Firms, policies, informality, and the labor market*

Camila Cisneros-Acevedo† Alessandro Ruggieri‡

University of Tübingen CUNEF Universidad

November 20, 2023

Abstract

This paper studies how taxes and regulations affect firms and labor market outcomes in the developing world. We build a general equilibrium model of firm dynamics subject to search frictions, taxes, and imperfectly enforced legislation. This setup leads to informal employment along the intensive and extensive margin. Estimated to match firm and worker-level data from Peru, the model sheds light on the effect of corporate income tax on informality and unemployment. On the one hand, a reduction in corporate taxes concentrates employment in larger and more productive firms, increasing efficiency and reallocating workers to formal jobs. On the other hand, employment shifts to a smaller mass of firms creating higher unemployment duration and higher income inequality. Holding tax revenues constant, we compare two simulated reforms: a reduction in corporate income tax and an equivalent reduction in payroll taxes. We find that contracting corporate income taxes can achieve 0.9% higher output gains, 1 p.p. higher formal employment, and a 1.3 p.p. lower unemployment rate. A cut in payroll taxes generates instead lower and more unequally distributed output gains. A revenue budget-neutral welfare-maximizing policy shifts the burden from corporate income to payroll taxes, reducing informality by 2.2 p.p. and increasing output per capita by 2.4%.

Keywords: firm dynamics, corporate taxes, payroll taxes, informality, unemployment, welfare

JEL Classification: H20, J46, J60, O17

---

*We are grateful to Francesco Amodio, Javier Fernandez Blanco, Nezih Guner, Luis Rojas, Sevi Rodriguez Mora, Aysegul Sahin, Cezar Santos, and Todd Schoellman for useful discussion and seminar participants at the CEPR Macro and Growth meeting, Spanish Macro Network, CUNEF Universidad, Universitat Autonoma of Barcelona, and University of Nottingham, for helpful comments. The usual disclaimers apply.

†University of Tübingen, Department of Economics, Mohlstr. 36, 72074 Tübingen, Germany, camila.cisneros-acevedo@uni-tuebingen.de.

‡CUNEF Universidad, Department of Economics, C. de Leonardo Prieto Castro 2, 28040, Madrid, Spain, E-mail: alessandro.ruggieri@cunef.edu.

1 Introduction

Over 60 percent of workers in the world operate in the informal economy. In developing countries, it accounts on average for 35% of GDP and 70% of the labor force (Perry, 2007). Since informal employment is a prominent cause of low-paying jobs and low aggregate productivity, common policy prescriptions aim to discourage it by improving regulatory governance and reducing the tax burden on firms (De Soto, 1989; Ohnsorge and Yu, 2022).

This paper studies the labor market consequences of such policy interventions in the developing world. On the one hand, the burden of heavy regulation encourages firms in developing countries to remain informal and it distorts employment decisions. On the other hand, workers in these countries face poorly functioning labor markets with relatively high search frictions. We shed light on the costs and benefits of taxes and regulations when frictions impede the correct functioning of the labor market.[1]

To study this issue, we focus on corporate income taxes, a widely common policy instrument. Corporate taxes are extensively used by governments in developing countries and are a key source of government revenues. Despite decreasing over the last two decades, corporate income taxes in 2018 accounted for 15.3% of all tax revenues in Africa, 15.4% in LACs, 10% in OECD countries, and more than 25% in several low-income countries.[2] Most importantly, high corporate taxes have been cited as a common reason for informal activity (Waseem, 2018), have been shown to negatively correlate with economic growth (Lee and Gordon, 2005), and policymakers 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. We show that these differences are associated with differential 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 this evidence, we build a model of firm dynamics that features search frictions in the labor market, corporate income and labor income taxes, and imperfectly enforced legislation. In the model, workers could be unemployed, self-employed,

[1] Labor market 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 and De Roux, 2021). See Poschke (2019), Donovan and Schoellman (2021), and Guner and Ruggieri (2022) for a review of the implications.

