Trade and Informality in the Presence of Labor Market Frictions and Regulations*

Rafael Dix-Carneiro   Pinelopi Koujianou Goldberg   Costas Meghir
Gabriel Ulyssea

May 12, 2018

Very Preliminary and Incomplete. Please do note circulate.

1 Introduction

Over the past 30 years, most developing countries have entered the world market. This process was most remarkable in Latin America: during the 1980's and early 1990's, many countries drastically cut tariff and non-tariff barriers and substantially increased their participation in international trade. These changes are often celebrated as contributing to economic growth, efficiency and overall welfare. However, those who oppose globalization argue that its benefits are not evenly distributed and that it may generate adverse effects on inequality and labor market performance in these countries. Amongst the potential adverse effects from globalization, the increase in the size of the informal sector is often pointed as a particularly important one (Goldberg and Pavcnik, 2007; Harrison et al., 2003; Harrison and Scorse, 2010).

On the firm side, informality implies that firms do not comply with taxes nor the relevant regulations (e.g. labor laws). This can be harmful to the economy for two main reasons. First, it implies substantial tax evasion thus hindering fiscal capacity and the provision of public goods. Second, it might entail substantial misallocation of resources and hamper growth, as non-productive firms can survive by evading taxes and avoiding compliance with labor market regulations. On the worker side, one can broadly define

*This project is currently being supported by award SES-1629124 from the National Science Foundation and by an Early Career Research Grant from the W.E. Upjohn Institute for Employment Research. We thank seminar participants at Duke and Yale for helpful comments in preliminary presentations.

informality in two ways: The first defines a worker as informal if she does not have permanent and stable employment associated with benefits such as health and social security. The second defines a worker as informal if, in addition to not receiving benefits, she is invisible to the tax authorities and her employer illegally evades labor market regulations (including minimum wages and firing rules). The first definition has become relevant even in developed countries in recent years with the emergence of companies such as Uber, Taskrabbit or Airbnb. The second definition applies primarily to developing countries where informality, and the tax evasion associated with it, is a first-order issue and has been shown to be associated with low productivity and a barrier to growth.

Although a substantial share of the labor force in developing countries is employed informally (for example, in Latin America, this share falls between 35% in Chile and 80% in Peru), trade models have typically abstracted from informality. Recent work has shown, in different contexts, that shifts into or out of informality and non-employment constitute important margins of labor market adjustment to trade (Dix-Carneiro and Kovak (2017a) and McCaig and Pavcnik (2018)). In addition, there is evidence that the magnitude of these effects depends on the intensity with which labor market regulations are enforced (Ponczek and Ulyssea (2018)). These facts imply that understanding and measuring the labor market and welfare effects of globalization within a model of trade with informality, unemployment and regulations is a first order question.

This paper studies the labor market and welfare effects of trade in an environment with burdensome regulations but with imperfect enforcement or monitoring of these regulations. The imperfect enforcement of regulations gives incentives for firms to operate in the informal sector. We anticipate that a trade-induced reallocation of resources towards the informal sector can have opposing effects on welfare. On one hand, it would constitute a reallocation of resources towards a less productive sector that does not contribute to the provision of public goods. This suggests that an expansion of the informal sector hurts welfare. On the other hand, recent work by Dix-Carneiro and Kovak (2017a) and Ponczek and Ulyssea (2018) suggests that the informal sector served as a fallback sector to trade-displaced workers. These papers focused on the Brazilian trade liberalization of the 1990s and exploited a difference-in-difference framework. Their results suggest that had the enforcement of labor market regulations been stricter in Brazil, the effect of import competition on trade-displaced workers' employment outcomes could have been much more adverse than it actually was. These facts motivate the set of questions we address in this paper: How do labor market regulations and policies directed towards the informal sector influence the labor market effects of globalization? More specifically, how are the labor market effects of trade shaped by the constellation of firing costs, minimum wages and enforcement of labor market regulations? What is the impact of the "costs of formality", such as payroll and sales taxes and the bureaucratic cost that comes with being a formal firm?

To shed light on these questions, we build on Cosar et al. (2016) and develop a structural equilibrium model with heterogeneous firms that choose whether to operate in the formal or in the informal sector.[1] The model features a rich institutional setting, where formal firms must comply with minimum wages, and are subject to firing costs as well as payroll and revenue taxes. Taxes and labor market regulations are imperfectly enforced by the government, giving rise to incentives for some firms to be informal. The labor market is characterized by labor market frictions and costs of hiring, features leading to unemployment. The economy is composed by tradable and non-tradable sectors, and tradable sector firms are able to export. We estimate the model using several data sources, including matched employer-employee data from formal and informal firms and workers in Brazil.

Brazil constitutes a relevant case study for several reasons. First, it has strict and burdensome labor regulations that are imperfectly enforced and a large informal sector: nearly two thirds of businesses, 40% of GDP and 35% of employees are informal (Ulyssea, 2018). Second, the Brazilian case is typical of developing countries, especially in Latin America, where the urban labor force employed informally averages over 50 percent, with this number varying from 35 percent in Chile to 80 percent in Peru (Perry et al., 2007). Third, it has unique data availability and quality, allowing the direct observation of informality for workers and firms. We define as informal workers those employees who do not hold a formal labor contract, which in Brazil is sharply defined as having a booklet (*carteira de trabalho*) that registers workers' entire employment history in the formal sector. We define as informal firms those not registered with the tax authorities, which means that they do not possess the tax identification number required for Brazilian firms (*Cadastro Nacional de Pessoa Juriidica* – CNPJ). We can observe both definitions directly from the data available (more details are provided in the Data section). Finally, even though Brazil experienced a relatively fast and intense trade liberalization episode in early 1990's (e.g. Dix-Carneiro and Kovak, 2017b), it remains a relatively closed economy. Therefore, our analyses in this paper are of great policy relevance.

This paper contributes to three different literatures. First, it contributes to the litera-

[1] Another promising framework for our purposes is Kaas and Kircher (2015). Understanding differences in implications across these two frameworks is an interesting topic for future work.

ture that seeks to identify the impact of globalization on labor market outcomes and welfare in developing countries. Several papers in this literature have empirically examined the effects of trade on informality using different countries, sectors and methodologies, yielding mixed conclusions(Goldberg and Pavcnik, 2003; Bosch et al., 2012; Menezes-Filho and Muendler, 2011; Dix-Carneiro and Kovak, 2017b). In particular, Dix-Carneiro and Kovak (2017a) focus on the dynamic behavior of unemployment and informality in the aftermath of the Brazilian trade liberalization. They document that in the short run, the reforms had large effects on unemployment and small effects on informality. However, in the long run, this pattern is reversed, suggesting a potentially important role of the informal sector in smoothing out the labor market trajectories of displaced workers. Second, this project contributes to an extensive literature on the causes and consequences of informality by developing a new framework to analyze firms' and workers' decisions regarding informality (De Soto, 1989; Perry et al., 2007; La Porta and Shleifer, 2008; Bacchetta et al., 2009; Ulyssea, 2010, 2018; Meghir et al., 2015). Third, it contributes to the literature on misallocation and the role of size-dependent distortions (e.g. Hsieh and Klenow, 2009; Guner et al., 2008; Adamopoulos and Restuccia, 2014; Garicano et al., 2016).

2 Model

2.1 Set Up

The economy is populated by homogeneous, infinitely-lived workers-consumers. Individuals derive utility from the consumption of a composite good of differentiated, tradable sector goods C and from the consumption of a composite good of differentiated, non-tradable sector goods S. Preferences are given by

$$U = \sum_{t=1}^{\infty} \frac{C_t^{\zeta} S_t^{1-\zeta}}{(1+r)^t}, \tag{1}$$

where

$$C_t = \left(\int_0^{N_{Ct}} c_t(n)^{\frac{\sigma-1}{\sigma}} dn \right)^{\frac{\sigma}{\sigma-1}} \tag{2}$$

$$S_t = \left(\int_0^{N_{St}} s_t(n)^{\frac{\sigma-1}{\sigma}} dn \right)^{\frac{\sigma}{\sigma-1}} \tag{3}$$

and $\zeta \in (0,1)$ is the fraction of expenditure on tradable sector goods, $\sigma > 1$ is the elasticity of substitution across varieties within sectors, N_{kt} denotes the measure of varieties available in sector $k = C, S$ at time t, and $n \in (0, N_{kt})$ indexes varieties. As we will focus on steady state equilibria, we henceforth drop the time subscript for notational convenience.

2.2 Firms

There is a continuum of firms in both tradable and non-tradable sectors. Formal and informal firms coexist in both sectors, and each firm produces a unique variety $n \in (0, N_k)$. Firms use labor as the single input in a constant returns to scale production function: $q(z, \ell) = z\ell$, where ℓ denotes firm's employment size. Firms' idiosyncratic productivity evolves over time following the AR(1) process below:

$$\ln z' = \rho_k \ln z + \sigma_k^z \varepsilon, \tag{4}$$

where $\rho_k \in (0,1)$, $\varepsilon \sim N(0,1)$ and σ_k^z is the standard deviation of the shocks. It will be convenient to denote $G_k(z'|z)$ the cumulative distribution function of z' conditional on z and $g_k(z'|z)$ its density.[2]

Monopolistic competition implies that revenues in sector $k = C, S$ are given by:

$$R_k(z, \ell) = \left(\frac{X_k}{P_k^{1-\sigma}} \right)^{\frac{1}{\sigma}} (z\ell)^{\frac{\sigma-1}{\sigma}} \tag{5}$$

where X_k is total expenditure in sector k goods, and $P_k = \left(\int_0^{N_k} p_k(n)^{1-\sigma} dn \right)^{\frac{1}{1-\sigma}}$ is the price index for sector $k = C, S$. For the tradable sector, $X_C = \zeta I$, where I is aggregate income. For the non-tradable sector, $X_S = (1-\zeta) I + R$, where R represents expenditures on service sector goods made by firms in order to cover hiring, fixed and export costs (which we discuss below). Aggregate income is determined by total wages, government transfers and aggregate firms' profits.

Timing

Every period, formal incumbent firms must choose whether to stay or exit their industry. If the firm decides to stay, it draws its new productivity shock and must decide

[2] This process is imposed to be the same across formal and informal firms within tradable and non-tradable sectors. Unfortunately, we do not have longitudinal data on firms in the informal sector, so that this process cannot be separately identified for formal and informal firms.

to adjust or not its labor force. Informal firms face a similar problem but also have one additional option, which is to formalize their businesses. If they decide to formalize, they will then be subject to all regulatory costs faced by formal firms, namely the payroll and revenue taxes, firing costs and minimum wages. After the informal firm decides to stay informal or migrate to the formal sector, it draws its new productivity shock and must also decide whether to adjust its labor force.

The timing of events and firms' behavior is illustrated in Figure 1. Consider an informal firm which starts period t with state (z, ℓ, i). There are three initial possibilities: (i) the firm decides to stay informal and draws a new shock z'; (ii) the firm exits because it decides to, or because it is hit with an exogenous death shock (with probability α_{ki}); or (iii) the firm registers with the authorities, becomes formal, and draws a new shock z'. If the firm decides to stay active (as informal or formal), it must choose how to adjust its workforce in response to the shock z'. To do so, it posts vacancies or fires workers and ends period t with ℓ' workers. At that point, it realizes profits and starts period $t+1$ with state (z', ℓ', i), if it decided to remain informal or with state (z', ℓ', f), if it decided to become formal.

Now, consider a formal firm which starts period t with state (z, ℓ, f). The timing and sequence of events is the same as for informal firms. The only difference is that we do not allow for formal firms to become informal, and the exogenous death shock arrives with probability α_{kf}.

Hiring and Firing Costs

When deciding employment levels, both formal and informal firms in tradable and non-tradable sectors face hiring costs. These are defined by the costs of posting vacancies, which are given by the following function:

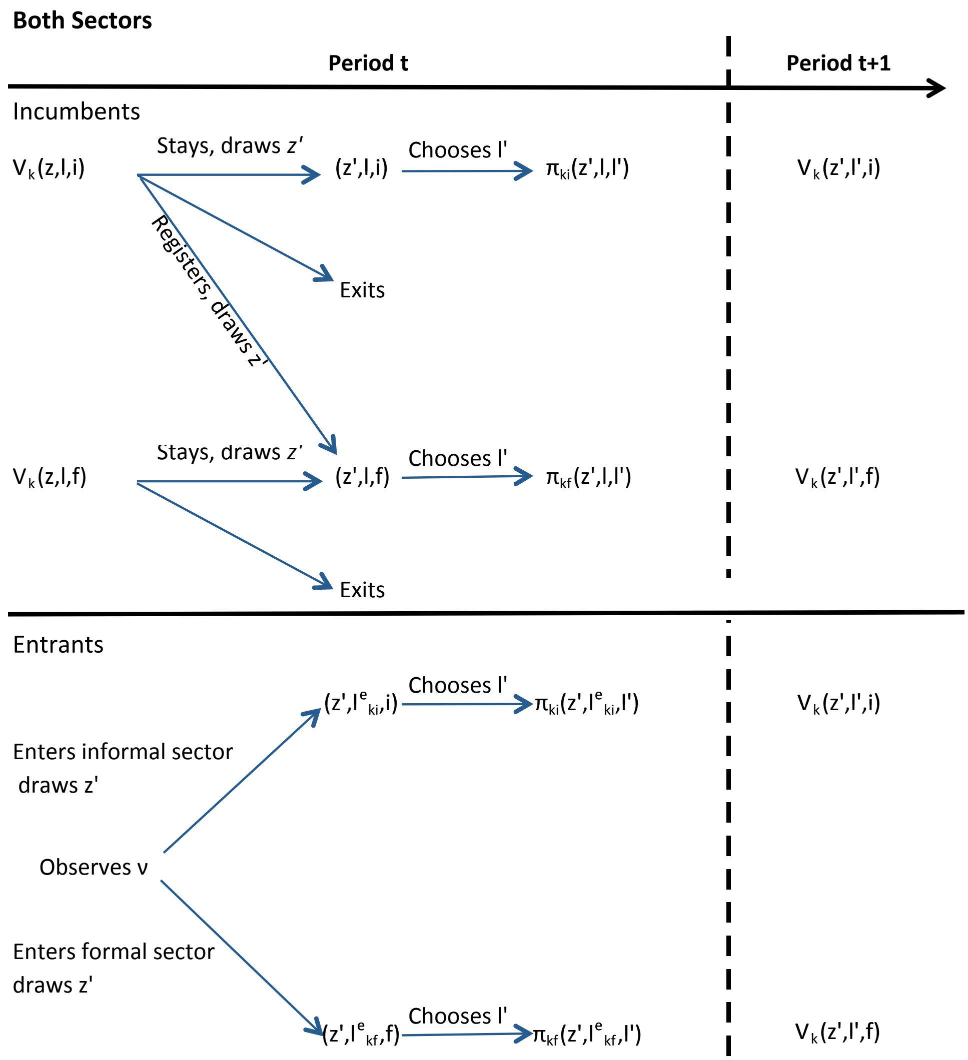
$$C^h_{kj}\left(\ell, v_{kj}\right) = \left(\frac{h_k}{\gamma_{k1}}\right)\left(\frac{v_{kj}}{\ell^{\gamma_{k2}}}\right)^{\gamma_{k1}} \tag{6}$$

where h_k, γ_{k1} and $\gamma_{k2} \in [0,1]$ for $k = C, S$ and $j = i, f$, are parameters. The γ_{kj1} determines the convexity in hiring costs and γ_{k2} captures economies of scale in hiring.[3]

Expanding from ℓ to ℓ' therefore requires posting $v_{kj} = \frac{\ell' - \ell}{\mu^v}$ vacancies, where μ^v is the probability of filling a vacancy. The cost of expanding from ℓ to ℓ' workers for a formal

[3]Note that the cost of hiring functions are the same for formal and informal firms within sector. Identification of h_k, γ_{k1} and γ_{k2} relies on longitudinal data, which is unavailable for informal firms.

