University of Nottingham
UK | CHINA | MALAYSIA

research paper series

Globalisation and Labour Markets Programme

Research Paper 2022/11

Firms, policies, informality and the labour market

Camila Cisneros-Acevedo and Alessandro Ruggieri

Firms, policies, informality, and the labor market*

Camila Cisneros-Acevedo† University of Tübingen

Alessandro Ruggieri‡ University of Nottingham

October 19, 2022

Abstract

In this paper we study how firm-level policies affect labor market outcomes in the developing world. For a large set of low- and medium-income countries, we document that high corporate income tax rate is associated to a higher rate of informal employment, lower GDP per worker and lower unemployment rate. To interpret this evidence, we build a general equilibrium model that features: 1) industry dynamics dictated by heterogeneous firms, 2) search and matching friction, 3) imperfectly enforced legislation, leading to informal employment along the intensive and the extensive margin. We estimate the model using firm and worker-level data from Peru — a country where more than 70 per cent of the population is employed in informal jobs, and use the model as a laboratory to assess the distributional consequences of firm-level taxes and policies. The model generates the observed cross-country relation between corporate income tax rate and labor market outcomes: a reduction in corporate taxes concentrates employment over a smaller mass of larger and more productive firms, increasing efficiency and reallocating workers to formal employment at the expense of a higher unemployment rate. Quantitatively, changes in tax rates account for about 60% of the difference in unemployment rate and 45% of the differences in GDP per worker observed across countries. The model provides a basis for the evaluation of various firm-level policy interventions.

Keywords: corporate tax, firm dynamics, informality, unemployment, welfare
JEL Classification: J46, F16, J31, O17

*All errors are ours.

†Department of Economics, Nauklerstr. 47, University of Tübingen, 72074 Tübingen, Germany, camila.cisneros-acevedo@uni-tuebingen.de.

‡School of Economics, Sir Clive Granger building, University of Nottingham, University Park, Nottingham, NG7 2RD, aruggierimail@gmail.com.

1 Introduction

Over 60 percent of workers in the world operate in the informal economy. Informal employment is particularly widespread in developing countries, where it accounts on average for about 35 percent of GDP and about 70 percent of the labor force (Loayza, 2016). Poor governance has been identified as one major cause of informality (De Soto, 1989) and policies aiming at reducing tax and regulatory burden on firms have been advocated to discourage the creation of informal jobs and boost aggregate income (Ohnsorge and Yu, 2022).

In this paper, we study the distributional consequence of such policies in the presence of informality. While the aggregate effects of these policies have been extensively emphasized (OECD, 2006; Card et al., 2010; McKenzie, 2017), the presence of frictions that limit the reallocation of labour across jobs in developing countries motivates us to focus on their implications across the distribution of workers and firms. On the one hand, firms in developing countries face the burden of a heavy regulation, which encourages them to remain informal and distort employment decisions (La Porta and Shleifer, 2014; Oviedo et al., 2009). On the other hand, workers in developing countries face poorly functioning labour markets with relatively high search frictions (Poschke, 2019; Rud and Trapeznikova, 2021). These frictions are largely due to geographical constraints (Lagakos, 2020), lack of job search support (Abebe et al., 2021) and firm market power (Brooks et al., 2021; Amodio et al., 2022). Understanding who reaps the benefits and who bears the costs of firm-level policies when frictions impede the correct functioning of the labor market is a first order question in economic development (see Donovan and Schoellman (2021) for a review).

To study this issue, we focus on the corporate income tax, a widely common policy instrument. Corporate taxes are extensively used by governments in developing countries and are a key source of government revenues (OECD, 2018).[1,2] Most importantly, high corporate taxes have been commonly cited as a major reason for informal activity (Perry, 2007; Waseem, 2018). They have been shown to negatively correlate with economy growth (Lee and Gordon, 2005) and policy makers have advised against them to reduce distortions on prices and the composition of consumption (Gordon and Li, 2009).

In this paper we document significant cross-country heterogeneity in the statutory tax rates on corporate income and show that these differences are associated to differential

[1]Despite decreasing over the last two decades, corporate income taxes in 2018 comprise on average 15.3% of all tax revenues in Africa, 15.4% in Latin America and the Caribbean, and 10% in OECD countries. They represent more than 25% of total tax revenues in many low-income countries like Bhutan, Chad, Democratic Republic of the Congo, Equatorial Guinea, Indonesia Nigeria and Papua New Guinea, among others.

[2]A parallel literature has shown that lack of state capacity and inability to collect taxes efficiently is at the heart of why low-income countries are as poor as they are (Dincecco and Katz, 2016). See Dzansi et al. (2022) for a discussion.

labor market outcomes. In particular, using a large sample of low- and middle-income countries, we show that countries with lower tax rates have a higher share of formal employment, higher GDP per worker, and a higher unemployment rate.

Motivated by these evidence we build a model of firm dynamics that features search frictions in the labor market and imperfectly enforced legislation, leading to informal employment. In the model, workers choose whether to search for firms in a frictional labor market or to secure a job in a frictionless outside sector. Among those who search, some match with firms, while others remain unemployed and keep looking for a job. Firms are ex-ante heterogeneous in their productivity and in the cost of setting up a formal business. First, they choose whether to enter the industry or not. Upon entry, they decide whether to register with the tax authority or to hide. Costly registration encourages firms to stay informal, generating informal employment along the *extensive margin*. Legislation is imperfectly enforced, hence informal firms are able to forego corporate income taxes subject to a monetary penalty if caught by the regulatory authority. While unregistered firms have access to informal workers only, registered firms can hire workers either formally or off-the-book, allowing them to trade-off payroll taxes and higher hiring expenses against a monetary fine if audited by the tax authority. When registered firms decide to hire workers off-the-books, they generate informal employment along the *intensive margin*. Through workers' and firms' dynamics, the model economy generates a collection of labor market outcomes that can be compared with the data.

We estimate the model using firm- and worker-level data for Peru. The choice of Peru as a benchmark economy reflect the following two considerations. First, Peru is a country with a very high informality rate: over 70 per cent of the population is employed informally, either along the intensive or the extensive margin. Second, the availability of data on informal firms and workers allow us to identify parameters governing expected costs of informality faced by either informal business or by registered companies. The estimated model closely matches basic features of Peruvian data. In particular, it replicates the size distribution of formal and informal firms, the share of informal workers within formal firms of different size, and different aggregate labor market outcomes. The model also reproduces the observed wage gaps between formal and informal workers.

We then turn to cross-country differences. We generate several counterfactual economies by applying alternative corporate tax rates to the baseline model, while keeping all the other parameters fixed at their estimated values. Quantitatively, corporate tax rates account for the entire difference in informality rate observed in the cross-country dataset, for about 60% of the observed differences in unemployment rate and for about 45% of the differences in GDP per worker.

The model delivers cross-country patterns in informality and unemployment through two major mechanisms: a *reallocation effect* and a *concentration effect*. The first effect operates through changes in firm-level registration decisions and general equilibrium

forces in the product market. A reduction in corporate income tax increases net revenues for the formal firms, relative to those informal. As a consequence, the share of registered firms in the economy increases, and, as they expand in size, the composition of the vacancies posted shifts towards formal jobs. These changes trigger a reallocation of workers from informal to formal jobs, reducing the overall informality rate.

In addition, lowering corporate income taxes allows formal businesses to charge a lower price for their varieties, forcing informal firms to leave the industry. Higher selection triggers a re-allocation of employment from low-productivity to high-productivity firms, and a reduction in aggregate price, which increase real output produced per worker employed.

The second effect operates instead through general equilibrium forces in the labor market. Because of improvements in allocative efficiency, a reduction in corporate income taxes increases the average wage of workers in industrial firms. This rises the expected value of searching for a job in the industry relative to the value of not searching. To restore the equilibrium in the labor market, jobs concentrate on a smaller mass of large and high-productivity firms, making jobs in the industry relatively scarce. These changes trigger an increase in labor market tightness and a reduction in the job finding rate, leading to higher unemployment rate.

Finally, we use the estimated model as a laboratory to compare the welfare properties of corporate income taxes to alternative policy interventions, including i) changes in the regulatory cost of being informal firms, ii) changes in workers payroll taxes, and iii) changes in the regulatory cost of hiring workers off-the-book for formal firms. Although each of these policies alter firms and labor market outcomes, their effects vary depending on whether they tackle formalization along the extensive margin — like policy i), or along the intensive margin — like policies ii) and iii).

Within the model, an increase in the regulatory cost for informal firms produces the same effects as those described for a reduction in corporate income taxes. However, the implications for welfare are quantitatively different. On the one hand, a monotonic trade-off between higher workers' welfare and lower unemployment rate arises as a result of both policies - the higher the gains in welfare, the lower the share of workers it accrues to. On the other hand, changes in the expected cost of informality do not produce as much welfare gains as corporate income taxes do: for the same increase of 1 percentage points in unemployment rate, a reduction in corporate taxes generates almost 4 times higher gains in welfare.

Policies that target formalization along the intensive margin have instead an ambiguous effect on unemployment rate. Although they both affect the value of being a formal firm, such policies have also an impact on the allocation of formal and informal workers within firms. As we lower the effective costs of informality along the intensive margin, either by cutting payroll taxes on formal workers, or by reducing the expected monetary

fine from hiring workers off-the-book, unemployment increases or declines depending on the relative effect of firm's employment decision versus registration decision. We find that equilibria with a higher share of workers in informal firms are welfare-dominated by those with higher share of workers hired off-the-books by registered businesses.

This paper contributes to different strands of literature. First, it speaks to the literature on the macroeconomic effects of corporate taxation in developing countries. De Paula and Scheinkman (2010) use firm-level data from Brazil to show that collecting value added taxes according to a credit scheme transmit informality over the supply chain. Narita (2020) estimates a life-cycle search model with self-employment and shows that a flat reduction in payroll taxes increases the share of formal sector workers mainly due to a drop in self-employment. Baumgartner et al. (2022) use administrative linked employer-employee data to study the effect of a large payroll tax reform in Brazil and find a long-run increase of formal employment due to both firm growth and firm entry. We complement this literature by studying the long-run effects of corporate income taxes on the labor market. We show that changes in corporate taxation have non trivial effects on labor market outcomes, increasing formal employment at the expense of a higher unemployment rate.

More generally, our analysis contributes to highlight the macro-implications of informal employment. A growing literature has showed that informality acts like a buffer that absorbs workers subject to labor market shocks (Dix-Carneiro and Kovak, 2019; Dix-Carneiro et al., 2021; Ponczek and Ulyssea, 2022). Ulyssea (2018) study the role of both margins of informality on output, TFP, and welfare. Erosa et al. (2021) develop a model of entrepreneurship to study the interaction between financial constraints and informality. However, both papers abstract from studying the distributional implications of search frictions and informality on labor market outcomes. Meghir et al. (2015) use a search model to study labor outcomes of formal and informal workers, but abstract from modeling firm dynamics and the role of tax policies. We contribute to the literature by focusing on the long-run consequence of firm regulations on labor market outcomes in the context of a developing country, using a model of frictional labor market where unemployment and informality endogenously respond to changes in firm-level policy interventions.

Finally, this paper also contributes to the literature that looks at labor market outcomes over development. Feng et al. (2018) use household survey data from countries of all income levels to document that the unemployment rate is increasing with GDP per capita. They rationalize this evidence within a two-sector model and show that as productivity of the modern-sector rises, the traditional sector shrinks, as progressively less-able workers enter the modern sector, leading to a rise in overall unemployment. Poschke (2019) documents that low-income countries have high rates of unemployment relative to wage employment, and that self-employment is particularly high where the

unemployment-wage employment ratio is high. He interprets this facts within a search model and show that labor market frictions can reduce aggregate output by pushing searchers into low-productivity own-account work. Donovan et al. (2020) documents that labor market flows such as job-finding rates, employment-exit rates, and job-to-job transition rates are significantly higher in the poorest countries, and show that this is consistent with theories of endogenous separation, like job ladder and learning models. We add to this literature by documenting how unemployment and informality rate vary with corporate income tax rates across low- and medium-income countries and use a structural model to study aggregate and distributional implications of various firm-level policies.

The remainder of the paper goes as follows. Section 2 documents cross-country evidence on corporate taxes and labor market outcomes. Section 3 describes our quantitative model. In Section 4 we introduce firm- and worker-level data and discuss the estimation strategy. We report our main quantitative results and counterfactual exercises in Section 5 and analyze alternative firm-level policies in Section 6. We conclude in Section 7.

2 Cross-Country Evidence

This section documents how labor market outcomes - such as informal employment and unemployment rate - and aggregate productivity vary across low- and medium-income countries with different corporate income tax rates.

The analysis draws from three data sources. Corporate income taxes are taken from the Tax Foundation (TF) database.[3] The dataset records the standard top statutory corporate income tax rates levied on domestic businesses for about 200 countries in the last 40 years. We merge this information to country-level data on informal employment and unemployment rates, sourced from the ILO-stat database. Informal employment is reported as a share of overall employment and comprises persons who in their main or secondary jobs were either i) own-account workers, or ii) contributing family workers, or iii) employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises, or as paid domestic workers by households. Informal jobs of employees are defined as those lacking of coverage by social security system, entitlement to paid annual or sick leave, or written employment contract. Unemployment comprises persons of working age who were not in employment, carried out activities to seek employment and were currently available to take up employment given a job opportunity. Both measures are constructed using the sample of workers with more than 25 years old. Finally, we use real GDP per worker as a measure of aggregate productivity and take it from the World-Bank Indicator database.

[3] https://taxfoundation.org/global-tax/corporate-income-taxes

Table 1: Cross-country summary

	Obs	Mean	St.dev.	Min	Max
GDP per capita, 2017 USD	367	5677.28	3897.49	370.301	16950.3
GDP per worker, 2017 USD	367	31124.1	16035.1	2583.41	72420.6
Corporate tax rate, %	367	24.9	7.36	10.0	38.5
Informality rate, %	367	60.4	21.6	9.90	96.9
Unemployment rate, %	367	6.88	6.22	0.21	29.3

Notes: Informal employment is expressed as percent of total employment and comprises persons who in their main or secondary jobs were own-account workers, contributing family workers, employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises, or as paid domestic workers by households. Informal jobs of employees are defined as those lacking of coverage by social security system, entitlement to paid annual or sick leave, or written employment contract. Unemployment rate is reported in percent of the labor force. Corporate tax rates refer to the standard top statutory corporate income tax rates levied on domestic businesses. GDP per worker is measured in 2017 USD and expressed in 1000 USD. Source: Tax Foundation, ILO-stat, World-Bank and authors' calculation.

Overall, we gather data for 75 countries, spanning the period 2010-2021, which makes our dataset an unbalanced panel of 367 country-year observations. Details on the data coverage are provided in Appendix A.

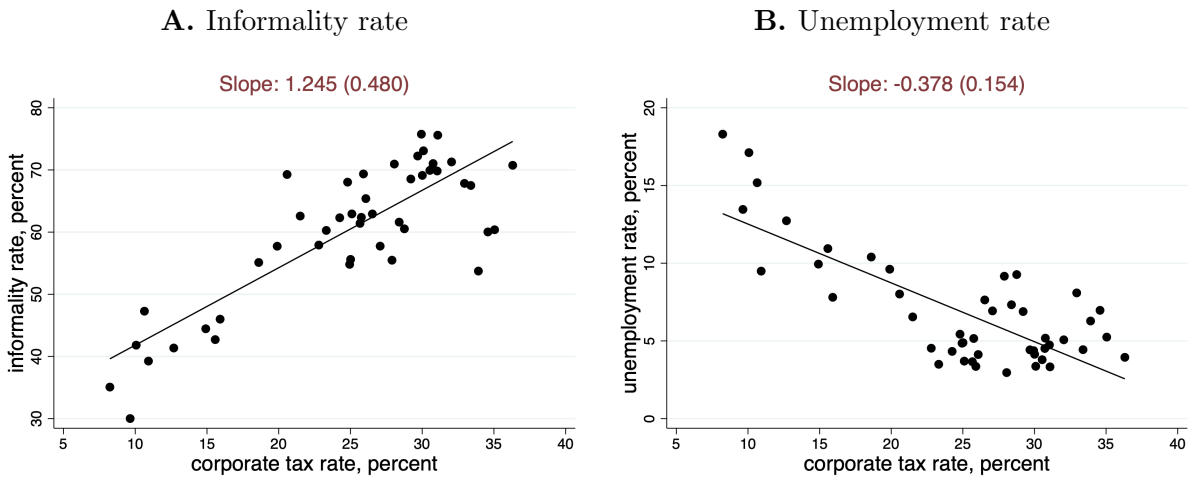
Table 1 reports a few summary statistics for the corporate tax rates, labor market outcomes and measures of real GDP. On average, countries in the sample have a yearly GDP per capita (at 2017 price level) of 5,677 USD: the poorest country in the sample is Malawi, with a GDP per capita of about 1 USD per day (370 USD yearly), while the richest country is the Barbados, with a yearly GDP per capita of 16,950 USD. On average, the GDP per worker, a standard measure of aggregate productivity, amounts to 31,124 USD. To place it context, the analogous measure for the US in 2021 was equal to 134,363 USD, a value about 4.3 time larger.

The average tax rate levied on corporate income is 24.9%, spanning a range that goes from a minimum of 10% to a maximum of 38%. Informal employment is large and widespread across countries in the sample: on average, about 60% of employment is informal, reaching more than 95% in sub-saharan countries (e.g. Benin, Chad and Mali). Finally, the unemployment rate amounts to 7% on average, although it is heterogeneous across countries and it is almost zero in Cambodia and Myanmar.

Figure 1 reports the cross-country relations between the statutory corporate income tax rates and i) the rate of informal employment (panel A), and ii) the unemployment rate (panel B). Each dot corresponds to the average outcome for countries in a given percentile of the corporate tax rates.[4] Outcomes are reported as residuals from a regression with

[4]All figures report 50 dots, each corresponding to a 2 percent interval in the distribution of corporate

Figure 1: Informality, unemployment and corporate income taxes



Notes: Informal employment is expressed as percent of total employment and comprises persons who in their main or secondary jobs were own-account workers, contributing family workers, employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises, or as paid domestic workers by households. Informal jobs of employees are defined as those lacking of coverage by social security system, entitlement to paid annual or sick leave, or written employment contract. Unemployment rate is reported in percent of the labor force. Corporate tax rates refer to the standard top statutory corporate income tax rates levied on domestic businesses. Source: ILOSTAT, Tax Foundation and authors' calculation.

year-fixed effects. On top of each panel, we report the slope of these relationships, and in parentheses robust standard errors clustered at country level.[5]

Panel A shows that as we move from low to high corporate tax rate countries, the rate of informal employment significantly increases. Countries with a corporate tax rate of about 10% have on average 40% of informal employment. On the other hand, in countries with a tax rate of 30%, almost 70% of employment is informal. The slope of this relation is large ($\hat{\beta}$= 1.245) and significant at 5% (s.e.= 0.540). To place it in context, this estimate implies that conditional on year fixed-effects, a 10 percent higher corporate tax rate is associated with a rate of informal employment 12.5 percentage points higher.

