. Given this, we shall consider equilibria displaying the following regimes.

1. **Inactivity**: For $\underline{p} \leq p < \underline{p}_2$, $\pi_1(p) < 0$, $\pi_2(p) < 0$, and $\Gamma_1(p) = \Gamma_2(p) = 0$. This is important to consider when discussing counterfactual policy experiments

2. **Informal sector only**: For $\underline{p}_2 \leq p \leq \underline{p}_1$, $\pi_1(p) < \pi_2(p)$, $\Gamma_1(p) = 0$, and $\Gamma_2(p) = \Gamma_0(p) - \Gamma_0(\underline{p}_2)$. It is possible that this interval is just zero, meaning that the first relevant interval is the next one. The existence of this interval depends on the relative importance of formal labor market regulation and the costs of informality.

3. **Overlapping region**: In this region formal and informal firms of identical productivity coexist and make the same profits: For $\underline{p}_1 \leq p \leq \overline{p}_2$, $\pi_1(p) = \pi_2(p)$, and
$$\Gamma_1(p) + \Gamma_2(p) = \Gamma_0(p) - \Gamma_0(\underline{p}_2).$$

4. **Formal sector only**: For all $p \geq \overline{p}_2$, $\pi_1(p) > \pi_2(p)$, $\Gamma_2(p) = \Gamma_2(\overline{p}_2)$, and
$$\Gamma_1(p) = \Gamma_0(p) - \Gamma_2(\overline{p}_2) - \Gamma_0(\underline{p}_2).$$

If there is a range of productivities where only formal firms operate, this will be in the higher range. Implicit in this assertion is that informality profits are increasing slower than formal profits, possibly because rapidly increasing costs of informality. For example, if the probability of detection as well as the fines increase fast enough with firm size, this will lead to convex costs of informality, making participation in that sector unprofitable. However, we cannot exclude the theoretical possibility that informal firms operate at all levels of productivity.

The computation of the equilibrium is not trivial. The way we calculate equilibrium distributions F_j and sectors sizes n_j is described in Appendix B. Note that, intuitively, what happens in the tail left to threshold $\underline{p}_2$ is irrelevant. Hence, many distributions Γ_0 are compatible with the same equilibrium. For any distribution Γ_0 that delivers an equilibrium (F_j, Γ_j, n_j), then distribution $\widetilde{\Gamma}_0(p) = \frac{\Gamma_0(p)-\Gamma_0(\underline{p}_2)}{1-\Gamma_0(\underline{p}_2)}$ delivers an equilibrium with the same distributions F_j, Γ_j and with sector sizes $\widetilde{n}_j = \frac{n_j}{n_1+n_2}$. This remark will take full sense when we discuss the identification and estimation of the productivity distribution.

The nature of this equilibrium has interesting implications because it can explain two seemingly contradictory assertions: first, we would expect compensating differentials to increase wages of the workers taking informal jobs: in the overlapping region the informal firms may have to offer higher wages than equivalent (in productivity) formal firms to make up for the lack of UI and severance pay and to account for the different labor market opportunities. However, there are more formal jobs at higher levels of productivity than at lower ones. This will imply that on average formal workers will be paid more than informal ones due to a composition effect. Hence the model can explain what is observed in the data and at the same time imply compensating differentials as we would expect.

3 Data

3.1 The Brazilian labour force survey (PME)

Our main source of data consists of a panel of individuals of working age, sampled by the labour force survey of Brazil, *Pesquisa Mensal de Emprego* (PME). PME was designed and conducted by the National Statistics Bureau to follow individuals of the six main metropolitan regions of Brazil. Each individual is interviewed during four consecutive months, then for another four consecutive months one year after

their entry into the sample. The sample period starts in January 2002 and goes until December 2007.[9]

For the purpose of this paper, we select workers from age 23, where the chance of returning to full time education is very low, up to age 65 who are found to be either unemployed[10] or working as an employee (registered or unregistered). Our definition of formal workers in this paper is thus whether the worker's current job is registered with the Ministry of Labour.[11] In Brazil, there is a federal minimum wage, which should be the minimum paid to all formal employees. The average legal minimum wage over the sample period is of 300 reais per month.[12] Workers under a formal contract found to earn less than the minimum wage were removed from the sample (8% of formal workers). We believe this is due to reporting error and we similarly discard the 5% lowest wages out of the informal workers sample, thus excluding mostly the zero-wage earners and some part-time jobs. We also trim the very top wages (0.01% highest of the sample).

Table 1 shows the proportions of workers unemployed, formal salaried and informal salaried, by year. The cross-sectional sample contains about 66% of formal salaried workers, 20% of informal salaried and 14% of unemployed. Over the period 2002-2007, we observe a large increase in the proportion of formal wage workers. In particular, substantial changes have taken place more recently with the proportion of formal workers increasing from 64% in 2004 to 68% in 2007. Over the same period, we observe a relatively large drop in the proportion unemployed.

[9]Due to methodological changes in the PME data with effect from 2002, we opted to use only PME from year 2002. The first reason is that we solve for the steady-state, which is an assumption hard to defend over a long period of time. The second reason is that PME from year 2002 contains retrospective information about duration of the actual employment, which we need to identify job-to-job transitions.

[10]We take out unemployed whose last job was not as an employee. By doing so, we exclude mostly unemployed who once was self-employed or inactive, e.g. individuals whose behavior deviate from the predictions of our model.

[11]The job is registered if the worker reports having a worker's card, which means that the workers is protected by the Employment laws.

[12]All wages are in reais of June of 2008.

TABLE 1
Working Status, by year

	2002	2003	2004	2005	2006	2007	Total
Unemployed	15.1	15.9	14.9	13.0	13.1	12.0	13.9
Formal salaried	64.7	63.6	63.9	65.7	66.5	68.4	65.6
Informal salaried	20.2	20.5	21.2	21.2	20.4	19.6	20.5

Note: Brazilian Labor Force Survey 2002-2007, individuals aged 23-65. The values are percentages of individuals according to their working status at the first interview.

Now, looking at our measure of informality (proportion of informal employees in the population 23 to 65 years old), we see that a significant fraction of the workforce is informal in the six largest metropolitan regions of Brazil, an average of 21% of the active workforce. As Table 1 shows, informality increased in our data until 2004 following the same trend observed since the 80s in the country. Thereafter informality decreased coinciding with an improvement in the business cycle. Our model does not distinguish across periods. However, one could estimate over different sub-periods to obtain a structural interpretation of the changes over time.

We follow individuals for up to four months or until their first move (if that is sooner). This can be job-to-job, unemployment-to-job or job-to-unemployment, where the job can be in the formal or in the informal sector.[13] At the date of the first interview, we observe the worker's employment status, the duration of the spell (time elapsed) and the wage earned. From the subsequent three months, we construct the censoring indicator (equal to one if the individual or data is missing in all three following months), the remaining time in the status and the transition indicators. We identify job-to-job transitions using the survey question on job duration.[14] For example, we classify a worker as a non-mover in the third month of the interview if she/he does not change status (e.g. remains formal) and declares that the current

[13]We do not use the entire sixteen-months window of PME due to attrition problems.
[14]This question is only available in PME after year 2002.

spell has lasted more than three months, i.e. more than the period that passed since the last interview.

Table 2 presents information on the transitions based on all sample and by region. The average exit rate from unemployment towards the formal sector is about 10% and towards the informal one 15% implying an overall duration of unemployment of 11 months. Exit from unemployment to an informal sector job is more frequent and counter-cyclical judging from the exit rates over the downturn years of 2003 and 2004. Exit to the formal sector is trending up.

Job to job mobility is much higher among informal workers than formal ones, both within the informal sector and from informal to formal. Relatively to all transitions which occur by sector, the transitions from the formal to the informal sector are quite high compared to the transitions from the informal to the formal sector. Finally, the transitions towards the formal sector have increased recently, as reflected in the decrease in the rate of informality.

When we break these down by region, Recife and Salvador which are less developed have a higher unemployment rate (18%) than the better off regions of Sao Paulo, Rio de Janeiro, Belo Horizonte and Porto Alegre (12%).[15] However, the level of development does not have an obvious relationship either to the degree of informality or to the turnover rates.

The way the model is set up, workers are homogeneous.[16] We thus focus on low education workers and estimate the model separately by sex. This implicitly assumes that the labour markets are segmented for these groups and they do not compete directly. We define low education to mean those with eight or less years

[15]Over the period of analysis (2002-2007), the average GDP per capita in 2008 prices for the Recife and Salvador regions were respectively 3.6 and 3.9 thousand dollars, whereas for Sao Paulo, Rio de Janeiro, Belo Horizonte and Porto Alegre the figures were about twice as much or more: 11.2, 9.8, 6.2 and 8.5 thousand dollars, respectively.

[16]Ridder and Van den Berg (2003) assume segmented markets with workers with same ability within market, but different across markets. Bontemps, Robin and Van den Berg (1999) include worker heterogeneity within market through differences in the value of leisure. Shephard (2009) uses this to consider the incidence of tax credits in a model with search frictions.

TABLE 2
Description of Data, all sample and by region

	All sample	Recife	Salvador	Belo Horizonte	Rio de Janeiro	Sao Paulo	Porto Alegre
Number of Individuals	441,249	61,822	56,873	83,278	64,544	107,592	67,140
Unemployed	58,004	10,338	10,687	8,959	7,566	13,875	6,579
Formal	290,243	36,238	35,156	57,367	43,500	70,009	47,973
Informal	93,002	15,246	11,030	16,952	13,478	23,708	12,588
Informality Rate (%)	24.3	29.6	23.9	22.8	23.7	25.3	20.8
Censored Observations (%)	24.4	33.8	21.6	25.3	17.4	22.6	26.6
Unemployed	34.5	45.8	28.7	39.9	24.2	31.0	38.3
Formal	20.9	28.7	18.7	21.1	15.1	19.7	23.2
Informal	29.0	37.8	23.6	31.7	20.7	26.5	33.3
Monthly transitions (% of workers by initial status)							
Unemployed-Formal	9.75	9.28	5.04	15.75	6.07	8.72	18.95
Unemployed-Informal	15.34	20.34	6.34	22.36	8.48	17.63	20.33
Formal-Formal	2.15	2.06	2.15	2.07	2.18	1.72	2.93
Formal-Unemployed	2.01	2.63	1.74	2.33	1.06	2.02	2.33
Formal-Informal	0.33	0.48	0.14	0.50	0.12	0.32	0.40
Informal-Informal	5.66	5.97	5.14	6.93	4.77	5.31	5.98
Informal-Unemployed	6.55	9.94	4.76	8.08	2.58	6.79	6.94
Informal-Formal	1.12	1.16	0.61	1.77	0.67	0.84	1.86
Mean Duration (in months)							
Unemployed	11.1	12.7	13.4	7.1	13.6	10.8	8.7
(std.dev)	12.9	14.7	14.6	9.1	13.3	11.9	10.4
Formal	70.0	71.9	70.8	64.8	76.9	70.4	67.7
(std.dev)	75.8	76.7	78.0	71.9	81.9	73.2	75.3
Informal	44.8	44.1	44.2	41.5	52.3	42.7	46.2
(std.dev)	65.3	64.2	65.1	62.6	72.3	62.0	67.8

Note: Brazilian Labor Force Survey 2002-2007, individuals aged 23-65. Transitions are the first move of individuals within four months, starting from the individuals' first interview.

of schooling. We also estimate the model separately for two regions with clearly distinct labour markets, namely Sao Paulo and Salvador. The former is a well developed low unemployment economy, while the latter is characterized by very high levels of unemployment. Separating these regions is important, because both the job destruction rates and the arrival rates are likely to be very different. Our empirical work treats these as independent local labor markets.