[2] Among others, Bhutan, Chad, the Democratic Republic of the Congo, Equatorial Guinea, Indonesia Nigeria, and Papua New Guinea (OECD, 2018).

or wage-employed for a firm. Firms are heterogeneous in their productivity and in the cost of setting up a formal business. On the one hand, imperfectly enforced legislation encourages firms to hide from the tax authority and save on corporate income taxes, thereby generating informal employment along the *extensive margin*. On the other hand, registered firms can hire workers either formally or off the books. If they choose the former, they save on labor taxes and generate informal employment along the *intensive margin*. Through workers' and firms' dynamics, the model economy produces 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 reflects the following three considerations. First, Peru is a country with a very high informality rate: over 70 percent of the population is employed informally, either along the intensive or the extensive margin. Second, firms in Peru are subject to large corporate income and labor tax rates, amounting to 29.5% and 22.0% respectively. Finally, the availability of data on informal firms and workers allow us to identify parameters governing the expected costs of informality faced by either informal business or registered companies. The estimated model closely matches the 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 sizes, 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 replicas of the Peruvian economy that differ only in their corporate tax rates 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 the unemployment rate and for about 45% of the differences in GDP per worker.

The model delivers cross-country patterns in informality and unemployment via two major mechanisms: a *reallocation effect* and a *scale 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 formal firms, relative to informal. As a consequence, the share of registered firms in the economy increases, and, as they expand in size, the composition of posted vacancies shifts toward 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 reallocation of employment from low- to high-productivity firms, and a reduction in aggregate price, which increases 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 formal firms. This raises the expected value of searching for a wage and salary job relative to the value of being self-employed. To restore the equilibrium in the labor market, jobs concentrate on large and high-productivity firms. Since there are fewer of these firms, labor market tightness and the job-finding rate decline. Hence, unemployment increases leading to higher labor income inequality.

An important contribution of this paper is to provide a structure for the evaluation of the efficiency-equity trade-offs of various firm-related policy interventions in the context of labor market frictions. Among the others, we compare changes in corporate income taxes to changes in workers' payroll taxes. Although both policies alter firms and labor market outcomes, their effects vary: While corporate income taxes tackle formalization along the extensive margin, labor taxes have a direct effect on the intensive margin. As a result of both policies, there is a monotonic trade-off between higher workers' welfare and a lower unemployment rate. On the other hand, changes in payroll taxes do not produce as much welfare gains as changes in corporate income taxes do. Compared to a similar change in payroll taxes, a reduction in corporate taxes that increases the unemployment rate by 3 p.p. generates almost 2 times higher gains in aggregate welfare,

Finally, we evaluate the cost-effectiveness of such policies. We compare the effectiveness of equivalent reductions in corporate and payroll tax. First, for the same drop in tax revenues, a simulated reform that reduces corporate income taxes can achieve 0.9% higher output gains, 1 p.p. higher formal employment, and 1.3 p.p. lower unemployment rate. Reducing payroll tax would instead generate lower and more unequally distributed output gains. Second, a simulated revenue budget-neutral policy that fully shifts the burden of firm taxation from wage payroll to corporate income increases informality by 3.1 p.p. and reduces output by 3.2%. Third, lowering corporate income tax and increasing payroll taxes would be Pareto optimal: keeping the aggregate tax revenue constant, aggregate welfare is maximized at a pair of corporate and payroll income tax rates of 22.5% and 42.1%, respectively. Under this policy tax rates, informality reduces by 2.2 p.p. while output per capita increases by 2.4%.

This paper contributes to different strands of literature. First, our analysis highlights the aggregate and distributional implications of informal employment. 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 overlook the joint effect of search frictions and corporate taxes on informality and unemployment. 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. More recently, Dix-Carneiro et al. (2021) developed a multi-sector model with formal and informal employers and showed that informality affects how the gains from trade are distributed across workers. We contribute to this literature by focusing on the long-run consequence of taxes and regulations on labor market outcomes and aggregate welfare.

More generally, 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. 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. 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. We add to this literature by documenting how unemployment and informality vary with corporate income tax rates across low- and medium-income countries. Moreover, we take advantage of a structural model to study the aggregate and distributional implications of various government policies.

Finally, this paper speaks also to the literature on the macroeconomic effects of dual labor markets (Bentolila et al., 2010; Pijoan-Mas and Roldan-Blanco, 2022; Ahn et al., 2023). We complement this literature by focusing on the duality between formal and informal jobs and studying the long-run effects of corporate income and payroll taxes on labor market outcomes.