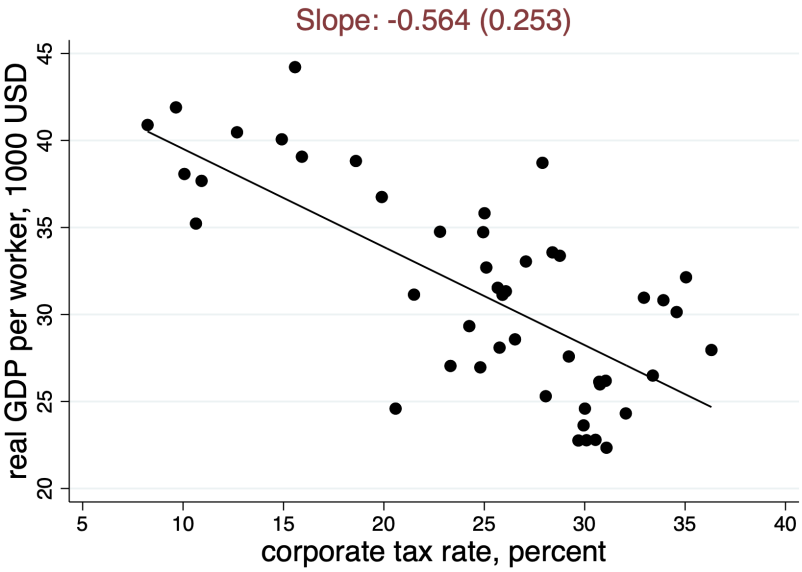
Panel B shows that the opposite pattern holds for the unemployment rate: high corporate tax rates are associated with lower unemployment. Countries with a tax rate of about 10% have on average a rate of unemployment of 15% while in countries with a tax rate of 30%, unemployment rate is about 5%. The slope is this relationship is also large in magnitude ($\hat{\beta}$=-0.378) and significant at 5% (s.e.= 0.154). Conditional on year fixed-effects, a 10 percent higher corporate tax rate is associated with a rate of unemployment 3.78 percentage points lower.

Figure 2 documents how GDP per person employed, expressed in 1000 USD, varies across countries with different corporate tax rates. Like Figure 1, each dot corresponds

income tax rates.

[5]We report the same scatter plots using the raw data in Appendix A, Figure A.1.

Figure 2: GDP per worker and corporate income taxes



Notes: GDP per worker is measured in 2017 USD and expressed in 1000 USD. Corporate tax rates refer to the standard top statutory corporate income tax rates levied on domestic businesses. Source: World-Bank, Tax Foundation and authors' calculation.

to the average values of the dependent variable for countries in a specific bin of corporate income tax, after removing year-fixed effects. GDP per worker declines significantly as countries increase their tax burden on firms. It drops from around 40,000 USD in countries with a tax rate of 10% to around 25,000 USD in countries with a tax rate of 35%. A 10% increase in corporate tax rate is associated with a decline in real GDP per worker of about 5,639 USD. The estimated slope ($\hat{\beta}$ =-0.564) is significant at 5% and implies that a 2% declines in corporate tax rate is associated with an increase in GDP of around 1,000 USD per employed worker.

To sum-up, this section documented three key cross-country patterns. As countries reduce their tax rates on corporate income, the share of informal employment out of total employment declines and the GDP per worker increases at the expense of a higher unemployment rate. In the next section, we develop a model of heterogeneous firms operating in a frictional labor market and use it to understand these patterns.

3 The Model

We consider a model economy that features 1) endogenous firm dynamics, 2) search frictions in the labor market, and 3) informality along the extensive and the intensive margin. We focus on a stationary equilibrium, hence aggregate outcomes are time-invariant.

Time is discrete. The economy is populated a unitary measure of workers-consumers and by an endogenous measure of firms. Workers are ex-ante homogeneous but differ in their employment status: they can be either wage employed in the industrial sector,

employed outside the industrial sector or unemployed. If wage employed, they can differ in their formality status — they can be formally employed, employed off-the-books by registered firms, or informally employed by unregistered firms.

Firms are ex-ante heterogeneous in productivity and in the cost of setting up a formal business. They can be formally registered or not. They post vacancy to hire workers formally (only if registered) and off-the-books, subject to a probability of being audited and receiving a monetary fine.

3.1 Preferences

Workers are infinitely lived and risk neutral. They live hand-to-mouth and derive utility from the consumption of a homogeneous good, s, and a CES bundle c of differentiated varieties $\omega \in [0, M]$, defined as follows:

$$c = \left(\int_0^M c(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}$$

where $\sigma > 1$ is the elasticity of substitution between varieties. The discounted individual utility at time t is equal to:

$$\mathcal{U}_t = \sum_{j=t}^{\infty} \frac{c_j^{\alpha} s_j^{1-\alpha}}{(1+r)^j}$$

where r is the discount rate, while $\alpha \in (0, 1)$ is the elasticity of the composite good in total consumption. Let the price of the homogeneous good be the numeraire of the economy, and let $p(\omega)$ denote the price of variety ω. Utility maximization for a worker i with income I_i yields a demand for the homogeneous good s and for variety ω equal to

$$s = (1-\alpha) I_i \qquad \text{and} \qquad c(\omega) = \alpha \frac{I_i}{P} \left(\frac{p(\omega)}{P} \right)^{-\sigma}$$

respectively, where P denotes the exact price index for the composite good, defined as

$$P = \left(\int_0^M p(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}.$$

3.2 Production

The homogeneous good is produced by a representative producer outside the industrial sector. Production requires labor, L_o as a unique input, homogeneous across suppliers. The representative firm in the outside sector generates A_o units of output per worker and face no frictions in the product and labor markets. Total production of the homogeneous good is then equal to

$$y_o = A_o L_o$$

Differentiated varieties are instead supplied by firms in the industrial sector, each of which produces a unique product $\omega \in [0, M]$. These firms are created through sunk investments and differ by their productivity levels z, which is drawn before entry from a distribution ψ_z, and kept until they exit. As in Melitz (2003) differences in productivity can equally well be thought of differences in product.

Firms also differ in the cost of setting up a formal business and by whether they are formally registered with the tax authority or not. To produce, unregistered firms only employ informal labor services, ℓ_i, in a linear production function:

$$y_i(z, \ell_i) = z\ell_i$$

Registered firms are allowed to combine informal and formal labor services, ℓ_i, and ℓ_f,

$$y_f(z, \ell_i, \ell_f) = z(\ell_i + \ell_f).$$

where ℓ_i, and ℓ_f are assumed to be perfectly substitute inputs.

3.3 Labor market

Every period jobless workers have the option of searching for a wage and salary job. If they choose not to search, they sustain themselves providing labor to the production of the homogeneous good in the outside sector. The labor market in the outside sector is frictionless; as a result, workers earn a wage which is equal to their marginal product, $w_o = A_o$.

If workers choose to search, they face search and matching frictions. Search is random. The total number of matches that are formed each period, $m(U, V)$, depends on the aggregate measure of workers searching for jobs, U, and the aggregate measures of vacancies posted, $V = V_{ii} + V_{fi} + V_{ff}$, where V_{ii}, V_{fi} and V_{ff} are measures of informal and formal vacancies posted by unregistered and registered firms, respectively, We assume the measure of matches are determined by the following function:

$$m(U, V) = \frac{UV}{(U^\eta + V^\eta)^{\frac{1}{\eta}}}$$

where $\eta > 0$ governs the elasticity with respect to the number of vacancies posted.

Let $\lambda(U, V) = \frac{m(U,V)}{UV}$ be a measure that summarizes the effect of market tightness in the labor market. The probability for a firm to meet a worker is proportional to the number of searchers and equal to

$$\tilde{\phi} = \lambda(U, V)U$$

while the probabilities for a worker to be hired in a formal or informal position depend on the relative measure of vacancy posted by registered and unregistered firms, and are equal respectively, to

$$\phi_{ii} = \phi\frac{V_{ii}}{V}, \quad \phi_{if} = \phi\frac{V_{if}}{V} \quad \text{and} \quad \phi_{ff} = \phi\frac{V_{ff}}{V}$$

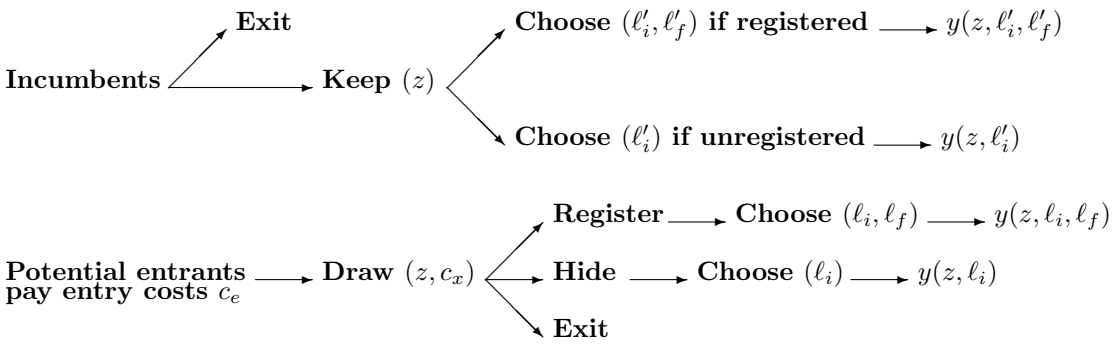
where $\phi = \lambda(U, V)V$.

Workers who get matched with a firm enter a bargaining stage to determine the wage rate, while workers who fail to match become unemployed. At the end of the matching process, the population of workers is split among those who are employed in the outside sector, L_o, those who are wage employed in formal and informal firms, L_e and those who are unemployed, L_u.

3.4 The problem of the industrial firms

Figure 3 shows the timing of firms' decisions in the model. At the beginning of each period, potential industrial firms pay an entry cost, observe their productivity level and their cost of operating formally, and decide whether to create a new business and whether to formally register. Once incumbent, firms choose their employment levels, produce, and pay wages. Each period, they face an exogenous probability of exiting the industry, and are subject to an expected cost of being audited, which depends on their registration status, and on how many informal workers are employed.

Figure 3: Firms' decisions



3.4.1 Revenues

Aggregating individual demand across consumers yields total demand for a variety ω,

$$q(\omega) = Dp(\omega)^{-\sigma} \quad \forall \omega \in [0, M]$$

where D is an aggregate demand shifter, common to all firms, and equal to

$$D = P^{\sigma-1}\alpha \int_0^1 I_i di$$

Notice that the population of worker-consumers is normalized to one. Given the aggregate demand, the total gross revenues of unregistered and registered firms can be written as:

$$R_i(z, \ell_i) = D^{\frac{1}{\sigma}} y_i(z, \ell_i)^{\frac{\sigma-1}{\sigma}} \quad \text{and} \quad R_f(z, \ell_i, \ell_f) = D^{\frac{1}{\sigma}} y_f(z, \ell_i, \ell_f)^{\frac{\sigma-1}{\sigma}}.$$

3.4.2 Employment decision

Unregistered firms choose how many informal workers to hire and post vacancies v_i at a cost c_v^i. The value of entering the industry for an unregistered firm with productivity z is then equal to

$$\begin{aligned} \mathcal{V}_i(z) = \max_{v_i} \quad & -c_v^i v_i + \frac{1-\delta_i}{1+r}\tilde{\mathcal{V}}_i(z,\ell_i) \\ \text{s.t.} \quad & \ell_i = \phi v_i \end{aligned}$$

δ_i is an exogenous exit probability for informal firms, while $\tilde{\mathcal{V}}_i(z,\ell_i)$ denotes the continuation value after entry, defined as follows:

$$\begin{aligned} \tilde{\mathcal{V}}_i(z,\ell_i) = \max_{v_i'} \quad & \pi_i(z,\ell_i) - c_v^i v_i' + \frac{1-\delta_i}{1+r}\tilde{\mathcal{V}}_i(z,\ell_i') \\ \text{s.t.} \quad & \ell_i' = (1-\delta_w)\ell_i + \phi v_i \end{aligned}$$

where $\pi_i(z,\ell_i)$ denote profits, equal to

$$\pi_i(z,\ell_i) = R(z,\ell_i) - w_i(z,\ell_i)\ell_i - \kappa_i(z)\ell_i$$

While unregistered firms do not incur any tax, they do face a per-worker expected cost of informality, $\kappa_i(z)$. This cost is reduced-form device that captures the probability of detection by the government and subsequent fines. It also includes a range of opportunity costs associated with informality such as limited access to formal credit markets, hampering the ability of firms to expand. Bigger firms are more visible to the government and therefore are inspected with higher probability, which entails higher costs in the form of monetary fines issued by the tax authority. Therefore, we pose the following specification:

$$\kappa_i(z) = \gamma_0 z^{\gamma_1} \quad \gamma_0 > 0, \gamma_1 > 0 \tag{1}$$

We depart from the functional form used in Dix-Carneiro et al. (2021), who specify the expected cost of informality as a fraction of gross revenues, and we assume the expected cost of informality per unit of worker to be increasing in the productivity of the firm. Under this formulation, everything else equal, more productive - hence larger - firms will find more costly to hire and extra informal worker and expand their size.

Registered firms choose how many formal and informal workers to hire and post vacancies for both types of workers, v_i and v_f, at a cost c_v^i and c_v^f, respectively. The value of entering the industry for a registered firm with productivity z is then equal to:

$$\begin{aligned} \mathcal{V}_f(z) = \max_{v_i,v_f} \quad & -\sum_{j\in\{i,f\}} c_v^j v_j + \frac{1-\delta_f}{1+r}\tilde{\mathcal{V}}_f(z,\ell_i,\ell_f) \\ \text{s.t.} \quad & \ell_j = \phi v_j \qquad \forall j \in \{i,f\} \end{aligned}$$

where δ_i is an exogenous exit probability for informal firms, while $\tilde{\mathcal{V}}_f(z, \ell_i, \ell_f)$ denotes the continuation value after entry, equal to

$$\begin{aligned}\tilde{\mathcal{V}}_f(z, \ell_i, \ell_f) = \max_{v'_i, v'_f} \quad & \pi_f(z, \ell_i, \ell_f) - \sum_{j \in \{i,f\}} c_v^j v'_j + \frac{1-\delta_f}{1+r} \tilde{\mathcal{V}}_f(z, \ell'_i, \ell'_f) \\ \text{s.t.} \quad & \ell'_j = (1-\delta_w)\ell_j + \phi v'_j \qquad \forall j \in \{i, f\}\end{aligned}$$

where $\pi(z, \ell_i, \ell_f)$ denotes profits of registered firms, equal to

$$\pi_f(z, \ell_i, \ell_f) = (1-\tau_y)R(z, \ell_i, \ell_f) - w_i(z, \ell_i, \ell_f)\ell_i - w_f(z, \ell_i, \ell_f)(1+\tau_w)\ell_f - \kappa_f(z, \ell_i, \ell_f)\ell_i$$

Registered firms are subject to taxes on corporate income, τ_y, and payroll taxes τ_w on their formal workers. Moreover, they face an expected cost of informality, $\kappa_f(z, \ell_i, \ell_f)$ defined as:

$$\kappa_f(z, \ell_i, \ell_f) = \gamma_2 z^{\gamma_3} \left(\frac{\ell_i}{\ell_i + \ell_f} \right)^{\gamma_4} \tag{2}$$

This function differ from the one considered by Ulyssea (2018), where all formal firms hire at most a fixed number $\bar{\ell}$ of informal workers, and the first $\bar{\ell}$ workers are always informal. Our function resembles instead the one used in Erosa et al. (2021), where the cost is increasing in the number of informal workers, and decreasing with the total number of workers. Under the current formulation, more productive firms and firms with a high share of informal employment will find more costly to hire an extra informal worker.

3.4.3 Entry and formalization decision

Every period, there is a large measure of potential employers draw their productivity, z, from distribution $\psi_z(z)$, and decide whether to start their business or not. After entry, employers draw an idiosyncratic cost, c_f, from a distribution ψ_c, and decide whether to pay the cost and operate as a formal business, or stay informal and forgo the cost. The value of operating, $\mathcal{V}(z)$ is therefore equal to:

$$\mathcal{V}(z) = \int_{c_f \in \mathcal{C}} \max\{\mathcal{V}_i(z), \mathcal{V}_f(z) - c_f\} \psi_c(c_f) dc_f$$

Let c_e denote a fixed cost of entry. In equilibrium, a free entry condition has to be satisfied, i.e.

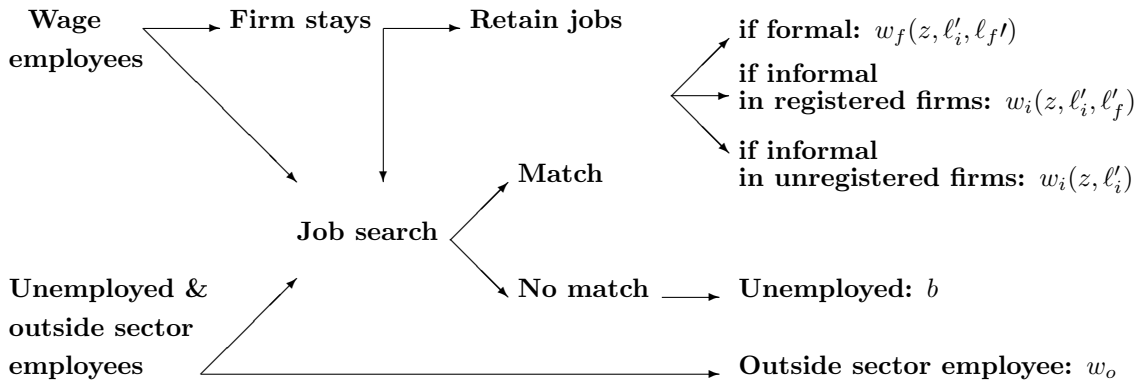
$$\mathcal{V}^e = \int_{z \in \mathcal{Z}} \max\{\mathcal{V}(z), 0\} \psi_z(z) dz \leq c_e$$

which holds with equality if the mass of entrants is strictly positive. A solution to this problem is a pair of thresholds, (z^*, c_f^*) which partitions the space of productivity and costs in three groups: firms who do not enter, firms entering without registering, firms entering and registering.

3.5 The problem of the workers

Figure 4 shows the timing of workers' decision in the model. Workers can be either employed in a wage and salary job, employed in the outside sector or unemployed. Only workers who are not already employed in a wage and salary job can look for it.