Table 3 displays the composition of workers at the date of the first interview by region and sex, the informality rate and turnover information. Informality is 3-4 percentage points higher among females, regardless of the region. Transitions out of unemployment in Salvador are much lower than in Sao Paulo, but within Salvador these transitions are much higher among males than females. Transitions out of formal jobs are similar for males and females in Sao Paulo, but again the turnover is larger among males than females in Salvador. On the contrary, the exit rate from informal sector jobs to formal ones is 2.6 times larger for males than for females in Sao Paulo and more similar across males and females in Salvador.

In Table 4 we show summary statistics of wages by region and sex and formal versus informal sector. On average, within each region and sector, males are paid more than females. Formal (informal) workers and those located in Sao Paulo (Salvador) earn more (less). The amount of wage dispersion (measured by the standard deviation of log wages) is larger for males than for females in both regions. The standard deviation of wages in the informal sector is larger than in the formal sector across all groups and more pronouncedly in Sao Paulo.

3.2 Specification and Estimation

The estimation problem is much more complicated than for the standard Burdett-Mortensen model, such as discussed in Bontemps *et al.* (2000). Because of the endogenous choice of sector activity, the market equilibrium sets two distributions of

TABLE 3
Description of Data, by region and sex

	Sao Paulo		Salvador	
	Males	Females	Males	Females
Number of Individuals	31,006	14,195	13,804	5,637
Unemployed	3,472	3,127	2,265	2,070
Formal	19,369	7,324	8,033	2,366
Informal	8,165	3,744	3,506	1,201
Informality Rate (%)	29.7	33.8	30.4	33.7
Censored Observations (%)	22.7	28.2	21.8	27.1
Unemployed	31.0	40.3	29.1	33.1
Formal	19.3	22.5	18.7	20.9
Informal	27.4	29.4	24.3	29.1
Transitions (% of workers by initial status)				
Unemployed-Formal	8.85	4.28	4.98	1.73
Unemployed-Informal	25.71	11.09	11.20	3.10
Formal-Formal	1.61	1.25	2.59	2.08
Formal-Unemployed	2.03	2.04	2.01	1.28
Formal-Informal	0.39	0.37	0.29	0.11
Informal-Informal	6.49	6.17	5.92	4.47
Informal-Unemployed	8.18	6.02	5.96	4.47
Informal-Formal	1.10	0.42	0.53	0.47
Mean Duration (in months)				
Unemployed	11.0	11.2	12.7	14.5
(std.dev)	12.8	12.7	14.5	15.8
Formal	74.2	64.6	69.5	76.3
(std.dev)	76.7	66.2	79.0	80.2
Informal	43.0	39.0	46.7	45.1
(std.dev)	64.8	61.8	70.0	66.9

Note: Brazilian Labor Force Survey 2002-2007, low education individuals aged 23-65. Transitions are the first move of individuals within four months, starting from the individuals' first interview.

TABLE 4
Description of log wages, by region, sex and whether a formal or an informal worker

	Sao Paulo		Salvador	
	Males	Females	Males	Females
Formal Sector Wages				
Mean	6.67	6.38	6.36	6.15
Std. Dev.	0.42	0.34	0.39	0.31
Obs.	18,631	6,688	5,897	1,214
Informal Sector Wages				
Mean	6.35	6.09	5.93	5.76
Std. Dev.	0.51	0.45	0.43	0.32
Obs.	7,669	3,397	2,945	926

Note: Brazilian Labor Force Survey 2002-2007, low education individuals aged 23-65.

labour contracts *and* two distributions for the productivities of firms operating in the formal and informal sectors. A new estimation strategy had to be devised such that the distributions of contract values have first to be guessed so as to verify equilibrium conditions given observables. Then, the distributions of firm productivities can be identified through profit-maximizing restrictions.

3.2.1 Offer Distributions

In Section 2.2, we have derived the way the offer distributions $F_j(W), j = 1,2$, are related to the accepted contract distributions $G_j(W), j = 1,2$. Adjusting these for the fact that they are defined in the contract space rather than in the wage space, the latter are observed and hence we can then estimate non-parametrically the offer distributions. However, we simplify the estimation problem by specifying a parametric

distribution as approximation, namely a non standard beta distribution:

$$F_j(x) \quad = \quad \text{betacdf}\left(\frac{x-\underline{W}_j}{\overline{W}_j-\underline{W}_j};\alpha_j,\beta_j\right), \underline{W}_j \leq x \leq \overline{W}_j,$$

where $\text{betacdf}(\cdot;\alpha,\beta)$ is the CDF of a beta distribution with parameters α and β.

Let F_1 and F_2 be two candidate offer distributions, defined on the spaces of contract present values. Let G_1^* and G_2^* denote the observable distributions of *wages* in both sectors. By construction, $G_1^*(w) = G_1(W_1(w))$, where $W_1(w)$ is the value of wage contract w derived in equation (2). A similar restriction holds for the informal sector. Given F_1 and F_2 we can use equations (6) and (7) to calculate G_1 and G_2. The estimation algorithm first aims at finding the couple of offer distributions (F_1,F_2) that best matches (G_1^*,G_2^*) with $(G_1 \circ W_1, G_2 \circ W_2)$. An important practical reason why a parametric specification is useful is that, in order to calculate the function Φ of equations (6) and (7), and other transition rates (see below), we need to calculate offer densities $f_1 = F_1'$ and $f_2 = F_2'$. Assuming a parametric specification guarantees the smoothness of both the distribution function and its derivative.

To estimate the parameters we use the method of moments. We match the distribution of wages for each sector and the transition rates implied by the model to those observed in the data. Given the above specification, we need to estimate the six arrival rates and the two job destruction rates all denoted by $\boldsymbol{\lambda} = (\lambda_{ij})_{i,j=0,1,2}$ and six further parameters $\boldsymbol{\theta} = (\underline{W}_1, \overline{W}_1, \underline{W}_2, \overline{W}_2, \alpha_1, \beta_1, \alpha_2, \beta_2)$ characterizing the offer distribution. Our algorithm estimates $\boldsymbol{\theta}$ given the $\boldsymbol{\lambda}$. We then update the latter. Although we could iterate on all parameters at the same time, this turned out to be a very quick procedure in practice.

Given the $\boldsymbol{\lambda}$ we estimate $\boldsymbol{\theta}$ by minimizing the quadratic distance

$$Q_1(\boldsymbol{\theta}|\boldsymbol{\lambda}) = \sum_{j=1}^{2}\sum_{k=0}^{M}\left(\widehat{G}_j^*(w_{jk}) - G_j(W_{jk})\right)^2, \tag{23}$$

where $W_{jk}, k = 0, ..., M$, defines a grid on the space of contract values, separately for both sectors $j = 1, 2$; $\widehat{G}_j^*(w_{jk})$ is the observed wage distribution for sector j estimated from the data and evaluated at the implied grids for wages: $w_{jk} \equiv w_j(W_{jk})$, using equations (15) and (16); and $G_j(W_{jk})$ is the distribution of contracts in the population of employed workers implied by the model and which depends on all parameters ($\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$).

We use Chebyshev nodes for the grid of contract values and we replace the integrals in contract spaces by Clenshaw-Curtis (CC) quadrature approximations. Computational details are provided in Appendix C.

3.2.2 Transition Rates

In a similar way as we estimate $\boldsymbol{\theta}$ given $\boldsymbol{\lambda}$, we can estimate $\boldsymbol{\lambda}$ given $\boldsymbol{\theta}$ by matching the appropriate moments. In the data we observe the proportion of workers in state $i = 0, 1, 2$ at the beginning of the survey moving to state $j = 0, 1, 2$ before the end of the survey, lasting T periods ($\widehat{D}_{ij}$). We can use the model to compute the theoretical counterparts to these proportions (D_{ij}) as we show in Appendix C. For example the proportion who were in a formal sector job and move to an alternative job within the same sector is given by

$$D_{11} = \int_{\underline{W}_1}^{\overline{W}_1} \frac{\lambda_{11} \overline{F}_1(x)}{d_1(x)} (1 - e^{-d_1(x)T}) \, dG_1(x).$$

Now, in equilibrium,

$$\ell_1(x) = \frac{1}{n_1} \frac{h_1(x)}{d_1(x)} = \frac{m_1}{n_1} \frac{dG_1(x)}{dF_1(x)},$$

allowing to replace the derivative of G_1 by that of F_1 inside the integral. Then CC-quadrature can be used to approximate the integral.

We thus estimate $\boldsymbol{\lambda}$ given $\boldsymbol{\theta}$ by minimizing the criterion

$$Q_2(\boldsymbol{\lambda}|\boldsymbol{\theta}) = \sum_{i,j=0,1,2} \left(\widehat{D}_{ij} - D_{ij}\right)^2,$$

where $\widehat{D}_{ij}$ is the empirical counterpart of D_{ij}.

3.2.3 Value of Leisure

As mentioned above we allow unemployment insurance to be determined endogenously: in Brazil about 8.5% of receipts from labour taxes fund UI. Hence we compute the implied amount using the government budget constraint

$$0.085\tau \int_{\underline{w}_1}^{\overline{w}_1} x \, d\widehat{G}_1^*(x) = UI \cdot D_{10}.$$

where D_{10} is the average transition probability from a formal sector job to unemployment and where $\widehat{G}_1^*$ is the estimated wage distribution. Remember that UI is paid to workers at the moment of transition into unemployment; hence this calculation is useful for constructing an amount that is consistent with the expected expenditure by Brazil and with the way we model UI.

Having estimated the contract values in both sectors and having set U to be equal to $\underline{W}_2$ (the legal minimum wage is not enforced in the informal sector and hence the minimum observed wage is the reservation wage) we can use the value function for the unemployed (3) to estimate the value of leisure, b.[17]

[17] An important issue here is measurement error. At present we have not allowed for wages to be measure with error. If we did, this would affect the estimation of the the distributions G and the value of leisure b.

3.2.4 Productivity Distribution

Up to this point, there has been no need to use the firm profit functions, the costs of informality, or indeed the distribution of productivities: the arrival rates, the job destruction rates and the wage offer distribution can be identified purely from the distribution of wages and the transition rates. The offer distribution implicitly depends on the costs of informality however. Thus, counterfactual policy simulations require an explicit specification and estimation of the costs of informality, which will allow us to compute the new equilibrium.

We specify the cost function as $C = c\ell_2(W)^\gamma$, with c and γ being the parameters to be estimated. Given values for c and γ, and for n_1 and n_2 such that $n_1 + n_2 \leq 1$, we solve for the labour force size in the formal sector ($\ell_1(W) = \frac{1}{n_1}\frac{h_1(W)}{d_1(W)}$) and in the informal sector ($\ell_2(W) = \frac{1}{n_2}\frac{h_2(W)}{d_2(W)}$). From the firm's maximization problem in each sector, we can derive the way contracts and productivities are related. We start by deriving the support of productivities in each sector and and we then show how we derive the entire distributions. The key point is that firm profits, given productivity are equalized, across sectors.

To derive the support note that the first order conditions for the firm's optimization problem (see (13), (14)) gives

$$K_1^{-1}(W) = (1 + \tau + \lambda_{10}s)\left[w_1(W) + w_1'(W)\frac{\ell_1(W)}{\ell_1'(W)}\right], \tag{24}$$

$$K_2^{-1}(W) = w_2(W) + w_2'(W)\frac{\ell_2(W)}{\ell_2'(W)} + C_1\gamma\ell_2(W)^{\gamma-1}, \tag{25}$$

where the expressions for $w_j'(W)$ and for $\ell_j'(W)$, as well as further details, are given in the appendix.