Figure 1: Diagram of Firms' Behavior



firm is therefore given by:

$$H_{kj}\left(\ell,\ell'\right)=\left(\mu^{v}\right)^{-\gamma_{k1}}\left(\frac{h_k}{\gamma_{k1}}\right)\left(\frac{\ell'-\ell}{\ell^{\gamma_{k2}}}\right)^{\gamma_{k1}} \tag{7}$$

The functional form of the hiring cost function is important for a couple of reasons. First, depending on the estimate of the scale parameter γ_{k2} , it is possible to generate the stylized fact that firm-level growth rates in employment decline with size. To obtain some intuition, suppose that $\gamma_{k2}=0$. In that case, all firms posting v vacancies face the same hiring costs, irrespective of their size. On the other hand, if $\gamma_{k2}=1$, then all firms face the same cost of a given employment growth rate. For values of γ_{k2} between 0 and 1, larger firms face a higher cost if they want to grow their employment by a particular rate. So, in the event of a positive shock, larger firms will grow less, and in the event of a negative shock, they will also downsize less (as they anticipate large hiring costs if they are hit with a positive shock in the future). Second, the parameter γ_{k1} governs the convexity of the hiring function. If $\gamma_{k1}>1$, then hiring costs are convex. Allowing for convexity is important for the model to be able to generate wage dispersion. In this type of model, linear hiring costs lead to no wage dispersion. This is because, as we show later when we discuss the wage determination process, in our framework wages are proportional to average revenue per worker, which is – by virtue of our assumptions - proportional to marginal worker revenue. Optimizing firms set marginal revenue equal to marginal cost of an additional worker. But with linear hiring costs, the marginal cost is constant and equal across firms, so that wages will also be equalized across firms. In contrast, with convex hiring costs, the marginal cost of an additional worker is increasing in the growth of employment, so that expanding firms will pay higher wages.

Regarding firing costs, since they are entirely driven by labor market regulation, we assume that only formal firms are subject to them and they are determined as follows:

$$F\left(\ell,\ell'\right)=\kappa\left(\ell-\ell'\right) \tag{8}$$

where $\kappa>0$ is the parameter governing the firing cost function. We assume that firing costs are equal across the C and S sectors, which is consistent with the Brazilian labor regulation. We also assume that firing costs are collected by the government and are rebated back to consumers, while the hiring costs are incurred in terms of the service sector composite good.

Profit and Value Functions

Formal firms are subject to payroll and revenue taxes, firing costs and the minimum wage regulation. The profit function of a formal firm in sector $k = C, S$ is given by:

$$\pi_{kf}\left(z', \ell, \ell'\right) = \left(1-\tau_y\right) R_k\left(z', \ell'\right) - C_{kf}\left(z', \ell, \ell'\right) - \bar{c}_{kf}, \tag{9}$$

where $R_k\left(z', \ell'\right)$ denotes the revenue function; $\bar{c}_{kf}$ denotes a per-period, fixed cost of operation, which we assume that it is incurred in terms of the service sector composite good; τ_y is a sales/revenue tax, collected by the government and rebated to consumers.

Due to hiring and firing costs, the total cost function for a formal firm adjusting from ℓ to ℓ' workers is given by the following expression:

$$C_{kf}\left(z', \ell, \ell'\right) = \begin{cases} \left(1+\tau_w\right)\max\left\{w_{kf}\left(z', \ell'\right), \underline{w}\right\}\ell' + H_{kf}\left(\ell, \ell'\right) & \text{if } \ell' > \ell \\ \left(1+\tau_w\right)\max\left\{w_{kf}\left(z', \ell'\right), \underline{w}\right\}\ell' + \kappa\left(\ell - \ell'\right) & \text{if } \ell' \leq \ell \end{cases} \tag{10}$$

where $w_{kf}\left(z', \ell'\right)$ denotes the wage of workers in a formal firm with productivity z' and size ℓ', $\underline{w}$ denotes the minimum wage and τ_w is the payroll tax, which is assumed to be collected by the government and rebated to consumers.

Since formal firms have to choose to stay or leave their industry, their value function is given by:

$$V_k\left(z, \ell, f\right) = \left(1-\alpha_{kf}\right)\max\left\{0, E_{z'|z}\max_{\ell'}\left\{\pi_{kf}\left(z', \ell, \ell'\right) + \frac{1}{1+r}V_k\left(z', \ell', f\right)\right\}\right\} \tag{11}$$

where α_{kf} denotes the exogenous death probability that firms face every period for $k = C, S$.

Even though informal firms do not have to incur in any of the regulatory costs (taxes, minimum wages, firing costs), they face a probability of detection by government authorities, which is (presumably) increasing in their size (measured by their number of employees). Therefore, we allow that the cost of being informal depends on firm size, which is a common formulation in the literature (see Ulyssea, 2018, and the references therein). The intuition for this assumption is that as firms grow larger, they become more visible to the government and therefore are inspected with higher probability, which entails costs in the form of fines and bribes, or can lead to the firm shutting down its operations. Similarly, this assumption captures the idea that the opportunity costs of informality increase as the firm becomes larger because it might want to access the formal financial market (e.g. credit lines), issue invoices and expand its costumers base.

Informal firm's profit function is thus given by:

$$\pi_{ki}(z',\ell,\ell') = (1-p_{ki}(\ell'))\,R_k(z',\ell') - C_{ki}(z',\ell,\ell') - \overline{c}_{ki}, \tag{12}$$

where $p_{ki}(\ell')$ summarizes the costs associated to informality, which are assumed to be proportional to firm's revenues. We impose that

$$p_{ki}(\ell') = \max\left\{\min\left\{a_k + b_k(\ell')^{c_k}, 1\right\}, 0\right\}. \tag{13}$$

Since informal firms are not subject to firing costs, their cost function is given by:

$$C_{ki}(z',\ell,\ell') = \begin{cases} w_{ki}(z',\ell')\,\ell' + H_{ki}(\ell,\ell') & \text{if } \ell' > \ell \\ w_{ki}(z',\ell')\,\ell' & \text{if } \ell' \leq \ell \end{cases} \tag{14}$$

where $w_{ki}(z',\ell')$ denotes the wage of workers in an informal firm with productivity z' and size ℓ'.

Informal firms' value functions are similar to formal firms', except that they have the additional option to formalize their businesses. The informal value functions are therefore given by:

$$V_k(z,\ell,i) = (1-\alpha_{ki})\max\left\{\begin{array}{c} 0, E_{z'|z}\max_{\ell'}\left\{\pi_{ki}(z',\ell,\ell') + \frac{1}{1+r}V_k(z',\ell',i)\right\}, \\ E_{z'|z}\max_{\ell'}\left\{\pi_{kf}(z',\ell,\ell') + \frac{1}{1+r}V_k(z',\ell',f)\right\} \end{array}\right\}. \tag{15}$$

Entry

Firm entry is illustrated in the lower panel of Figure 1. Every period there is a pool of potential entrants into the tradable and non-tradable sectors. These potential entrants observe a pre-entry signal of how productive they will be if they decide to enter, denoted by ν, which is drawn from the ergodic distribution of z'. They can choose to enter as a formal or an informal firm, and the decision to enter is made solely based on ν. Once they enter, they draw their actual productivity, z', from:

$$\ln z' = \rho_k \ln \nu + \sigma_k^z \varepsilon.$$

which is analog to incumbents' productivity process, described in expression (4).

Once entry occurs and entrants draw their actual productivity, z', they start behaving as incumbents. Formal and informal entrants start their first period with workforce 1 and we assume that the recruitment costs of these initial workforces are included in the fixed entry costs. The value functions for entrants in either sector are given by:

$$V_k^e(\nu, j) = E_{z'|\nu} \max_{\ell' \geq 1} \left\{ \pi_{kj}(z', 1, \ell') + \frac{1}{1+r} V_k(z', \ell', j) \right\} \tag{16}$$

where $j = i, f$.

The entry conditions into the informal and formal sectors are given by the following inequalities, respectively:

$$V_k^e(\nu, i) \geq \max\{0, V_k^e(\nu, f)\} \tag{17}$$

$$V_k^e(\nu, f) \geq \max\{0, V_k^e(\nu, i)\} \tag{18}$$

and if there is positive entry in both sectors, these conditions hold with equality at $\overline{\nu}_{ki}$ and $\overline{\nu}_{kf}$, where $\overline{\nu}_{kf} > \overline{\nu}_{ki}$. These thresholds denote the marginal entrants in the informal and formal sectors, respectively.

2.3 Labor Market Frictions

Formal and informal labor markets are characterized by search and matching frictions, which prevent unemployed workers to immediately find open vacancies. We assume undirected search, and therefore unemployed workers form a unique pool of individuals who are randomly matched with formal or informal firms in one of the sectors $k = C, S$. Thus, formal and informal firms operating in tradable and non-tradable sectors compete for workers in the labor market. Given the total number of vacancies posted in the economy, $V = V_{Cf} + V_{Ci} + V_{Sf} + V_{Si}$, and the mass of unemployed workers searching for jobs, U, the total number of matches that are formed is given by:[4]

$$m(V, U) = \frac{VU}{(V^\theta + U^\theta)^{1/\theta}}, \tag{19}$$

and matches are split across sectors proportionally, $m_{kj} = \frac{V_{kj}}{V} m(V, U)$. This implies that firms in all sectors face the same probability of filling a vacancy, which is given by

$$\mu_{kj}^v = \frac{m_{kj}}{V_{kj}} = \frac{U}{(V^\theta + U^\theta)^{1/\theta}} \equiv \mu^v, \tag{20}$$

[4] This functional form ensures that matching probabilities are bound between 0 and 1 (see Cosar et al., 2016, and the references therein).

which highlights the fact that formal firms directly compete with informal firms in the labor market. In contrast, unemployed workers face job finding probabilities that differ across sectors:

$$\mu^{e}_{kj}(\psi_{kj}) = \frac{m_{kj}}{U} = \psi_{kj}\mu^{v}, \tag{21}$$

where $\psi_{kj} \equiv \frac{V_{kj}}{U}$ denotes the sector specific labor market tightness.

2.4 Wages

We assume that workers collectively bargain with their employer, after hiring costs are sunk and matching has taken place. More concretely, we assume that workers collectively bargain with their firms in a "all in or all out" fashion. To simplify exposition, we refer to workers as "unions". The surpluses of a formal firm in sector k, and the union it faces are given by, respectively:

$$S^{e}_{kf}(z,\ell) = (1-\tau_y) R_k(z,\ell) - (1+\tau_w) w_{kf}(z,\ell)\ell - \overline{c}_{kf} + \frac{1}{1+r} V_k(z,\ell,f) \tag{22}$$

$$S^{u}_{kf}(z,\ell) = \left[w_{kf}(z,\ell) + \frac{1}{1+r} J^{e}_{k}(z,\ell,f) - \left(b + b^{u} + \frac{1}{1+r} J^{u} \right) \right] \ell \tag{23}$$

where b denotes the utility flow from being unemployed and b^u the value of unemployment benefits, which are only received by formal workers.

We assume that if all workers leave, the firm exits, and that fixed operating costs are incurred after the bargaining process. Let β_f be the bargaining power of workers in the formal sector, the outcome of bargaining is given by:

$$(1-\beta_f) S^{e}_{kf}(z,\ell) = \beta_f S^{u}_{kf}(z,\ell) \tag{24}$$

Substituting expressions (22) and (23) into (24), and assuming that the current surplus is shared the same way as future surpluses (as in, for example, Bertola and Garibaldi, 2001; Cosar et al., 2016), one obtains the following wage functions for formal workers:

$$w_{kf}(z,\ell) = \frac{(1-\beta_f)(b+b^u)}{1+\beta_f\tau_w} + \frac{\beta_f(1-\tau_y)}{1+\beta_f\tau_w}\frac{R_k(z,\ell)}{\ell} - \frac{\beta_f}{1+\beta_f\tau_w}\frac{\overline{c}_{kf}}{\ell} \tag{25}$$

Wages in the informal sector are determined in a similar way. Let the bargaining power parameter be denoted by β_i, where we allow the bargaining power of formal and informal workers to be different. These could differ due to institutional reasons, such as the existence of a centralized union or labour courts, or because informal workers and firms have greater flexibility to negotiate wages. Since these will be directly estimated, the question of whether these bargaining power parameters are indeed different is an empirical one. Following the same steps as above, it is straightforward to obtain:

$$w_{ki}(z,\ell) = (1-\beta_i)\,b + \beta_i\,(1-p_{ki}(\ell))\,\frac{R_k(z,\ell)}{\ell} - \beta_i\frac{\overline{c}_{ki}}{\ell} \tag{26}$$

where the major differences relatively to expression (25) are the absence of unemployment benefits (b^u), payroll and revenue taxes (τ_w and τ_y, respectively); and the presence of the cost of informality function, $p_{ki}(\ell)$.