Figure 4: Workers' decisions



A worker who is not employed in a wage and salary job at the beginning of the period chooses whether to search for it or not, and solves the following problem

$$\mathcal{J}^n = \max\{\mathcal{J}^o, \mathcal{J}^s\}$$

where $\mathcal{J}^o$ is the value is being employed in the outside sector, equal to

$$\mathcal{J}^o = w_o + \frac{1}{1+r}\mathcal{J}^n,$$

$\mathcal{J}^s$ is the value of searching for a job, equal to

$$\mathcal{J}^s = (1-\tilde{\phi})\mathcal{J}^u + \tilde{\phi}\mathbf{E}\mathcal{J}^e,$$

and $\mathcal{J}^u$ is the value of being unemployed at the end of the period, equal to,

$$\mathcal{J}^u = b + \frac{1}{1+r}\mathcal{J}^n$$

Workers who choose the outside sector earns a wage w_o and have the option of searching again next period. Workers who choose to search fail to get matched to a firm with probability $1-\tilde{\phi}$ and get a flow value of unemployment b. The expected value of matching to a firm $\mathbf{E}\mathcal{J}^e$ reads as follows:

$$\mathbf{E}\mathcal{J}^e = \left[\frac{V_{ii}}{V}\int_z\int_{\ell_i}\mathcal{J}_i^e(z,\ell_i)\nu_{ii}(z,\ell_i)dzd\ell_i + \frac{V_{if}}{V}\int_z\int_{\ell_i}\int_{\ell_f}\mathcal{J}_i^e(z,\ell_i,\ell_f)\nu_{if}(z,\ell_i,\ell_f)dzd\ell_i d\ell_f \right.$$
$$\left. + \frac{V_{ff}}{V}\int_z\int_{\ell_i}\int_{\ell_f}\mathcal{J}_f^e(z,\ell_i,\ell_f)\nu_{ff}(z,\ell_i,\ell_f)dzd\ell_i d\ell_f\right]$$

where ν_{ii}, ν_{if} and ν_{ff} are distributions of informal vacancies in unregistered and registered firms, and formal vacancies, respectively, over firm productivity and number of employees.

Finally, it remains to specify the values of being employed, $\mathcal{J}_i^e(z,\ell_i)$, $\mathcal{J}_i^e(z,\ell_i,\ell_f)$ and $\mathcal{J}_f^e(z,\ell_i,\ell_f)$. These values are equal to:

$$\mathcal{J}_i^e(z,\ell_i) = w_i(z,\ell_i) + \frac{1}{1+r}\left[(\delta_w + (1-\delta_w)\delta_i)\mathcal{J}^n + (1-\delta_w)(1-\delta_i)\mathcal{J}_i^e(z,\ell_i)\right]$$

$$\mathcal{J}_i^e(z,\ell_i,\ell_f) = w_i(z,\ell_i,\ell_f) + \frac{1}{1+r}\left[(\delta_w + (1-\delta_w)\delta_f)\mathcal{J}^n + (1-\delta_w)(1-\delta_i)\mathcal{J}_i^e(z,\ell_i,\ell_f)\right]$$

and

$$\mathcal{J}_f^e(z,\ell_i,\ell_f) = w_f(z,\ell_i,\ell_f) + \frac{1}{1+r}\left[(\delta_w + (1-\delta_w)\delta_f)\mathcal{J}^n + (1-\delta_w)(1-\delta_f)\mathcal{J}_f^e(z,\ell_i,\ell_f)\right]$$

Employed workers are paid $w_i(z,\ell_i)$ if informal in unregistered firms, $w_i(z,\ell_i,\ell_f)$ if informal in registered firms and $w_f(z,\ell_i,\ell_f)$ if formal. They lose their job either because of an exogenous separation shock, δ_w, common across workers, or because of firm exit, which differs between unregistered and registered firms, δ_i and δ_f.

3.6 Wage bargaining

Search frictions generate a surplus between firms and each worker that is shared through a bargaining protocol. We assume that workers collectively bargain with their employer ex post, meaning after matching has taken place and the labor market has already closed. At the time of negotiation, vacancy posting costs are already sunk and workers who walk away from the bargaining table cannot be replaced in the current period. Similarly, if an agreement between the firm and the worker is not reached, the worker remains unemployed in the current period. However, neither party has incentive to break the match. Following Binmore et al. (1986), production delay constitutes the only credible threat in the negotiation.

Consider the bargaining problem between an unregistered firm and its employees. The surpluses accruing to the firm and to the collective of informal employees are given by, respectively:

$$\Pi_i^{\text{firm}}(z,\ell_i) = R_i(z,\ell_i) - w_i(z,\ell_i)\ell_i$$

$$\Pi_i^{\text{worker}}(z,\ell_i) = [w_i(z,\ell_i) - b]\ell_i$$

Failing to reach an agreement generates a loss for the employers equal to the per-period aggregate revenues net of the wage bills, and a loss for workers equal to their labor earnings net of the unemployment transfer. Let ζ_i be the bargaining power of informal workers. The outcome of the bargaining is given by a standard Nash splitting rule:

$$\zeta_i \Pi_i^{\text{firm}}(z,\ell_i) = (1-\zeta_i)\Pi_i^{\text{worker}}(z,\ell_i)$$

A solution to this problem is given by the following wage schedule:

$$w_i(z, \ell_i) = (1 - \zeta_i)b + \zeta_i \frac{R_i(z, \ell_i)}{\ell_i}$$

Informal workers get paid a ζ_i share of the average product and a share $1 - \zeta_i$ of their outside option, b.

Consider now the bargaining problem between a registered firm and its employees. The aggregate surplus is equal to

$$\Pi_f^{\text{firm}}(z, \ell_i, \ell_f) = (1 - \tau_y)R_f(z, \ell_i, \ell_f) - w_i(z, \ell_i, \ell_f)\ell_i - (1 + \tau_w)w_f(z, \ell_i, \ell_f)\ell_f$$

We assume formal and informal employees bargain separately with their employer over the average surplus they generate.[6] For instance, the surplus shared by registered firms and the collective of informal employees are given by

$$\Pi_f^{\text{firm}}(z, \ell_i, \ell_f) = \frac{\ell_i}{\ell_i + \ell_f}(1 - \tau_y)R_f(z, \ell_i, \ell_f) - w_i(z, \ell_i, \ell_f)\ell_i$$

$$\Pi_f^{\text{worker}}(z, \ell_i, \ell_f) = [w_i(z, \ell_i, \ell_f) - b]\, \ell_i$$

Using the same Nash splitting rules as above, we obtain the following wage function for informal workers in registered firms:

$$w_i(z, \ell_i, \ell_f) = (1 - \zeta_i)b + \zeta_i(1 - \tau_y)\frac{R_f(z, \ell_i, \ell_f)}{\ell_i + \ell_f}$$

Similarly, the surplus shared by registered firms and the collective of formal employees are equal to:

$$\Pi_f^{\text{firm}}(z, \ell_i, \ell_f) = \frac{\ell_f}{\ell_i + \ell_f}(1 - \tau_y)R_f(z, \ell_i, \ell_f) - (1 + \tau_w^f)w_f(z, \ell_i, \ell_f)\ell_f$$

$$\Pi_f^{\text{worker}}(z, \ell_i, \ell_f) = [w_f(z, \ell_i, \ell_f) - b]\, \ell_f$$

which give the following wage function for formal workers in registered firms:

$$(1 - \beta\tau_w^f)w_f(z, \ell_i, \ell_f) = (1 - \zeta_f)b + \zeta_f(1 - \tau_y)\frac{R_f(z, \ell_i, \ell_f)}{\ell_i + \ell_f}.$$

[6] An alternative would be to use the infra-marginal bargaining protocol extended to accommodate heterogeneous agents as in Cahuc et al. (2008). However, this protocol allows us to avoid the counterfactual prediction of a negative firm size-wage premium. See Elsby and Michaels (2013) for a discussion.

3.7 Stationary Equilibrium

A stationary equilibrium for this economy is a list of value functions and policy functions, values for the job finding probability and the job filling probability, measures of informal workers employed in unregistered firms, informal and formal workers in registered firms, unemployed workers, and workers employed in the outside sectors, wages, measure of firms, share of unregistered firms, distribution of firms across productivity values, and distribution of workers over employment status, firm productivity and size, such that:

- policy functions solve the problem of workers and firms, and value functions attain their maximum;
- workers non employed in a wage and salary job are indifferent between searching for a wage and salary job or not, i.e.

$$\mathcal{J}^n = \mathcal{J}^s = \mathcal{J}^o = \frac{1+r}{r} w_o \tag{3}$$

- the measure of entrants is such that the free entry condition holds with equality:

$$\mathcal{V}^e = \int_{z \in \mathcal{Z}} \max\{\mathcal{V}(z), 0\} \psi_z(z) dz = c_e \tag{4}$$

- wages are determined as the solution of the bargaining problems;
- the distributions of firms over productivity and size replicate themselves through entry and registration decisions and exit shocks;
- the distributions of workers over employment status and firm characteristics replicates themselves through hiring decisions and separation shocks;
- the labor market for wage and salary jobs clears;
- the product market for the outside good clears.

In Appendix C we report detailed equilibrium conditions and describe the numerical algorithm employed to find a solution to this model.

4 Bringing the model to the data

In this section we numerically quantify our model economy. In what follows, we first highlight a number of empirical facts on informal employment and, more generally, on labor market outcomes in Peru. We then describe our model estimator, discuss estimates, and how the model fits firm- and worker-level data.

4.1 Informality in Peru

Informal employment is a significant feature of in the Peruvian economy. We employ three datasets containing information on formal and informal firms and workers to describe it. An overview of these datasets and their main features is provided in Table 2.

Table 2: Summary of datasets

Datasets	Years	Source
National Household Survey (ENAHO)	2007-2014	Peruvian National Institute of Statistics (INEI)
Enterprise Survey (ES)	2006, 2010, 2017	World-Bank
Informal Enterprise Survey (IFS)	2010	World-Bank

Data on informal and formal workers is taken from the Peruvian National Household Survey (ENAHO). The ENAHO is a continuous cross-sectional survey representative of the Peruvian population. The survey is conducted in all regions in Peru and it is divided in sections. In what follows, we focus on the section on independent workers and the section that provides information on employment and income for all members of the household older than 14 years old. With these two sections, we have information on individuals demographic characteristics such as age, gender, education and region of residency. Moreover, the survey collects information on industry (4 digit ISIC), ownership, and number of workers of the employers whom individuals work for. Last but not least, surveyed jobs characteristics allow us to classify all employed workers in the sample in three categories: extensive-informal workers, intensive-informal workers and formal workers.

We consider extensive-informal workers those employees who declare to be employed by a firm that does not keep books in the online platform or software required by the Peruvian Tax Collection Agency (SUNAT) when filling tax declaration.

All individuals who declare to be employed in firms that we classify as "registered" are evaluated further. We classify as intensive-informal workers all individuals working in a registered family firm as a non-paid family worker. Regarding salaried workers, we follow Cisneros-Acevedo (2021) and rely on two different questions for the periods 2007-2011 and 2012-2017. Between 2007 and 2011, salaried workers in registered firms who declare SUNAT does not deduct their income in any way are classified as intensive-informal. Between 2012 and 2014, intensive-informal workers are those who declare that their employees, contrary to Peruvian laws, do not pay health insurance on their behalf.

We restrict our sample to women and men between 25 and 60 years old employed in non-military occupations, reporting positive hours worked, and who are not self-employed. Table B.1 in Appendix B describes the final sample of workers.

Data on formal firms is taken from the World-Bank Enterprise Survey (WB-ES). The

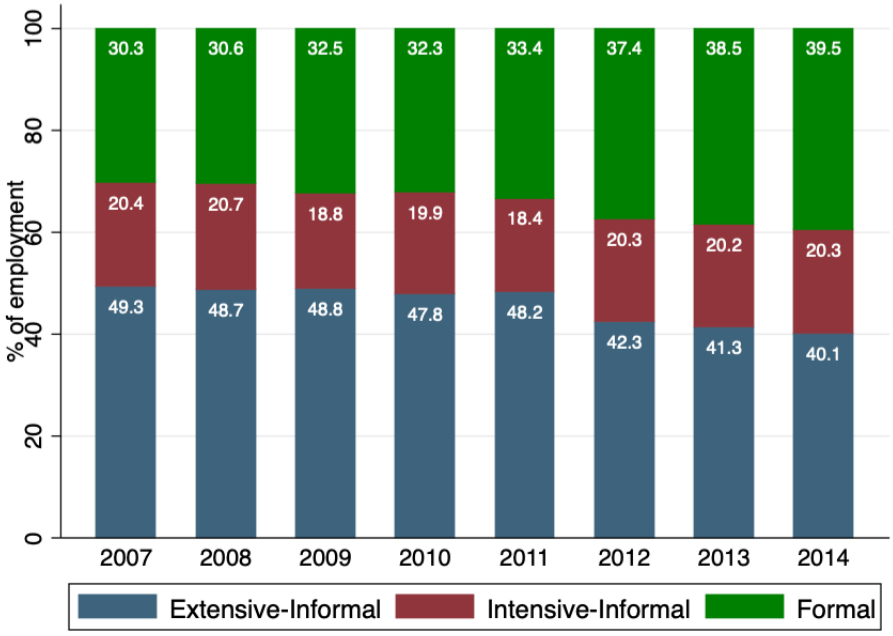
WB-ES is a cross-country survey of a representative sample of private sector firms. The survey covers several topics regarding the business environment and business performance, including general firm demographics (age, number of employees, ownership), sales and input purchases. Peruvian firms are surveyed in the years 2006, 2010 and 2017. Formal companies, defined as those registered with the SUNAT, with 5 or more employees are targeted.

Data on informal firms is taken from the World-Bank Informal Enterprise Survey (WB-IFS). The WB-IFS is a cross-country survey implemented in parallel to the WB-ES. It collects data on similar business topics although it targets informal business activities across countries. In its implementation, IFS equates informality with non-registration. For specific case of Peru, informal firms are defined as those not registered with the SUNAT, which makes it consistent to definition used in the ENAHO. Tables B.2 and B.3 in Appendix B describes the samples of formal and informal firms.

We now highlight four important facts on formal and informal workers and firms in Peru, which we will target in the estimation procedure.

More than 60% of wage and salary employment in Peru is informal. One third of it is made of informal workers employed in registered firms. Figure 5 reports the share of formal employment and informal workers along both margins on total employment from 2007 to 2014.

Figure 5: Employment composition

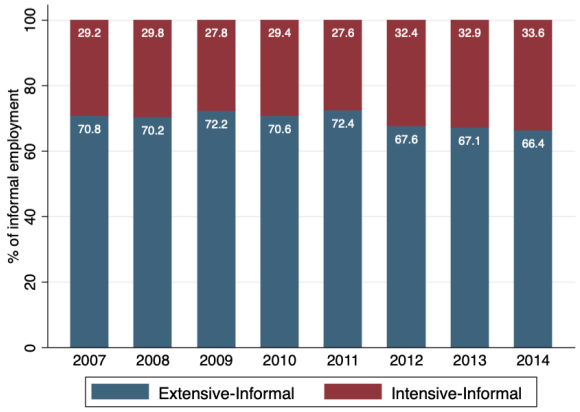


Notes: This figure reports the percentage of wage and salary employees who are informally employed in unregistered firms (blue bar), informally employed in registered firms (red bars) and formally employed (green bars). Source: ENAHO and authors' calculation.

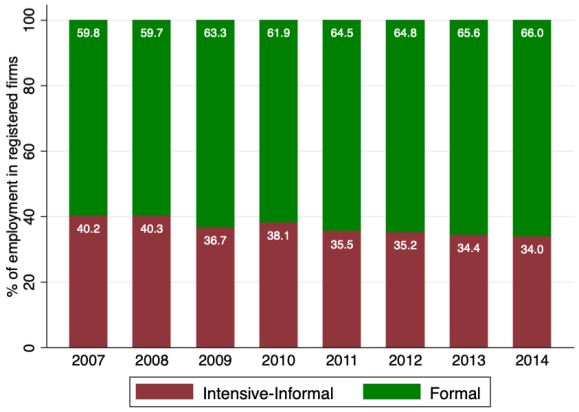
During this period, more than 40 per cent of workers were employed in non-registered firms, while around 20 per cent of workers were employed "off the books" by firms that were registered with the Tax Collection Agency. Combining intensive and extensive margins, between 60 and 70 per cent of workers in Peru were employed without any safety net such as retirement, paid holidays or sick leave.

Figure 6: Composition of formal and informal employment

A. Informal employment



B. Employment in registered firms



Notes: Panel A reports the percentage of informal wage and salary employees employed in registered firms (blue bars) and unregistered firms (red bars). Panel B reports the percentage of wage and salary employees employed in registered firms who are informal (red bars) and formal (green bars). Source: ENAHO and authors' calculation.

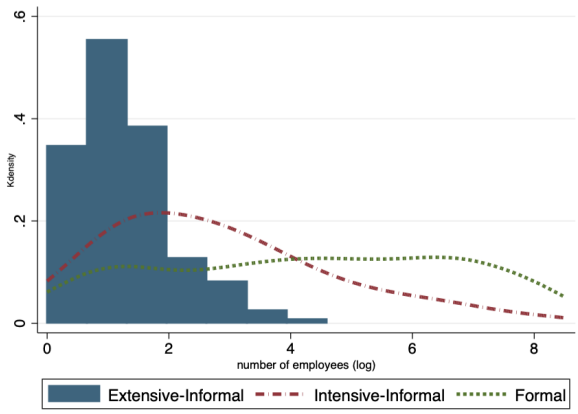
While formal employment has increased in this decade, the share of informal workers in the intensive margin has remained constant. Informal workers employed in formal firms accounts for 30% of overall informal employment in Peru (Figure 6A). Shares of 47% and 40% have been reported for Mexico (Samaniego de la Parra, 2017), and Brazil (Ulyssea, 2018), respectively. Moreover, informal workers account for 35% and 40% of the total employment in formal firms (Figure 6B). In the Appendix B we discuss the composition of informal employment for workers separately by their education level (college and non-college), gender (male and female) and sector (manufacturing and non-manufacturing). We show that incidence of informal employment along the intensive margin persists and it is large within each of these groups.

Informal workers are more likely to be employed in smaller firms. The share of informal workers in registered firms decline with size. Figure 7A reports the distribution of formal and informal workers across the size of the firm they work for. We measure size using the total number of employees. Informal workers are clustered in smaller firms. More than 70% of these workers are employed in unregistered firms with at most two employees. Informal workers employed in registered firms are instead more prevalent in medium-size companies.[7]

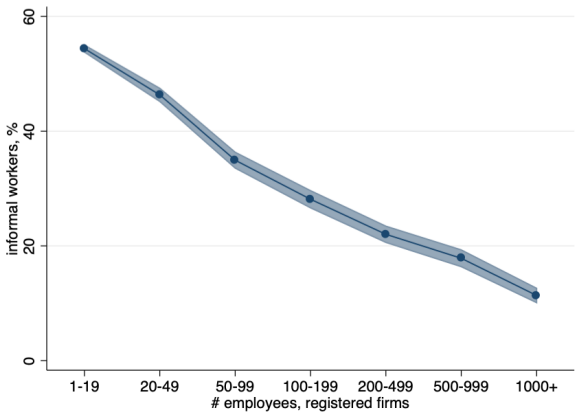
[7]In the Appendix we show that the distribution of firm size remains qualitatively the same when we

Figure 7: Firm size across formal and informal workers

A. Distribution of firm-size



B. Informal workers in registered firms



Notes: Panel A reports the distribution of firm-size for workers who are informally employed in unregistered firms (blue bars), informally employed in registered firms (red dashed line), and formally employed (green dotted line). Panel B reports the average share of informal workers employed in registered firms for different firm size. Source: ENAHO and authors' calculation.