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 Corporate income taxes around the world

This section documents how labor market outcomes 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 standard statutory corporate income tax rates levied on domestic businesses for about 200 countries in the last 40

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

years.[4] We merge this information with 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 holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises or as paid domestic workers by households.[5] Informal jobs of employees are defined as those lacking coverage by the social security system, entitlement to paid annual or sick leave, or written employment contracts. Unemployment comprises people 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 a sample of workers with more than 25 years old. Finally, we proxy aggregate productivity using real GDP per worker and a production-side measure of total factor productivity.[6]

Overall, we gather data for 75 countries in the period 2010-2021 and construct an unbalanced panel of 326 country-year observations. Details on the data coverage are provided in Appendix A.

Table 1 reports 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. While the poorest country in the sample is Malawi, with a GDP per capita of about 1 USD per day (370 USD yearly), the richest country is 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 in context, the analogous measure for the US in 2021 was equal to 134,363 USD, a value about 4.3 times larger. Similarly, real TFP averages 60% of the value for the US, and it is as low as 19% in the poorest countries of the sample.

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 wage employment is large and widespread across countries in the sample: on average, about 17% of wage employment is informal, reaching more than 45% 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

[4]Where a progressive (as opposed to flat) rate structure applies, the top marginal rate is reported. See Appendix A for a comparison of the statutory corporate income tax rates with the average profit taxes reported by the World Bank Doing Business database, https://data.worldbank.org/indicator/IC.TAX.PRFT.CP.ZS.

[5]Our measure of informality excludes the self-employed. The ILO reports the informality rate as the sum of both self-employed (i.e. own-account workers) and informal wage employees. To focus on wage employment we subtract the self-employment rate from the overall informality rate. See Appendix A for additional results on self-employment.

[6]Real GDP per worker is taken from the World Bank Indicator database and the total factor productivity measure is taken from the Penn World Table v.10.0 (variable *ctfp*).

Table 1: Cross-country summary

	Obs	Mean	St.dev.	Min	Max
GDP per capita, 2017 USD	326	5677.28	3897.49	370.301	16950.3
GDP per worker, 2017 USD	326	31124.1	16035.1	2583.41	72420.6
TFP, PPP (US=100)	326	59.1	19.1	23.3	124.9
Corporate tax rate, %	326	24.9	7.36	9.21	38.5
Informality rate, %	326	17.0	11.1	0	47.4
Unemployment rate, %	326	6.88	6.22	0.21	29.3

Notes: Informal employment is expressed as a percent of total employment and comprises persons who in their main or secondary jobs were 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 coverage by the social security system, entitlement to paid annual or sick leave, or written employment contracts. The unemployment rate is reported in percent of the labor force. Corporate tax rates refer to the standard statutory corporate income tax rates levied on domestic businesses. GDP per worker is measured in 2017 USD and expressed in 1000 USD. TFP is constructed following Feenstra et al. (2015), expressed in PPP, and reported as a percent of the value for the US (=100). Source: Tax Foundation, ILO-stat, World Bank, Penn-World Table v.10.0 and authors' calculation.

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.[7] Outcomes are reported as residuals from a regression with 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.[8]

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 10% of informal wage employment. On the other hand, in countries with a tax rate of 30%, almost 20% of wage employment is informal. The slope of this relation is large ($\hat{\beta}$=0.371) and significant at 5% (s.e.= 0.09). 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 3.71 p.p. higher rate of informal wage employment.

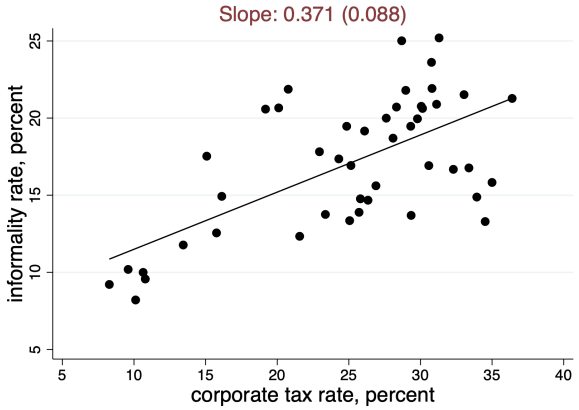
Panel B shows that the opposite pattern holds for the unemployment rate: high corporate tax rates are associated with lower unemployment. Countries with a tax rate of about 10% have on average a rate of unemployment of 15% while in countries with a tax rate of 30%, the unemployment rate is about 5%.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

[7] All figures report 50 dots, each corresponding to a 2 percent interval in the distribution of corporate income tax rates.