Expressions (25) and (26) are intuitive: wages are directly increasing with sales per worker, and the slope is larger if bargaining power is larger. An alternative to this wage setting would be to assume a somewhat more common structure *a la* Stole and Zwiebel (1996), where firms bargain with all of their workers simultaneously and continuously in a one-to-one basis, treating each worker as the marginal one. However, the present formulation generates a richer wage distribution that fits much better the degree of wage dispersion found in the data. Frameworks *a la* Stole and Zwiebel (1996) tend to generate less realistic distributions, as they imply that, for example, all firms that are willing to downsize pay the same wage to all workers (which is equal to workers' reservation wage). Additionally, the present wage setting framework implies wage schedules that are very close to those in the rent sharing literature (e.g. Card et al., 2018) and commonly found in trade models, such as Helpman and Itskhoki (2010).

2.5 Open Economy

We now extend the model to the open economy case. We assume that the home country is small relative to the rest of the world and therefore foreign conditions do not react to its policies. In the following analysis, we drop the formal/informal qualifier in order to simplify notation, as we assume throughout that informal firms cannot export.[5] In what follows, it will be convenient to re-write domestic revenues (Equation (5)) as $R_k(z,\ell) = D_{H,k}^{\frac{1}{\sigma}} q(z,\ell)^{\frac{\sigma-1}{\sigma}}$, where $k = C, S$, $q(z,\ell) = z\ell$, and $D_{H,k} = \frac{X_k}{P_k^{1-\sigma}}$. Since the focus in this section lies on the tradable sector only, and for the sake of notation simplicity, we drop the subscript $k = C, S$ for the reminder of this subsection.

[5] This assumption comes from the fact that firms that are not registered cannot undertake the necessary legal and bureaucratic procedures to export.

Price Indices and Aggregates

The price index in the non-tradable sector remains the same, but in the tradable sector it is modified to account for trade. First, we characterize the price index of imports denominated in home-currency:

$$P_F = \epsilon\tau_a\tau_c \left(\int_0^{N_F} p^*(n)^{1-\sigma}\, dn\right)^{\frac{1}{1-\sigma}} = \epsilon\tau_a\tau_c$$

where $p^*(n)$ is the free on board (FOB) price of imported variety n, denominated in foreign currency; N_F denotes the mass of imported varieties; ϵ is the exchange rate, $\tau_a - 1 > 0$ is the ad-valorem tariff and $\tau_c > 1$ the iceberg trade cost. The second equality in the above expression comes from the normalization $\left(\int_0^{N_F} p^*(n)^{1-\sigma}\, dn\right)^{\frac{1}{1-\sigma}} \equiv 1$. This is without loss of generality, as this term is exogenous to our model given the small open economy assumption. The price index of domestically produced varieties $n \in (N_F, N]$ is given by:

$$P_H = \left(\int_{N_F}^{N} p(n)^{1-\sigma}\, dn\right)^{\frac{1}{1-\sigma}}$$

and the price index for the composite tradable sector good is given by

$$P = \left[P_H^{1-\sigma} + P_F^{1-\sigma}\right]^{\frac{1}{1-\sigma}} = \left[\int_{N_F}^{N} p(n)^{1-\sigma}\, dn + (\epsilon\tau_a\tau_c)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

The foreign market price index for exported goods, denominated in foreign currency, is given by $P_x^* = \left(\int_{N_F}^{N} \mathcal{I}^x(n)\, p_x^*(n)^{1-\sigma}\, dn\right)^{\frac{1}{1-\sigma}}$, where $p_x^*(n)$ is the price of domestic variety n in the foreign country, denominated in foreign currency, and $\mathcal{I}^x(n)$ is an indicator function that equals one if variety n is exported.

The domestic demand for domestically produced goods is given by $Q_H(n) = D_H p(n)^{-\sigma}$, for $n \in (N_F, N]$; and the domestic demand for foreign produced goods is given by $Q_H(n) = D_H(\epsilon\tau_a\tau_c p^*(n))^{-\sigma}$, for $n \in [0, N_F]$. Finally, foreign demand for domestically produced goods is given by $Q_F(n) = D_F^*(p_x^*(n))^{-\sigma}$, for $n \in (N_F, N]$. Thus. we have that the value of aggregate imports (before import tariffs) and exports are given by the following expressions:

$$Imports = \frac{D_H}{\tau_a}\int_0^{N_F} (\epsilon\tau_a\tau_c p^*(n))^{1-\sigma}\, dn = \frac{D_H P_F^{1-\sigma}}{tau_a} = \frac{D_H(\epsilon\tau_a\tau_c)^{1-\sigma}}{\tau_a} \quad (27)$$

$$Exports = D_F^*\epsilon\int_{N_F}^{N} \mathcal{I}^x(n)\, p_x^*(n)^{1-\sigma}\, dn = \epsilon D_F^* P_x^{*1-\sigma} \quad (28)$$

Exporters

Given the expression of foreign demand for home variety n just described, $Q_F(n)$, revenues from exports are given by $\epsilon D_F^{*\frac{1}{\sigma}}(q_x/\tau_c)^{\frac{\sigma-1}{\sigma}}$, where q_x is the total quantity exported. If a firm exports, it must decide which fraction η of its product to sell abroad. Conditional on being an exporter, total gross revenue is given by

$$\begin{aligned} R^x(z,\ell,\eta) &= D_H^{\frac{1}{\sigma}}\left[(1-\eta)q(z,\ell)\right]^{\frac{\sigma-1}{\sigma}} + \epsilon D_F^{*\frac{1}{\sigma}}\left(\frac{\eta q\left(z,\ell\right)}{\tau_c}\right)^{\frac{\sigma-1}{\sigma}} \\ &= q(z,\ell)^{\frac{\sigma-1}{\sigma}}\exp\left(d_H + d_F(\eta)\right) \end{aligned} \tag{29}$$

where $d_H = \ln\left(D_H^{\frac{1}{\sigma}}\right)$ and $d_F(\eta) = \ln\left((1-\eta)^{\frac{\sigma-1}{\sigma}} + \epsilon\left(\frac{D_F^*}{D_H}\right)^{\frac{1}{\sigma}}\left(\frac{\eta}{\tau_c}\right)^{\frac{\sigma-1}{\sigma}}\right)$.

The optimal share of exports is given by:

$$\eta^o = \arg\max_{\eta} d_F(\eta) = \left(1 + \frac{\tau_c^{\sigma-1}}{\epsilon^\sigma}\frac{D_H}{D_F^*}\right)^{-1} \tag{30}$$

which shows that, conditional on exporting, all firms choose to export the same share of their output. The revenue functions for non-exporters and exporters are then given by, respectively:

$$\begin{aligned} R^d(z,\ell) &= (z\ell)^{\frac{\sigma-1}{\sigma}}\exp\left(d_H\right) \\ R^x(z,\ell) &= R^d(z,\ell)\Delta(z,\ell) \end{aligned}$$

where $\Delta(z,\ell) = \exp\left(d_F(\eta^o)\right)$, and $d_F(\eta^o)$ is obtained by substituting the expression of the optimal η^o into $d_F(\eta)$.[6] The export policy is then given by:

$$I_C^x\left(z,\ell\right) = \left\{\begin{array}{ll} 1 & \text{if } R_C^x\left(z,\ell\right) - f_x > R_C^d\left(z,\ell\right) \\ 0 & \text{otherwise} \end{array}\right\} \tag{31}$$

where $f_x > 0$ denotes the fixed cost of exporting, which is denominated in terms of the service composite good.

Since $\Delta(z,\ell) > 1$, being an exporter magnifies firms' revenues and also makes them more sensitive to productivity shocks, for any given state (z,ℓ). Thus, as in Cosar et al. (2016), reducing trade costs will produce two opposing forces: (i) there will be a reallocation of workers toward larger and higher productivity firms, which tend to

[6] When one substitutes η^o into $d_F(\eta)$, one obtains $d_F(\eta^o) = \ln\left(\left(\frac{D_F^*}{D_H}\epsilon^\sigma\tau_c + \tau_c^\sigma\right)^{\frac{1-\sigma}{\sigma}}\left[\tau_c^{\sigma-1} + \frac{D_F^*}{D_H}\epsilon^\sigma\right]\right)$.

be more stable and have lower worker turnover (e.g. they face larger costs of growing the workforce); (ii) due to the term $\Delta(z, \ell)$, both new and old exporters become more sensitive to idiosyncratic shocks, which tends to increase turnover. We follow Cosar et al. (2016) and refer to these two forces as the "distribution effect" and "sensitivity effect", respectively. Turnover is tightly linked to unemployment, as workers who are fired must spend at least one period in unemployment. In turn, workers transition from unemployment to formal and informal sector jobs. In addition to these forces, we also have "Melitz effects", where trade liberalization affects the "productivity/size threshold" for firms to export, but it will also affect the thresholds for operating formally, informally and exit. An attractive feature of this model is that it can accommodate both a increase or a decrease in informality. The net effect of the forces in the model is ultimately an empirical question.

2.6 Equilibrium

In equilibrium, firms act optimally and make entry and exit decisions and post vacancies according to equations (11), (15), and (16). If entry is positive in both sectors, entry conditions (17) and (18) hold with equality. Wages solve the bargaining problem between workers and the firm, as in equations (25) and (26). Labor markets clear, that is, the sum of employment levels across sectors and the number of unemployed workers must be equal to the total labor force. The government runs a balanced budget, the tradable and non-tradable markets must clear, and trade is balanced. Government's revenues come from tax collection and firing costs, while it pays unemployment benefits to all unemployed who come from formal employment. We assume that any surplus is directly rebated to consumers. Aggregate income is given by the sum of wages, unemployment benefits, profits and government transfers. Expenditure on nontradable goods is divided between final goods expenditure – given by $(1 - \zeta)$ – and intermediate goods expenditure R (hiring costs, fixed operational costs and the fixed costs of exporting).

We focus on steady state equilibria, where all aggregates remain constant. In particular, no sector can be expanding or contracting, which implies that: (i) the flow of workers out of unemployment and into the formal/informal and tradable/non-tradable sectors must be the same as the flow out of these sectors and into unemployment; (ii) the mass of firms entering the informal sector must be equal to the mass of informal firms that decide to exit or to formalize their businesses in either sector $k = C, S$; and (iii) the sum of the number of firms entering the formal sector and those formalizing their businesses must be equal to the mass of formal firms that decide to exit either sector $k = C, S$. In the Appendix A.1 we provide a detailed discussion of the equilibrium conditions.

3 Background: The cost of labor regulations in Brazil

The relevant laws and regulations that apply to formal labor relations in Brazil are contained in the the Brazilian Labor Code (*Consolidação das Leis Trabalhistas* – CLT), which dates back to 1943. In 1988, the new Federal Constitution was enacted and extended the range of labor regulations and workers' benefits, which substantially increased both the variable labor costs associated to formal employment and firing costs (De Barros and Corseuil, 2004).[7] As a result of the changes in 1988, the regulatory framework of the Brazilian labor market became quite burdensome and costly, and that has remained unaltered since then. According to the employment index in Botero et al. (2004), the cost of labor regulations in Brazil is around 20 percent above the mean and median of 85 countries and more than 2.5 times as large as in the United States.

The main aspects of the labor regulations in Brazil, in terms of their magnitude and potential impacts on labor market functioning, are the following: the presence of a national minimum wage, sizeable payroll taxes, unemployment insurance that is only available to formal workers, and substantial firing costs. Since these play an important role in our model and counterfactuals, we provide a brief background discussion on each of them individually and refer the reader to existing studies that provide a more in depth analysis of these different institutional aspects.

Starting by the national minimum wage, since 1995 (with the end of hyper-inflation) its nominal value is determined by the federal government once a year and is typically quite binding. In 2003, for example, the minimum wage corresponded to 49 percent of the national average wage and 81.3 percent of the national median wage.[8] As for the unemployment insurance, its rules remained unaltered from 1994 to 2015 but substantial changes have been implemented since then. Since our empirical analysis focuses on the period prior to the UI reforms, we discuss the rules in place until 2015.[9] In terms of

[7]Among the changes introduced by the new Constitution, one can highlight the following: regular working hours went from 48 to 44 hours per week; overtime premium increased from 20 to 50 percent; maternity leave increased from three to four months; and the value of paid vacations increased from one to, at least, 4/3 of the regular monthly wage (see De Barros and Corseuil, 2004, for a more detailed description of the changes).

[8]The mean and the median wages are computed using micro data from the National Household Survey (PNAD) and pooling together all formal and informal employees who are between 18 and 64 years old and worked at least 20 hours per week.

[9]See Carvalho et al. (2018) for a discussion of the reform, which substantially changed the eligibility criteria of unemployment benefits, and its impacts on layoffs in Brazil.

eligibility, generally a formal worker who is laid off and who has at least 6 months of job tenure is eligible to receive UI benefits for up to 5 months.[10] The actual duration of the benefit depends on the worker's accumulated tenure across her formal jobs in the 36 months prior to layoff. In practice, most workers receive between 4 and 5 months of UI benefits, with the mean and median number of monthly payments per UI spell equal to 4.3 and 4.7 months, respectively. Finally, the value of the benefit depends on the worker's average wage in the three months prior to layoff and the replacement rate is 100% for individuals who earn one minimum wage, with an average replacement rate of 64 percent (all data comes from Gerard and Gonzaga, 2018).