While formal employment can be found in firms of almost any size, the largest share of formal workers has a job in big companies. Figure 7B reports the percent of informal workers in registered firms for different employers ranked by their firm size. Larger firms are composed by a significantly higher share of formal workers.

Formal firms are more productive than informal firms. A large literature has already documented that formal firms are on average more productive than informal ones. Our data confirm this evidence for the case of Peru.

Figure 8A reports the distribution of yearly log sales per employee, for formal (registered) and informal (unregistered) firms. Figure 8B reports the distribution of yearly payroll expenditure per employee incurred by either types of firms. Both variables are expressed in Peruvian local currency and expressed in 2010 price level.

On average, sales per employee of formal firms are 2.3 log-points higher compared to informal firms. Similarly, the labor payroll of formal firms is on average 0.85 log-points higher than that of informal firms.

Formal workers are paid on average higher wages than informal workers. We compare labor earnings across workers and we estimate the following equation:

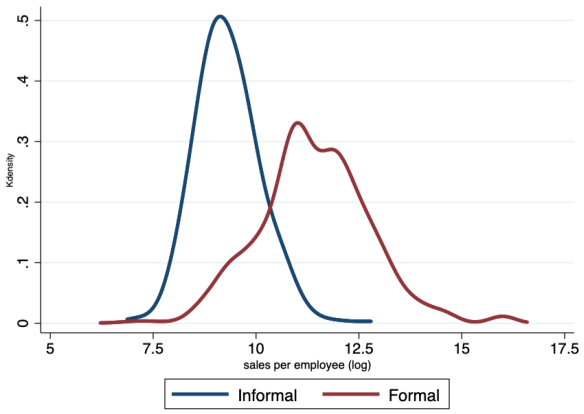
$$\log w_{it} = \alpha \mathbf{1}[\text{Formal}]_{it} + \beta \mathbf{1}[\text{Int.Mg.Inform}]_{it} + \mu_t + \gamma X_{it} + \epsilon_{it} \tag{5}$$

where w_{it} is the real monthly earnings of worker i at time t, $\mathbf{1}[\text{Formal}]_{it}$ and $\mathbf{1}[\text{Int.Mg.Inform}]_{it}$ are indicator variables denoting whether the worker is employed formally and informally

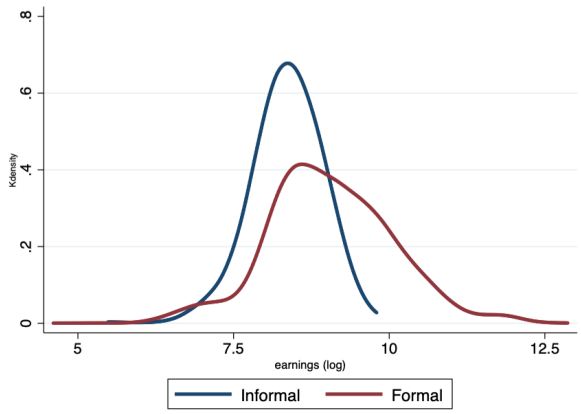
restrict the sample to workers in the manufacturing and non-manufacturing sectors, male and female, college and non-college educated workers. See Figure B.2.

Figure 8: Productivity of formal and informal firms

A. Sales per employee



B. Payroll per employee



Notes: Panel A reports the distribution of yearly sales per employee (in logs) for formal (red line) and informal firms (blue line). Panel B reports the distribution of yearly average payroll (in logs) for formal (red line) and informal firms (blue line). Data are expressed in 2010 Peruvian local currency (Nuevo sol). Source: WB-ES, WB-IFS and authors' calculation.

in a registered firm respectively, μ_t are time fixed effects and X_{it} are various worker- and job-level controls.

Table 3: Earnings gap of informal workers

	Log monthly earnings			
	(1)	(2)	(3)	(4)
$\mathbf{1}[\text{Formal}]_{it}$	0.984	1.129	0.583	0.828
	(0.004)	(0.006)	(0.006)	(0.009)
$\mathbf{1}[\text{Int.Mg.Inform}]_{it}$		0.316		0.335
		(0.007)		(0.009)
Observations	127,640	127,640	67,253	67,253
R-squared	0.3145	0.3297	0.5635	0.5743
Time F.E.	✓	✓	✓	✓
Controls			✓	✓

Notes: Earnings are expressed in 2010 Peruvian local currency (Nuevo sol). $\mathbf{1}[\text{Formal}]_{it}$ is a dummy variable for formal workers. $\mathbf{1}[\text{Int.Mg.Inform}]_{it}$ is a dummy variable for informal workers in registered firms. Controls include dummies for gender, education, age, ethnicity, civil status, geographical areas, ISIC-4 Rev.3 industry, firm size and firm ownership. Standard errors in parentheses. Source: ENAHO and authors' calculation

Table 3 reports the OLS estimates of equation (5). Since the omitted group is made of informal workers in unregistered firms, these estimates can be interpreted as the conditional wage premia for formal workers against the entire pool of informal workers (columns 1 and 3) and for formal and informal workers in registered firms against workers employed

in unregistered firms (columns 2 and 4).

First, the earnings premium from being a formal worker is large: formal workers earn on average twice as much as informal workers (column 1). Second, earnings of informal workers depend on whether workers are employed in a registered firm or not. Informal workers in registered firm they face a wage premium of about 0.3 log-points relative to those employed in unregistered firms, and a wage penalty of more than 1.13 log-points on average relative to formal workers (column 2). These results persist even after conditioning on worker- and job-level controls, including dummies for gender, education, age, ethnicity, civil status, geographical areas, ISIC-4 Rev.3 industry, firm size and firm ownership (columns 3 and 4), or if we focus on log hourly earnings (see Table B.4 in Appendix B).

4.2 Estimation

The estimation strategy proceed in two steps. We first select a subset of parameters without solving the model. Some of these parameters are not identified by the model and are taken from the literature, while some others are either calibrated to directly match specific targets, or set to their statutory values, like for the case of tax rates. Next, we estimate the remaining parameters of the model using the method of simulated moments, which allows us to combine information from the different data sources discussed in the previous section.

Table 4: Parameters Calibrated Without Solving the Model

Parameters	Description	Value	Source/Targets
r	Interest rate, %	1.08	Real lending rate= 13.80%
σ	Elasticity of substitution	6.40	Anderson and Van Wincoop (2001)
δ_f	Exit rate, % formal firm	5.68	Average age= 17.62 y.o.
δ_i	Exit rate, % informal firm	10.4	Average age= 9.61 y.o.
δ_s	Workers' separation rate, %	7.60	Monthly E-U rate= 7.6%
τ_y	Corporate tax rate, %	29.5	SUNAT (2016)
τ_w	Payroll tax rate, %	22.0	SUNAT (2016)
b	Unemployment benefits	0	OECD (2016)

Table 4 reports the parameters that are calibrated without solving the model. A model period is a month, hence the interest rate, r is set equal to 1.08% to match an annual real lending rate of 13.8% (WB-IMF). The elasticity of substitutions, σ is taken from Anderson and Van Wincoop (2001) and set equal to 6.4. The firm exit probabilities, δ_f and δ_i, are calibrated to match the average age of formal and informal firms in the economy, which are equal to 17.62 and 9.61 years, respectively (WB-ES). The separation rate, δ_s is chosen to have an monthly E-U rate of 7.6% (Reynaga and Ramírez-Rondán, 2021). Corporate income tax rate is set equal to 29.5% as reported by the Social Security (SUNAT) in

2016, while the payroll tax rate is set to 22% of the compensation paid to employees, inclusive of health contribution payments (9%) and pension funds contributions (13%). Finally, the unemployment benefit is set to 0 (OECD).

Firms productivity is drawn from a log-normal distributions, $z \sim \log\mathcal{N}(0, \sigma_z)$, with $\sigma_z > 0$, while the formality cost comes from a uniform distribution, given by $c_f \sim \mathcal{U}(0, \overline{c}_f)$, with $\overline{c}_f > 0$. These parametric assumptions leave us with 15 parameters to estimate, collected in the following vector

$$\vartheta := \{A_o, c_e, \overline{c_f}, c_v^i, c_v^f, \gamma_0, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \alpha, \varphi_z, \zeta_i, \zeta_f, \eta\}$$

These parameters are estimated using the method of simulated moments. The estimator ϑ is the minimizer of the following objective function:

$$\hat{\vartheta} = \arg\min_{\vartheta} d(\vartheta) W' d(\vartheta)$$

where $d(\vartheta)$ denotes the absolute distance between a vector of empirical targets, $\bar{g}$ and their model counterpart, $g(\vartheta)$, while W is a diagonal matrix with entries equal to the inverse squared of each empirical moments.[8]

The vector of empirical targets, $\bar{g}$ is constructed using firm- and worker-level statistics discussed in the previous section. Table 5 reports selected empirical moments and their model counterparts, grouped according to the type of information they convey. The first group includes average log-revenues, average and dispersion of log-size for formal and informal firms, plus different percentiles of the log-size distribution for formal firms. Firm revenues are sales expressed in 2010 Peruvian local currency. Firm size is measured using number of employees. The last three groups include a few labor market outcomes like the share wage employment and unemployed workers over working age population, the shares of informal wage employment along the extensive and the intensive margins, the overall job finding rate and the finding rate for informal jobs, and measure of wage inequality, like the wage gap between formal and informal workers in registered firms, and the gap between informal workers in registered versus unregistered firms.

Figure 9 completes the list of targeted moments. The upper panels report the shares of informal and formal firms across different firm-size bins, while the lower panels report the percentile of the size distribution for formal firms, and the share of informal workers employed in registered firms across different firm-size bins.

The model does not provide a direct map between parameters in ϑ and the list of moments in $\bar{g}$. Yet each moment plays a more important roles in identifying a particular parameter. Entry cost c_e and formality costs $\overline{c}_f$ are identified by average log-revenues of formal and informal firms, while the vacancy costs c_v^i and c_v^f are informed by average log-size, through their effects on vacancy posted.

[8] After experimenting with the efficient weighting matrix, we opted for this to ensure stability of our estimator while maintaining consistency and keeping it independent of units of measurement.

Figure 9: Selected targeted moments

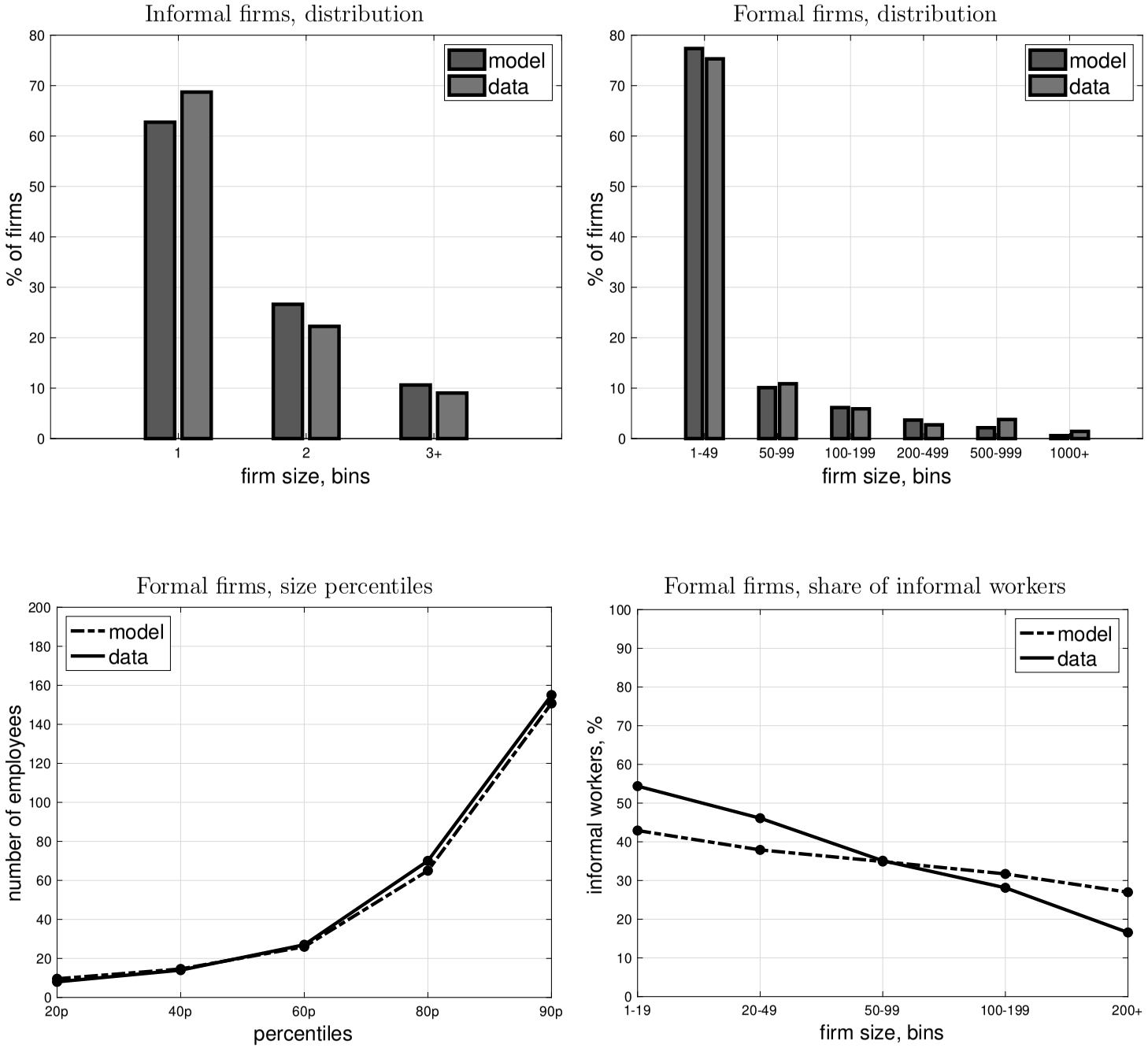


Table 5: Selected Targeted Moments

Moment	Data	Model	Moment	Data	Model
Firm-level moments			*Worker-level moments*		
Informal firms			*Labor market outcomes*		
Average log-revenues, $\mathbf{E}[\log R_i]$	7.061	8.146	(Wage employment+unemployed)/population	0.450	0.444
Average log-size, $\mathbf{E}[\log \ell_i]$	0.266	0.186	Wage employment rate, extensive-informal	0.436	0.395
Log-size dispersion, $\%\mathbf{std}[\log \ell_i]$	0.425	0.295	Wage employment rate, intensive-informal	0.221	0.189
Formal firms			*Wage gaps*		
Average log-revenues, $\mathbf{E}[\log R_f]$	11.97	11.76	Formal vs informal intensive	1.130	1.231
Average log-size, $\mathbf{E}[\log(\ell_i + \ell_f)]$	3.227	3.186	Informal intensive vs extensive	0.316	0.240
Log-size dispersion, $\mathbf{std}[\log(\ell_i + \ell_f)]$, %	1.303	1.187			
Log-size, 20th cutoff	2.079	2.257	*Aggregate outcomes*		
Log-size, 40th cutoff	2.639	2.678	Job finding rate (overall)	0.437	0.437
Log-size, 60th cutoff	3.296	3.256	Job finding rate (informal)	0.283	0.260
Log-size, 80th cutoff	4.249	4.173			

The parameters governing the expected costs of informality for informal and formal firms, λ_0, λ_1, λ_2, λ_3 and λ_4 are identified by the distribution of both types of firms, and by the share of informal workers in formal firms of different size.

We interpret employment in the outside sector as composed of those who are either non-employed or self-employed workers, hence the sum between wage-employment and unemployed shares of working age population pins the consumption share of the industrial good, α. The job finding rates for formal and informal jobs in the industry help identify the productivity of the outside sector, A_o and the elasticity of matching function η, respectively. Finally, the standard deviation of productivity, σ_z and the bargaining powers, ζ_i and ζ_f, map into dispersion of log size and wage gaps, respectively.

4.3 Estimates and model fit

Overall, the model is able replicate all the major features of the data. At the estimated values, the average percentage deviation between data- and model-based moments is 12%. In particular, the model generate the observed difference in firm size between registered and unregistered firms, it captures different percentiles of the firm-size distribution, and the share of firms across size groups. Informal firms are significantly smaller, the majority being composed by one or two workers. Formal firms are larger, and more than 10 percent of those have more than 100 employees.

The model also generates within-firm informality share that is declining in firm size as observed in the data. Finally, the model captures the differences in wages across formal and informal workers, and it captures the wage gap of informal workers employed in registered firms relative to formal workers. While part of this gap is generated exogenously by differences in bargaining power, the remaining is endogenously generated by workers allocation across firms. Since formal workers are more likely to be employed in large firms, and larger firms are those with higher productivity, they enjoy a firm productivity premium.

Table 6: Parameters Estimated with Simulated Method of Moments

Parameters	Description	Estimates	C.I. (± S.E.)	
c_e	Entry cost	3832.66	3780.66	3884.66
$\overline{c_f}$	Formal entry cost, upper bound	98010.8	13144.7	182876
c_v^i	Vacancy cost, informal workers.	10425.8	8491.78	12359.9
c_v^f	Vacancy cost, formal workers	18532.0	14305.8	22758.2
A_o	Productivity of the outside sector	1051.92	1040.40	1063.44
γ_0	Informality cost, informal firms	44.553	38.025	51.080
γ_1	Informality cost, informal firms	1.1603	1.1148	1.2059
γ_2	Informality cost, formal firms	96.482	77.698	115.27
γ_3	Informality cost, formal firms	1.6464	1.4793	1.8135
γ_4	Informality cost, formal firms	0.9486	0.9105	0.9866
α	Share of industrial goods	0.5516	0.3128	0.7904
φ_z	Productivity dispersion	0.9795	0.9549	1.0041
η	Elasticity of the matching function	2.1119	1.8970	2.3267
ζ_f	Bargaining power, formal workers	0.5065	0.3929	0.6201
ζ_i	Bargaining power, informal workers	0.2062	0.1603	0.2521

Table 6 reports our estimates and confidence intervals. Standard errors are constructed using the standard asymptotic variance expression. The parameters A_o, c_e, $\bar{c}_f$, c_v^i, c_v^f are measured in terms of our numeraire, the price of the outside good, which is expressed in 2010 Peruvian local currency.[9] In equilibrium, the earnings of the those employed in the outside sector, w_o equals the productivity of the outside good, A_o. We calculate this to be $S/1{,}051.92\times0.353$ = \$371.33 per month \$4,455.93 per year. This estimate implies the earnings of the those employed outside the industry are on average 89% of the earnings of those employed in the industry.

Expressed in dollars, the sunk cost of creating a new firm is estimated to be $S/3{,}832.66\times0.353$ = \$1352,92, while the costs of operating formally vary uniformly between 0 and $S/98{,}010.8\times0.353$ = \$34,597.81. The estimates imply an average entry costs for formal firms equal to \$18,652. Using Colombian micro-data on formal manufacturing firms, Coşar et al. (2016) estimate an entry cost of \$27,532, net of operating fixed cost.[10] Fajgelbaum (2016) uses official tax records of the manufacturing sector of Argentina and estimate the entry cost net of operating costs to be \$25,000.