For each point of the contract grids, W_{jk}, one can thus calculate a point $p_{jk} = K_j^{-1}(W_{jk})$ on a productivity grid, with $\underline{p}_2 = p_{20}$, $\underline{p}_1 = p_{10}$, $\overline{p}_2 = p_{2N}$ and $\overline{p}_1 = p_{1N}$,

allowing the calculation of the productivity distributions as

$$\Gamma_j(p_{jk}) = n_j \cdot F_j(W_{jk}),\ j = 1,2, k = 0,...,N.$$

To estimate the remaining parameters, c and γ, and n_1 and n_2, we use the equilibrium conditions requiring that $\pi_2(\underline{p}_2) = 0$, and $\pi_1(p) = \pi_2(p) > 0$ for $p \in [\underline{p}_1, \overline{p}_2]$. We thus estimate c and γ, and n_1 and n_2 such that $n_1 + n_2 = 1$, so as to minimize

$$\pi_2(p_{20})^2 + \sum_{k,k'=0}^{M} \mathscr{K}(p_{1k} - p_{2k'})[\pi_1(p_{1k}) - \pi_2(p_{2k'})]^2,$$

where $\mathscr{K}$ is a kernel density.

Lastly, the unconditional productivity distribution Γ_0 follows as

$$\Gamma_0(p) = \begin{cases} \Gamma_2(p), & \forall p \in [\underline{p} = \underline{p}_2, \underline{p}_1], \\ \Gamma_1(p) + \Gamma_2(p), & \forall p \in [\underline{p}_1, \overline{p}_2], \\ n_2 + \Gamma_1(p), & \forall p \in [\overline{p}_2, \overline{p}_1 = \overline{p}], \end{cases}$$

with the additional restriction: $\Gamma_0(\underline{p}_2) = 0$. This distribution is the exogenous distribution of productivity truncated below at $\underline{p}_2$. There is no way one can identify the portion of the distribution below $\underline{p}_2$.

3.3 Endogenous Arrival Rates: Estimating a Matching Function

Counterfactual simulations in a general equilibrium framework require accounting for the impact of policy on arrival rates, as the number of job seekers and the number of firms in the two sectors react to policy changes. To estimate the relationship

between arrival rates and search activity we specify a matching function $f(\theta)$, describing the flow of matches as a function of market tightness (the effective number of firms divided by the effective number of job seekers).[18] This is combined with an assumption about the way these contacts are allocated between the formal and the informal sector. We then identify the parameters of this model by imposing the restriction implied by the matching functions on the job arrival rates separately for each submarket using minimum distance.

Define market tightness as

$$\theta = \frac{n_1 + \alpha n_2}{u + s_1 m_1 + s_2 m_2}. \tag{26}$$

where s_1 and s_2 are the search effort of those employed in the formal and the informal sector respectively relative the search effort of the unemployed s_0, which we normalize to one. Define the flow of contacts by the matching function $f(\theta) = \mu\theta^{\eta}$. We assume that the probability of an offer from the formal sector is $n_1/(n_1 + \alpha n_2)$ and the informal one $\alpha n_2/(n_1 + \alpha n_2)$, where α denotes relative *visibility* of informal vacancies in the market. Given these specifications, the job offer arrival rates to workers in state $i = 0, 1, 2$ from the formal and the informal sector, can be written respectively as

$$\lambda_{i1} = \frac{n_1}{(n_1 + \alpha n_2)} s_i f(\theta); \tag{27}$$

$$\lambda_{i2} = \frac{\alpha n_2}{(n_1 + \alpha n_2)} s_i f(\theta). \tag{28}$$

We do not attempt to estimate η, but we try alternatives based on the range of

[18] Since firms will hire anyone they meet offering them their posted wage the effective number of firms relative to the effective number of job seekers is equivalent to the number of vacancies per worker searching.

estimates from the literature.[19]

4 Results

We focus our estimation for low education individuals, for whom individual heterogeneity is probably less important. We present estimates separately for males and females and for two contrasting regions of Brazil: wealthy and dynamic Sao Paulo and the poorer region of Salvador, allowing us to see how the welfare effects of policy towards informality may change with the context.

4.1 The Model Fit

Table 5 presents evidence on the fit of the model. The model is capable of replicating well the proportions of workers in the formal and informal sectors and the unemployed and particularly well all the transitions between sectors. The distribution of wages is also very well replicated, although the fit is not always perfect, particularly at the lower tail.

4.2 Frictional Parameters and the Level of Informality

Table 6 shows the estimates for the job destruction and the job arrival rates with associated standard errors obtained using 500 bootstrap replications.[20] The estimated job destruction rates are three to five times as high in the informal sector as in the formal one. However, even informal jobs seem to be very stable, with an expected duration of nearly five years in the absence of job to job mobility. Unemployed

[19]The elasticity of the matching function with respect to vacancies η is usually estimated in the range 0.3-0.5 (see, Petrongolo and Pissarides, 2001).

[20]The unit of time is a month. Subscript 0 refers to unemployment, 1 refers to the formal sector and 2 to the informal. The arrival rates λ_{ij} denote an offer arriving from sector j to someone currently in sector i.

TABLE 5
Model Fit

	Sao Paulo				Salvador			
	Males		Females		Males		Females	
	Actual	Model	Actual	Model	Actual	Model	Actual	Model
Formal Employment (m_1)	0.625	0.579	0.516	0.495	0.582	0.486	0.420	0.473
Informal Employment (m_2)	0.263	0.325	0.264	0.299	0.254	0.336	0.213	0.207
Unemployment (u)	0.112	0.097	0.220	0.207	0.164	0.178	0.367	0.320
Transitions								
D_{01}	0.088	0.089	0.043	0.043	0.050	0.050	0.017	0.017
D_{02}	0.257	0.257	0.111	0.111	0.112	0.112	0.031	0.031
D_{10}	0.020	0.020	0.020	0.020	0.020	0.020	0.013	0.013
D_{11}	0.016	0.016	0.013	0.013	0.026	0.026	0.021	0.021
D_{12}	0.004	0.004	0.004	0.004	0.003	0.003	0.001	0.001
D_{20}	0.082	0.082	0.075	0.075	0.060	0.060	0.045	0.045
D_{22}	0.065	0.065	0.062	0.062	0.059	0.059	0.045	0.045
D_{21}	0.011	0.011	0.008	0.008	0.005	0.005	0.005	0.005
Formal Wages (log)								
P10	6.28	6.24	5.89	6.08	5.95	5.59	5.72	5.59
P25	6.42	6.51	6.27	6.29	6.09	6.09	5.90	5.91
Median	6.65	6.80	6.41	6.47	6.30	6.28	6.03	6.12
P75	6.93	6.95	6.58	6.63	6.57	6.50	6.25	6.27
P90	7.24	7.15	6.87	6.77	6.89	6.63	6.48	6.40
Informal Wages (log)								
P10	5.87	5.20	5.86	5.50	5.59	5.32	5.41	5.23
P25	6.07	5.86	5.96	5.89	5.70	5.68	5.57	5.47
Median	6.34	6.42	6.16	6.18	5.88	5.94	5.69	5.74
P75	6.67	6.68	6.42	6.50	6.17	6.13	5.81	5.88
P90	7.04	6.98	6.75	6.67	6.51	6.30	6.04	6.01

TABLE 6
Transition Parameters

	λ_{10}	λ_{20}	λ_{01}	λ_{02}	λ_{11}	λ_{22}	λ_{12}	λ_{21}
Sao Paulo								
Males	0.0052	0.0223	0.0271	0.0789	0.0172	0.0652	0.0553	0.0039
	(0.0003)	(0.0010)	(0.0019)	(0.0032)	(0.0020)	(0.0058)	(0.0148)	(0.0004)
Females	0.0052	0.0203	0.0116	0.0301	0.0115	0.0629	0.0376	0.0028
	(0.0005)	(0.0013)	(0.0013)	(0.0021)	(0.0022)	(0.0117)	(0.0194)	(0.0005)
Salvador								
Males	0.0051	0.0159	0.0136	0.0305	0.0383	0.0673	0.0619	0.0019
	(0.0004)	(0.0013)	(0.0015)	(0.0022)	(0.0077)	(0.0105)	(0.0389)	(0.0005)
Females	0.0033	0.0117	0.0044	0.0080	0.0364	0.0482	0.0661	0.0015
	(0.0006)	(0.0019)	(0.0009)	(0.0012)	(0.0282)	(0.0175)	(0.0146)	(0.0007)

Note: The unit of time is a month. Bootstrap standard errors in parentheses.

workers receive twice to three times as many job offers relative to those employed in both regions. Interestingly, the arrival rates of offers from other informal jobs is higher for individuals already working in either sector than for those who are unemployed. It is also easier to locate formal jobs once working in the formal sector. However, obtaining formal job offers while working in the informal sector is much harder than when unemployed.

Comparing across regions, Sao Paulo has higher destruction rates than Salvador in the informal sector, while for both regions the destruction rates in the formal sector are very small. Effectively formal jobs last a very long time. In the more dynamic Sao Paulo informal jobs in particular, seem to be created and destroyed at a much higher rate. Within sector offer rates are similar in both regions; however in Sao Paulo the chance of obtaining an offer from the formal sector when in an informal job, although low, is substantially higher than in Salvador. However the key differences between the regions seems to be in job destruction rates and in offers received when unemployed.

These differences reflect themselves in the implied unemployment rates: that of Salvador is twice that of Sao Paulo (Table 5), which mirrors the data. The implied proportion of formal and informal firms are essentially the same across markets

TABLE 7
Proportion of Formal Firms by market

Sao Paulo		Salvador	
Males	Females	Males	Females
0.34	0.31	0.32	0.30
(0.04)	(0.03)	(0.04)	(0.07)

Note: Bootstrap standard errors in parentheses.

TABLE 8
Matching Function Estimates

				μ		
	s_1	s_2	α	$\eta = 0.3$	$\eta = 0.5$	θ
Sao Paulo						
Males	0.668	0.485	3.937	0.091	0.067	4.586
	(0.116)	(0.041)	(0.630)	(0.004)	(0.003)	(0.682)
Females	1.119	1.167	4.206	0.046	0.037	2.896
	(0.352)	(0.224)	(2.158)	(0.008)	(0.005)	(1.581)
Salvador						
Males	2.422	1.173	6.104	0.070	0.058	2.555
	(0.808)	(0.195)	(2.032)	(0.033)	(0.031)	(0.837)
Females	8.597	3.628	7.308	0.063	0.066	1.143
	(3.150)	(1.299)	(5.707)	(0.468)	(0.051)	(0.878)

Note: Bootstrap standard errors in parentheses.

(about 30%; see Table 7) – there is no data counterpart to this. The estimates of the matching function that relates these transition rates to market tightness are shown in Table 8. According to these estimates informal jobs are 4-6 times as visible as formal ones. Moreover, search intensity is higher among the employed than the unemployed in all submarkets but those for males in Sao Paulo.