As for the firing costs, the Brazilian labor regulation states that all formal workers "dismissed with no just cause" should receive a monetary compensation paid by the employer. Since labour courts are extremely favourable to workers, *de facto* all workers are entitled to receive this compensation upon an involuntary separation. The magnitude of this compensation is determined proportionally to the funds accumulated in the worker's *Fundo de Garantia por Tempo de Serviço* (FGTS), which is a job security fund accumulated while the worker remains employed at a given firm. This is a private and individual fund that is specific to the worker, and to which employers must contribute, every month, the equivalent of 8 percent of worker's monthly wage. Hence, the worker's FGTS funds are proportional to her tenure and accumulate at a rate of roughly one monthly wage per year. Although these resources are owned by the worker, the fund is run by the government and the real return rates are typically below market rates, when not negative. Moreover, workers only have access to their own fund when they are laid off or upon retirement. In addition to the totality of their fund, workers who are laid off also receive a penalty, paid by their employer, which amounts to 40 percent of total resources accumulated in their fund during the duration of the job they are being laid off from. Firms must also pay an additional 10 percent of the FGTS in fines, which go directly to the federal government. In addition to this severance payment of 50 percent (40 plus 10 percent) of the FGTS, firms must provide a one-month advance notice, which *de facto* means that workers receive an additional monthly wage and are dismissed immediately.[11]

Finally, Brazil has a burdensome tax system, which is not only characterized by high tax rates but also by a complex structure that implies large compliance costs. For example, the estimated cost in terms of time required to comply with the tax system in

[10] There are some nuances to eligibility that depend upon the elapsed time since worker's last successful application to UI benefits. See Gerard and Gonzaga (2018) for a more detailed discussion of the UI program in Brazil.

[11] Gonzaga et al. (2003) provide an in depth discussion of the legislation on dismissal costs in Brazil.

Brazil is 2,600 hours, which is the highest in the world, and more than 8 times larger than the cost that a firm faces in the U.S. Even though a substantial part of this cost is not due to the payment of labor taxes, the time required to comply with labor taxes in Brazil is almost 5 times higher than in the U.S. (491 and 100 hours, respectively).[12] In terms of the tax rate, even though we use the statutory values for both payroll and revenue taxes in our estimation, it is useful to provide a comparison to other countries, which is done in Doing Business (2007): The labor tax computed as a share of commercial profits amounts to 42.1 percent in Brazil, while it is 12.9 percent in Canada and 10 percent in the U.S. Hence, not only labor taxes seem to be quite high in Brazil, but also they imply substantial compliance costs.

4 Data and Facts

4.1 Firms

In this paper we make use of 6 datasets that contain information on formal and informal firms and their workers. The first is the *Relação Anual de Informações Sociais* (RAIS), which is a matched employer-employee dataset assembled by the Brazilian Ministry of Labor every year since 1976. RAIS is a high quality panel that contains the universe of formal firms and workers.[13] It provides information on firms' 5-digit industry, location and ownership (i.e public vs. private enterprises), among others. At the worker level, the main variables are gender, age, level of education, monthly wage, number of hours in the contract, tenure at the firm, occupation, month of accession into the job (if accession occurred during the current year), and month of separation (if any). We use the matched employer-employee structure to compute firm size and firm-level average wages over time.

We also use three economic surveys that cover the *formal* manufacturing, retail and service sectors: *Pesquisa Industrial Anual* (PIA), *Pesquisa Anual de Comércio* (PAC), and *Pesquisa Anual de Serviços* (PAS), respectively. These surveys collect detailed information about firms' inputs, output and revenues, and are a combination of a census for larger firms and a representative sample for smaller firms. In the manufacturing sector (PIA), all firms with at least 30 employees are part of the census and are surveyed every

[12]These data come from Doing Business (2007), which is the earliest report available on paying taxes in the Doing Business Initiative that provides comparability across a comprehensive set of countries.

[13]The RAIS data set has been increasingly used in different applications. For recent examples see Dix-Carneiro (2014), Helpman et al. (2017), Alvarez et al. (2018), Ulyssea (2018), among others.

year, while firms with 5 to 29 employees are randomly sampled.[14] The PAC (retail sector) and PAS (services) follow similar designs, although they have lower size thresholds for firms to be included in the census: firms with 20 employees or more are part of the census, while firms with up to 19 employees are randomly sampled.

The fifth data source used is Customs data from *Secretaria de Comércio Exterior* (SECEX), which give us the list of every export and import transaction (and values) made from and by Brazilian firms every year since 1990 and until 2007. Importantly for this study, there is a unique firm identifier across these 5 data sets, which allows us to merge the production information from PIA, PAS and PAC with the information about firms' labor and wages coming from RAIS, and the customs data from SECEX.

These 5 data sets provide a comprehensive coverage of the formal sector, but are completely silent about the informal sector (by design). We therefore use a sixth data source, which is the *Pesquisa de Economia Informal Urbana* (ECINF). This survey was collected by the Brazilian Bureau of Statistics (IBGE) in 1997 and 2003, and was designed to be representative of the universe of *urban* firms with up to five employees (both formal and informal). It is a matched employer-employee data set that contains information on entrepreneurs, their businesses and employees. Firms are directly asked whether they are registered with the tax authorities and whether each of their workers has a formal labor contract. Thus, it is possible to directly observe both firms' and workers' formal status. Given that the formality/informality statuses are self-reported, one could have concerns about measurement error and under-reporting. However, the IBGE has a long tradition of accurately measuring labor informality, and it has very strict confidentiality clauses, so the information cannot be used for auditing purposes by other government branches, in particular those responsible for enforcing the relevant laws and regulations. These characteristics, associated to the high levels of informality observed in the data, make us confident that respondents are not systematically underreporting their informality status.[15]

In all 6 data sets we exclude public sector firms and those in agriculture, mining, coal, oil and gas industries. We do so because our focus lies on private sector, urban firms. Moreover, our model is not well suited to describe sectors with very large economies of scale and dominated by few very large firms, such as oil and gas. In the data, as well

[14] The main source of information used by IBGE to design its sample is the RAIS data set described above.

[15] Additionally, Ulyssea (2018) shows that the ECINF reproduces very well the RAIS in all the dimensions that are common to both data sets (e.g. size and sectoral distributions), which is reassuring of ECINF's quality.

as in the model, the tradable sector is comprised by manufacturing firms and the non-tradable sector is comprised by services and retail firms. In sum, information on the formal tradable sector comes from RAIS, PIA and SECEX; on the non-tradable formal sector comes from RAIS, PAS and PAC; and data on both tradable and non-tradable informal sectors come from the ECINF survey.

Since 2003 is the last year available for the ECINF survey, we use it as the reference year for all other data sets. Table 1 shows the size distribution (measured as number of employees) in the tradable and non-tradable sectors for formal and informal firms. As expected, the number of observations is much larger for formal firms, as these come from a census (the RAIS data). Nevertheless, the share of tradable sector firms is quite similar in the formal and informal sectors (13.1 and 14.2 percent, respectively). The size difference between formal exporters and non-exporters in the tradable sector is quite remarkable, with exporting firms being more than 8 times larger than non-exporting firms, on average. Figure 2 shows this fact from a different angle, as the share of exporters increases steeply moving up in the size distribution.

Table 1: Firm Size Distribution in Number of Employees

	Formal		Informal	
	Sector C	Sector S	Sector C	Sector S
All Firms				
Mean (Log-Employment)	1.78	1.18	0.10	0.10
Variance (Log-Employment)	1.82	1.26	0.09	0.08
Exporters				
Mean (Log-Employment)	3.9	–	–	–
Variance (Log-Employment)	2.7	–	–	–
Employment Distribution				
Pct. 20	2	1	1	1
Pct. 40	4	2	1	1
Pct. 60	7	4	1	1
Pct. 80	17	8	1	1
Pct. 90	35	14	2	2
Pct. 95	67	25	2	2
Pct. 99	298	109	4	3
# Observations	216,467	1,430,633	1,069	6,192

Notes: To compute the moments for the formal tradable sector, we use the PIA; for non-tradable formal sector, the PAS and PAC data sets; and for both tradable and non-tradable informal sectors, we use the ECINF survey.

Formal firms in the tradable sector are also larger on average than those in the non-tradable sector and the distribution is more skewed to the left. The size difference between informal firms in tradable and non-tradable sectors is almost null, which is expected: the ECINF survey has a size cap, which mechanically limits the size differential. More substantially, informal firms cannot grow much without becoming too visible to the authorities and cannot export either, which limits their ability to grow.

Figure 2: Share of Exporters by Firm Size Percentiles

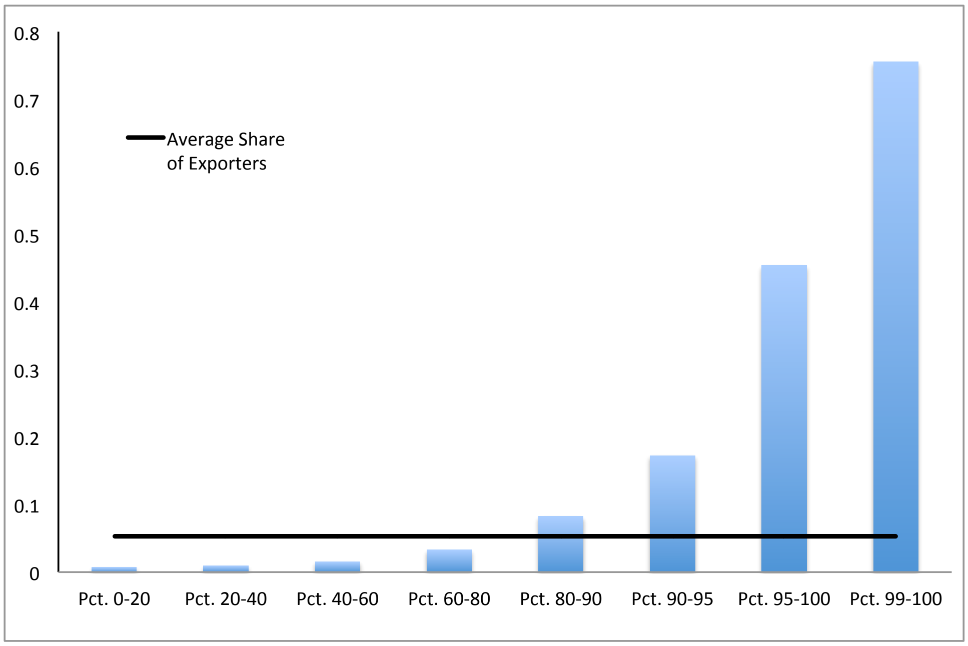


Table 2 shows the same information as in Table 1 but focusing on firms' revenues. The same patterns found in Table 1 arise, but it is worth noting that the size differences across percentiles are much larger when one uses revenues instead of employment as the size measure. For example, the 99th percentile of the size distribution measured as number of employees is nearly three times larger in the formal tradable than in the formal non-tradable. The same ratio is more than 30 when one uses revenues. Interestingly, this relationship is inverted in the informal sector, where firms in the non-tradable sector earn higher revenues than firms in the tradable sector. This is intuitive, as one would expect that the penalty for remaining small (and informal) is lower in the non-tradable sector.

Figure 3 shows that there is a substantial size-wage premium in both tradable and non-tradable formal sectors, but the same is not true for informal firms. This is somewhat mechanical, as most informal firms have only one employee. As for employment, wage and revenue growth, Tables 3 and 4 show different patterns moving up the firm size distribution. Table 3 shows that, on average, expanding firms tend to present higher wage growth, but this relationship is not constant across different percentiles of the size

Table 2: Revenue Distribution

	Formal		Informal	
	Tradable	Non-Tradable	Tradable	Non-Tradable
All Firms				
Mean (Log-Revenue)	12.73	10.81	8.53	8.95
Variance (Log-Revenue)	3.51	2.07	1.44	1.30
Exporters				
Mean	15.46	–	–	–
Variance	4.45	–	–	–
Revenue Distribution (in 2003 R$)				
Pct. 20	77,962	15,897	1,920	3,600
Pct. 40	166,110	31,102	4,200	6,000
Pct. 60	407,595	59,492	6,600	9,600
Pct. 80	1,143,359	137,162	13,428	19,200
Pct. 90	4,038,112	288,717	24,000	32,160
Pct. 95	12,494,325	558,989	36,000	49,200
Pct. 99	103,287,792	3,229,837	72,000	108,000

Notes: To compute the moments for the formal tradable sector, we use the PIA; for non-tradable formal sector, the PAS and PAC data sets; and for both tradable and non-tradable informal sectors, we use the ECINF survey.

distribution (for none of the groups considered in the table). On the contrary, Table 4 shows a clear pattern that is in line with other available evidence in the literature: yearly employment and revenues growth rates decrease with size, except at the very top of the distribution (top 5 and one percent of the size distribution).

4.2 Workers

In order to complement the information on firms, we use the Monthly Employment Survey (PME – *Pesquisa Mensal de Emprego*) to obtain information on worker allocations and labor market flows. This is a rotating panel with a similar design to that of the Current Population Survey in the U.S.: individuals in a given household are interviewed for 4 consecutive months, they "rest" for 8 months and are then re-interviewed for additional 4 consecutive months, which implies a maximum panel length of 16 months. This employment survey covers the six main metropolitan areas in Brazil and contains detailed information on individuals' socio-demographic characteristics and labor market outcomes, including informal employment and self employment.[16]

We exploit the panel structure of PME to estimate one-year labor market transitions between formal employment, informal employment (in both tradable and non-tradable

[16]See Meghir et al. (2015) for a more detailed description of the PME data.