Combining formal and informal firms, the average entry costs amounts to \$1,901, a value comparable to the estimates of Dix-Carneiro et al. (2021). Expressed in 2003 Brazilian Reals, they estimate an average sunk cost of entry for firms in the manufacturing

[9] In 2010, there were 2.83 Peruvian soles per dollar. We use a rate of $1/2.83 = 0.353$ to convert our estimates in 2010 USD.

[10] Within this model the operating fixed costs cannot be separately identified from the entry costs and set to zero by assumption. Hence, the estimate for the entry costs also embeds the discounted sum of future operating costs.

and the service sectors equal to R$5,332 and R$2,067 respectively, which corresponds to $1,818 and $705 in 2010 USD. Finally, the cost of posting formal and informal vacancies amount to $S/10{,}425.8 \times 0.353 = \$3{,}680.3$ and $S/18{,}532 \times 0.353 = \$6{,}541.8$, respectively.

For an unregistered firm with average productivity, the estimates γ_0 and γ_1 map into a monthly expected cost of informality equal to $S/184.87 \times 0.353 = \$65.26$ per employee. The values of γ_2, γ_3 and γ_4 generate a monthly expected cost of informality for a registered firm with average productivity and shares of informal workers of 10%, 50% and 90% equal to $S/81.80 \times 0.353 = \$28.87$, $S/376.51 \times 0.353 = \$132.90$ and $S/657.55 \times 0.353 = \$232.11$ per employee, respectively.

Finally, the matching function parameters, $\eta = 2.11$ is close to the value calibrated by Coşar (2013) using aggregate labor market statistics from Brazil (2.22) and to the value estimated by Coşar et al. (2016) using Colombia micro-data (1.84), whereas the bargaining power of formal and informal workers are estimated to be 0.5 and 0.2, respectively, suggesting formal workers are largely more protected than informal ones.

Table 7: Additional Statistics

Moment	Data	Model
Wage dispersion std[$\log w$]	0.875	0.517
Unemployment rate	0.037	0.042

4.4 Non-targeted statistics

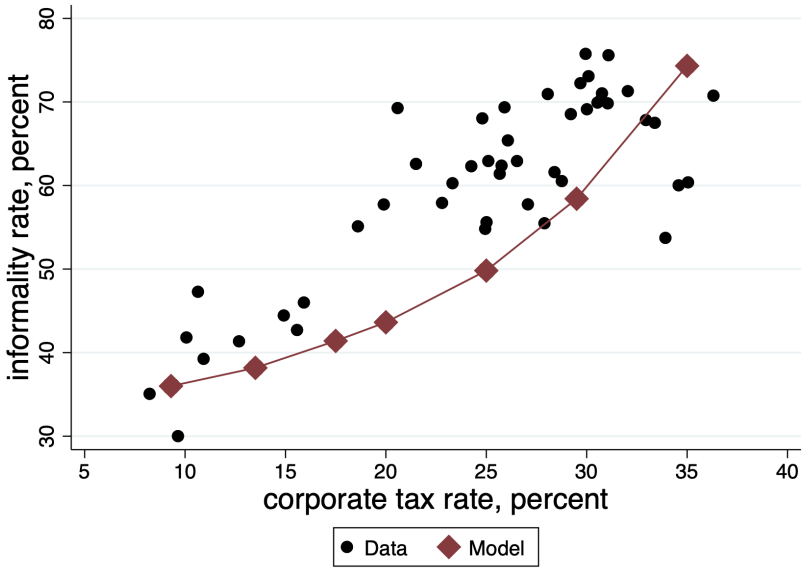
Table 7 compares data and model-based measures of wage inequality. Despite being non-targeted, the models accounts for more than 60% of the observed wage dispersion across workers. Even though workers are ex-ante homogeneous, the model generates wage dispersion between firms - because of differences productivity, and because of differences in bargaining power through the allocation of informal workers. At the estimated values, the model also generates an unemployment rate of 4.2%, a value which is very close to what is observed in the reference period, although not targeted.

5 Corporate income tax reforms

We are now ready to study the long-term effect of corporate tax reforms on labor market outcomes. To this purpose, we construct counterfactual economies that differ from the benchmark only in their corporate tax rate, keeping all the other parameters fixed at their benchmark values. Each of these economies also provides us with measures of informal employment, unemployment and GDP per worker that we can compare to the data.

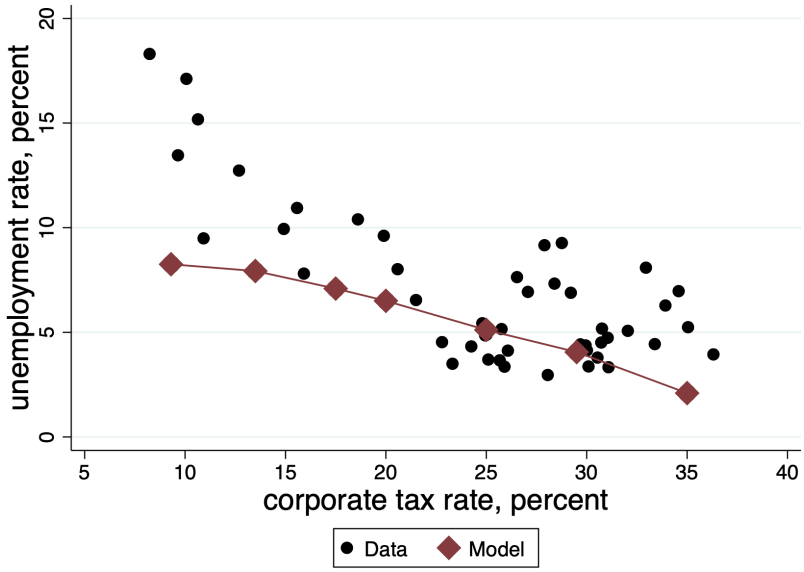
Figure 10 plots the informality rate measured in the data across countries (black dots) against the model counterparts (red diamonds). The model predictions are very much in

Figure 10: Taxes and Informality: Model vs. Data



Notes: This figure shows the rate of informal employment for countries with different corporate income tax rates. The black dots represent the data and red diamonds the model.

Figure 11: Taxes and Unemployment: Model vs. Data

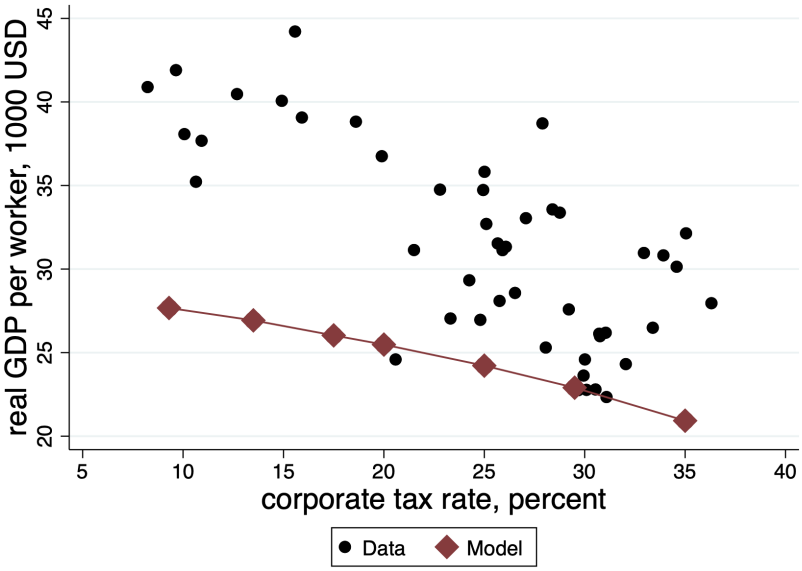


Notes: This figure shows the unemployment rate for countries with different corporate income tax rates. The black dots represent the data and red diamonds the model.

line with the data. As we move from low- to high-corporate tax environments, the share on informal workers employed increases from 36% to 74% of aggregate employment. This is achieved by a reduction in the share of registered firms and a shift in the composition of vacancies towards informal jobs (see next section for a discussion).

Figures 11 and 12 report data and model predictions for the unemployment rate and real GDP per worker. The model reproduces the pattern for the unemployment rate that

Figure 12: Taxes and GDP per worker: Model vs. Data



Notes: This figure shows the GDP per employed worker for countries with different corporate income tax rates. GDP per worker is measured in 2017 USD and expressed in 1000 USD. The black dots represent the data and red diamonds the model.

we observe across countries: as we lower the corporate income tax rate, the unemployment rate increases by 5 percentage points, from 3% to 8%. The magnitudes are similar to the ones in the data, although the model somewhat under-predicts the steepness of this relationship for very low tax rate countries.

Table 8: Slope Coefficients: Data vs. Model

Moment	Data	Model
Informality rate	1.245	1.437
	(0.480)	(0.244)
Unemployment rate	-0.378	-0.244
	(0.154)	(0.023)
Real GDP per worker	-0.564	-0.262
	(0.253)	(0.017)

Notes: The table reports estimated slope coefficients from regressions of the statistics in each row on corporate income tax rates. Data regressions include year fixed effects. The first column reports the slopes from our cross-country database. The second column reports the slopes from the quantitative model.

Finally, consistent with the data, a reduction in corporate income tax rates increases GDP per worker in the model. This is achieved through reallocation of jobs from to low-productivity informal firms to high-productivity formal firms (see next section). Everything else equal, a model reduction of 25 percentage points in corporate tax rates generates an increase in GDP per worker of about 7.000 USD.

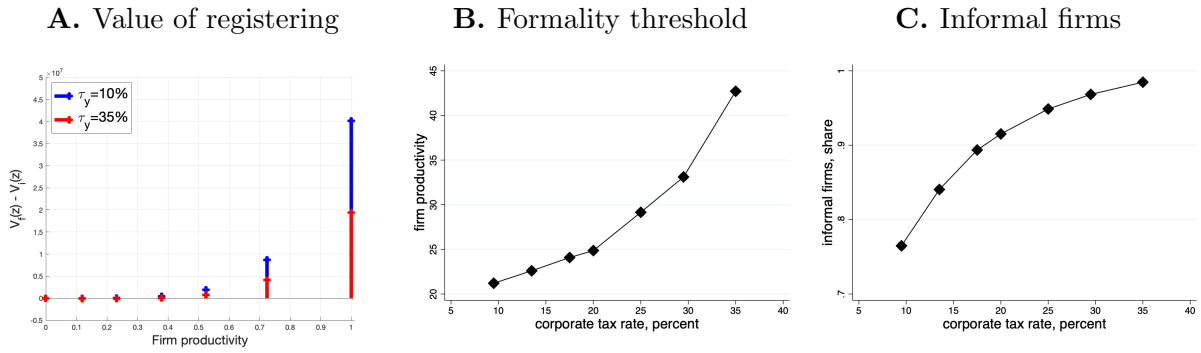
Table 8 reports the slope coefficients from regressing the informality rate, the unemployment rate and the GDP per worker on corporate tax rates and a constant, in our model and in the data.[11] For the informality rate, the model generates a slope of 1.437, very much aligned to the one in the data (1.245). Thus, the model accounts for almost all the empirical relationship between informality rate and corporate tax rate.

For the aggregate unemployment rate, the model yields a slope of -0.244 compared to -0.378 in the data. Hence the model accounts for around 60 percent of the empirical relationship between unemployment and corporate tax rates. Finally, the model generates a slope of -0.262 for GDP per worker, which is about 45 percent of what is predicted by the data (-0.564).

5.1 Mechanisms

How does the model generate these facts? Two major mechanisms are at play. The first mechanism operates through changes in the composition of firms in the industry, and reallocation of workers across jobs. To describe them, Figure 13.A reports the value of being a registered business relative to being informal in two selected counterfactual economies, one with a relatively low corporate income tax ($\tau_y = 10\%$, blue line) versus one with a relatively high tax ($\tau_y = 35\%$, red line).

Figure 13: Reallocations of firms



Notes: Panel A shows the relative value of forming a formal business ($V_f(z) - V_i(z)$) for firms with different productivity levels when the corporate income tax rate is equal to 10% (blue line) and 35% (red line). Panel B reports the average productivity thresholds for formal firms. Panel C reports the share of informal firms across various counterfactual economies with different corporate income tax rate.

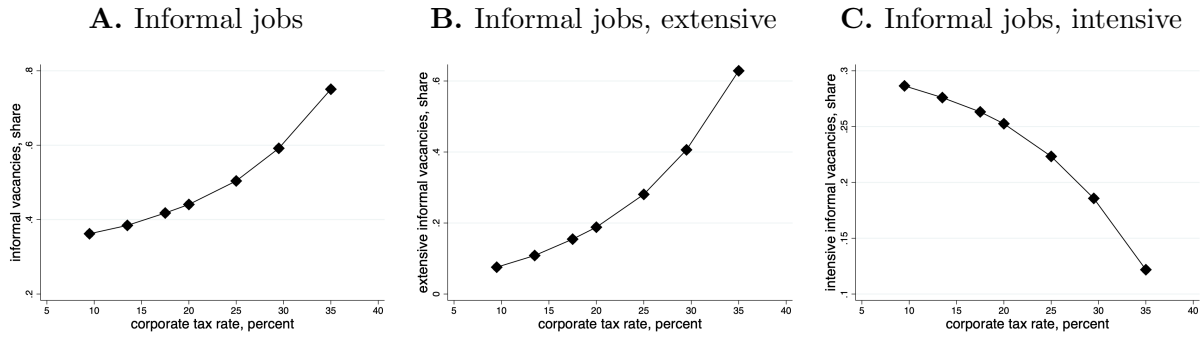
Corporate tax rate acts as a distortion on firms output which forces them to hide from regulation. This mechanism functions as in Ulyssea (2018). A reduction in corporate income taxes increases the value of operating as a registered business against the value of operating informally (Figure 13.A). This happens across the board of all firms and, although high-productivity formal businesses gains the most, a tax relief allows low-productivity to cover the cost of formalization and push them out of informality. As a

[11] The regressions using our dataset include year fixed effects. See Section 2 for more details.

result, as we move from a 35% to a 10% corporate income tax rate, the average productivity threshold above which firms become formal drops (Figure 13.B) and the share of informal firms declines by more than 20 percentage points (from 98% to 77%, Figure 13.C)

This force has two consequences. First of all, it reallocates workers from informal to formal jobs. Figure 14 scatter the overall share of informal vacancies posted in the economy (panel A), and the shares of informal vacancies along the extensive and the intensive margins (panels B and C, respectively), against corporate income tax rates, for all the simulated counterfactual economies. Moving from a 35% to a 10% corporate income tax rate reduces the share of informal vacancies by more than 50 percentage points (from 75% to 36%, Figure 14.A). This is achieved through formalization of jobs along the extensive margin, as opposed to intensive margin, whose share increases with lower corporate tax rates, but not enough the overturn the trend. A higher share of informal vacancies translates into lower informality rate, as documented in Section 2.

Figure 14: Reallocations of jobs



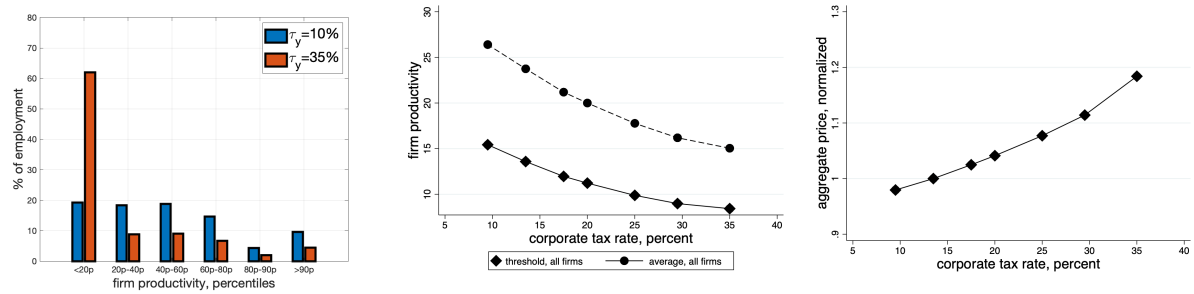
Notes: This figure reports the share of informal vacancies, on average (panel A), in the extensive margin (panel B), and in the intensive margin (panel C), across various counterfactual economies with different corporate income tax rate.

Second, the reallocation of firms increases also aggregate efficiency. Figure 15.A reports the employment distribution across firm productivity percentiles in two selected counterfactual economies, one with a relatively low corporate income tax ($\tau_y = 10\%$, blue bars) versus one with a relatively high tax ($\tau_y = 35\%$, red bars). Panels B and C of Figure 15 display the average productivity and productivity threshold across counterfactual economies (panel B), and the aggregate price index (panel C).

High-productivity firms take advantage of lower taxes to charge lower price and expand. This drives low-productivity firms out of the industry and let workers reallocate to high-productivity firms (Figure 15.A). The distribution of employment with respect productivity becomes more left-skewed, and the share of workers in the top 10 percentile of the productivity distribution doubles from 5% to 10%. Because of higher firm selection, the productivity threshold for incumbent firms rises, which makes aggregate productivity increase (Figure 15.B). Efficiency gains are expressed in the form of lower aggregate prices

Figure 15: Efficiency and prices

A. Employment distribution **B.** Firm productivity **C.** Aggregate price index



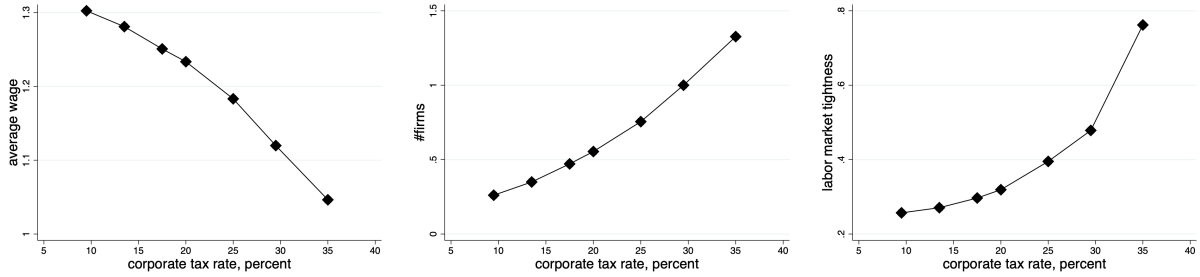
Notes: Panel A shows the distribution of employment across firms with different productivity levels when the corporate income tax rate is equal to 10% (blue bar) and 35% (red bar). Panel B and C report average productivity and productivity thresholds, and aggregate price index, for various counterfactual economies with different corporate income tax rate.

(Figure 15.C). This effect is similar to Melitz (2003), where lower trade costs increase foreign competition and push low-productivity firms out of the industry. In this economy competition increases among domestic firms as a result of corporate tax relief that favors only formal high-productivity businesses. Higher efficiency and lower prices translate into higher real GDP per worker employed, as documented in Section 2.

The second effect operates through general equilibrium forces in the labor market. To illustrate them, Figure 16 reports the average wage and salary earnings, expressed as a share of the earnings in the outside sector (panel A) the measure of firms operating in the economy (panel B) and the labor market tightness (panel C).