4.2.1 Informality Cost and the Value of Leisure

Table 9 presents the implied cost to the firm of remaining informal. This cost arises from random monitoring and imposition of fines. We report the parameters of the cost function, $C = c\ell_2(W)^\gamma$, and the mean cost per unit of profit. The costs are linear in firm size for Sao Paulo and strictly convex for Salvador. In both cases the gradient

TABLE 9
Cost of Informality and Value of Leisure

	c	γ	mean cost-profit ratio	b
Sao Paulo				
Males	54.5	1.0	0.148	-704.3
	(14.2)	(0.25)	(0.046)	(74.3)
Females	62.0	1.0	0.278	78.4
	(11.4)	(0.22)	(0.088)	(40.3)
Salvador				
Males	70.5	1.5	0.243	7.7
	(9.3)	(0.46)	(0.081)	(40.5)
Females	72.5	1.6	0.282	182.7
	(12.6)	(0.66)	(0.292)	(21.3)

Note: Bootstrap standard errors in parentheses.

c is very high for both males and females. These features imply that informality will be concentrated among smaller firms as we would expect. This is an interesting result particularly given that we do not observe firms directly.

In the last column of Table 9 we present the estimated flow value of leisure. These are lower in Sao Paulo for each gender. Moreover, women value leisure much more than for men, possibly reflecting the demands of families and home production. The differences across regions are not significant in this case.

4.3 Formal and Informal Sector Productivity and Wages

A key feature of the equilibrium we describe is that given productivity, both formal and informal firms can coexist. This can have important policy implications. Policies that reduce informality will not necessarily shut down all jobs in this part of the productivity distribution; on the other hand this should not be taken to imply that such an exercise will be costless, because lower levels of productivity may be able to sustain only smaller and fewer formal firms, given the amount of competition for workers and the overall regulatory costs. We consider these issues by first describing

the equilibrium that results from our estimates and subsequently by counterfactual simulations.

Based on the estimates we can back out the implied allocation of workers to the formal and the informal sector for different levels of productivity, as well as the pay structure. The results are presented in Tables 10 and 11 for low education males, and in Tables 12 and 13 for low education women, in Sao Paulo and Salvador, respectively.

For males the lowest point of support of the productivity distribution is similar for both Sao Paulo and Salvador. However, all other percentiles are lower in Salvador, reflecting lower productivity and lower wages. In Sao Paulo there is less than six percent of firms in the formal sector below the 10th percentile of the productivity distribution. In Salvador formal firms start operating at a level of productivity above the 25th percentile. In both markets, informality is to be found at decreasing rates at all levels of productivity, but the size of formal firms increases rapidly.

One of the most interesting features of the model is the implied wage structure. First, comparing wages and productivities the implied rents are quite high and particularly so for low productivity firms. In both labor markets frictions imply substantial rents accruing to firms, which of course can motivate welfare improving policies.

Second, the results justify two seemingly contradictory statements. Wages are on average higher in the formal sector than in the informal one, because the formal firms become increasingly large as productivity increases and wages increase with productivity (as indeed in the standard Burdett and Mortensen model): this is a composition effect. However, given productivity, for the most part formal firms pay less than informal ones: this is a compensating differential for the UI, pension entitlements and severance pay, although frictions and different job arrival rates will imply that the relationship is not one-to-one with these benefits. This differential declines and even gets reversed at the highest levels of productivity.

The overall picture is similar for women with some small differences: first formal firms in Salvador start operating at a higher point of the distribution of productivity like in the male market; second the wage structure is different and the distribution of productivities do have different shapes. Comparing the wage structures is not straightforward because of the differing productivities of the jobs they tend to work and the resulting changes in composition. However, male wages in the formal sector are more dispersed than those of females in both regions.

Table 14 presents male and female wages for the two regions by sector and overall, at the same productivity level. In all cases, but the informal sector of Salvador, women are paid more conditional on productivity, for lower productivity levels. This is reversed at the higher productivity levels. Thus women in most cases seem to work in more competitive labour markets with lower monopsony power for firms. However, on average women are paid less than men because most of them work in lower productivity (and hence lower paid) jobs. In other words the model interprets discrimination as being due to the type of jobs in the female labour market.

5 Policy Analysis

The model aims at providing a framework for understanding the impact of reducing or eliminating informality in an equilibrium setting. The latter is crucial here, because we need to know how the wage structure will change and what will be the overall welfare loss from such policies after allowing firms to relocate sector in response to the policy change. Note that we can only simulate reforms that decrease the number of operating firms (increase $\underline{p}_2$), as we have not identified the part of Γ_0 below the inferred minimum productivity in the sample.

We simulate an increase in UI which we briefly discuss below. We then consider a policy that increases the cost of informality (possibly as a result of improved

TABLE 10
Sao Paulo, Males - Estimates by productivity

Productivity		Cumulative	Fraction of	Fraction of	Wage (log)		Value (log)		Firm size	
Percentiles	(log)	workforce	formal firms	formal workers	Formal	Informal	Formal	Informal	Formal	Informal
10th	6.289	0.114	0.059	0.121	-	4.702	11.499	11.476	1.6	1.6
25th	6.417	0.129	0.084	0.136	4.080	5.203	11.517	11.485	2.0	2.0
50th	6.705	0.219	0.181	0.302	5.843	6.081	11.594	11.525	4.9	4.3
75th	7.039	0.400	0.321	0.499	6.390	6.557	11.695	11.583	12.2	9.7
90th	7.306	0.613	0.378	0.669	6.715	6.784	11.813	11.630	28.0	14.6
99th	8.388	0.971	0.968	0.975	7.214	7.529	12.116	11.894	73.4	32.8

TABLE 11
Salvador, Males - Estimates by productivity

Productivity		Cumulative	Fraction of	Fraction of	Wage (log)		Value (log)		Firm size	
Percentiles	(log)	workforce	formal firms	formal workers	Formal	Informal	Formal	Informal	Formal	Informal
10th	5.713	0.185	0.000	-	-	4.484	-	11.009	-	1.0
25th	5.900	0.210	0.000	-	-	5.079	-	11.023	-	1.5
50th	6.080	0.253	0.148	0.205	5.316	5.517	11.111	11.047	2.2	2.7
75th	6.406	0.389	0.270	0.305	5.595	5.938	11.197	11.098	5.8	7.8
90th	6.643	0.552	0.330	0.486	6.087	6.133	11.299	11.140	17.0	14.0
99th	7.353	0.941	0.582	0.815	6.625	6.522	11.567	11.271	98.2	27.1

TABLE 12
Sao Paulo, Females - Estimates by productivity

Productivity		Cumulative	Fraction of	Fraction of	Wage (log)		Value (log)		Firm size	
Percentiles	(log)	workforce	formal firms	formal workers	Formal	Informal	Formal	Informal	Formal	Informal
10th	6.110	0.214	0.008	0.007	4.662	5.013	11.297	11.301	1.4	1.2
25th	6.301	0.259	0.091	0.103	5.354	5.714	11.333	11.326	2.5	2.4
50th	6.508	0.328	0.190	0.237	5.787	6.049	11.377	11.353	4.6	4.5
75th	6.737	0.478	0.296	0.507	6.192	6.298	11.468	11.389	11.9	8.4
90th	7.141	0.781	0.402	0.720	6.553	6.589	11.624	11.454	32.6	15.7
99th	7.963	0.967	0.945	0.960	6.839	7.063	11.801	11.617	51.1	25.5

TABLE 13
Salvador, Females - Estimates by productivity

Productivity		Cumulative	Fraction of	Fraction of	Wage (log)		Value (log)		Firm size	
Percentiles	(log)	workforce	formal firms	formal workers	Formal	Informal	Formal	Informal	Formal	Informal
10th	5.584	0.326	0.000	-	-	4.987	-	10.863	-	0.6
25th	5.733	0.341	0.000	-	-	5.232	-	10.875	-	0.9
50th	5.892	0.372	0.119	0.226	4.250	5.466	10.976	10.897	1.6	1.8
75th	6.156	0.457	0.269	0.327	5.333	5.739	11.045	10.942	3.4	5.2
90th	6.382	0.586	0.322	0.629	5.914	5.883	11.168	10.981	15.8	8.8
99th	6.877	0.924	0.552	0.907	6.334	6.139	11.401	11.074	131.0	15.2

TABLE 14
Comparing male and female wages, by productivity

	Sao Paulo				Salvador			
	Formal		Informal		Formal		Informal	
Productivity	Males	Females	Males	Females	Males	Females	Males	Females
6.00	-	3.889	2.423	4.786	4.096	4.948	5.318	5.567
6.25	-	5.047	4.702	5.503	5.316	5.590	5.820	5.813
6.50	5.063	5.787	5.571	6.049	5.800	6.116	6.041	5.950
6.75	5.843	6.192	6.081	6.298	6.285	6.267	6.218	6.077
7.00	6.390	6.473	6.557	6.499	6.437	6.334	6.299	6.320
7.25	6.715	6.629	6.784	6.675	6.566	6.399	6.449	6.604
Mean	6.765	6.477	6.437	6.226	6.275	6.096	5.946	5.720

detection rates or fines). We then simulate a bold counterfactual, where we close down completely the informal sector. Table 15 reports estimates of the effects of these changes on the composition of the workforce, firm size and welfare, for the San Paolo region and for low-education males. Table 16 reports the effects on the wage distributions and Table 17 the effects on the distributions of firm productivities. Tables in Appendix D show results for other groups. Here we summarize the implications.

5.1 Unemployment Insurance and Severance

We consider an increase of UI by 100%: although this sounds a lot, UI in Brazil is quite low particularly because it is time limited: we increase it from one to two minimum wages per month, payable for three months.[21] In our model such a policy encourages employment, rather than the opposite, because it is payable upfront and because once employed individuals have no endogenous incentive to quit – the only way to claim again is to be laid off due to exogenous job destruction. In reality claiming UI after expiration requires six months legal work. Changing UI will change the

[21]UI benefit ranges from 1 to about 2 minimum wages monthly, depending on the average of the three last wages received from last job, and are payable up to 5 months, depending on the last job spell. The majority of low education workers are entitled to 1 minimum wage per month during about 3 months.

equilibrium distribution because it will increase the relative attractiveness of formal jobs, it will increase the cost of formal employment and it will increase corporation taxes, which is the source of funding - all our simulations keep government revenue constant.

We only summarize the results. As we expect the effects of increasing UI from such a low base are small but interesting: the increase in UI decreases overall welfare because it decreases firm profits without increasing worker welfare. This is because it increases the supply of workers to formal firms, which now become a bit larger, although some lower productivity formal firms become informal. The resulting shift increases the profits in the formal sector but decreases informal profits, with the net effect being no change in worker's welfare and an overall drop in firm profits. Increasing severance pay by 5 percentage points has a very small **negative** effect on welfare which can be related mainly to a small decline in formal profits.

5.2 Increasing the Cost of Informality

We now consider a 10% increase in the costs of informality. The reform is revenue neutral; any increase in revenue from fines is redistributed by reducing corporation tax. As seen in Table 15 this increases the proportion of formal firms, reduces unemployment and increases the proportion of workers in the formal sector. The resulting average firm size declines.

The movements also lead to changes in the distribution of wages: as seen in Table 16, mean wages in the informal sector in Sao Paulo decline by 8% and the distribution (within that sector) becomes more compressed. Formal sector wages increase above the 25th percentile. Firms that are relocating to the formal sector tend to be the higher productivity informal firms. Thus competition at the higher levels of productivity increases and leads to more rents being captured by the workers. Moreover, with the increase in revenues from fines in the informal sector, the corporation

tax decreases. The net effect is an increase in welfare overall (Table 15) and for all concerned (formal and informal workers and firms as well as the unemployed). In particular, the welfare of formal workers increases because their wages go up, due to the increased competition; informal workers and the unemployed are also better off because the value of a formal sector job, that they may move to, has increased. This more than counteracts the decline in informal sector wages.