Figure 3: Average Log-Wages by Firm Size Percentiles

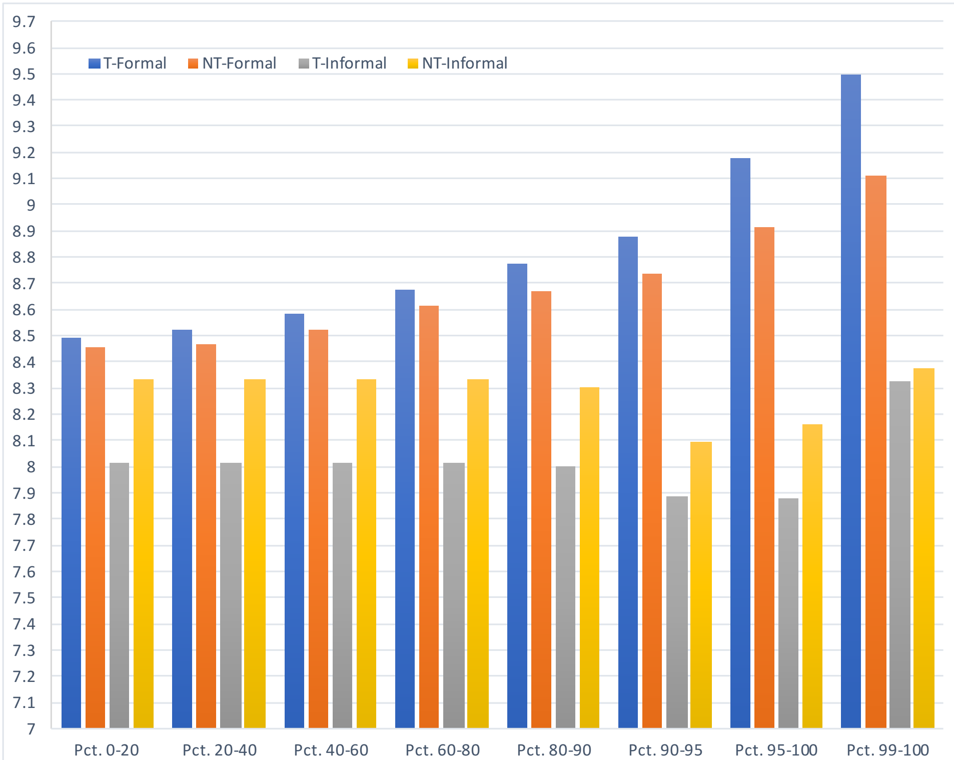


Table 3: Formal Firms' Average Wage Growth

	Surviving Firms		Expanding Firms		Contracting Firms	
	T	N-T	T	N-T	T	N-T
All Firms	0.051	0.044	0.063	0.059	0.042	0.037
Pct. 0-20	0.043	0.032	0.065	–	0.040	0.032
Pct. 20-40	0.055	0.042	0.072	0.075	0.048	0.038
Pct. 40-60	0.050	0.049	0.066	0.064	0.040	0.042
Pct. 60-80	0.051	0.042	0.062	0.057	0.039	0.032
Pct. 80-90	0.048	0.042	0.060	0.053	0.033	0.031
Pct. 90-95	0.055	0.047	0.072	0.056	0.032	0.036
Pct. 95-100	0.074	0.060	0.050	0.061	0.109	0.059
Pct. 99-100	0.059	0.123	0.029	0.094	0.110	0.168

Notes: We compute firm-level yearly average growth using the RAIS data for the years of 2002 and 2003. Surviving firms are those that are alive in 2002 and 2003. Expanding firms are those for which employment in 2003 is strictly larger than 2002. Contracting firms are those whose labor force remains constant or decreases between 2002 and 2003. *T* and *N-T* denote the tradable and non-tradable sectors, respectively.

sectors) and unemployment statuses.[17] As in the firm-level data, we exclude individuals

[17] A worker is defined to be unemployed if she is not working – regardless of whether she is searching

Table 4: Formal Firms' Employment and Revenue Growth

	Employment Growth		Revenue Growth	
	Tradable	Non-Tradable	Tradable	Non-Tradable
All Firms	0.156	0.121	0.201	0.229
Pct. 0-20	0.362	0.318	0.216	0.242
Pct. 20-40	0.155	0.231	0.212	0.227
Pct. 40-60	0.096	0.073	0.210	0.235
Pct. 60-80	0.072	0.036	0.201	0.236
Pct. 80-90	0.072	0.031	0.178	0.215
Pct. 90-95	0.071	0.036	0.169	0.212
Pct. 95-100	0.088	0.046	0.146	0.149
Pct. 99-100	0.101	0.060	0.148	0.112

Notes: We compute firm-level yearly employment and revenue growth using the years of 2003 and 2004. We compute employment growth using the RAIS data set. For revenue growth in the formal tradable sector, we use the PIA data set; for the non-tradable formal sector, we use the PAS and PAC data sets.

employed in the public sector, agriculture, mining, coal, oil and gas industries. In addition to these filters, we also exclude individuals younger than 17 and older than 65 years old.

Panel A of Table 5 shows worker allocations in 2003. It is noteworthy that 15% of the working age population is unemployed (or more precisely, not employed), and that approximately 20% of employed workers are in the C-sector. These numbers also indicate that 48% of the labor force is employed in the informal sector. In addition, 35% of C-sector workers are informal, whereas 51% of S-sector workers are. Panel B of Table 5 shows the relevant transition matrix for the purposes of our model. Even though we estimate the full transition matrix in the data, we only report the transitions that are accounted for in our model, while the remaining ones are omitted (such as from the formal tradable sector to the informal non-tradable sector). We start by noting that the table confirms two well-known facts: (i) most of the labor force is in the non-tradable sector (69.3 percent); and (ii) informality is very high in Brazil, accounting for 41 percent of the labor force. As for the probabilities of transition, the rate of retention (main diagonal in the transition matrix) is highest in the non-tradable formal sector (68.5 percent) and is lowest in the informal tradable sector (27 percent). Unemployed workers are most likely to exit to a non-tradable informal sector (38 percent), while the formal tradable sector is the least likely destination of those who are unemployed.

for a job or not.

Table 5: Sectoral Shares and 12-month Transition Rates

Panel A: Workers Allocation

	Unemp.	Tradable Inf.	Tradable Form.	Non-Trad. Inf.	Non-Trad. Form.
Shares†	0.137	0.058	0.112	0.352	0.341

Panel B: 12-month Transition Rates

	Unemp.	Tradable Inf.	Tradable Form.	Non-Trad. Inf.	Non-Trad. Form.
Unemp.	0.335	0.062	0.049	0.381	0.159
Tradable - Inf.	0.087	0.270	0.135	0.369	0.119
Tradable - Form.	0.039	0.044	0.530	0.074	0.300
Non-Tradable - Inf	0.086	0.062	0.035	0.610	0.177
Non-Tradable - Form.	0.044	0.016	0.100	0.132	0.685

Notes: Authors' own calculations from the Monthly Employment Survey (PME), years 2003 and 2004. We use the first and 4th interviews to compute 4-months transition rates for the full transition matrix, M. We then annualize M by computing M^3. † We use sampling weights to compute these shares using the entire sample.

5 Estimation

Attention: Estimation is still not complete. The results discussed here are still not optimal. Think of these initial results as a "calibration" exercise, but still not final.

Our estimation procedure follows two steps. First, we fix a subset of parameters using a combination of aggregate data, estimates from previous papers and the statutory value of institutional parameters, such as revenue and payroll taxes. Then, we estimate the remaining parameters of the model using an Indirect Inference estimator, which allow us to combine information from the different data sources discussed in the previous section.

As discussed in Section 3, labor regulations are quite costly and cumbersome in Brazil, so we need to make a few simplifying assumptions. We follow Ulyssea (2018) and set τ_w so that it reflects the main taxes that are proportional to firms' wage bill, namely, employer's social security contribution (20 percent), payroll tax (9 percent), and severance contributions to FGTS (8.5 percent). τ_y includes only the federal VAT taxes, IPI (20 percent) and PIS/COFINS. We exclude state level value added taxes because these vary greatly across states and there is a cumbersome system of tax substitution across the production chain, which would be impossible to properly capture.[18]

Firing costs are set following Heckman et al. (2000), which compute the expected discounted cost of dismissing a worker for several Latin American countries, including Brazil. This is done taking into account the main characteristics of dismissal costs in Brazil, as discussed in Section 3, and the expected cost is expressed as a multiple of the monthly wage. To make this parameter compatible with our model, we convert it to a fixed monetary value using the average formal wage found in the data in 2003. The minimum wage corresponds to the annualized value of the national monthly minimum wage in 2003. Finally, the unemployment benefit is set assuming that all workers receive the maximum number of benefits (5 monthly payments), which is very close to both the mean and median number of benefits (Section 3), while we use the mean monthly value paid in 2003 reported by the Ministry of Labor, which is denominated in multiples of the minimum wage. Table 6 shows these parameter values and their sources.

Table 6: Calibrated Parameters

Parameter	Description	Source	Value
σ_C	CES parameter	Cosar, Guner and Tybout (2016)	6.667
σ_S	CES parameter	Cosar, Guner and Tybout (2016)	6.667
τ_c	Iceberg Trade Costs	Cosar, Guner and Tybout (2016)	2.50
ζ	Share of expend. C	World Input-Output Database	0.283
r	Interest rate	Ulyssea (2010)	0.08
τ_y	Sales Tax	Ulyssea (2017)	0.293
τ_w	Payroll Tax	Ulyssea (2017)	0.375
τ_a	Import Tariff	TRAINS	1.12
κ	Firing Costs	Heckman and Pages (2000) (in R$)	1,956.7
$\underline{w}$	Min. Wage	Annualized 2003 value (in R$)	2,880
b_u	Unemp. Benefit	$1.37 \times 5 = 6.85$ monthly MW	1,644

In a second step, we take the parameters described in Table 6 as given and estimate the remaining parameters using an Indirect Inference estimator with equilibrium constraints.[19] The estimation algorithm is described in details in Appendix B.1. In this step we estimate 31 parameters using 99 data moments and auxiliary parameters.

[18]As discussed in Ulyssea (2018), these taxes can be large in some states, which would imply that we underestimate the overall tax burden that firms face. However, we do not include intermediate inputs, which implies that we might be overestimating the actual tax burden faced by some firms. The net effect of these forces is *a priori* unclear.

[19]This is the usual Indirect Inference estimator (e.g. Gourieroux and Monfort, 1996; Smith, 2008), but we also penalize deviations from the model's equilibrium constraints in the objective function.

5.1 Identification

The choice of parameters from auxiliary regressions (and moments) to be matched by the model is crucial to achieve identification. Given the high non-linearity and dimension of the model, it is not possible to provide a direct proof of identification. Nevertheless, we provide a heuristic discussion of which variation in the data provides information on different sets of parameters to be estimated.

We start by noting that even though one can directly use micro data to estimate the parameters of the AR(1) processes for productivity (ρ_k and σ_k^z), we estimate them within our Indirect Inference procedure and use the persistence and volatility of firm revenues and labor force sizes to obtain information about these parameters. This information is obtained with PIA (Manufacturing Survey), PAS (Services Survey), PAC (Retail-trade Survey) and RAIS (all sectors). We choose to proceed in this way because the production functions typically assumed by the existing estimators (e.g. Olley and Pakes, 1996; Ackerberg et al., 2007, 2015) are not compatible with our setting where firms use labor as their only input and there is no investment decision.[20]

The parameters of the hiring costs function (h, γ_1 and γ_2 in equation 6) are identified using information on growth rates of formal firms, and how these depend on firm size. The convexity in hiring is also important for the model to generate dispersion in wages across firms. Therefore, the relationship between wages and size provide useful information on γ_1. The matching function's parameters are identified from worker transitions out of unemployment and into formal and informal employment in the tradable and non-tradable sectors. We estimate those from the monthly employment survey (PME) and annualize the transitions to make them compatible with the model's period.

The exogenous death shocks for formal firms can be identified off the exit rates of very large firms. Because we cannot observe exit rates of informal firms, we set the exogenous death shocks to be equal for formal and informal firms within sectors. The fixed costs of operation of formal firms ($\bar{c}_{kf}$) are disciplined by how exit rates decline with firm size. In addition, average firm-level revenues help the identification of fxed operating costs of formal firms but also those of informal firms ($\bar{c}_{ki}$). Average firm-level revenues are available from PIA, PAS and PAC for formal firms and ECINF (for informal firms). Larger fixed costs force low-revenue firms to exit, and thereby increase average revenues among survivors.

[20]As a cross-check, we use Olley and Pakes (1996)'s estimator to obtain a measure of firm-level productivity for manufacturing firms and use it to estimate a simple AR(1) process. The estimate for the persistence parameter,ρ_C, is remarkably close to the one we obtain in our Indirect Inference estimator. These results are available upon request.

The fixed cost of exporting, f_x, is identified by the fraction of exporters among formal firms, which is available merging information from RAIS and SECEX. The foreign demand shifter d_F is identified using information on the average size and revenue of exporters (RAIS, PIA and SECEX), fraction of revenues in the tradable sector coming from exports (PIA and SECEX), and the log-wage premium (regression of firm-level log-wages on log-size and exporter indicator, using RAIS and SECEX).

The identification of bargaining power parameters β_f and β_i is straightforward in light of equations (25) and (26). We target the coefficients of a linear regression of firm-level log-wages on revenue per worker (for both the formal and informal sectors, using data from RAIS and ECINF respectively). Lastly, the probability of detection is identified off the size distribution of firms in the informal sector (ECINF), and the share of informal firms by employment size (ECINF).

5.2 Estimates

Table 7 shows our preliminary estimation results. We now discuss the magitude and plausibility of some of these estimates. First, note that our estimate of θ, the matching function parameter, is reasonably close to CGT's estimate of 1.8. Next, notice that the value of leisure is estimated at $b = 3,119$, a little over the annualized value of the minimum wage of R\$ 2,880. This relatively high value is necessary for a good fit of wages. Also important for the fit of wages are the bargaining parameters $\beta_f = 0.11$ and $\beta_i = 0.89$. Our estimates point toward a larger bargaining power of informal sector workers. Table 11 shows that informal sector workers tend to capture a larger fraction of revenue per worker, and this is interpreted as a larger bargaining power. Perhaps surprising, this result is not necessarily unreasonable. Revenues (and revenues per worker) in the informal sector are much lower in the informal sector. The only way the model can assign wages in the informal sector that are closer to those in the data is by assigning a high value of the bargaining power parameter of informal workers. The value of $\beta_f = 0.11$ is somewhat lower than the value CGT estimate in Colombia, which amounts to 0.4.

The probability of detection (or more accurately, the expected penalty of being informal) is large and increases steeply with firm size. For example, we estimate p_{Ci} to be approximately 0.51 for C-sector firms of size 1, and 1 for firms of size 2 or above. In the S-sector, p_{Si} is approximately 0.17 for firms of size 1 and 1 for firms of size 2 or above.

The hiring cost function presents very large convexity in both sectors ($\gamma_{1C} = 8.4$ and $\gamma_{1S} = 6.1$) and a fair degree of scale economies ($\gamma_{2C} = 0.49$ and $\gamma_{2S} = 0.29$). For comparison, CGT obtain estimates of $\gamma_1 = 3.1$ and $\gamma_2 = 0.39$. To illustrate the magnitude of hiring costs, consider a firm of size 10 in the C-sector. It will cost this firm $R\$48$ to expand to 11 employees, or 0.003 times the annual average wage in the formal sector (which is of $R\$12,230$), $R\$16,830$ to expand to 12, and a most likely prohibitive $R\$515,790$ (or 42 times the annual average wage) to expand to 13. On the other hand, it will cost $R\$462$ for a firm with size 100 to expand to 104 employees.