Figure 16: Concentration in the labor market

A. Wage ratio **B.** Measure of firms **C.** Market tightness



Notes: Panel A shows the average wage and salary earnings relative to earnings in the outside sector, for various counterfactual economies with different corporate income tax rate. Panel B and C report measure of firms, and labor market tightness, for various counterfactual economies with different corporate income tax rate.

An equilibrium in the labor market is governed by the no-arbitrage condition between the value of searching for a wage and salary jobs and the value of securing a job in the outside sector, as described by Equation (3). Lowering corporate taxes increases the average labor earnings for workers with a wage and salary job, relative to the value of earnings in the outside sector (Figure 16.A). Which makes the value of searching for a

wage and salary jobs more appealing. To re-establish an equilibrium, jobs concentrate on a smaller measure of firms (Figure 16.B). This lowers market tightness (Figure 16.C), hence making harder for workers to find wage and salary jobs. Lower job finding rates translates into higher unemployment, as documented in Section 2. This mechanism operates similarly in Feng et al. (2018) where changes in productivity of a modern sector increase the value of searching for jobs away from a traditional sector, leaving workers with a higher chance of becoming unemployed.

6 Firm-level policy interventions

We now use the estimated model to study the labor market outcomes of three alternative firm-policy interventions and compare them to corporate income tax reforms. The first policy is a change in expected cost of hiding for informal firms, captured by the parameter γ_0 in the cost equation (1). The second policy is a change in payroll taxes for formal workers in registered firms, τ_w.[12] The last policy is a change in expected fine from hiring workers off-the-book for formal firms, which we implement through the parameter γ_2 in the cost equation (2).[13]

Expected cost for informal firms. Table 9 reports firm-level and aggregate outcomes for different counterfactual values of κ_0. We include the share of informal firms, the share of informal vacancies and the average firm size within the first group. For the second group, we report the rates of informality, the measure of firms, the labor market tightness, unemployment rate, real average wage and the real GDP per employee.

The counterfactual outcomes following an increase in the expected informality cost for informal firms mirrors those obtained by reducing corporate income tax rates. Stricter regulations for informal firms are qualitatively analogous to lowering corporate income taxation for formal firms. As regulation becomes more costly for informal firms, the share of informal firms and the share of informal vacancies decline. As a result, the informality rate declines, although driven only by the extensive margin. Reallocation of firms triggers productivity improvements that lead to higher average wages, higher GDP per worker and higher unemployment rate. Quantitatively, doubling the expected cost of informality (from 33.41 to 66.83, columns 1 and 5 of Table 9) increases real wages in the industrial sector by 29.6% (from 1.0158 to 1.3123) and real GDP per worker by 11.5% (from 0.9308 to 1.0386).

[12] We report the counterfactual outcomes for payroll tax rate reforms in Appendix D, Table D.1.

[13] In addition to firm-level policies, we explore the effects of labor market policy intervention as unemployment benefits and minimum wages. See Tables E.1 and E.2 in Appendix E.

Table 9: Expected informality cost for informal firms

Informality cost, κ_0	33.41	41.66	44.55*	55.69	66.83
Firm-level outcomes					
Informal firms, share	0.9930	0.9771	0.9683	0.9322	0.8198
Informal vacancies, share	0.8698	0.6623	0.5918	0.4756	0.3863
Average firm size	2.7679	2.9469	3.2498	4.3123	8.1875
Aggregate Outcomes					
Informality rate	0.8652	0.6546	0.5842	0.4702	0.3835
- , extensive margin	0.7946	0.4916	0.3948	0.2252	0.1015
- , intensive margin	0.0706	0.1630	0.1894	0.2450	0.2820
Measure of firms	0.1563	0.1401	0.1243	0.0868	0.0436
Market tightness	1.1452	0.6012	0.4785	0.4145	0.3426
Unemployment rate	0.0108	0.0295	0.0406	0.0463	0.0586
Average wage	1.0158	1.0783	1.1198	1.2336	1.3123
Real GDP per worker	0.9308	0.9856	1	1.0279	1.0386

Notes: * refers to the baseline outcomes. Average wage is expressed as function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

Expected cost of hiring informal workers. Table E.1 reports the firm-level and aggregate counterfactual outcomes for different values of κ_2. The implications of this policy are qualitatively different from the previous ones.

Table 10: Expected cost of hiring informal workers for registered firms

Informality cost, κ_2	48.24	72.36	144.72	289.45	385.93
Firm-level outcomes					
Informal firms, share	0.9259	0.9587	0.9780	0.9863	0.9884
Informal vacancies, share	0.6264	0.5966	0.6175	0.6706	0.7032
Average firm size	4.2281	3.4523	2.8811	2.5350	2.4539
Aggregate Outcomes					
Informality rate	0.6222	0.5902	0.6092	0.6618	0.6943
- , extensive margin	0.2484	0.3425	0.4819	0.5958	0.6451
- , intensive margin	0.3739	0.2477	0.1273	0.0660	0.0493
Measure of firms	0.0989	0.1182	0.1389	0.1597	0.1676
Market tightness	0.6415	0.5206	0.4506	0.4985	0.5744
Unemployment rate	0.0271	0.0364	0.0434	0.0386	0.0318
Average real wage	1.0603	1.0973	1.1105	1.0950	1.0933
Real GDP per worker	1.0060	1.0029	0.9830	0.9625	0.9567

Notes: Average wage is expressed as function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

First, informality rate does not react monotonically to changes in the regulation faced by formal firms: for low values of κ_2, informal jobs in registered firms expands enough to overturn the pattern of formalization driven by changes in the extensive margin. Non-

monotonic job formalization makes the average wage in the industry follow an inverted-U shape as we lower the expected cost of informality. This dynamics is mirrored by the value of searching for a job in the industry, which affects the overall measure of firms, the labor market tightness and the unemployment rate: as formal jobs starts reducing, the industry becomes less concentrated, the labor market thickens again and the unemployment rate declines. The real GDP per worker, which reflects both gains in the industrial wages and workers composition across labor market states, monotonically increases as we lower the expected informality cost although it does it at a diminishing rates, as a result of higher informal jobs in registered business.[14]

6.1 Welfare gains

We finally assess the welfare properties and the efficiency-equity trade-off generated by each of these policies. To do so, we measure workers' aggregate welfare, $\mathcal{J}$, as a weighted average of the end-of-period value of being employed in the industry, $\mathbf{E}\mathcal{J}^e$, the end-of-period value of being employed in the outside sector, $\mathcal{J}_o$, and the end-of-period value of being unemployed, $\mathcal{J}_u$, i.e.

$$\mathcal{J} = L_o\mathcal{J}_o + L_u\mathcal{J}_u + L_e\mathbf{E}\mathcal{J}^e$$

where L_o, L_e, and L_e are the shares of workers employed in the outside sector, employed in the industry, and unemployed.

To study how inclusive are the welfare gains from firm-level policies, we define and study a policy possibility frontier.[15] The frontier confronts two feasible outcomes for firm-level policies. Specifically, Figure 17 scatters the unemployment rate against workers' average welfare for different levels of corporate income taxes and regulatory cost for informal firms (panel A), and for different payroll tax rates and regulatory cost of formal firms (panel B).[16] Each dot in the figure corresponds to a different counterfactual economy and welfare in the estimated economy is normalized to one.

We start by focusing on policies that directly tackle formalization along the extensive margins, i.e. corporate income tax and regulatory costs for informal firms (Figure 17.A). Both policies admit a monotonic trade-off between workers aggregate welfare and employment rate. However, the elasticity of workers' aggregate welfare to unemployment

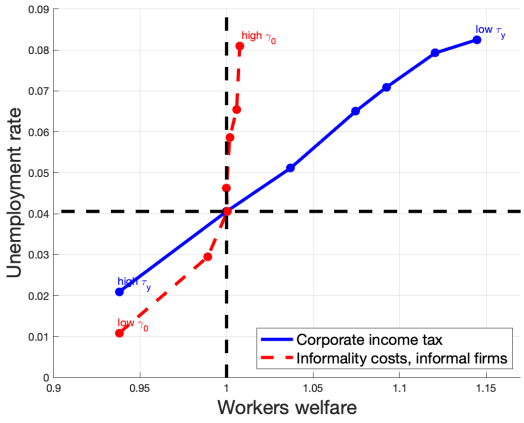
[14] By targeting informality along the intensive margin, changes in payroll taxes could lead to a qualitatively similar dynamics in the labor markets. On the other hand, in our quantitative exercise, a reform than eliminates payroll taxes is not strong enough to make job formalization non-monotonic. See Table D.1 in Appendix D.

[15] In alternative, we could pose the existence of a social welfare function that maps the vector of agents' individual welfare into a single real number (Antràs et al., 2017). However, this strategy allows us to be agnostic on the nature of the social welfare function. See Ruggieri (2019) for a recent example of such a strategy in the context of welfare gains from trade.

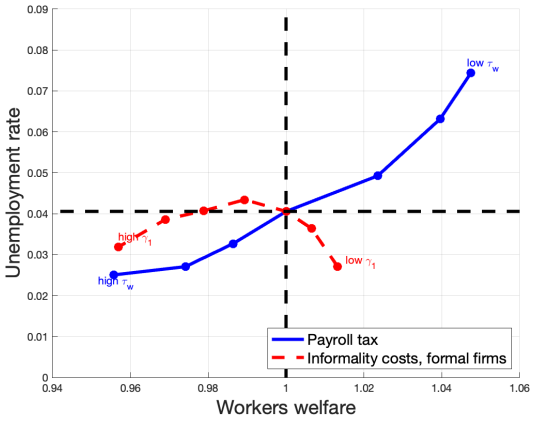
[16] As highlighted in Tables 9 and E.1, results would not change if we reported real GDP per worker instead of workers' aggregate welfare.

Figure 17: Efficiency-equity trade-off

A. Extensive-margin policies



B. Intensive-margin policies



rate is lower for the latter compared to former. Fixing a change of 1 percentage points in the unemployment rate, changes in corporate tax rates are associated with almost 4 times higher changes in welfare gains. This implies that neither policies are dominated by the other. While low values of corporate tax rates ensures higher welfare for the same level of unemployment rate (right-upper block of Figure 17.A), a low costs of informality for unregistered firms generates lower unemployment rate for the same level of welfare (lower-left block of Figure 17.A).

Figure 17.B replicates the same plot for the intensive margin policies, i.e. payroll tax rates (blue line) and informality costs for registered firms (red line). In this case, low regulatory fines for hiring workers off-the-book unambiguously dominate low payroll taxes. By ensuring relatively higher welfare and relatively lower unemployment rate than any other combination of policies, lower informality costs move the possibility frontier to lower-right block of the efficiency-equity space. This is achieved through enough allocative efficiency gains from firms formalization, and enough expansion of informal jobs along the intensive margins that makes unemployment reducing rather than expanding. Overall, economies with a larger share of registered firms and a larger share of informal workers hired by formal firms are welfare dominants in the efficiency-equity space.

7 Conclusion

In this paper we study the distributional consequence of firm-level policies in developing countries. We first document that countries with higher corporate income tax rates have higher informality rates, lower GDP per workers and lower unemployment rate. We then build a general equilibrium model of firm dynamics and show that a reduction in corporate tax rates generates these cross country patterns through two major mechanisms, i.e. 1) a reallocation of jobs from low- to high-productivity firms, and 2) a higher concentration

of employers in the labor market.

While our environment borrows several features from the literature in macro-development, such as corporate taxation, state-dependent distortions, entry barriers and informal employment, we push the state of the art and merge these elements within a search framework. This allows us to study the distributional implications of various firm- and labor-market policy interventions aiming at tackling informality along the extensive and the intensive margins.

Understanding how growth-oriented reforms can influence income distribution remains a first-order question for developing countries (Lagarde, 2017). This research contributes to this debate, being a first step to highlight the existing trade-offs between gains in welfare and equity across workers for alternative firm-level policies, and the importance of considering both margins of job informality for policy evaluation.

References

Abebe, G., Caria, A. S., Fafchamps, M., Falco, P., Franklin, S., and Quinn, S. (2021). Anonymity or distance? job search and labour market exclusion in a growing african city. *The Review of Economic Studies*, 88(3):1279–1310.

Amodio, F., Medina, P., and Morlacco, M. (2022). Labor market power, self-employment, and development. *IZA Discussion Paper*, 15477.

Anderson, J. E. and Van Wincoop, E. (2001). Borders, trade and welfare. Technical report, National Bureau of Economic Research.

Antràs, P., De Gortari, A., and Itskhoki, O. (2017). Globalization, inequality and welfare. *Journal of International Economics*, 108:387–412.

Baumgartner, E., Raphael Corbi, R. N., et al. (2022). Payroll tax, employment and labor market concentration. Technical report, University of São Paulo (FEA-USP).

Binmore, K., Rubinstein, A., and Wolinsky, A. (1986). The nash bargaining solution in economic modelling. *The RAND Journal of Economics*, pages 176–188.

Brooks, W. J., Kaboski, J. P., Li, Y. A., and Qian, W. (2021). Exploitation of labor? classical monopsony power and labor's share. *Journal of Development Economics*, 150:102627.

Cahuc, P., Marque, F., and Wasmer, E. (2008). A theory of wages and labor demand with intra-firm bargaining and matching frictions. *International Economic Review*, 49(3):943–972.

Card, D., Kluve, J., and Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The economic journal*, 120(548):F452–F477.

Cisneros-Acevedo, C. (2021). Unfolding trade effect in two margins of informality. the peruvian case. *The World Bank Economic Review*, 00(0):1 – 49.

Coşar, A. K., Guner, N., and Tybout, J. (2016). Firm dynamics, job turnover, and wage distributions in an open economy. *American Economic Review*, 106(3):625–63.

Coşar, K. (2013). Adjusting to trade liberalization: Reallocation and labor market policies. Technical report, University of Virginia.

De Paula, A. and Scheinkman, J. A. (2010). Value-added taxes, chain effects, and informality. *American Economic Journal: Macroeconomics*, 2(4):195–221.

De Soto, H. (1989). *The Other Path: The Invisible Revolution in the Third World*. Harper & Row New York.

Dincecco, M. and Katz, G. (2016). State capacity and long-run economic performance. *The Economic Journal*, 126(590):189–218.

Dix-Carneiro, R., Goldberg, P. K., Meghir, C., and Ulyssea, G. (2021). Trade and informality in the presence of labor market frictions and regulations. Technical report, National Bureau of Economic Research.

Dix-Carneiro, R. and Kovak, B. K. (2019). Margins of labor market adjustment to trade. *Journal of International Economics*, 117:125 – 142.

Donovan, K., Lu, W. J., and Schoellman, T. (2020). Labor market dynamics and development. Technical report, Federal Reserve Bank of Minneapolis.

Donovan, K. and Schoellman, T. (2021). The role of labour market frictions in structural transformation. Technical report, Yale University.

Dzansi, J., Jensen, A., Lagakos, D., and Telli, H. (2022). Technology and local state capacity: Evidence from ghana. Technical report, National Bureau of Economic Research.

Elsby, M. W. and Michaels, R. (2013). Marginal jobs, heterogeneous firms, and unemployment flows. *American Economic Journal: Macroeconomics*, 5(1):1–48.

Erosa, A., Fuster Pérez, L., and Martinez, T. R. (2021). Public financing with financial frictions and underground economy. Technical report, University of Madrid Carlos III.

Fajgelbaum, P. (2016). Labor market frictions, firm growth, and international trade. *NBER Working Paper*, 19492.

Feng, Y., Lagakos, D., and Rauch, J. E. (2018). Unemployment and development. Technical report, National bureau of economic research.

Gordon, R. and Li, W. (2009). Tax structures in developing countries: Many puzzles and a possible explanation. *Journal of public Economics*, 93(7-8):855–866.

La Porta, R. and Shleifer, A. (2014). Informality and development. *Journal of economic perspectives*, 28(3):109–26.

Lagakos, D. (2020). Urban-rural gaps in the developing world: Does internal migration offer opportunities? *Journal of Economic perspectives*, 34(3):174–92.

Lagarde, C. (2017). The fruits of growth: Economic reforms and lower inequality. *IMF Blog. January*, 26:2017.

Lee, Y. and Gordon, R. H. (2005). Tax structure and economic growth. *Journal of public economics*, 89(5-6):1027–1043.

Loayza, N. V. (2016). Informality in the process of development and growth. *The World Economy*, 39(12):1856–1916.

McKenzie, D. (2017). How effective are active labor market policies in developing countries? a critical review of recent evidence. *The World Bank Research Observer*, 32(2):127–154.

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

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

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

OECD (2006). *OECD Employment Outlook 2006: Boosting Jobs and Incomes*. Organisation for Economic Co-operation and Development, Paris, France.

OECD (2018). *Corporate Tax Statistics: Third Edition*. Organisation for Economic Co-operation and Development, Paris, France.

Ohnsorge, F. and Yu, S. (2022). *The long shadow of informality: Challenges and policies*. World Bank Publications.

Oviedo, A. M., Thomas, M. R., and Karakurum-Ozdemir, K. (2009). *Economic informality: Causes, costs, and policies a literature survey*. World Bank Publications.

Perry, G. (2007). *Informality: Exit and exclusion*. World Bank Publications.

Ponczek, V. and Ulyssea, G. (2022). Enforcement of labour regulation and the labour market effects of trade: Evidence from brazil. *The Economic Journal*, 132(641):361–390.

Poschke, M. (2019). Wage employment, unemployment and self-employment across countries. *IZA Discussion Paper*, 12367.

Reynaga, N. C. and Ramírez-Rondán, N. R. (2021). Job finding and separation rates in an economy with high labor informality. In *Workplace Productivity and Management Practices*. Emerald Publishing Limited.

Rud, J. P. and Trapeznikova, I. (2021). Job creation and wages in least developed countries: Evidence from sub-saharan africa. *The Economic Journal*, 131(635):1331–1364.

Ruggieri, A. (2019). Trade and labor market institutions: A tale of two liberalizations. *GEP working paper*, (15).

Samaniego de la Parra, B. (2017). Formal firms, informal workers, and household labor supply in mexico. Technical report, University of Chicago.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from brazil. *American Economic Review*, 108(8):2015–47.

Waseem, M. (2018). Taxes, informality and income shifting: Evidence from a recent pakistani tax reform. *Journal of public economics*, 157:41–77.