For females in Sao Paulo, (Tables D.1 and D.2) the proportion of formal firms do not increase but the proportion of formal workers increases by 3 percentage points (pp) and there is a decline in unemployment. The increased supply of workers in that sector forces contract values and wages down, particularly at the lower end. On average, there is a small decrease in the values offered in the formal sector, following an also slight decrease in wages in that sector. However, overall welfare still goes up, mostly due to an increase in formal sector profits.

The results above were for Sao Paulo. For males in Salvador, Tables D.4 and D.5 show that increasing the cost of informality has a positive but much smaller impact on the overall welfare of workers and no effect on firms profits. This follows from a 2-3% increase in wages in the formal sector, despite a 10% decline of wages for the informal sector at all percentiles. As for females, Tables D.7 and D.8 show that overall welfare increases; the decline in wages in the informal sector by about 4% at the median and more at lower percentiles is counteracted with an increase in informal wages at higher percentiles. This occurs due to relocation of some low productivity informal firms to the formal sector. Moreover, informal firm size goes up by 2 percentage points, which leads to an increase in profits in the informal sector.

While there are differences in the results implied by variation in preference and technology parameters across markets (regions and genders), one thing stands out: reducing informality increases welfare overall. This is because the presence of informal firms limits the size of the more productive formal firms and at the same

TABLE 15

Effects on the composition of workforce, firm size and welfare, of changes in the informality cost, and of eliminating the informal sector - Sao Paulo, Low Education Males

		Increase in	No Informal Sector		
	Benchmark	C	Exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
m_1	0.58	0.60	0.84	0.95	0.93
m_2	0.32	0.31	-	-	-
u	0.10	0.09	0.16	0.05	0.07
n_1	0.34	0.46	0.82	0.82	0.82
n_2	0.66	0.54	-	-	-
Formal firm size (Mean)	17.1	13.1	10.3	11.6	11.4
Informal firm size (Mean)	4.9	5.7	-	-	-
Welfare (Reais($) per month)					
Formal worker $[rE(W_1)]$	708.2	779.1	474.6	698.1	689.9
Informal worker $[rE(W_2)]$	532.4	577.2	-	-	-
Unemployed $[rU]$	476.6	529.4	175.8	521.8	482.9
Average worker $[r(uU + m_1E(W_1) + m_2E(W_2))]$	628.7	694.2	426.8	689.1	676.3
Formal firm $[E(\pi_1)]$	1,671.5	1,425.7	1,564.3	817.8	967.3
Informal firm $[E(\pi_2)]$	239.9	363.0	-	-	-
Average firm $[N_1E(\pi_1) + N_2E(\pi_2)]$	726.6	851.8	1279.7	669.1	791.3
Total (Workers + Firms)	1,355.3	1,546.0	1,706.5	1,358.1	1,467.5
Government Revenue (formal sector)	610.4	572.6	638.7	636.8	637.5
Government Revenue (informal sector)	26.8	65.5	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

time allows the latter to keep more rents per worker. We now ask the question of what would happen if we could abolish completely the informal sector. Clearly, this eventuality is far outside the sample and we do not wish to imply that all effects are captured here. However, this will give an idea of the direction of the effects.

5.3 Abolishing Informality

We now present an experiment where we close down the informal sector. The results for Sao Paulo are in Tables 15 and 16. The Tables for the other markets are in Appendix D. All simulations are revenue neutral, which is achieved by adjusting the corporation tax. Note that in the absence of an informal sector the corporation

TABLE 16

Effects on wages and overall wage inequality - Sao Paulo, Low Education Males

		Increase in	No Informal Sector		
	Benchmark	C	Exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Wages (log)					
P10	6.24	6.24	5.11	5.93	5.64
P25	6.51	6.64	5.57	6.44	6.29
Median	6.80	6.83	6.50	6.72	6.72
P75	6.95	7.05	6.88	6.86	6.94
P90	7.15	7.18	7.17	7.03	7.11
Mean	6.80	6.86	6.48	6.65	6.66
Informal Wages (log)					
P10	5.20	5.32	-	-	-
P25	5.86	5.91	-	-	-
Median	6.42	6.44	-	-	-
P75	6.68	6.68	-	-	-
P90	6.98	6.89	-	-	-
Mean	6.51	6.43	-	-	-
Overall Wage Inequality					
P75/P25	1.71	1.77	3.69	1.52	1.91
P90/P10	3.43	3.77	7.82	3.00	4.35

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

tax is non-distortionary because it is imposed on rents and hence can never affect the decision of a firm either to hire or to operate. We present three different scenarios: one in which the contact rates are kept fixed and two where they are endogenised as shown in subsection 3.3, each with a different elasticity for the matching function.

With fixed contact rates (the $\boldsymbol{\lambda}$) unemployment increases in all markets. Workers welfare always declines, while overall welfare only increases for the male markets. However, the results look very different once we allow the contact rates to adjust. In this case unemployment declines in all markets and overall welfare increases as does workers' welfare. In Sao Paulo about 18% of firms now close down and the average firm size declines to 11-12: many of the previously informal firms become relatively small formal ones, reducing the average firm size. Despite the increased competition among firms the increase in the number of workers has meant a decline in wages, with respect to those in the formal sector and indeed a decrease in overall inequality. With greater sensitivity of the matching function to labor market tightness wages still decline but inequality goes up. The overall effect is an increase in workers' welfare. However, what happens to firm profits depends on the the elasticity of the matching function. At a low elasticity firm profits decline. When the elasticity is higher profits increase. Most of these effects are common (to various degrees across markets). In all cases overall welfare increases with the abolition of informality and so does workers' welfare, driven by a decline in unemployment, a repositioning in the higher paying and more productive formal sector and somewhat counteracted by a decline in formal sector wages relative to the baseline.

Thus the key result that is found across all markets we look at is that abolishing informality redistributes wealth towards workers. However, the extent to which this happens varies with the specific conditions (reflected in the estimated parameters) and does rely on allowing the contact rates to adjust as the number of job seekers and firms change. Part of this redistribution occurs because workers are shifted to the

TABLE 17
Effects on the distribution of productivity - Sao Paulo, Low Education Males

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Productivity (log)					
P10	6.48	6.48	6.14	6.17	6.11
P25	6.71	6.75	6.34	6.46	6.37
Median	6.93	7.08	6.72	6.82	6.83
P75	7.19	7.31	7.22	7.09	7.09
P90	7.47	7.57	7.61	7.31	7.32
Mean	7.18	7.29	7.23	6.89	6.94
Informal Productivity (log)					
P10	6.20	6.23	-	-	-
P25	6.38	6.40	-	-	-
Median	6.59	6.59	-	-	-
P75	6.81	6.74	-	-	-
P90	7.18	7.95	-	-	-
Mean	6.71	6.79	-	-	-
Minimum Thresholds					
$\underline{p}_2$	3.84	4.29	-	-	-
$\underline{p}_1$	5.99	5.88	6.11	5.81	5.82

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

formal sector, with an associated decrease in unemployment. Although wages in the formal sector decline on average (Sao Paulo) or increase just marginally (Salvador - depending on η) the formal sector wages remain much higher than the wages in the informal sector. In terms of productivity formal firms still start operating at the same level; so all low productivity informal firms that did not have formal counterparts just close down and do not switch to the formal sector. However the density of lower productivity formal firms increases as some of the informal firms on the overlapping range switch to become formal.

6 Conclusions

Informality is extremely common in developing countries. While the phenomenon is well recognized its effects are highly disputed and policy makers tend to be hesitant in addressing the issue one way or another. With this paper we wish to contribute to this debate.

On the one hand informal firms are portrayed as regulation busters that offer a much needed competitive fringe. Hence they are considered job creators and an indirect way by which employment protection legislation can be relaxed without governments being accused of siding in favor of business and against the workers. Moreover, informal firms have low productivity; an interpretation is that these jobs, which would not have existed in a tightly regulated economy are allowed to exist and hence increase employment. On the other hand workers in the informal sector are often denied access to the benefits of modern societies, such as unemployment insurance and public pensions (except at a minimum level) as well as a proper health and safety framework.

To understand the balance between the pros and cons of informality we set up a model with search frictions, costs of informality (because of penalties when caught) and with endogenous decisions by both workers and firms as to where to work and locate jobs respectively. Clearly a competitive framework would necessarily imply that informality is welfare improving, at least with risk neutral agents. Our results show that search frictions are very important and without these elements in the model it would be very hard to understand the role of informality. Our model is motivated by the empirical observation that low skill workers are observed working both in formal and informal jobs; there is little empirical support of the idea of comparative advantage that leads to segmentation on the basis of skill - at least in the observable dimension. In our model segmentation is endogenously determined by an interplay between search frictions, the institutional requirements for formal firms

and the penalties of informality. This generates profit making opportunities by firms selecting in either the formal or informal sector. In equilibrium profits are equalized.

Using the simulations from our model we draw two sets of important conclusions. First, marginal increases in regulation, such as UI, in the presence of an informal sector have little or no perceptible effect on the economy; they also have little effect in the distribution of activity between the formal and informal sector. However, increasing the cost of informality by 10% actually improves overall welfare and either reduces unemployment or leaves it unchanged in all markets we looked at. The way the gains a redistributed really depends on the extent of search frictions. In the more dynamic economy of Sao Paulo this also meant an increase in workers welfare; in Salvador workers loose out, but only very marginally.

If we go as far as abolishing informality we obtain increases in welfare in all markets we consider, so long as we allow the arrival rates of job offers to adjust endogenously. In addition workers' welfare also always increases and unemployment declines. While the details of how the gains are distributed and what precisely happens to wages differ by market the main result that informality decreases welfare, given the costs structure, remains. Viewing informality in the light of search frictions, can alter fundamentally our views of its role: rather than being a benign regulation busting mechanism it acts to redistribute welfare to firms and to reduce overall welfare. Of course the result is predicated on the assumption that informal firms incur detection costs.

References

[1] Albrecht, James, Lucas Navarro, and Susan Vroman. 2009. "The Effects of Labour Market Policies in an Economy with an Informal Sector." *The Economic Journal*, 119(539): 1105–1129.

[2] Barros, Ricardo P., Guilherme L. Sedlacek, and Simone Varandas. 1990. “Segmentação e Mobilidade no Mercado de Trabalho: A Carteira de Trabalho em São Paulo.” *Pesquisa e Planejamento Economico,* 20(1): 87–103.

[3] Boeri, Tito, and Pietro Garibaldi. 2005. "Shadow Sorting." NBER Chapters, in: NBER International Seminar on Macroeconomics 2005 National Bureau of Economic Research, Inc.

[4] Bontemps, Christian, Jean-Marc Robin, and Gerard J. van den Berg. 1999. "An Empirical Equilibrium Job Search Model with Search on the Job and Heterogeneous Workers and Firms." *International Economic Review*, Department of Economics, University of Pennsylvania and Osaka University Institute of Social and Economic Research Association, 40(4): 1039-1074, November.

[5] Bontemps, Christian, Jean-Marc Robin, and Gerard J. van den Berg. 2000. "Equilibrium Search with Continuous Productivity Dispersion: Theory and Nonparametric Estimation." *International Economic Review*, 41(2): 305-358.

[6] Bosch, Mariano. 2006. "Job Creation and Job Destruction in the Presence of Informal Labour Markets." mimeo, LSE.