Finally, we note that the fixed costs of operation in the formal sector ($\bar{c}_{Cf} = 37,344$ and $\bar{c}_{Sf} = 6,358$) are 11 to 17 times larger than the fixed costs of operation in the informal sector ($\bar{c}_{Ci} = 2,176$ and $\bar{c}_{Si} = 541$). This is expected, given the large perceived costs of operating a firm in the formal sector (compliance with regulations, bureaucracy, bribes, etc.). In particular, the magnitude of the fixed costs of operation in the formal C-sector amounts to $37,344$, or 3 times annual average wages.

Finally, remember that μ_v, d_{HC}, d_{HS} are actually endogenous objects and not parameters. As we discuss in Appendix B.1, we treat these objects as parameters in the estimation procedure, but penalize the deviations from their equilibrium values in the objective function. What is noteworthy is that in equilibrium, the vacancy filling rate $\mu_v = 0.39$, so that firms need to post 2.6 vacancies to be able to hire 1 worker.

Table 7: Estimated Parameters

Parameter	Description	Estimate	Parameter	Description	Estimate
μ^v	vacancy filling rate	0.391	ρ_C	AR(1): persistence, C-sector	0.969
b	non-pecuniary value of unemployment	3,119	σ_C^z	AR(1): volatility, C-sector	0.372
β_f	bargaining power, formal sector	0.111	α_{Cf}	Exogenous death prob., formal C-sector	0.107
β_i	bargaining power, informal sector	0.886	$\bar{c}_{Cf}$	Fixed operating cost, formal C-sector	37,344
θ	Matching function	1.297	α_{Ci}	Exogenous death prob., formal C-sector	0.107
a_C	Prob. Detection: intercept, Sector C	0.324	$\bar{c}_{Ci}$	Fixed operating cost, informal C-sector	2,176
b_C	Prob. Detection: slope, Sector C	0.185	$d_{H,S}$	Domestic demand shifter, S-sector	7.932
c_C	Prob. Detection: exponent, Sector C	2.446	ρ_S	AR(1): persistence, S-sector	0.955
a_S	Prob. Detection: intercept, Sector S	0.013	σ_S	AR(1): volatility, S-sector	0.471
b_S	Prob. Detection: slope, Sector S	0.160	α_{Sf}	Exogenous death prob, formal S-sector	0.084
c_S	Prob. Detection: exponent, Sector S	2.806	$\bar{c}_{Sf}$	Fixed operating cost, formal S-sector	6,358
h_{fC}	Hiring Cost Function, level: Sector C	1,868	α_{Si}	Exogenous death prob., formal S-sector	0.084
γ_{1C}	Hiring Cost Function, convexity: Sector C	8.441	$\bar{c}_{Si}$	Fixed operating cost, formal S-sector	541.40
γ_{2C}	Hiring Cost Function, scale: Sector C	0.486	d_F	Foreign demand shifter	0.328
h_{fS}	Hiring Cost Function, level: Sector S	1,844	f_x	fixed cost of exporting	725,101
γ_{1S}	Hiring Cost Function, convexity: Sector S	6.054	$d_{H,C}$	Domestic demand shifter, C-sector	8.661
γ_{2S}	Hiring Cost Function, scale: Sector S	0.290			

5.3 Model Fit

Tables 8 through 14 compare the moments and statistical relationships generated by the model (under the parameterization described in Tables 6 and 7) with those found in the data. Several features of the data are well matched by the model, while there is still room for improvement in some dimensions.

Table 8 shows worker transitions from unemployment to each of the four sectors. While the model is able to replicate these flows reasonably well, our model is currently overestimating flows from unemployment to the formal S-sector. We are also currently overestimating the exit rate from sector-C formal firms, but we are matching well the fact that the probability of exit is smaller for larger firms.

Table 8: Model Fit: Transition Rates

	Source	Model	Data
Unemployment to Informal C	PME	0.068	0.062
Unemployment to Formal C	PME	0.085	0.050
Unemployment to Informal S	PME	0.350	0.380
Unemployment to Formal S	PME	0.238	0.159
Sector C – Formal			
Firm Exit Rate	RAIS	0.137	0.096
Regression Exit – Constant	RAIS	0.186	0.185
Regression Exit – log($size$)	RAIS	-0.026	-0.050
Sector S – Formal			
Firm Exit Rate	RAIS	0.134	0.113
Regression Exit – Constant	RAIS	0.198	0.178
Regression Exit – log($size$)	RAIS	-0.054	-0.055

Table 9 shows that our model is able to fit the size distributions across all sectors very well. Perhaps, we are underestimating the average size of informal firms a bit. As we discuss later, our model cannot (currently) generate informal firms with size of 3 or larger.

Table 10 shows statistics that are important for the relationship between globalization and informality and unemployment in our model. Note that employment growth tends to be smaller in larger firms, as they tend to have less volatile employment levels. However, conditional on size, exporters tend to have larger growth rates, consistent with the sensitivity effect discussed in section 2.

Table 11 shows that the model fits quite well wage moments and how wages vary with firm-level size and revenue per worker. Table 12 shows how our model fits moments

Table 9: Model Fit: Size Distributions

	Source	Model	Data
Sector C – Formal			
Percentile 20	RAIS	2	2
Percentile 40	RAIS	4	4
Percentile 60	RAIS	8	7
Percentile 80	RAIS	19	17
Mean log-size	RAIS	1.898	1.779
Var log-size	RAIS	1.434	1.821
Mean log-size Exporters	RAIS + SECEX	4.022	3.936
Var log-size Exporters	RAIS + SECEX	0.555	2.747
Sector S – Formal			
Percentile 20	RAIS	1	1
Percentile 40	RAIS	2	2
Percentile 60	RAIS	4	4
Percentile 80	RAIS	8	8
Mean log-size	RAIS	1.184	1.178
Var log-size	RAIS	1.076	1.262
Sector C – Informal			
Percentile 20	ECINF	1	1
Percentile 40	ECINF	1	1
Percentile 60	ECINF	1	1
Percentile 80	ECINF	1	1
Mean log-size	ECINF	0.002	0.105
Variance log-size	ECINF	0.001	0.092
Sector S – Informal			
Percentile 20	ECINF	1	1
Percentile 40	ECINF	1	1
Percentile 60	ECINF	1	1
Percentile 80	ECINF	1	1
Mean log-size	ECINF	0.004	0.097
Variance log-size	ECINF	0.003	0.075

related to revenues. Although most of the moments and relationships are reasonably close to those in the data, the model systematically underestimates the dispersion in revenues across firms.

Table 13 shows that the model cannot fit well the share of informal firms by size. We predict that the overwhelming majority of firms of size 1 are informal. However, we underestimate the share of size 2 firms that are informal, and we are predicting zero informal firms with size larger than 2. This is an aspect of the model that needs improvement, but since the estimation procedure is not done yet, it may be the case that our optimization algorithm is just stuck in a local optimum. Finally, Table 14 shows that

Table 10: Model Fit: Moments Related to Firm Growth

	Source	Model	Data
Sector C – Formal			
$Corr\left(\log\left(\ell_t\right), \log\left(\ell_{t+1}\right)\right)$	RAIS	0.979	0.92
Mean Growth	RAIS	0.216	0.156
Regression growth rate: Constant	RAIS	0.356	0.303
Regression growth rate: log-size	RAIS	-0.074	-0.082
Regression growth rate: Export Status	RAIS	0.059	0.148
Sector S – Formal			
$Corr\left(\log\left(\ell_t\right), \log\left(\ell_{t+1}\right)\right)$	RAIS	0.937	0.91
Mean Growth	RAIS	0.116	0.121
Regression growth rate: Constant	RAIS	0.153	0.217
Regression growth rate: log-size	RAIS	-0.035	-0.076

the model fits well statistics related to exporters, and how firm-level revenues correlate with size and with previous period revenue.

Table 11: Model Fit: Wages

	Source	Model	Data
Sector C – Formal			
Mean log-Wages	RAIS	8.664	8.637
Mean log-Wages, Exporters	RAIS + SECEX	9.314	9.276
Regression 1: Constant	RAIS	8.454	8.443
Regression 1: log-Size	RAIS	0.098	0.094
Regression 1: Exporter	RAIS	0.465	0.462
Regression 2: Constant	IBGE	4.647	6.334
Regression 2: log(Rev/Worker)	IBGE	0.385	0.235
Sector S – Formal			
Mean log-Wages	RAIS	8.557	8.562
Regression 1: Constant	RAIS	8.433	8.434
Regression 1: log-Size	RAIS	0.105	0.108
Regression 2: Constant	IBGE	5.533	7.417
Regression 2: log(Rev/Worker)	IBGE	0.310	0.109
Sector C – Informal			
Mean log-Wages	ECINF	8.244	8.014
Regression 2: Constant	ECINF	-3.242	3.777
Regression 2: log(Rev/Worker)	ECINF	1.215	0.397
Sector S– Informal			
Mean log-Wages	ECINF	8.075	8.415
Regression 2: Constant	ECINF	0.691	3.912
Regression 2: log(Rev/Worker)	ECINF	0.878	0.379

Notes: Regression 1 is a firm-level regression of log-wages onto a constant and log-size, where size is measured as number of employees; for sector C firms it also includes an exporter dummy; Regression 2 is a firm-level regression of log-wages onto a constant and the logarithm of revenue per worker.

Table 12: Model Fit: Revenues

	Source	Model	Data
Sector C – Formal			
Mean log Revenue	IBGE	12.333	12.726
Variance log Revenue	IBGE	1.595	3.511
Mean log Revenue – Exporters	IBGE	15.401	15.465
Variance log Revenue – Exporters	IBGE	0.258	4.448
Sector S – Formal			
Mean log Revenue	IBGE	10.928	10.814
Variance log Revenue	IBGE	1.893	2.074
Sector C – Informal			
Mean log Revenue	ECINF	9.439	8.533
Variance log Revenue	ECINF	0.405	1.444
Sector S – Informal			
Mean log Revenue	ECINF	8.394	8.952
Variance log Revenue	ECINF	0.588	1.298

Table 13: Model Fit: Fraction of Informal Firms by Firm Size

	Source	Model	Data
Size = 1 Employee	ECINF	0.951	0.933
Size = 2 Employees	ECINF	0.340	0.711
Size = 3 Employees	ECINF	0.000	0.491
Size = 4 Employees	ECINF	0.000	0.261
Size = 5 Employees	ECINF	0.000	0.372

Table 14: Model Fit: Miscellaneous

	Source	Model	Data
Related to Export Status			
$Corr\,(\log \ell_t, \mathcal{I}_t^x)$	RAIS + SECEX	0.411	0.378
Fraction of Firms that Export	RAIS + SECEX	0.051	0.053
Share of Export Revenue	SECEX + IBGE	0.136	0.136
Revenue and Size			
$Corr\,(\log R_t, \log \ell_t)$, Informal C-Sector	ECINF	0.202	0.339
$Corr\,(\log R_t, \log \ell_t)$, Informal S-Sector	ECINF	0.248	0.318
$Corr\,(\log R_t, R_{t+1})$, Formal C-Sector	IBGE	0.928	0.929
$Corr\,(\log R_t, R_{t+1})$, Formal S-Sector	IBGE	0.862	0.845

6 Counterfactual Simulations

In this section we conduct preliminary counterfactual analyses. We vary τ_a between 1 (no import tariff) and 1.9 (import tariff rate of 90%). For each value of the import tariff rate we compute the informality rate and the unemployment rate in the economy. We also compute the informality rate within the C sector and within the S sector. Remember that the baseline scenario is a tariff rate of 12% ($\tau_a = 1.12$). It is also instructive to highlight that the average tariff rate in Brazil at the onset of the trade liberalization of the 1990's was of 32% ($\tau_a = 1.32$). Note that the relationship between informality rates and τ_a is a bit "noisy", due to simulation error. Therefore, we add linear or polynomial fits to all of the figures below to make the relationships clearer.

Figure 4 shows that as import tariffs decrease (smaller τ_a), the informality rate in the economy increases. In particular, for $\tau_a = 1.9$ the informality rate is of 41.7%, for $\tau_a = 32\%$ it is 46.6% and for $\tau_a = 1.12$ it is 48.2%. Figures 5 and 6 unpack how this relationship is heterogeneous within sectors. Within the C-sector, the relationship is increasing (increases in tau_a are reflected in a larger share of $C-$sector workers in informal jobs), but apparently concave. In the S-sector, the relationship is very similar to the relationship for the whole economy, but the magnitude of the relationship between τ_a and informality is stronger.

Interestingly, Figure 7 shows no visible relationship between import tariffs and unemployment. Together, Figures 4 and 7 suggest that, in response to the Brazilian trade liberalization (τ_a going from 1.32 to 1.12), informality rates would increase but unemployment rates would not respond (in the long run). This is precisely the story told by Dix-Carneiro and Kovak (2017a), but within a difference-in-difference framework, which cannot identify the absolute effects as the structural analysis in this paper allows.