Appendices

A Data

Table A.1: Cross-country data

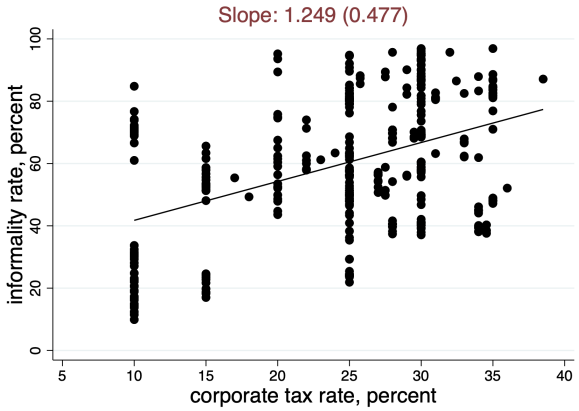
Country	Years	Country	Years
Angola	2011	Albania	2012-2019
Argentina	2010-2020	Armenia	2011-2021
Benin	2011, 2018	Burkina Faso	2018
Bangladesh	2010, 2013, 2017	Bosnia and Herzegovina	2010-2021
Bolivia	2011-2019	Brazil	2011-2021
Barbados	2016	Botswana	2019-2020
Chile	2018-2021	Cameroon	2014
Colombia	2010-2019, 2021	Costa Rica	2010-2021
Djibouti	2017	Dominican Republic	2010-2020
Ecuador	2010-2019, 2021	Ethiopia	2021
Fiji	2016	Georgia	2019-2020
Ghana	2013, 2015	Guinea-Bissau	2018
Guatemala	2010-2019	Guyana	2018-2019
Honduras	2019-2017	Indonesia	2016-2019
India	2010, 2012, 2018-2020	Jamaica	2016-2020
Jordan	2017-2020	Kenya	2019
Cambodia	2012, 2019	Lebanon	2019
Liberia	2017	Sri Lanka	2010-2019
Lesotho	2019	Madagascar	2015
Maldives	2016, 2019	Mexico	2010-2021
North Macedonia	2010-2021	Mali	2013-2018, 2020
Myanmar	2015, 2017-2020	Mongolia	2010-2020
Mozambique	2015	Mauritania	2012, 2017
Mauritius	2012-2019	Malawi	2013
Niger	2011, 2017	Nicaragua	2012
Nepal	2017	Pakistan	2010-2011, 2013-2015
Panama	2010-2021		2018-2019, 2021
Paraguay	2010-2019, 2021	Peru	2010-2021
Sudan	2011	Rwanda	2017-2020
Sierra Leone	2018	Senegal	2015-2019
Serbia	2010-2021	El Salvador	2014-2020
Eswatini	2016	Suriname	2016
Togo	2017	Chad	2018
Timor-Leste	2013	Thailand	2014-2018
Tunisia	2015	Tonga	2018
Uruguay	2010-2020	Uganda	2012, 2017
South Africa	2010-2021	Samoa	2012, 2017
Zimbabwe	2011, 2014, 2019	Zambia	2017-2020

Notes: This table reports countries and years covered by our cross-country dataset.

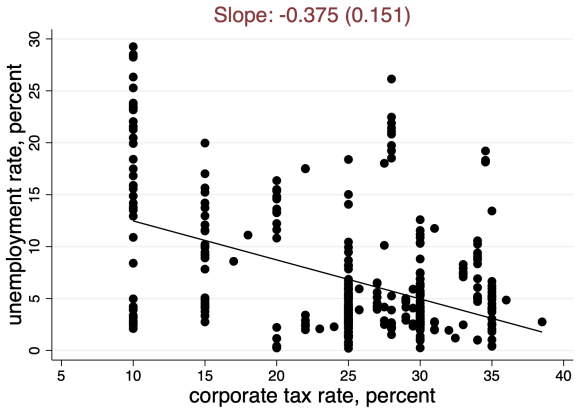
In Figure A.1 we report raw data for informality rate (panel A), unemployment rate (panel B) and real GDP per worker (panel C) across countries. Panel D reports the share of out-of-labor force and self-employed workers out of total working age population. In red we report the estimated slope of each relationship and robust standard errors, clustered at country level.

Figure A.1: Labor market outcomes and corporate income taxes, raw data

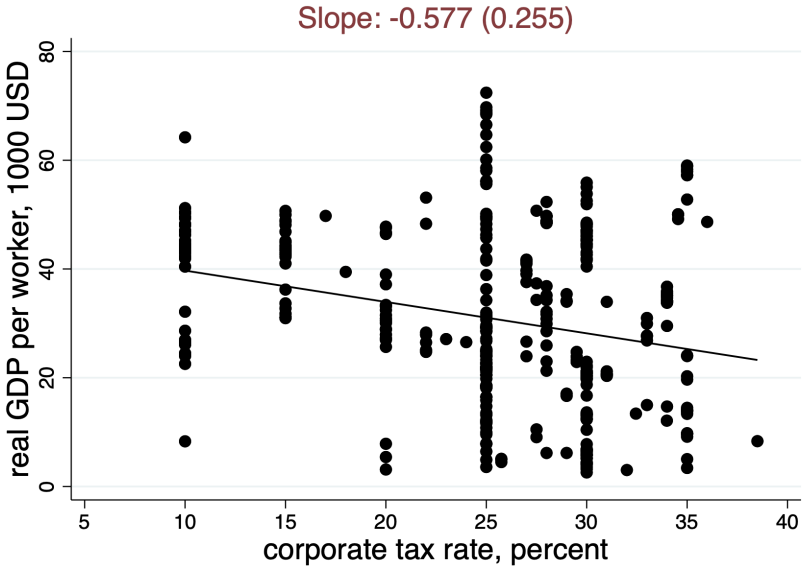
A. Informality rate



B. Unemployment rate



C. GDP per worker



Notes: Informal employment is expressed as percent of total employment and comprises persons who in their main or secondary jobs were own-account workers, contributing family workers, employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises, or as paid domestic workers by households. Informal jobs of employees are defined as those lacking of coverage by social security system, entitlement to paid annual or sick leave, or written employment contract. Unemployment rate is reported in percent of the labor force. GDP per worker is measured in 2017 USD and expressed in 1000 USD. Corporate tax rates refer to the standard top statutory corporate income tax rates levied on domestic businesses. Source: ILOSTAT, Tax Foundation and authors' calculation.

B Supplementary Evidence

Tables B.1, B.2 and B.3 report summary statistics for workers, formal firms and informal firms in Peru, respectively.

Table B.1: Summary statistics - Workers in Peru

	N.Obs.	Avg.	St.Dev.
Age	123554	40.058	9.819
Female	123554	0.456	0.498
Household Head	123554	0.423	0.494
College	123554	0.198	0.399
Urban	123554	0.699	0.459
Manufacturing	123554	0.088	0.283
Quechua ethnicity	123554	0.151	0.358
Monthly earnings	123554	405.572	595.235

Notes: This table reports selected statistics for workers in Peru. Earnings are reported in 2010 Peruvian local currency (Nuevo sol). Source: ENAHO, 2007-2014

Table B.2: Summary statistics - Formal firms in Peru

	N.Obs.	Avg.	St.Dev.
# employees	2583	163.1	542.2
Age	2628	21.99	17.51
Annual sales	2365	4.46e+07	1.92e+08
Annual average payroll	2042	21775.3	31334.6

Notes: This table reports selected statistics for formal firms in Peru. Sales and average payroll are reported in 2010 Peruvian local currency (Nuevo sol). Source: WB-ES, 2006, 2010, 2017

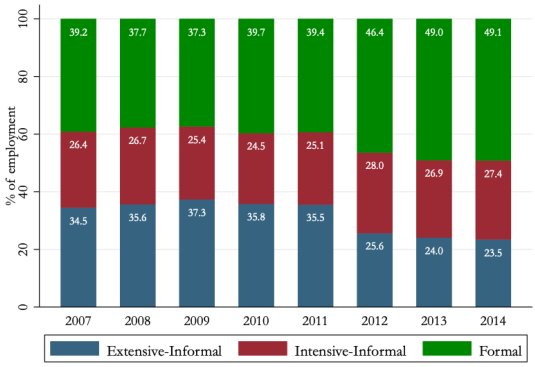
Table B.3: Summary statistics - Informal firms in Peru

	N.Obs.	Avg.	St.Dev.
# employees	454	1.456	0.867
Age	453	9.614	9.780
Annual sales	454	22393.3	31515.5
Annual average payroll	453	4892.98	2571.19

Notes: This table reports selected statistics for informal firms in Peru. Sales and average payroll are reported in 2010 Peruvian local currency (Nuevo sol). Source: WB-IFS, 2010

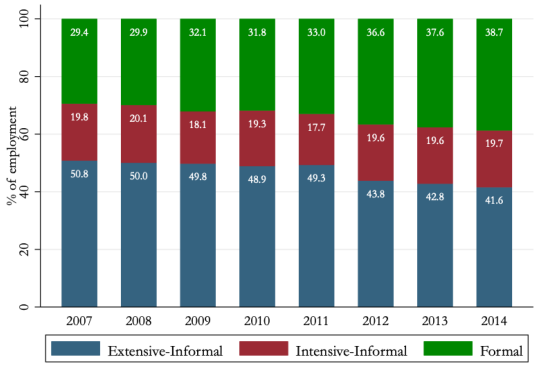
Figure B.1: Employment composition

A. manufacturing



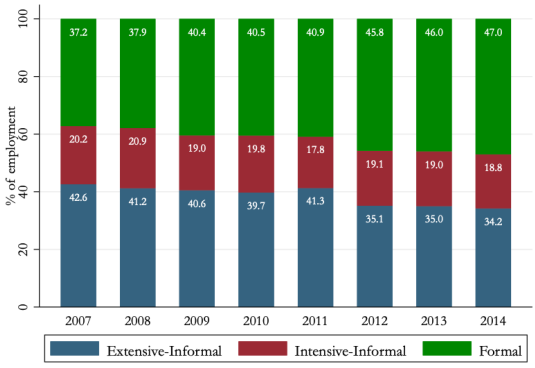
Source: Peruvian National Household Survey (ENAHO) 2007-2014.

B. non-manufacturing



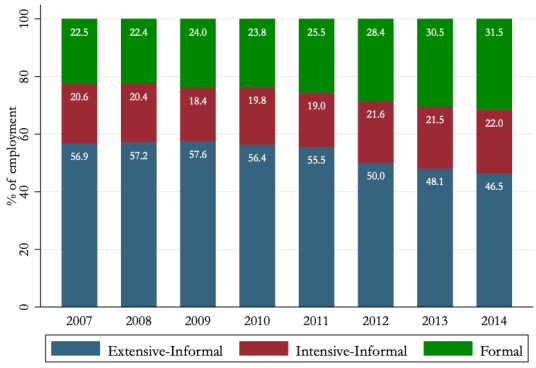
Source: Peruvian National Household Survey (ENAHO) 2007-2014.

C. male



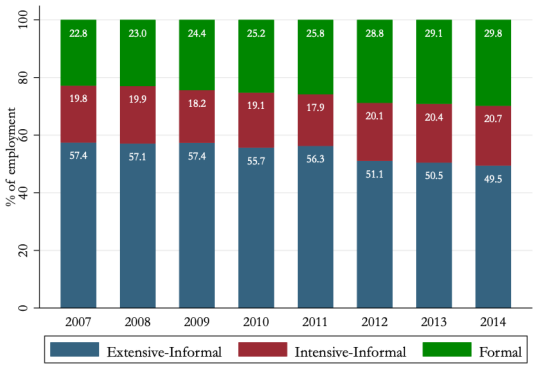
Source: Peruvian National Household Survey (ENAHO) 2007-2014.

D. female



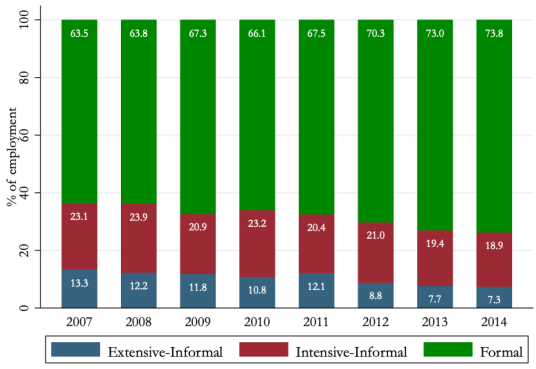
Source: Peruvian National Household Survey (ENAHO) 2007-2014.

E. non-college



Source: Peruvian National Household Survey (ENAHO) 2007-2014.

F. college

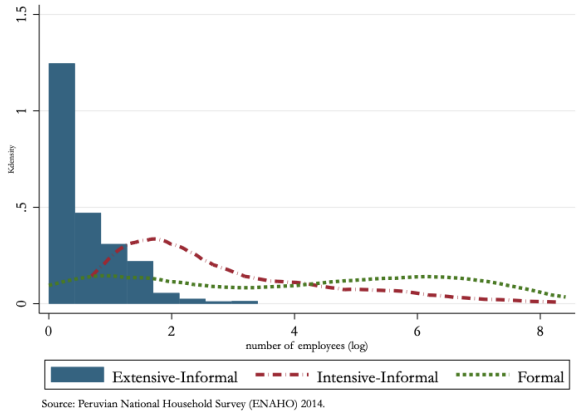


Source: Peruvian National Household Survey (ENAHO) 2007-2014.

Notes: This figure reports the shares of formal and informal workers separately for manufacturing and non-manufacturing workers, males and females, college and non-college workers. Source: ENAHO and authors' calculation

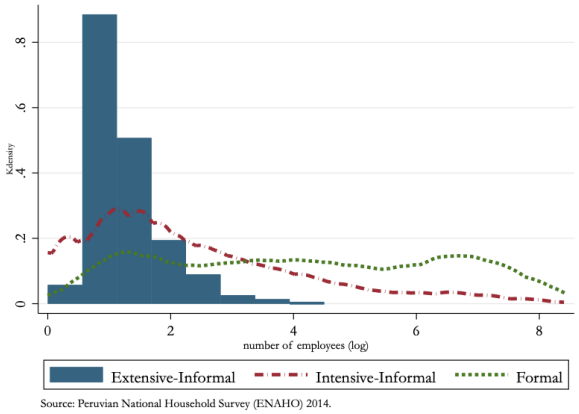
Figure B.2: Firm size distribution across workers

A. manufacturing



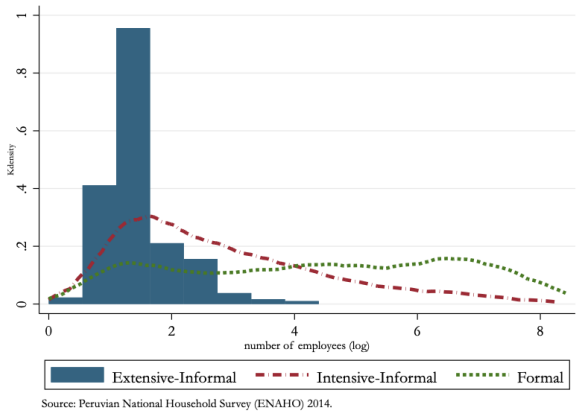
Source: Peruvian National Household Survey (ENAHO) 2014.

B. non-manufacturing



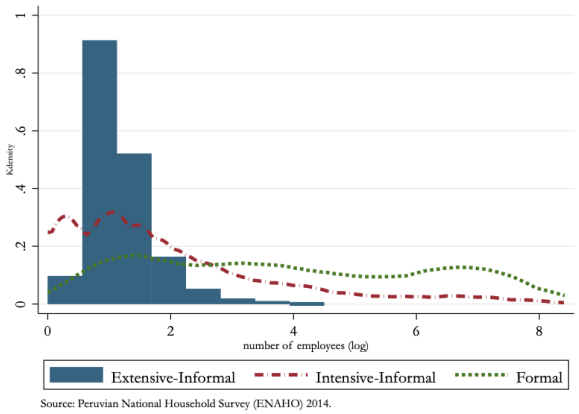
Source: Peruvian National Household Survey (ENAHO) 2014.

C. male



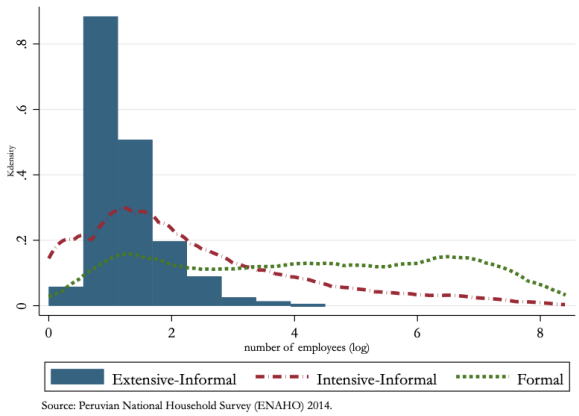
Source: Peruvian National Household Survey (ENAHO) 2014.

D. female



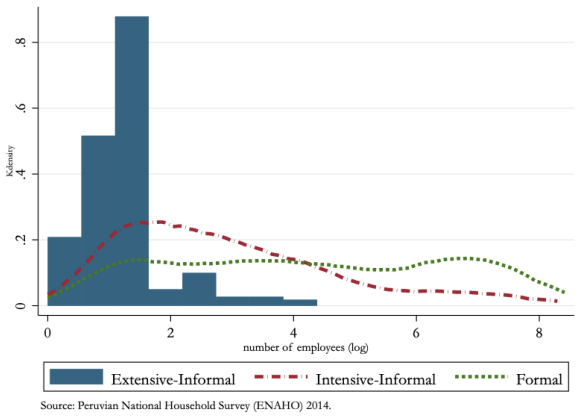
Source: Peruvian National Household Survey (ENAHO) 2014.

E. non-college



Source: Peruvian National Household Survey (ENAHO) 2014.

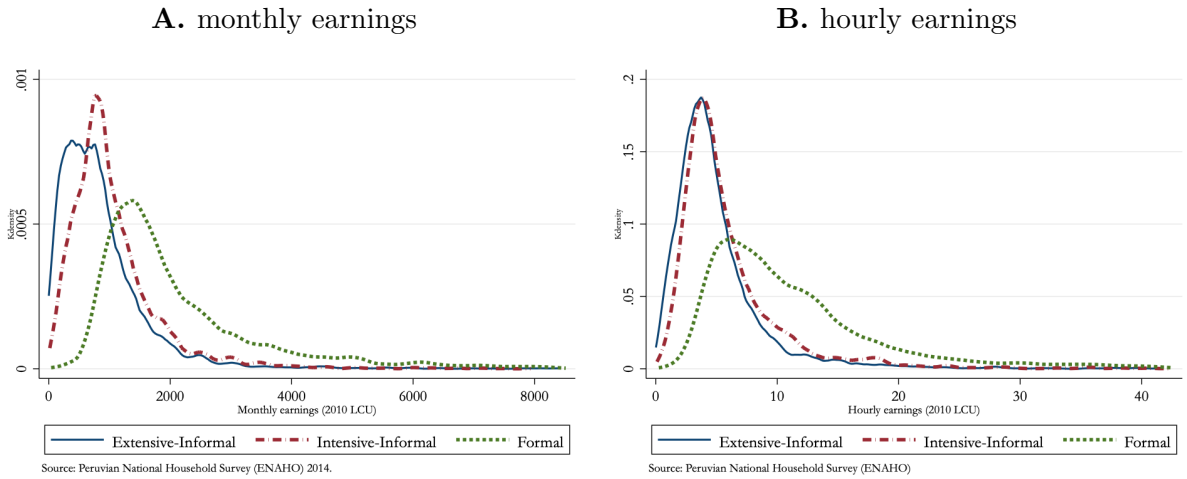
F. college



Source: Peruvian National Household Survey (ENAHO) 2014.