[7] Burdett, Kenneth, and Dale T. Mortensen. 1998. "Wage Differentials, Employer Size, and Unemployment." *International Economic Review*, 39(2): 257-273.

[8] Curi, Andréa Z., and Naercio A. Menezes-Filho. 2006. "O Mercado de Trabalho Brasileiro é Segmentado? Alterações no Perfil da Infor-

malidade e nos Diferenciais de Salários nas Décadas de 1980 e 1990." *Estudos Economicos*, 36(4): 867-899.

[9] El-Badaoui, Eric Strobl, and Frank Walsh. 2010. “The Formal Sector Wage Premium and Firm Size.” *Journal of Development Economics*, 91(1): 37-47.

[10] Fugazza, Marco, and Jean-Francois Jacques. 2003. "Labor Market Institutions, Taxation and Underground Economy." *Journal of Public Economics*, 88(1-2): 395–418.

[11] Jolivet, Gregory, Fabien Postel-Vinay, and Jean-Marc Robin. 2006. "The Empirical Content of a Job Search Model: Labor Mobility and Wage Dispersion in Europe and the US." *European Economic Review*, 50(4): 877-907.

[12] Lazear, Edward P. 1990. “Job Security Provisions and Employment.” *Quarterly Journal of Economics*, 105(3): 699-726.

[13] Maloney, William F. 1999. "Does Informality Imply Segmentation in Urban Labor Markets? Evidence from Sectoral Transitions in Mexico." *The World Bank Economic Review*, 13(2): 275–302.

[14] Mortensen, Dale T., and Christopher A. Pissarides. 1994. "Job Creation and Job Destruction in the Theory of Unemployment." *The Review of Economic Studies*, 61(3): 397-415.

[15] Perry, Guillermo, William F. Maloney, Osmar S. Arias, Pablo Fajnzylber, Andrew D. Mason, and Jaime Saavedra-Chanduvi. 2007. Informality: Exit and Exclusion. The World Bank.

[16] Petrongolo, Barbara, and Christopher A. Pissarides. 2001. "Looking into the Black Box: A Survey of the Matching Function." *Journal of Economic Literature*, 39(2): 390-431.

[17] Pratap, Sangeeta, and Erwan Quintin. 2006. "Are Labor Markets Segmented in Developing Countries? A Semiparametric Approach." *European Economic Review*, 50(7): 1817–1841.

[18] Ridder, Geert, and Van den Berg, Gerard J. 1998. "An Empirical Equilibrium Search Model of the Labor Market." *Econometrica*, 66(5): 1183–1221.

[19] Shephard, Andrew. 2009. "Equilibrium Search and Tax Credit Reform." mimeo, Princeton University.

[20] Ulyssea, Gabriel. 2010. "Regulation of Entry, Labor Market Institutions and the Informal Sector." *Journal of Development Economics*, 91(1): 87–99.

[21] Van den Berg, Gerard J. 2003. "Multiple Equilibria and Minimum Wages in Labor Markets with Informational Frictions and Heterogeneous Production Technologies." *International Economic Review*, 44(4): 1337-1357.

[22] Waldvogel, J. (2006): “Fast Construction of the Fejér and Clenshaw-Curtis Quadrature Rules,” BIT *Numerical Mathematic*, 46, 195–202.

APPENDIX [FOR ONLINE PUBLICATION]

A Equilibrium Offer and Accepted Contract Distributions

In this section, we derive G_1 and G_2 from F_1 and F_2.

By equation (4), for any $W \in [\underline{W}_1, \overline{W}_1]$,

$$\left[\lambda_{10} + \lambda_{11}\overline{F}_1(W)\right] m_1 G_1(W) + \lambda_{12} m_1 \int_{\underline{W}_1}^{W} \overline{F}_2(x)\, dG_1(x) = \lambda_{01} u F_1(W) + \lambda_{21} m_2 \int_{\underline{W}_2}^{W} [F_1(W) - F_1(x)]\, dG_2(x).$$

Making use of the identities (integration by parts):

$$\int_{\underline{W}_1}^{W} \overline{F}_2(x)\, dG_1(x) = \overline{F}_2(W) G_1(W) + \int_{\underline{W}_1}^{W} G_1(x)\, dF_2(x),$$
$$\int_{\underline{W}_2}^{W} [F_1(W) - F_1(x)]\, dG_2(x) = \int_{\underline{W}_2}^{W} G_2(x) dF_1(x),$$

we can rewrite this equation as

$$d_1(W)\frac{m_1}{u} G_1(W) = \lambda_{01} F_1(W) - \Phi(W), \tag{A.1}$$

where $d_1(W) = \lambda_{10} + \lambda_{11}\overline{F}_1(W) + \lambda_{12}\overline{F}_2(W)$, and

$$\Phi(W) = \lambda_{12} \int_{\underline{W}_1}^{W} \frac{m_1}{u} G_1(x)\, dF_2(x) - \lambda_{21} \int_{\underline{W}_2}^{W} \frac{m_2}{u} G_2(x)\, dF_1(x). \tag{A.2}$$

Turning to the informal sector, equation (5) indicates that for $W \in [\underline{W}_2, \overline{W}_2]$,

$$\left[\lambda_{20} + \lambda_{22}\overline{F}_2(W)\right] m_2 G_2(W) + \lambda_{21} m_2 \int_{\underline{W}_2}^{W} \overline{F}_1(x)\, dG_2(x) = \lambda_{02} u F_2(W) + \lambda_{12} m_1 \int_{\underline{W}_1}^{W} [F_2(W) - F_2(x)]\, dG_1(x).$$

Using the same integrations by part, we obtain that

$$d_2(W)\frac{m_2}{u} G_2(W) = \lambda_{02} F_2(W) + \Phi(W), \tag{A.3}$$

where $d_2(W) = \lambda_{20} + \lambda_{21}\overline{F}_1(W) + \lambda_{22}\overline{F}_2(W)$.

Next, multiplying equation (A.1) by $\frac{\lambda_{12} f_2(W)}{d_1(W)}$ (with $f_2 = F_2'$) and equation (A.3) by

$-\frac{\lambda_{21} f_1(W)}{d_2(W)}$, and adding the two resulting equations, we obtain the first-order differential equation

$$\Phi' = A - B\Phi, \tag{A.4}$$

where

$$A = \lambda_{01} F_1 \frac{\lambda_{12} f_2}{d_1} - \lambda_{02} F_2 \frac{\lambda_{21} f_1}{d_2},$$
$$B = \frac{\lambda_{12} f_2}{d_1} + \frac{\lambda_{21} f_1}{d_2},$$

with boundary condition $\Phi(U) = 0$ (in fact $\Phi(W) = 0, \forall W \leq \max\{\underline{W}_1, \underline{W}_2\}$)**.**

The solution of differential equation (A.4) is given by

$$\Phi(W) = \frac{\int_U^W e^{\int_U^x B(x')dx'} A(x)\,dx}{e^{\int_U^W B(x)\,dx}}. \tag{A.5}$$

Substituting this solution back into equations (A.1) and (A.3) we obtain the equilibrium relationship between the distribution of offered (F) and accepted (G).

B Computing the Equilibrium

In this section we describe the computation of the equilibrium.

1. Define contract value offer distribution F_1 and F_2, with supports bounds $\underline{W}_2 = U < \underline{W}_1 < \overline{W}_2 < \overline{W}_1$. Note that. from equation (3),

$$\underline{W}_2 = U = \frac{b + \lambda_{01}\mu_1 + \lambda_{02}\mu_2}{r + \lambda_{01} + \lambda_{02}}.$$

   Define the numbers of firms in each sector n_1, n_2, with $n_1 + n_2 \leq 1$.

2. Use steady-state flow condition (see Appendix (A)) to derive m_1, m_2, u and G_1, G_2 from F_1, F_2.

3. Profit maximization then implies that optimal decision rules satisfy

$$p = K_1^{-1}(W) = (1 + \tau + \lambda_{10} s)[w_1(W) + w_1'(W)\frac{\ell_1(W)}{\ell_1'(W)}],$$
$$p = K_2^{-1}(W) = w_2(W) + w_2'(W)\frac{\ell_2(W)}{\ell_2'(W)} + C_1 \gamma \ell_2(W)^{\gamma - 1},$$

   with

$$(1 + \lambda_{10} s) w_1(W) = (r + \lambda_{10}) W - \lambda_{10}(U + UI) - \lambda_{11} \int_W^{\overline{W}_1} \overline{F}_1(x)\,dx - \lambda_{12} \int_W^{\overline{W}_2} \overline{F}_2(x)\,dx,$$
$$(1 + \lambda_{10} s) w_1'(W) = r + \lambda_{10} + \lambda_{11} \overline{F}_1(W_1(w)) + \lambda_{12} \overline{F}_2(W_1(w)),$$
$$\ell_1(W) = \frac{1}{n_1}\frac{h_1(W)}{d_1(W)} = \frac{1}{n_1}\frac{\lambda_{01} u + \lambda_{11} m_1 G_1(W) + \lambda_{21} m_2 G_2(W)}{\lambda_{10} + \lambda_{11}\overline{F}_1(W) + \lambda_{12}\overline{F}_2(W)},$$
$$h_1'(W) = \lambda_{11} n_1 \ell_1(W) F_1'(W) + \lambda_{21} n_2 \ell_2(W) F_2'(W),$$

with similar expressions for the informal section.

4. Then calculate productivity distributions

$$\Gamma_1(K_1^{-1}(W)) = n_1 F_1(W),$$
$$\Gamma_2(K_2^{-1}(W)) = n_2 F_2(W).$$

5. Consistency with the predetermined distribution of productivity Γ_0 requires that

$$\Gamma_0(p) = \begin{cases} \Gamma_0(\underline{p}_2) + \Gamma_2(p), & \forall p \in [\underline{p}_2, \underline{p}_1], \\ \Gamma_0(\underline{p}_2) + \Gamma_1(p) + \Gamma_2(p), & \forall p \in [\underline{p}_1, \overline{p}_2], \\ \Gamma_0(\underline{p}_2) + n_2 + \Gamma_1(p), & \forall p \in [\overline{p}_2, \overline{p}_1], \end{cases}$$

with $\Gamma_0(\underline{p}_2) + n_1 + n_2 = 1$.

6. If this consistency restriction is not satisfied, reiterate that sequence with another guess of F_1, F_2 and n_1, n_2.

In practice we discretise functions and approximate integrals as described in the estimation section, and we search for discrete approximations of F_1 and F_2, as well as shares n_1, n_1 so as to minimize a distance between Γ_0 and its prediction. The dimensionality of the optimization problem can be reduced by using simple parametric approximations for F_1, F_2 such as the beta distribution used in the estimation section.

C Estimation

Let F_1 and F_2 be two candidate offer distributions, defined on the spaces of contract present values. Although we could implement this procedure nonparametrically, we use non standard beta distributions as approximations:

$$F_j(x) \quad = \quad \text{betacdf}\left(\frac{x - \underline{W}_j}{\overline{W}_j - \underline{W}_j}; \alpha_j, \beta_j \right) \; j = 1,2; \; \underline{W}_j \leq x \leq \overline{W}_j$$

where $\text{betacdf}(\cdot; \alpha, \beta)$ is the CDF of a beta distribution with parameters α and β. An important practical reason why a (flexible) parametric specification is useful is that, in order to calculate the function Φ and other transition rates (see below) we need to calculate offer densities $f_1 = F_1'$ and $f_2 = F_2'$. Assuming a parametric specification guarantees the smoothness of both the distribution function and its derivative.