Figure 4: Informality Rate vs τ_a

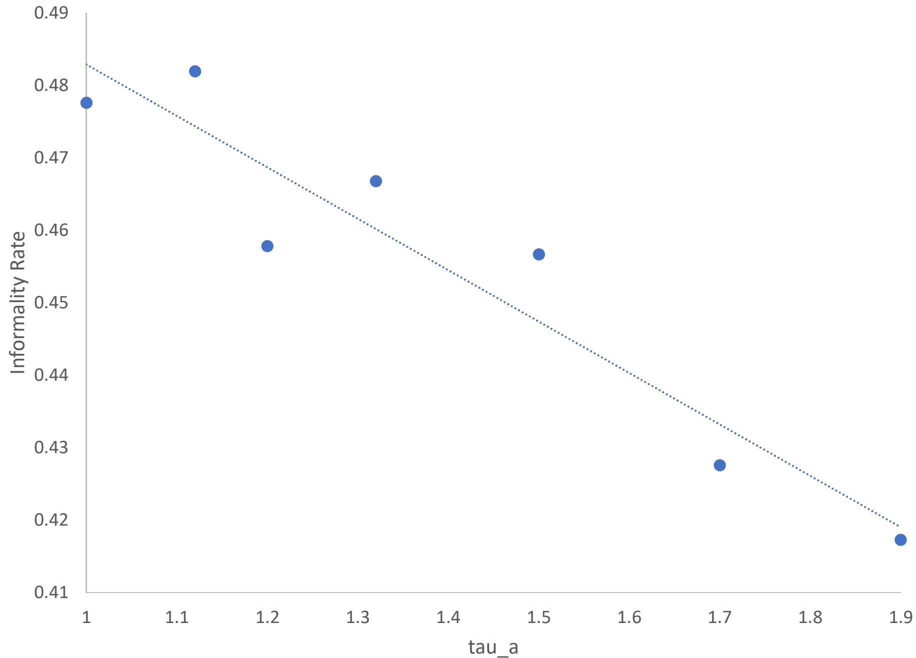


Figure 5: Share of Informality Within the C-Sector vs τ_a

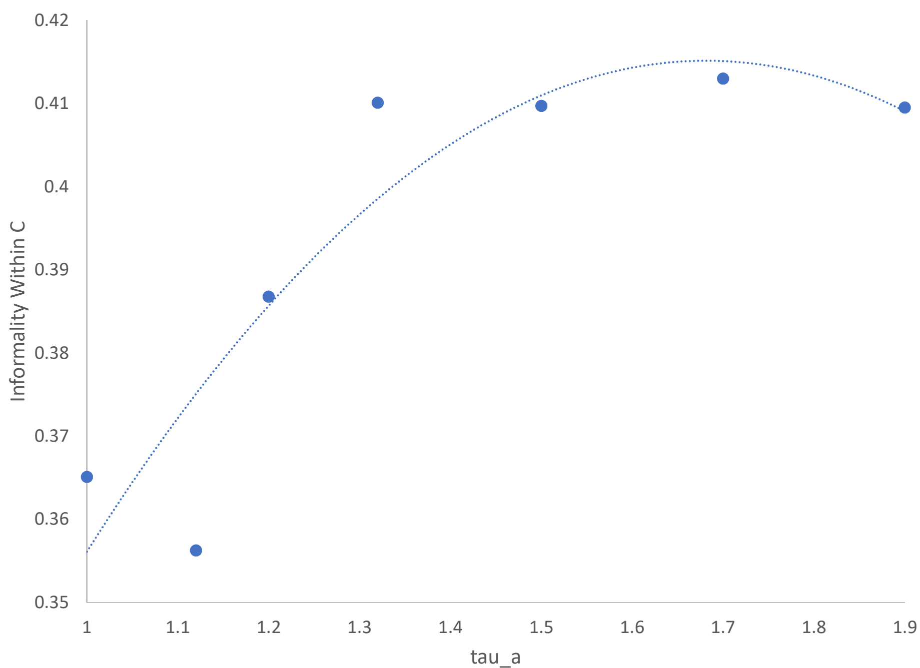


Figure 6: Share of Informality Within the S-sector vs τ_a

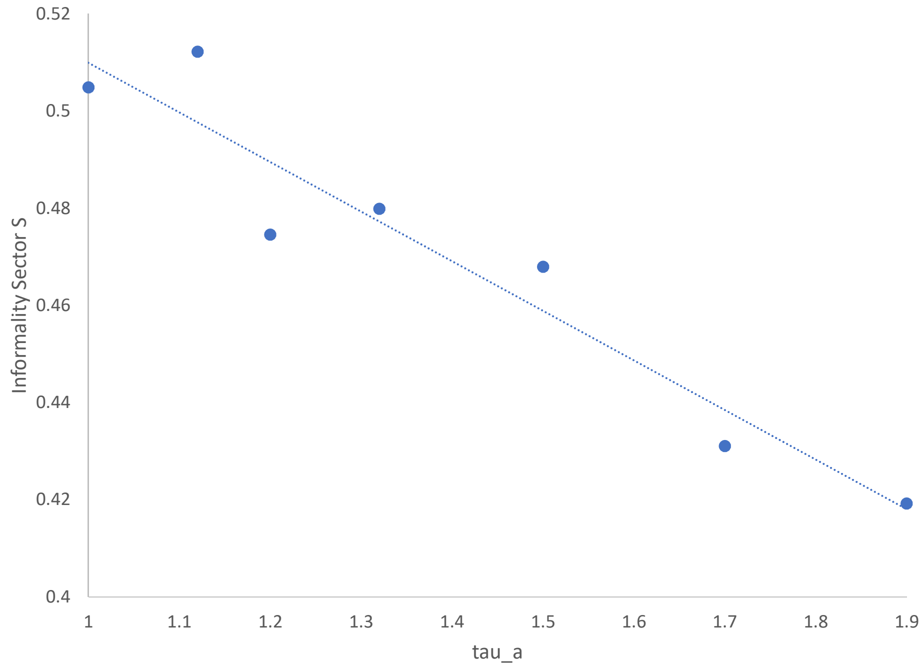
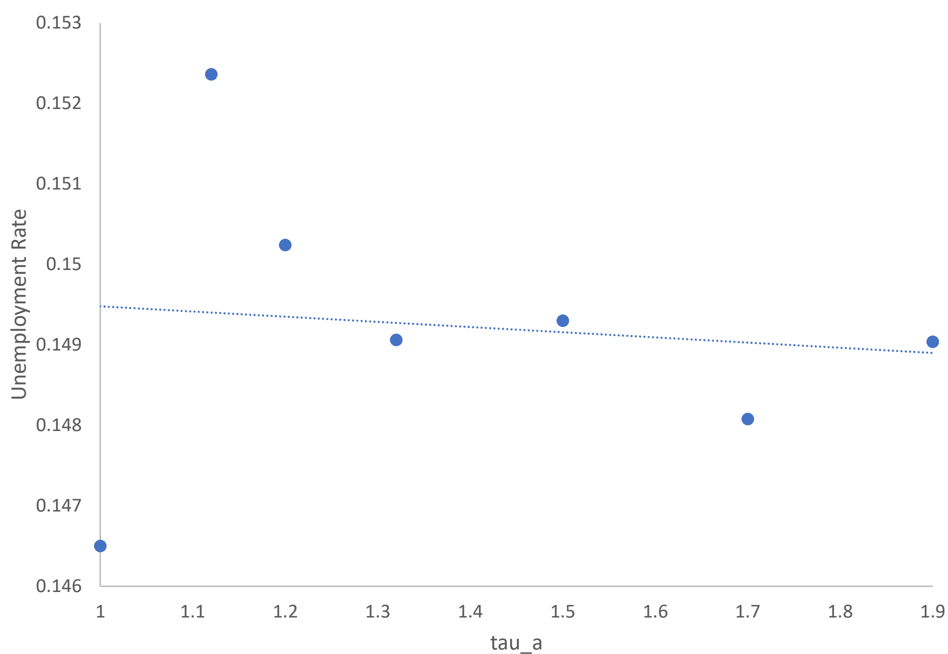


Figure 7: Unemployment Rate vs τ_a



7 Next Steps

- Conclude the estimation (add standard errors!).
- Extend the counterfactual experiments:
  - Redo Figures 4 through 7 when we shut down the informal sector (by greater enforcement).
  - Redo Figures 4 through 7 when we reduce the costs of formality by: (i) reducing fixed costs of operation in the formal sector; (ii) reducing payroll taxes; and (iii) reducing the minimum wage.
- Extend the model to incorporate: (i) risk aversion; (ii) public goods; (iii) intensive margin of informality; (iv) worker heterogeneity and (v) intermediate inputs.

References

Ackerberg, D., C. L. Benkard, S. Berry, and A. Pakes (2007). Econometric tools for analyzing market outcomes. *Handbook of econometrics 6*, 4171–4276.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Adamopoulos, T. and D. Restuccia (2014). The size distribution of farms and international productivity differences. *The American Economic Review 104*(6), 1667–1697.

Alvarez, J., F. Benguria, N. Engbom, and C. Moser (2018). Firms and the decline in earnings inequality in brazil. *American Economic Journal: Macroeconomics 10*(1), 149–89.

Bacchetta, M., E. Ernst, and J. P. Bustamante (2009). *Globalization and informal jobs in developing countries*. International Labour Organization, Geneva.

Bertola, G. and P. Garibaldi (2001). Wages and the size of firms in dynamic matching models. *Review of Economic Dynamics 4*(2), 335–368.

Bosch, M., E. Goñi-Pacchioni, and W. Maloney (2012). Trade liberalization, labor reforms and formal–informal employment dynamics. *Labour Economics 19*(5), 653–667.

Botero, J. C., S. Djankov, R. L. Porta, F. Lopez-de Silanes, and A. Shleifer (2004). The regulation of labor. *The Quarterly Journal of Economics 119*(4), 1339–1382.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Carvalho, C. C., R. Corbi, and R. Narita (2018). Unintended consequences of unemployment insurance: Evidence from stricter eligibility criteria in brazil. *Economics Letters 162*, 157–161.

Cosar, A., N. Guner, and J. Tybout (2016). Firm dynamics, job turnover, and wage distributions in an open economy. *The American Economic Review 106*(3), 625–663.

De Barros, R. P. and C. H. Corseuil (2004). The impact of regulations on brazilian labor market performance. In *Law and Employment: Lessons from Latin America and the Caribbean*, pp. 273–350. University of Chicago Press.

De Soto, H. (1989). *The Other Path*. Harper e Row, New York.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica 82*(3), 825–885.

Dix-Carneiro, R. and B. K. Kovak (2017a). Margins of labor market adjustment to trade. Technical report, National Bureau of Economic Research.

Dix-Carneiro, R. and B. K. Kovak (2017b). Trade liberalization and regional dynamics. *American Economic Review 107*(10), 2908–46.

Doing Business (2007). Paying taxes - the global picture. The World Bank: Washington, DC.

Garicano, L., C. Lelarge, and J. Van Reenen (2016). Firm size distortions and the productivity distribution: Evidence from france. *American Economic Review 106*(11), 3439–79.

Gerard, F. and G. Gonzaga (2018). Informal labor and the efficiency cost of social programs: Evidence from the brazilian unemployment insurance program. Mimeo.

Goldberg, P. K. and N. Pavcnik (2003). The response of the informal sector to trade liberalization. *Journal of Development Economics 72*(2), 463–496.

Goldberg, P. K. and N. Pavcnik (2007). Distributional effects of globalization in developing countries. *Journal of economic literature 45*(1), 39–82.

Gonzaga, G., W. F. Maloney, and A. Mizala (2003). Labor turnover and labor legislation in brazil [with comments]. *Economia 4*(1), 165–222.

Gourieroux, C. and A. Monfort (1996). *Simulation-Based Econometric Methods*. Oxford University Press.

Guner, N., G. Ventura, and Y. Xu (2008). Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics 11*(4), 721–744.

Harrison, A. and J. Scorse (2010). Multinationals and anti-sweatshop activism. *American Economic Review 100*(1), 247–73.

Harrison, A., J. Scorse, S. Collins, and K. A. Elliott (2003). Globalization's impact on compliance with labor standards [with comments and discussion]. In *Brookings trade forum*, pp. 45–96.

Heckman, J. J. et al. (2000). The cost of job security regulation: evidence from latin american labor markets. Technical report, National bureau of economic research.

Helpman, E. and O. Itskhoki (2010). Labour market rigidities, trade and unemployment. *The Review of Economic Studies 77*(3), 1100–1137.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. J. Redding (2017). Trade and inequality: From theory to estimation. *The Review of Economic Studies 84*(1), 357–405.

Hsieh, C.-T. and P. J. Klenow (2009). Misallocation and manufacturing tfp in china and india. *Quarterly Journal of Economics 124*(4), 1403 – 1448.

Kaas, L. and P. Kircher (2015, October). Efficient firm dynamics in a frictional labor market. *American Economic Review 105*(10), 3030–60.

La Porta, R. and A. Shleifer (2008). The unofficial economy and economic development. *Brooking Papers on Economic Activity 105*(3), 473–522.

McCaig, B. and N. Pavcnik (2018). Export markets and labor reallocation in a poor country. *American Economic Review Forthcoming.*

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

Menezes-Filho, N. A. and M.-A. Muendler (2011). Labor reallocation in response to trade reform. NBER Working Paper No. 17372.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

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

Ponczek, V. and G. Ulyssea (2018, May). Is informality an employment buffer? evidence from the trade liberalization in brazil. Unpublished manuscript.

Smith, A. (2008). Indirect inference. *The New Palgrave Dictionary of Economics, 2nd Edition (forthcoming).*

Stole, L. A. and J. Zwiebel (1996). Intra-firm bargaining under non-binding contracts. *The Review of Economic Studies 63*(3), 375–410.

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

Ulyssea, G. (2018). Firms, informality and development: Theory and evidence from brazil. *American Economic Review Forthcoming.*

Appendix

A Model Appendix

A.1 Equilibrium Conditions

Firm Flows

Iin order to state the equilibrium flow conditions for firms, we first define the relevant quantities:

- Fraction of formal firms exiting sector k:

$$\varrho_{kf}^{exit} = \alpha_{kf} + (1 - \alpha_{kf}) \int_z \int_\ell I_k^{exit}(z, \ell, f)\, \psi_{kf}(z, \ell)\, d\ell dz$$

- Fraction of informal firms exiting sector k:

$$\varrho_{ki}^{exit} = \alpha_{ki} + (1 - \alpha_{ki}) \int_z \int_\ell \left(I_k^{exit}(z, \ell, i) + I_k^{change}(z, \ell, i) \right) \psi_{ki}(z, \ell)\, d\ell dz$$

- Fraction of informal firms changing status in sector k:

$$\varrho_{ki}^{change} = (1 - \alpha_{ki}) \int_z \int_\ell I_k^{change}(z, \ell, i)\, \psi_{ki}(z, \ell)\, d\ell dz$$

With these quantities, it is straightforward to write the equilibrium flow conditions for formal and informal firms, that characterize the steady state equilibrium:

$$\begin{aligned} \varrho_{kf}^{exit} N_{kf} &= M_{kf} + \varrho_{ki}^{change} N_{ki} \\ \varrho_{ki}^{exit} N_{ki} &= M_{ki} \end{aligned}$$

where M_{kj} is the number of entrants and N_{kj} is the number of firms in the beginning of the period in sector $k = C, S$ and $j = i, f$.