Notes: This figure reports the distribution of formal and informal workers over the size of their employers, separately for manufacturing and non-manufacturing workers, males and females, college and non-college workers. Source: ENAHO and authors' calculation

Figure B.3: Earnings distribution



Notes: This figure reports the distribution over monthly and hourly earnings for formal and informal workers. Source: ENAHO and authors' calculation

Table B.4: Earnings gap of informal workers

	Log hourly earnings			
	(1)	(2)	(3)	(4)
$\mathbf{1}[\text{Formal}]_{it}$	0.838	0.925	0.416	0.522
	(0.005)	(0.006)	(0.006)	(0.009)
$\mathbf{1}[\text{Int.Mg.Inform}]_{it}$		0.186		0.145
		(0.006)		(0.009)
Observations	102,355	102,355	54,254	54,254
R-squared	0.3012	0.3978	0.5402	0.5429
Time F.E.	✓	✓	✓	✓
Controls			✓	✓

Notes: Hourly earnings refer to monthly earnings divided by the number of hours worked in a week times 4.2, and are expressed in 2010 Peruvian local currency (Nuevo sol). $\mathbf{1}[\text{Formal}]_{it}$ is a dummy variable for formal workers. $\mathbf{1}[\text{Int.Mg.Inform}]_{it}$ is a dummy variable for informal workers in registered firms. Controls include dummies for gender, education, age, ethnicity, civil status, geographical areas, ISIC-4 Rev.3 industry, firm size and firm ownership. Standard errors in parentheses. Source: ENAHO and authors' calculation

C Model

C.1 Stationary equilibrium

A stationary equilibrium for this economy is a list of value functions and policy functions for employment decisions $L_i(z,\ell_i)$, $L_i(z,\ell_i,\ell_f)$ and $L_f(z,\ell_i,\ell_f)$, entry decision $\mathbf{1}^e(z)$, registration decision $\mathbf{1}^f(z,c_f)$, values for the job finding probability, ϕ and the job filling probability, $\tilde{\phi}$, measures of informal workers employed in unregistered firms, informal and formal workers in registered firms, unemployed workers, and workers employed in the outside sectors, wages, measure of firms M, share of unregistered firms s_i, distribution of firms across productivity values and size, $\psi_i(z,\ell_i)$ and $\psi_f(z,\ell_i,\ell_f)$, and distribution of workers over employment status, firm productivity and size, such that the following conditions hold:

1. **Aggregate consistency.** In equilibrium the distribution of formal and informal firms, $\psi_i(z,\ell_i)$ and $\psi_f(z,\ell_i,\ell_f)$ reproduce themselves through entry and registration decisions and exit shocks. Since all entering firms start the interim period a productivity draw from $\psi_z(z)$ and with a registration cost drawn from $\psi_c(c_f)$, we can measures formal and informal firms in their respective states as

$$\psi_i(z,\ell_i) = \delta_i \int_{c_f} [\mathbf{1}^e(z)][1-\mathbf{1}^f(z,c_f)]\psi_z(z)\psi_c(c_f)dc_f + (1-\delta_i)\psi_i(z,\ell_i)$$

and

$$\psi_f(z,\ell_i,\ell_f) = \delta_f \int_{c_f} [\mathbf{1}^e(z)][\mathbf{1}^f(z,c_f)]\psi_z(z)\psi_c(c_f)dc_f + (1-\delta_f)\psi_f(z,\ell_i,\ell_f)$$

2. **Equilibrium in the outside sector**. Demand for the outside sector good comes from consumers, who spend a fraction $1-\alpha$ of their income on it, and from firms, who demand it to pay entry costs, registration costs and vacancy costs. Aggregate income itself is the sum of wage income earned by wage and salary employees, and earnings of workers in the outside sector, i.e.

$$I = L_o w_o + L_e \bar{w}$$

where w is the average earnings of workers employed in the industry. The market clearing condition is then given by:

$$A_o L_o = (1-\alpha)I + M^e(c_e + \bar{c}_x) + M\bar{c}_v$$

where $\bar{c}_f$ is the average formalization costs paid by firms that register their business, while $\bar{c}_v$ is the average vacancy costs paid by formal and informal firms.

3. **Equilibrium in the labor market**. At the beginning of each period,the total number of wage and salary jobs is

$$L_e = M\bar{\ell}$$

where ℓ is the average employment in the industry, equal to

$$\bar{\ell} = s_i \int_z \int_{\ell_i} \ell_i \psi_i(z, \ell_i) dz d\ell_i + (1 - s_i) \int_z \int_{\ell_i} \int_{\ell_f} (\ell_i + \ell_f) \psi_f(z, \ell_i, \ell_f) dz d\ell_i d\ell_f$$

where s_i is the share of unregistered firms. Some of these jobs are destroyed as firms exit for exogenous reasons or because of exogenous separation. Summing these sources of job destruction, we obtain our measure of industrial workers who are thrown into unemployment before the interim period

$$\begin{aligned}\tilde{U} =&(\delta_i + \delta_s) M s_i \int_z \int_{\ell_i} \ell_i \psi_i(z, \ell_i) dz d\ell_i + \\ &(\delta_f + \delta_s) M (1 - s_i) \int_z \int_{\ell_i} \int_{\ell_f} (\ell_i + \ell_f) \psi_f(z, \ell_i, \ell_f) dz d\ell_i d\ell_f\end{aligned}$$

hence the associated destruction rate is equal to $\mu_d = \tilde{U}/L_e$. In the steady state equilibrium there are no net flows of workers out of the outside sector. Accordingly, the total number of wage and salary job seekers each period includes those who just lost their jobs ($\tilde{U}$), and those who searched unsuccessfully for jobs last period (L_u),

$$U = \tilde{U} + L_u$$

Since $L_u = (1 - \tilde{\phi})U$, then

$$\tilde{\phi} U = \mu_d L_e$$

That is, the number of workers flowing into wage and salary jobs $\tilde{\phi}U$ must match the number of wage and salary jobs that are turning over. Finally, at the end of each period, workers either must have jobs in one of the sectors or be unsuccessful job seekers:

$$1 = L_o + L_u + L_e$$

4. **No arbitrage condition**. Workers non employed in a wage and salary job are indifferent between searching for a wage and salary job or not, i.e.

$$\mathcal{J}^n = \mathcal{J}^s = \mathcal{J}^o = \frac{1 + r}{r} w_o$$

C.2 Solution algorithm

To solve the model in general equilibrium, we implement the following algorithm:

- Guess a firm's probability of filling a vacancy, ϕ^0
- Use the matching function to compute worker's probability of finding a a job, $\tilde{\phi}^0$ as follows:

$$\tilde{\phi}^0 = (1 - (\phi_f^0)^\eta)^{\frac{1}{\eta}}$$

  - Compute wages of formal and informal workers in registered and unregistered firms using the solution of the bargaining problem
  - Solve the problem of the formal and informal firms and store policy functions for hiring of formal and informal workers, firm registration and firm entry
- Use the firm's policy function to simulate a panel of firms and compute shares and distribution of informal vacancies posted by unregistered firms, informal vacancies posted by registered firms and formal vacancies
- Solve the problem of the workers and compute the expected value of being employed, $\mathbf{E}[\mathcal{J}^e]$
- Compute the value of searching and, $\mathcal{J}^s$ and evaluate convergence by comparing it to the value of being out of labor force, $\mathcal{J}^o = w_o/r$
  - if $|\mathcal{J}^s - \mathcal{J}^o| > \epsilon$, update the guess for domestic shifter:
    - set $\phi^1 < \phi^0$ if $J^s < J^0$
    - set $\phi^1 > \phi^0$ otherwise

    and go back till convergence
  - if $|\mathcal{J}^s - \mathcal{J}^o| < \epsilon$, stop here, store $\phi^* = \phi$ and $\tilde{\phi}^* = \tilde{\phi}$ and go ahead
- Use the converged value of ϕ^* and $\tilde{\phi}^*$, the definition of matching function, and the market clearing for the outside sector, to obtain a solution for the endogenous measure of incumbent firms M^* and a measure of workers searching for wage and salary jobs, U^*.
- Use the labor market identities to compute the measure of wage and salary employed L_e^*, unemployed L_u^* and employed in the outside sector, L_o^*.

The problem of the firm is solved with value function iteration using a 50-points grid for productivity, 350-point grid for informal employment and formal employment, and 500-point grid for the cost of registration. We set the maximum number of formal

workers to 3000 and the maximum number of informal workers to 50. In the steady state, a negligible fraction of firms reaches the maximum size, and this is also the case in the data.

C.3 Estimation algorithm

In the calibration algorithm, we exploit the worker's no-arbitrage condition between value of searching and value of non-participating

$$J^o = J^s = J^n$$

and the solution for the earnings of those employed in the outside sector,

$$w_o = A_o$$

to treat the job filling probability, ϕ, as a parameters to estimate, and to treat the productivity in the outside sector, A_o, as equilibrium objects. Moreover we exploit the free-entry condition to treat the aggregate demand shifter D as a parameter to estimate, and treat the sunk cost of entry, c_e as an equilibrium object.

Hence, we start by guessing the following set of parameters,

$$\vartheta^0 := \{\phi^0, D^0, \overline{c_f}^0, c_v^{i0}, c_v^{f0}, \gamma_0^0, \gamma_1^0, \gamma_2^0, \gamma_3^0, \gamma_4^0, \alpha^0, \varphi_z^0, \zeta_i^0, \zeta_f^0, \eta^0\}$$

Then we solve the model as follows:

- Use the matching function to compute worker's probability of finding a a job, $\tilde{\phi}^0$ as follows:

$$\tilde{\phi}^0 = (1 - (\phi^0)^{\eta^0})^{\frac{1}{\eta^0}}$$

  - Solve the problem of the formal and informal firms and store policy functions for hiring of formal and informal workers, firm registration and firm entry
  - Store $c_e^0 = V^e(\vartheta^0)$

- Use the firm's policy function to simulate a large panel of firms and compute shares and distribution of informal vacancies posted by unregistered firms, informal vacancies posted by registered firms and formal vacancies

- Solve the problem of the workers and compute the expected value of being employed in the industry, $\mathbf{E}[\mathcal{J}^e(\vartheta^0)]$

- Compute the value of searching for a job in the industry, $\mathcal{J}^s(\vartheta^0)$

- Using the no-arbitrage condition, set w_o such that $\mathcal{J}^s(\vartheta^0) = \mathcal{J}^o(\vartheta^0) = \mathcal{J}^u(\vartheta^0)$, i.e.

$$w_o^0 = \frac{r}{(r + \tilde{\phi}^0)} \left(\tilde{\phi}^0 \mathbf{E}[\mathcal{J}^e(\vartheta^0)] + (1 - \tilde{\phi}^0) b \right)$$

- Use ϕ^0 and $\tilde{\phi}^0$, the definition of matching function, and the market clearing for the outside sector, to obtain a solution for the endogenous measure of incumbent firms $M(\vartheta^0)$ and a measure of workers searching for wage and salary jobs, $U(\vartheta^0)$.

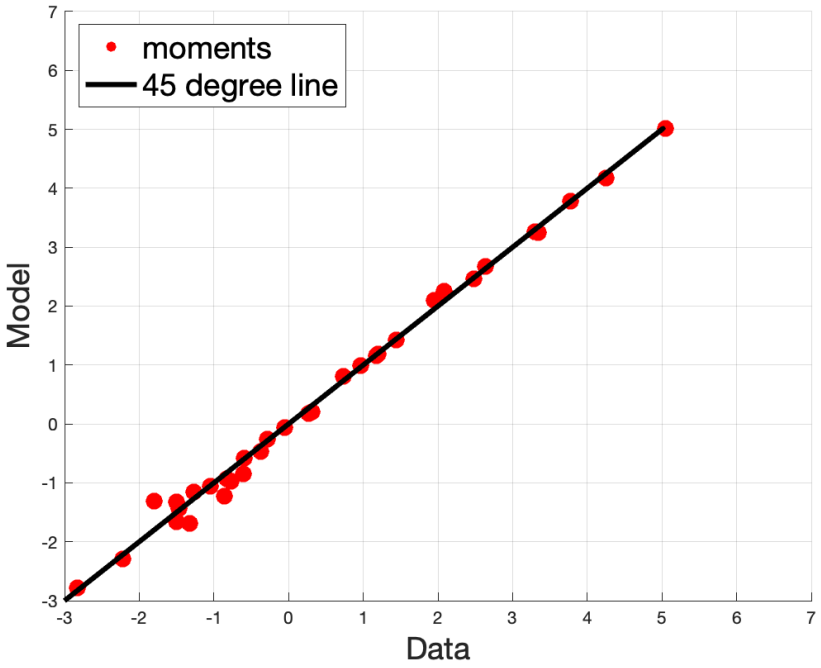
- Use the labor market identities to compute the measures of wage and salary employed $L_e(\vartheta^0)$, unemployed $L_u(\vartheta^0)$ and those employed in the outside sector, $L_o(\vartheta^0)$.

We use worker's and firm's policy functions to simulate a large panel of workers and to compute a vector of model-based moment conditions, $g(\vartheta^0)$. Therefore, we iterate on parameters to minimize the following objective function:

$$d(\vartheta)W'd(\vartheta)$$

where $d(\vartheta)$ denotes the absolute distance between a vector of empirical targets, $\bar{g}$ (discussed in Section 4.2) and their model counterpart, $g(\vartheta)$, while W is a diagonal matrix with entries equal to the inverse squared of each empirical moments. We follow a genetic algorithm to update the vector of guesses. Figure C.1 scatters empirical versus simulated moments. At the obtained minimum, the log deviation between empirical and simulated moments is 0.12.

Figure C.1: Estimation fit



D Payroll tax reform

Table D.1 reports a set of counterfactual outcomes for different values of payroll tax rates, τ_w. We include the share of informal firms, the share of informal vacancies and the average firm size. For the second group, we report the rates of informality, the measure of firms, the labor market tightness, unemployment rate, real average wage and the real GDP per employee.

Table D.1: Payroll taxes on formal workers for registered firms

Payroll tax rate, τ_w	0	0.10	0.20	0.30	0.40
Firm-level outcomes					
Informal firms, share	0.9513	0.9614	0.9671	0.9748	0.9790
Informal vacancies, share	0.4765	0.5326	0.5778	0.6585	0.7097
Average firm size	4.1359	3.6054	3.3072	2.8946	2.7012
Aggregate Outcomes					
Informality rate	0.4706	0.5255	0.5702	0.6511	0.7025
- , extensive margin	0.2647	0.3265	0.3944	0.4766	0.5435
- , intensive margin	0.2060	0.1990	0.1920	0.1745	0.1590
Measure of firms	0.0897	0.1071	0.1200	0.1420	0.1549
Market tightness	0.2885	0.4040	0.4619	0.6319	0.6726
Unemployment rate	0.0744	0.0493	0.0419	0.0271	0.0250
Average real wage	1.2126	1.1721	1.1313	1.0913	1.0388
Real GDP per worker	1.0406	1.0309	1.0080	0.9778	0.9433

Notes: Average wage is expressed as function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

Counterfactual outcomes following a change in payroll tax qualitatively resemble those obtained after changes in corporate income tax reported in Section 5.

E Labor market policy interventions

Unemployment benefits. Table E.1 reports a set of counterfactual outcomes for different values of the unemployment benefits, b. We include the share of informal firms, the share of informal vacancies and the average firm size. For the second group, we report the rates of informality, the measure of firms, the labor market tightness, unemployment rate, real average wage and the real GDP per employee.

Table E.1: Unemployment benefits

Unemployment benefits, b	0^*	$0.05w_o$	$0.10w_o$	$0.15w_o$	$0.20w_o$
Firm-level outcomes					
Informal firms, share	0.9683	0.9680	0.9665	0.9663	0.9641
Informal vacancies, share	0.5918	0.5862	0.5713	0.5680	0.5546
Average firm size	3.2498	3.2745	3.4204	3.4115	3.5672
Aggregate Outcomes					
Informality rate	0.5842	0.5785	0.5642	0.5609	0.5480
- , extensive margin	0.3948	0.3875	0.3687	0.3653	0.3486
- , intensive margin	0.1894	0.1910	0.1954	0.1956	0.1995
Measure of firms	0.1243	0.1186	0.1090	0.1054	0.0960
Market tightness	0.4785	0.4345	0.3876	0.3360	0.2769
Unemployment rate	0.0406	0.0448	0.0506	0.0594	0.0728
Average wage	1.1198	1.1630	1.2217	1.2638	1.3197
Real GDP per worker	1	1.0150	1.0357	1.0501	1.0700

Notes: * refers to the baseline outcomes. Average wage is expressed as function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

We express the unemployment benefits as a share of earnings in the outside sector, w_o, and we assume it is financed with lump-sum taxes paid by all workers. An increase in unemployment benefits has the same qualitatively implications as a reduction in corporate income taxes, although firm-level and aggregate outcomes change by a much smaller magnitude.

Minimum wage. Table E.2 reports a set of counterfactual outcomes for different values of the minimum wage, $\underline{w}$. We express the minimum wage as a multiple of earnings in the outside sector, w_o, and we assume only formal firms in registered firms are subject to it. This implies the following wage schedule for formal firms in registered firms:

$$w_f(z, \ell_i, \ell_f) = \max \left\{ \underline{w}, \frac{(1-\zeta_f)b + \zeta_f(1-\tau_y)\frac{R_f(z,\ell_i,\ell_f)}{\ell_i+\ell_f}}{(1-\beta\tau_w^f)} \right\}$$

The introduction of a minimum wage does not produce any effect unless large enough. For values of larger than two times the average earnings in the outside sector, the minimum wage moves firms and jobs out of formality, reallocates workers from high- to low-productivity firms, reducing allocative efficiency, real wages and income per worker.

Table E.2: Minimum wage on formal workers for registered firms

Minimum wage, $\underline{w}$	0^*	$1w_o$	$1.5w_o$	$2w_o$	$2.5w_o$	$3w_o$
Firm-level outcomes						
Informal firms, share	0.9683	0.9683	0.9683	0.9782	0.9860	0.9905
Informal vacancies, share	0.5918	0.5918	0.5918	0.7316	0.8572	0.9159
Average firm size	3.2498	3.2498	3.2498	2.3329	2.0616	2.0246
Aggregate Outcomes						
Informality rate	0.5842	0.5842	0.5842	0.7241	0.85202	0.9127
- , extensive margin	0.3948	0.3948	0.3948	0.5918	0.76641	0.8540
- , intensive margin	0.1894	0.1894	0.1894	0.1323	0.0856	0.0587
Measure of firms	0.1243	0.1243	0.1243	0.1772	0.2088	0.2148
Market tightness	0.4785	0.4785	0.4785	0.6043	0.7619	0.9986
Unemployment rate	0.0406	0.0406	0.0406	0.0294	0.0215	0.0139
Average wage	1.1198	1.1198	1.1198	1.0601	1.0053	1.0017
Real GDP per worker	1	1	1	0.9545	0.8960	0.8610

Notes: * refers to the baseline outcomes. Average wage is expressed as function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

On the other hand, lower allocative efficiency reduces competition among firms, lowers concentration in the labor market, and raises the probability of finding a wage and salary job for workers.