To estimate the parameters we use the method of moments. We match the distribution of wages for each sector and the transition rates implied by the model to those observed in the data. Given the above specification, we need to estimate the six arrival rates and the two job destruction rates all denoted by $\boldsymbol{\lambda} = (\lambda_{ij})_{i,j=0,1,2}$ and six further parameters $\boldsymbol{\theta} = (\underline{W}_1, \overline{W}_1, \underline{W}_2, \overline{W}_2, \alpha_1, \beta_1, \alpha_2, \beta_2)$ characterizing the offer distribution. Our algorithm estimates $\boldsymbol{\theta}$ given the $\boldsymbol{\lambda}$, then $\boldsymbol{\lambda}$ given $\boldsymbol{\theta}$, and iterates until convergence. Although we could

estimate all parameters at the same time, this turned out to be a very quick procedure in practice.

C.1 Contract Offer Distributions

We start by taking the arrival rates $\boldsymbol{\lambda}$ as given to estimate $\boldsymbol{\theta}$ as follows. Let $z_k = \cos(k\pi/M), k = 0,...,M$, be $M+1$ Chebychev nodes on $[-1,1]$. These nodes allow to define grids on $[\underline{W}_1, \overline{W}_1]$ and $[\underline{W}_2, \overline{W}_2]$ as

$$W_{jk} = \frac{\underline{W}_j + \overline{W}_j}{2} + \frac{\underline{W}_j - \overline{W}_j}{2} z_k, \ j = 1,2, \ k = 0,...,M.$$

For each point on the grids, one can calculate a corresponding wage w_{jk} using equations (15) and (16), and replacing integrals by quadrature approximations. The appropriate quadrature for Chebychev nodes is the Clenshaw-Curtis (CC) quadrature, whose weights ω_k can be easily calculated using Fast Fourier Transform (FFT) (see Waldvogel, 2006). For example, we have

$$\begin{aligned}(1+\lambda_{10}s)w_{1k} = (r+\lambda_{10})W_{1k} - \lambda_{10}(\underline{W}_2 + UI) \\ - \lambda_{11}\frac{\underline{W}_1 - \overline{W}_1}{2}\sum_{n=0}^{N}\omega_n \mathbf{1}_{(W_{1n}>W_{1k})}\overline{F}_1(W_{1n}) \\ - \lambda_{12}\frac{\underline{W}_2 - \overline{W}_2}{2}\sum_{n=0}^{N}\omega_n \mathbf{1}_{(W_{2n}>W_{1k})}\overline{F}_2(W_{2n}),\end{aligned}$$

where $\mathbf{1}_{(\cdot)}$ is the indicator function. A similar expression can be obtained to determine wage nodes for the informal sector, w_{2k}.

Then we search for $\boldsymbol{\theta}$ minimising

$$Q_1(\boldsymbol{\theta}|\boldsymbol{\lambda}) = \sum_{j=1,2}\sum_{k=0}^{M}\left(\widehat{G}_j^*(w_{jk}) - G_j(W_{jk})\right)^2,$$

where $G_j(W_{jk})$ is calculated using equations (6) and (7), and replacing integrals by CC-quadrature approximations, and $\widehat{G}_j^*$ is an estimate of wage distribution functions,.

Note that, assuming that $U = \underline{W}_2 \leq \underline{W}_1$ and $\overline{W}_2 \leq \overline{W}_1$, we have

$$(1+\lambda_{10}s)\underline{w}_1 = (r+\lambda_{10})\underline{W}_1 - \lambda_{10}(\underline{W}_2 + UI) - \lambda_{11}(\mu_1 - \underline{W}_1) - \lambda_{12}\int_{\underline{W}_1}^{\overline{W}_2}\overline{F}_2(x)\,dx, \quad \text{(C.1)}$$

$$(1+\lambda_{10}s)\overline{w}_1 = (r+\lambda_{10})\overline{W}_1 - \lambda_{10}(\underline{W}_2 + UI), \quad \text{(C.2)}$$

$$\underline{w}_2 = r\underline{W}_2 - \lambda_{21}(\mu_1 - \underline{W}_2) - \lambda_{22}(\mu_2 - \underline{W}_2), \quad \text{(C.3)}$$

$$\overline{w}_2 = (r+\lambda_{20})\overline{W}_2 - \lambda_{20}\underline{W}_2 - \lambda_{21}\int_{\overline{W}_2}^{\overline{W}_1}\overline{F}_1(x)\,dx, \quad \text{(C.4)}$$

where $[\underline{w}_1, \overline{w}_1]$ and $[\underline{w}_2, \overline{w}_2]$ are the observed wage supports in the formal and informal sec-

tors, respectively, and with

$$\mu_1 = \underline{W}_1 + (\overline{W}_1 - \underline{W}_1)\frac{\alpha_1}{\alpha_1+\beta_1},$$
$$\mu_2 = \underline{W}_2 + (\overline{W}_2 - \underline{W}_2)\frac{\alpha_2}{\alpha_2+\beta_2}.$$

Hence, we can simplify the estimation problem slightly by using equations (C.2) and (C.3) to substitute observed wage bounds $\underline{w}_2$ and $\overline{w}_1$ for $\underline{W}_2 = U$ and $\overline{W}_1$ (given the α, β and $\underline{W}_1, \overline{W}_2$).

C.2 Transition Rates

In a similar way as we estimate $\boldsymbol{\theta}$ given $\boldsymbol{\lambda}$, we can estimate $\boldsymbol{\lambda}$ given $\boldsymbol{\theta}$. Natural counterparts to the theoretical transition rates can be calculated from observed flows between states (0: unemployment; 1: working in the formal sector; and 2: working in the informal sector).

From the labour force survey, we calculate the intensity of transitions from unemployment to job ($\widehat{D}_{0j}$; $j = 1,2$), from a formal sector job to unemployment, to another job in the same sector or to the informal sector ($\widehat{D}_{1j}$; $j = 0,1,2$) and similar ones for a workers initially in the informal sector ($\widehat{D}_{2j}$; $j = 0,1,2$). We then estimate our transition parameters using the method of moments. We choose the parameters so as to match the observed transition rates between sectors.

Consider first the workers who are unemployed at the date of the first interview, that we follow over T periods. Workers are not heterogeneous in this model and hence the remaining unemployment duration is exponentially distributed. Thus the implied proportion of those who move out of unemployment and into a job in sector j over the time period of observation T is

$$D_{0j} = \frac{\lambda_{0j}}{\lambda_{01}+\lambda_{02}}(1 - e^{-(\lambda_{01}+\lambda_{02})T}),\ j = 1,2 \tag{C.5}$$

Now consider workers in the formal sector. Over T periods the proportion making a transition to an alternative job in the same sector, to a job in the informal sector or to unemployment is, respectively

$$\begin{aligned} D_{11} &= \int_{\underline{W}_1}^{\overline{W}_1} \frac{\lambda_{11}\overline{F}_1(x)}{d_1(x)}(1 - e^{-d_1(x)T})dG_1(x), \\ D_{12} &= \int_{\underline{W}_1}^{\overline{W}_1} \frac{\lambda_{12}\overline{F}_2(x)}{d_1(x)}(1 - e^{-d_1(x)T})dG_1(x), \\ D_{10} &= \int_{\underline{W}_1}^{\overline{W}_1} \frac{\lambda_{10}}{d_1(x)}(1 - e^{-d_1(x)T})dG_1(x). \end{aligned} \tag{C.6}$$

where $d_1(W) = \lambda_{10} + \lambda_{11}\overline{F}_1(W) + \lambda_{12}\overline{F}_2(W)$. Now, in equilibrium,

$$\ell_1(x) = \frac{1}{n_1}\frac{h_1(x)}{d_1(x)} = \frac{m_1}{n_1}\frac{dG_1(x)}{dF_1(x)},$$

allowing to replace the derivative of G_1 by that of F_1 inside the integral. This is useful as we have parametrised F_1, F_2 using a continuous distribution. Then CC-quadrature can be used to approximate the integral.

Similarly the corresponding transition rates for those observed working initially in the informal sector are

$$\begin{aligned} D_{22} &= \int_U^{\overline{W}_2} \frac{\lambda_{22}\overline{F}_2(x)}{d_2(x)}(1-e^{-d_2(x)T})dG_2(x), \\ D_{21} &= \int_U^{\overline{W}_2} \frac{\lambda_{21}\overline{F}_1(x)}{d_2(x)}(1-e^{-d_2(x)T})dG_2(x), \\ D_{20} &= \int_U^{\overline{W}_2} \frac{\lambda_{20}}{d_2(x)}(1-e^{-d_2(x)T})dG_2(x). \end{aligned} \tag{C.7}$$

with $d_2(W) = \lambda_{10} + \lambda_{11}\overline{F}_1(W) + \lambda_{12}\overline{F}_2(W)$.

These are the model counterparts for these empirical moments as functions of the arrival rates, the job destruction rates, the offers distributions F_i and as a function of the equilibrium contract values distributions G_i $(i = 1,2)$. Contract offers and equilibrium distributions are related by a complex function as explained in Appendix A.

We can thus estimate $\boldsymbol{\lambda}$ given $\boldsymbol{\theta}$ by minimising the criterion

$$Q_2(\boldsymbol{\lambda}|\boldsymbol{\theta}) = \sum_{i,j=0,1,2} \left(\widehat{D}_{ij} - D_{ij}\right)^2,$$

where $\widehat{D}_{ij}$ is the empirical counterpart of D_{ij}.

We could minimize the two criteria Q_1 and Q_2 jointly but it is numerically faster to use a nested algorithm.

C.3 Productivity Distribution

Up to this point, there has been no need to use the firm profit functions, or indeed the distribution of productivities. To complete estimation we need to estimate the cost function of informality. This will allow us to characterize the choice of firms to locate in either sector and ultimately to carry out counterfactual simulations.

We specify the cost function as $C = c\ell_2(W)^\gamma$, with C_1 and γ being the parameters to be estimated. Given values for c and γ, and for n_1 and n_2 such that $n_1 + n_2 = 1$, we solve for the labour force size in the formal sector ($\ell_1(W) = \frac{1}{n_1}\frac{h_1(W)}{d_1(W)}$) and in the informal sector ($\ell_2(W) = \frac{1}{n_2}\frac{h_2(W)}{d_2(W)}$). From the firm's maximization problem in each sector, we next derive the support of the distribution of formal and informal productivities, i.e. $p_1 = K_1^{-1}(W)$ and $p_2 = K_2^{-1}(W)$ respectively. The first order conditions for the firm's optimization problem

(see (13), (14)) gives

$$p_1 = K_1^{-1}(W) = (1+\tau+\lambda_{10}s)[w_1(W) + w_1'(W)\frac{\ell_1(W)}{\ell_1'(W)}], \tag{C.8}$$

$$p_2 = K_2^{-1}(W) = w_2(W) + w_2'(W)\frac{\ell_2(W)}{\ell_2'(W)} + C_1\gamma\ell_2(W)^{\gamma-1}, \tag{C.9}$$

where the expressions for $w_1'(W)$, $i = 1,2$, are given by

$$\begin{aligned} w_1'(W) &= \frac{r+\lambda_{10}+\lambda_{11}\overline{F}_1(W_1(w))+\lambda_{12}\overline{F}_2(W_1(w))}{1+\lambda_{10}s}, \\ w_2'(W) &= r+\lambda_{20}+\lambda_{21}\overline{F}_1(W_2(w))+\lambda_{22}\overline{F}_2(W_2(w)), \end{aligned}$$

and where firm sizes can be differentiated using

$$\begin{aligned} h_1'(W) &= \lambda_{11}m_1G_1'(W)+\lambda_{21}m_2G_2'(W) \\ &= \lambda_{11}n_1\ell_1(W)F_1'(W)+\lambda_{21}n_2\ell_2(W)F_2'(W), \end{aligned}$$

with a similar expression for $h_2'(W)$.