Service Sector Market Clearing

Service sector goods are used for final consumption (consumers spend $(1 - \zeta) I$ on it), and as inputs for hiring costs, fixed costs of operation, and fixed costs of exporting. We first define the following key objects:

- Average hiring costs:

$$\overline{H}_{kj} = \int_{z'} \int_\ell H_{kj}(\ell, L_k(z', \ell, j))\, I^{hire}(z', \ell, j)\, \widetilde{\psi}_{kj}(z', \ell)\, d\ell dz', \text{ for } k = C, S; j = i, f$$

- Fraction of tradable-sector goods firms that export

$$\mu_x = \int_z \int_\ell \psi(z,\ell) I^x(z,\ell) \, d\ell dz$$

With these objects, one can write the total expenditures on service goods, which is given by:

$$X_S = (1-\zeta) I + \sum \left(N_{kj} \left(\overline{H}_{kj} + \overline{c}_{kj} \right) + M_{kj} K_{kj} \right) + N_{Cf} \mu_x f_x$$

and in equilibrium these expenditures must be equal to sector S total production.

Government Budget

As mentioned in the text, we assume that all government revenue G (taxes and firing costs) that are not spent in unemployment benefits – T – are rebated to consumers. The government's budget constraint is thus given by

$$T = G - \underbrace{\left(b^u \times \underbrace{\sum_k \left(W_{kf}^{DS} + W_{kf}^{EE} + W_{kf}^{D} \right)}_{\text{mass of } \textbf{formal} \text{ workers who transition to unemployment}} \right)}_{\text{Total Expenditure with Unemployment Benefits}}$$

where government's revenue is given by

$$\begin{aligned} G = \; & \sum_k N_{kf} \tau_y \int_z \int_\ell R_k(z,\ell) \psi_{kf}(z,\ell) \, d\ell dz + \\ & \sum_k N_{kf} \tau_w \int_z \int_\ell w_{kf}(z,\ell;\underline{w}) \ell \psi_{kf}(z,\ell) \, d\ell dz + \\ & \sum_k N_{ki} \int_z \int_\ell p_{ki}(\ell) R_k(z,\ell) \psi_{ki}(z,\ell) \, d\ell dz' + \\ & \sum_k N_{kf} \kappa \int_{z'} \int_\ell \widetilde{\psi}_{kf}(z',\ell) (\ell - L_k(z',\ell,f)) \left(1 - I^{hire}(z',\ell,f)\right) d\ell dz' + \\ & (\tau_a - 1) \frac{D_{H,C} (\epsilon \tau_a \tau_c)^{1-\sigma}}{\tau_a} \end{aligned}$$

B Estimation Appendix

B.1 Estimation Algorithm

In this section we describe the estimation algorithm in detail, which we break down into several steps for expositional clarity.

Step 1: Fix employment L_{kj} ($k = C, S$ and $j = i, f$) and unemployment L_u to data values (PME). The estimation procedure will perfectly target these numbers.

Step 2: $d_{H,C}$, $d_{H,S}$, and μ^v are treated as parameters to be estimated, along with the remaining ones. These are endogenous variables, so we will penalize deviations between these "parameters" and the equilibrium quantities that arise (see Step 10).

Step 3: Compute wage schedules.

$$w_{kf}(z,\ell) = \frac{(1-\beta_f)(b+b^u)}{1+\beta_f\tau_w} + \frac{\beta_f(1-\tau_y)}{1+\beta_f\tau_w}\frac{R_k(z,\ell)}{\ell} - \frac{\beta_f}{1+\beta_f\tau_w}\frac{\overline{c}_{kf}}{\ell}$$

$$w_{kf}\left(z,\ell;\underline{w}_f\right) = \max\left\{w_{kf}(z,\ell), \underline{w}_f\right\}$$

$$w_{ki}(z,\ell) = (1-\beta_i)\,b + \beta_i(1-p_{ki}(\ell))\frac{R_k(z,\ell)}{\ell} - \beta_i\frac{\overline{c}_{ki}}{\ell}$$

$$w_{ki}(z,\ell;\underline{w}_i) = \max\left\{w_{ki}(z,\ell), \underline{w}_i\right\}$$

Where $\underline{w}_f$ is the minimum wage in the formal sector (which is observed and fixed throughout estimation). $\underline{w}_i$ is the first percentile of the distribution of informal wages in PME and fixed throughout the estimation procedure. This is to avoid zero or negative informal wages.

Step 4: Compute firms' value functions. Obtain firms' policy functions. Solve firms' entry decisions and obtain thresholds for entry in formal and informal sectors.

Step 5: Compute steady state distribution of states. For informal firms, start with a guess for ψ_{ki}. Then, compute

$$\psi^e_{ki}(z') = \frac{\int_{\overline{\nu}_{kf}>\nu\geq\overline{\nu}_{ki}} g_k(z'|\nu)\, g^e_k(\nu)\, d\nu}{\int_{z'}\int_{\overline{\nu}_{kf}>\nu\geq\overline{\nu}_{ki}} g_k(z'|\nu)\, g^e_k(\nu)\, d\nu dz'}$$

$$\begin{aligned}
\varrho^{exit}_{ki} &= \alpha_{ki} + (1-\alpha_{ki})\int_z\int_\ell \left(I^{exit}_k(z,\ell,i) + I^{change}_k(z,\ell,i)\right)\psi_{ki}(z,\ell)\, d\ell dz \\
&= \frac{M_{ki}}{N_{ki}}
\end{aligned}$$

$$\widetilde{\psi}_{ki}(z',\ell) = \mathbf{1}[\ell=1] \times \varrho_{ki}^{exit} \times \psi_{ki}^{e}(z') + \mathbf{1}[\ell \geq 1] \times (1-\alpha_{ki}) \times \left(\int_z \psi_{ki}(z,\ell)\, I_k^{stay}(z,\ell,i)\, g_k(z'|z)dz \right)$$

$$\psi_{ki}(z',\ell') = \frac{\int_\ell \widetilde{\psi}_{ki}(z',\ell)\, I(L_k(z',\ell,i)=\ell')\, d\ell}{\int_{z'}\int_\ell \widetilde{\psi}_{ki}(z',\ell)\, I(L_k(z',\ell,i)=\ell')\, d\ell dz'}$$

And repeat until convergence of ψ_{ki}.
For formal firms, start with guess for ψ_{kf} and compute

$$\psi_{kf}^{e}(z') = \frac{\int_{\nu\geq\bar{\nu}_{kf}} g_k(z'|\nu)\, g_k^e(\nu)\, d\nu}{\int_{z'}\int_{\nu\geq\bar{\nu}_{kf}} g_k(z'|\nu)\, g_k^e(\nu)\, d\nu dz'}$$

$$\frac{N_{ki}}{N_{kf}} = \frac{L_{ki}}{L_{kf}} \frac{\int_z\int_\ell \ell\psi_{kf}(z,\ell)\, d\ell dz}{\int_z\int_\ell \ell\psi_{ki}(z,\ell)\, dz d\ell}$$

$$\varrho_{kf}^{exit} = \alpha_{kf} + (1-\alpha_{kf}) \int_z\int_\ell I_k^{exit}(z,\ell,f)\, \psi_{kf}(z,\ell)\, d\ell dz$$

$$\varrho_{ki}^{change} = (1-\alpha_{ki}) \int_z\int_\ell I_k^{change}(z,\ell,i)\, \psi_{ki}(z,\ell)\, d\ell dz$$

$$\frac{M_{kf}}{N_{kf}} = \varrho_{kf}^{exit} - \varrho_{ki}^{change} \frac{N_{ki}}{N_{kf}}$$

$$\widetilde{\psi}_{kf}(z',\ell) = \begin{array}{c} \mathbf{1}[\ell=1] \times \frac{M_{kf}}{N_{kf}} \times \psi_{kf}^{e}(z') + \\ \mathbf{1}[\ell\geq 1] \times \left(\begin{array}{c} (1-\alpha_{kf}) \times \left(\int_z \psi_{kf}(z,\ell)\, I_k^{stay}(z,\ell,f)\, g_k(z'|z)dz\right) + \\ (1-\alpha_{ki}) \frac{N_{ki}}{N_{kf}} \times \left(\int_z \psi_{ki}(z,\ell)\, I_k^{change}(z,\ell,i)\, g_k(z'|z)dz\right) \end{array} \right) \end{array}$$

$$\psi_{kf}(z',\ell') = \frac{\int_\ell \widetilde{\psi}_{kf}(z',\ell)\, I(L_k(z',\ell,f)=\ell')\, d\ell}{\int_{z'}\int_\ell \widetilde{\psi}_{kf}(z',\ell)\, I(L_k(z',\ell,f)=\ell')\, d\ell dz'}$$

$$= \int_\ell \widetilde{\psi}_{kf}(z',\ell)\, I(L_k(z',\ell,f)=\ell')\, d\ell$$

And repeat until convergence of ψ_{kf}.

Step 6 Obtain aggregate vacancies. Armed with firms' policy functions and steady state distributions, we can compute vacancies per firm and the mass of firms in each sector and type of firm using:

$$N_{ki} \int_z\int_\ell \ell\psi_{ki}(z,\ell)\, dz d\ell = L_{ki}$$

$$N_{kf}\int_z\int_\ell \ell\psi_{kf}\left(z,\ell\right)dzd\ell = L_{kf}$$

So, we can obtain aggregate vacancies V_{kj}.

$$V_{kj} = \int_{z'}\int_\ell v_{kj}\left(z',\ell\right)\widetilde{\psi}_{kj}\left(z',\ell\right)d\ell dz' + M_{kj}\int_{z'}\int_\ell \frac{1}{\mu^v}I\left(\ell=1\right)\psi^e_{kj}\left(z'\right)dz'$$

Step 7 Obtain job finding rates μ^e_{kj} and all moments to be matched with the data.

$$\mu^e_{kj} = \frac{V_{kj}}{V_{Cf}+V_{Ci}+V_{Sf}+V_{Si}}\frac{m\left(V_{Cf},V_{Ci},V_{Sf},V_{Si},L_u\right)}{L_u}$$

Step 8 Equilibrium restriction imposed in the loss function.

$$L_1 = LargeWeight \times abs\left(\frac{\mu^v - \frac{L_u}{\left(\left(V_{Cf}+V_{Ci}+V_{Sf}+V_{Si}\right)^\theta + L_u^\theta\right)^{1/\theta}}}{\mu^v}\right)$$

Step 9: Equilibrium restriction that the mass of entrants in the formal sector is non-negative

$$L_2 = LargeWeight \times abs\left(\min\left(\varrho^{exit}_{Cf} - \varrho^{change}_{Ci}\frac{N_{Ci}}{N_{Cf}},0\right)\right)$$
$$L_3 = LargeWeight \times abs\left(\min\left(\varrho^{exit}_{Sf} - \varrho^{change}_{Si}\frac{N_{Si}}{N_{Sf}},0\right)\right)$$

Step 10: Equilibrium restrictions on demand shifters. Compute aggregate income I, exchange rate ϵ, service sector expenditures R, and price indices P_C and P_S. Use balanced trade to obtain the exchange rate ϵ (which enters P_C) and tax revenues, which enter I. At this point, we can obtain total exports generated by the model.

$$imports = \frac{D_{H,C}\left(\epsilon\tau_a\tau_c\right)^{1-\sigma}}{\tau_a} = exports$$
$$\Rightarrow \epsilon = \frac{1}{\tau_a\tau_c}\left(\frac{\tau_a \times exports}{exp(d_{H,C}\times\sigma)}\right)^{\frac{1}{1-\sigma}}$$

$$Tax\ Revenue = \left(\tau_a - 1\right)imports = \left(\tau_a - 1\right)exports$$

$$L_4 = LargeWeight \times abs\left(\frac{d_{H,C} - \log\left(\left(\frac{\zeta I}{P_C^{1-\sigma}}\right)^{\frac{1}{\sigma}}\right)}{d_{H,C}}\right)$$

$$L_5 = LargeWeight \times abs\left(\frac{d_{H,S} - \log\left(\left(\frac{(1-\zeta)I+R}{P_S^{1-\sigma}}\right)^{\frac{1}{\sigma}}\right)}{d_{H,S}}\right)$$

Step 11: Equilibrium restrictions in the allocation of labor. Note that

$$L'_{kf} = \left(1 - \chi_{kf}^{layoff}\right) L_{kf} + L_{ki}\chi_{ki\to f}^{change} + L_u \mu_{kf}^{e}$$

$$L'_{ki} = \left(1 - \chi_{ki}^{leave}\right) L_{ki} + L_u.\mu_{ki}^{e}$$

$$L'_u = L - \sum L'_{ki} - \sum L'_{kf}$$

So that

$$L_6 = LargeWeight \times abs\left(\frac{L'_{Cf} - L_{Cf}}{L_{Cf}}\right)$$

$$L_7 = LargeWeight \times abs\left(\frac{L'_{Ci} - L_{Ci}}{L_{Ci}}\right)$$

$$L_8 = LargeWeight \times abs\left(\frac{L'_{Sf} - L_{Sf}}{L_{Sf}}\right)$$

$$L_9 = LargeWeight \times abs\left(\frac{L'_{Si} - L_{Si}}{L_{Si}}\right)$$

$$L_{10} = LargeWeight \times abs\left(\frac{L'_u - L_u}{L_u}\right)$$

$$L_{11} = LargeWeight \times abs\left(\min\left\{L'_u, 0\right\}\right)$$

Step 12: Compute Loss Function. Add Model/Data deviations to equilibrium restrictions L_1 through L_{12}. The objective function is therefore given by

$$L = L_{mom} + \sum_{k=1}^{12} L_k$$

Where L_{mom} penalizes deviations between moments in the data and $\{L_k\}_{k=1}^{12}$ penalize deviations from equilibrium restrictions.

Step 13: Optimization routine picks new parameter vector. Go back to Step 3.

Step 14 (Post-estimation): Obtain D_F^* (this is what we need for the counterfactuals...

d_F is endogenous)

$$d_F = \ln\left(\left(\frac{D^*_{F,C}}{exp(d_{H,C}\times\sigma)}\epsilon^\sigma\tau_c + \tau_c^\sigma\right)^{\frac{1-\sigma}{\sigma}}\left[\tau_c^{\sigma-1} + \frac{D^*_{F,C}}{exp(d_{H,C}\times\sigma)}\epsilon^\sigma\right]\right)$$

Solve above equation to obtain $D^*_{F,C}$.