For each point of the contract grids, W_{jk}, one can thus calculate a point p_{jk} on a productivity grid, with $\underline{p}_2 = p_{20}$, $\underline{p}_1 = p_{10}$, $\overline{p}_2 = p_{2N}$ and $\overline{p}_1 = p_{1N}$. This in turn allows to tabulate productivity distributions as

$$\Gamma_j(p_{jk}) = n_j \cdot F_j(W_{jk}),\ j = 1,2, k = 0,...,N.$$

Equilibrium conditions require that $\pi_2(\underline{p}_2) = 0$, and $\pi_1(p) = \pi_2(p) > 0$ for $p \in [\underline{p}_1, \overline{p}_2]$. We thus estimate C_1 and γ, and n_1 and n_2 such that $n_1 + n_2 \leq 1$, so as to minimize

$$\pi_2(p_{20})^2 + \sum_{k,k'=0}^{M} \mathscr{K}(p_{1k} - p_{2k'})[\pi_1(p_{1k}) - \pi_2(p_{2k'})]^2,$$

where $\mathscr{K}$ is a kernel density.

Lastly, the unconditional productivity distribution Γ_0 follows as

$$\Gamma_0(p) = \begin{cases} \Gamma_2(p), & \forall p \in [\underline{p}_2, \underline{p}_1], \\ \Gamma_1(p)+\Gamma_2(p), & \forall p \in [\underline{p}_1, \overline{p}_2], \\ n_2+\Gamma_1(p), & \forall p \in [\overline{p}_2, \overline{p}_1], \end{cases}$$

with the additional restriction: $\Gamma_0(\underline{p}_2) = 0$.

D Simulation Results

TABLE D.1

Effects on the composition of workforce, firm size and welfare, of changes in the informality cost, of eliminating the informal sector - Sao Paulo, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
m_1	0.49	0.52	0.69	0.89	0.87
m_2	0.30	0.28	-	-	-
u	0.21	0.20	0.31	0.11	0.13
n_1	0.31	0.31	0.94	0.94	0.94
n_2	0.69	0.69	-	-	-
Formal firm size (Mean)	16.0	16.8	7.4	9.6	9.3
Informal firm size (Mean)	4.3	4.1	-	-	-
Welfare (Reais($) per month)					
Formal worker $[rE(W_1)]$	545.8	548.0	341.6	561.5	519.5
Informal worker $[rE(W_2)]$	440.5	434.7	-	-	-
Unemployed $[rU]$	402.7	401.6	231.0	452.3	405.6
Average worker $[r(uU + m_1E(W_1) + m_2E(W_2))]$	484.8	486.7	307.5	550.0	504.6
Formal firm $[E(\pi_1)]$	1,252.1	1,309.1	657.9	544.2	523.7
Informal firm $[E(\pi_2)]$	133.7	121.5	-	-	-
Average firm $[N_1E(\pi_1) + N_2E(\pi_2)]$	480.4	489.7	616.1	509.6	490.4
Total (Workers + Firms)	965.2	976.4	923.6	1,059.6	995.0
Government Revenue (formal sector)	457.7	438.6	483.2	484.6	483.2
Government Revenue (informal sector)	26.8	45.8	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.2

Effects on wages and overall wage inequality - Sao Paulo, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Wages (log)					
P10	6.08	6.07	5.14	5.88	5.58
P25	6.29	6.29	5.49	6.16	6.13
Median	6.47	6.47	5.91	6.44	6.36
P75	6.63	6.63	6.24	6.58	6.54
P90	6.77	6.77	6.51	6.75	6.69
Mean	6.51	6.49	6.02	6.41	6.35
Informal Wages (log)					
P10	5.50	5.54	-	-	-
P25	5.89	5.88	-	-	-
Median	6.18	6.14	-	-	-
P75	6.50	6.35	-	-	-
P90	6.67	6.60	-	-	-
Mean	6.29	6.18	-	-	-
Overall Wage Inequality					
P75/P25	1.62	1.61	2.10	1.51	1.50
P90/P10	2.79	2.42	3.92	2.39	3.03

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.3

Effects on the distribution of productivity - Sao Paulo, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Productivity (log)					
P10	6.40	6.37	6.20	6.28	6.25
P25	6.57	6.58	6.35	6.44	6.44
Median	6.76	6.78	6.52	6.62	6.64
P75	6.95	7.03	6.83	6.85	6.82
P90	7.29	7.26	7.22	7.06	7.00
Mean	7.02	7.04	7.06	6.77	6.76
Informal Productivity (log)					
P10	6.11	6.11	-	-	-
P25	6.18	6.17	-	-	-
Median	6.34	6.31	-	-	-
P75	6.62	6.43	-	-	-
P90	6.84	7.29	-	-	-
Mean	6.48	6.45	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.4

Effects on the composition of workforce, firm size and welfare, of changes in the informality cost, and of eliminating the informal sector - Salvador, Low Education Males

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
m_1	0.49	0.50	0.73	0.90	0.86
m_2	0.33	0.32	-	-	-
u	0.18	0.18	0.27	0.10	0.14
n_1	0.32	0.40	0.63	0.63	0.63
n_2	0.68	0.60	-	-	-
Formal firm size (Mean)	13.7	11.3	10.4	13.0	12.3
Informal firm size (Mean)	4.4	4.9	-	-	-
Welfare (Reais($) per month)					
Formal worker $[rE(W_1)]$	446.5	445.9	393.4	569.5	459.2
Informal worker $[rE(W_2)]$	334.8	329.8	-	-	-
Unemployed $[rU]$	300.6	298.5	215.9	430.9	318.8
Average worker $[r(uU + m_1E(W_1) + m_2E(W_2))]$	383.0	382.3	344.7	556.0	439.1
Formal firm $[E(\pi_1)]$	726.8	570.5	845.4	507.4	495.1
Informal firm $[E(\pi_2)]$	117.8	145.9	-	-	-
Average firm $[N_1E(\pi_1) + N_2E(\pi_2)]$	312.7	315.8	529.2	317.6	309.9
Total (Workers + Firms)	695.7	698.0	873.9	873.6	749.0
Government Revenue (formal sector)	304.6	274.4	338.5	338.3	338.7
Government Revenue (informal sector)	34.0	63.6	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.5

Effects on wages and overall wage inequality - Salvador, Low Education Males

		Increase in	No informal sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Wages (log)					
P10	5.59	5.61	4.27	5.71	5.09
P25	6.09	6.09	5.70	6.27	5.93
Median	6.28	6.29	6.22	6.53	6.30
P75	6.50	6.51	6.54	6.66	6.53
P90	6.63	6.63	6.71	6.76	6.65
Mean	6.32	6.28	6.17	6.44	6.25
Informal Wages (log)					
P10	5.32	5.31	-	-	-
P25	5.68	5.64	-	-	-
Median	5.94	5.89	-	-	-
P75	6.13	6.16	-	-	-
P90	6.30	6.31	-	-	-
Mean	6.00	5.91	-	-	-
Overall Wage Inequality					
P75/P25	1.75	1.78	2.31	1.47	1.81
P90/P10	3.53	3.48	5.47	2.85	4.79

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.6

Effects on the distribution of productivity - Salvador, Low Education Males

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Productivity (log)					
P10	5.79	5.77	5.37	5.33	5.29
P25	6.06	6.06	5.63	5.71	5.59
Median	6.39	6.39	6.12	6.26	6.09
P75	6.63	6.62	6.50	6.59	6.44
P90	6.78	6.77	6.84	6.75	6.69
Mean	6.47	6.46	6.42	6.30	6.23
Informal Productivity (log)					
P10	5.71	5.74	-	-	-
P25	5.89	5.90	-	-	-
Median	6.09	6.08	-	-	-
P75	6.30	6.18	-	-	-
P90	6.53	6.95	-	-	-
Mean	6.16	6.19	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.7

Effects on the composition of workforce, firm size and welfare, of changes in the informality cost, and of eliminating the informal sector - Salvador, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
m_1	0.47	0.46	0.58	0.80	0.72
m_2	0.21	0.22	-	-	-
u	0.32	0.32	0.42	0.20	0.28
n_1	0.30	0.34	0.52	0.52	0.52
n_2	0.70	0.66	-	-	-
Formal firm size (Mean)	14.2	12.2	10.0	13.8	12.5
Informal firm size (Mean)	2.7	3.0	-	-	-
Welfare (Reais($) per month)					
Formal worker $[rE(W_1)]$	394.7	374.5	377.8	436.0	372.3
Informal worker $[rE(W_2)]$	284.2	264.0	-	-	-
Unemployed $[rU]$	259.8	250.6	238.7	291.3	247.7
Average worker $[r(uU + m_1E(W_1) + m_2E(W_2))]$	328.6	310.3	319.0	406.7	337.4
Formal firm $[E(\pi_1)]$	457.9	435.8	297.7	461.4	325.1
Informal firm $[E(\pi_2)]$	43.2	69.5	-	-	-
Average firm $[N_1E(\pi_1) + N_2E(\pi_2)]$	167.6	194.1	154.5	239.5	168.7
Total (Workers + Firms)	496.2	504.3	473.5	646.2	506.1
Government Revenue (formal sector)	221.0	209.5	233.6	235.8	236.4
Government Revenue (informal sector)	15.4	26.5	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.8

Effects on wages and overall wage inequality - Salvador, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Wages (log)					
P10	5.59	5.43	5.49	5.35	5.36
P25	5.91	5.81	5.84	6.01	5.78
Median	6.12	6.04	6.06	6.22	6.03
P75	6.27	6.21	6.30	6.37	6.21
P90	6.40	6.35	6.37	6.50	6.36
Mean	6.14	6.04	6.05	6.16	6.00
Informal Wages (log)					
P10	5.23	4.78	-	-	-
P25	5.47	5.19	-	-	-
Median	5.74	5.61	-	-	-
P75	5.88	5.80	-	-	-
P90	6.01	5.98	-	-	-
Mean	5.76	5.57	-	-	-
Overall Wage Inequality					
P75/P25	1.54	1.78	1.58	1.43	1.53
P90/P10	2.60	3.16	2.39	3.16	2.72

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.

TABLE D.9

Effects on the distribution of productivity - Salvador, Low Education Females

		Increase in	No Informal Sector		
	Benchmark	C	exogenous λ's	$\eta = 0.3$	$\eta = 0.5$
Formal Productivity (log)					
P10	5.78	5.43	6.04	4.28	4.58
P25	5.94	5.81	5.96	4.83	5.30
Median	6.15	6.04	6.10	5.83	5.76
P75	6.38	6.21	6.26	6.16	6.06
P90	6.53	6.35	6.51	6.46	6.37
Mean	6.23	6.04	6.31	5.82	5.82
Informal Productivity (log)					
P10	5.58	4.78	-	-	-
P25	5.64	5.19	-	-	-
Median	5.79	5.61	-	-	-
P75	6.06	5.80	-	-	-
P90	6.27	5.98	-	-	-
Mean	5.92	5.57	-	-	-

Note: In all simulations government revenue is held constant through adjustments in corporate taxes. The cost of informality (C) is raised by 10 percent.