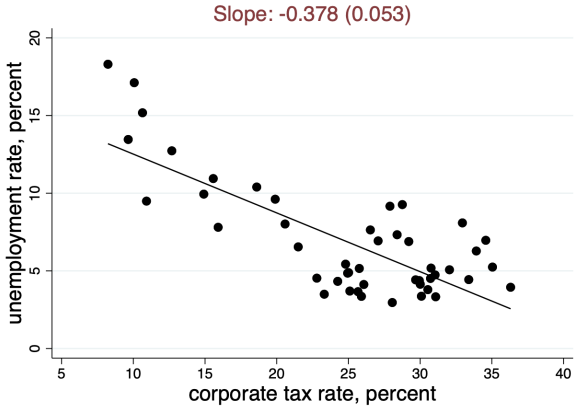
[8] We report the same scatter plots using the raw data in Appendix A, Figure **??**.

Figure 1: Informality, unemployment and corporate income taxes

(A) Informality rate



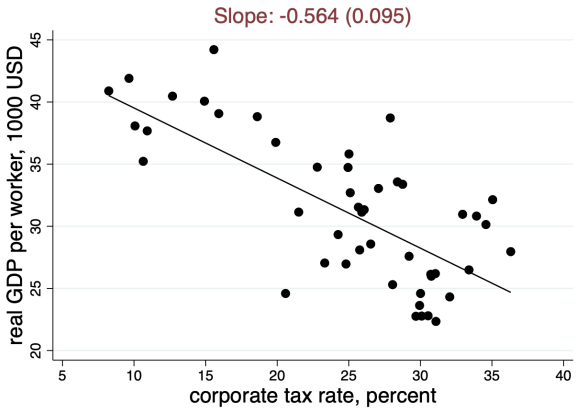
(B) Unemployment rate



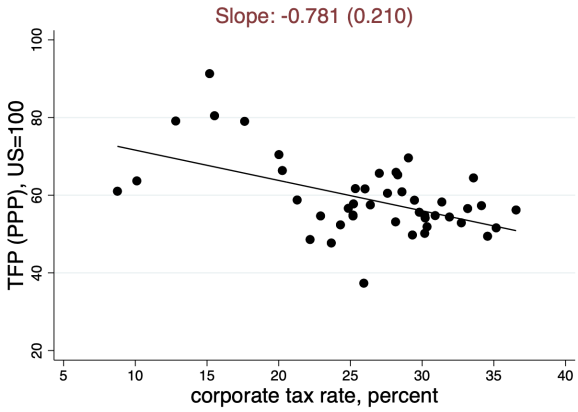
Notes: Informal employment is expressed as a percent of total employment and comprises persons who in their main or secondary jobs were employees holding informal jobs, whether employed by formal sector enterprises, informal sector enterprises, or as paid domestic workers by households. Informal jobs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

Figure 2: Aggregate productivity and corporate income taxes

(A) Real GDP per worker



(B) Total Factor Productivity



Notes: Real GDP per worker is measured in 2017 USD and expressed in 1000 USD. TFP is constructed following Feenstra et al. (2015), reported in PPP and expressed relative to the value for the US (=100). Corporate tax rates refer to the standard statutory corporate income tax rates levied on domestic businesses. Source: World Bank, Penn-World Table v.10.0, Tax Foundation and authors' calculation.

on year fixed-effects, a 10 percent higher corporate tax rate is associated with a rate of unemployment 3.78 p.p. lower.

Figure 2 documents how aggregate productivity varies across countries with different corporate tax rates. Panel A scatters GDP per person employed, expressed in 1000 USD. Panel B reports total factor productivity, expressed in PPP, and reported as a percent of the value in the US. Like Figure 1, each dot corresponds to the average

values of the dependent variable for countries in a specific bin of corporate income tax, after removing year-fixed effects. Both GDP per worker and TFP decline significantly as countries increase their tax burden on firms. GDP per worker 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% decline in corporate tax rate is associated with an increase in GDP of around 1,000 USD per employed worker.

TFP displays the same pattern as GDP per worker. As we move from countries with a 10% corporate tax rate to countries with a 35% rate, it declines by around 20 percent, relative to the US. The estimated slope ($\hat{\beta}$ =-0.781) is significant at 5% and implies that a 2% decline in corporate tax rate is associated with an increase in TFP of 1.5 percent, relative to the US.

In summary, this section unveils three key cross-country patterns. As countries reduce their tax rates on corporate income, the proportion of informal employment out of wage employment declines. Simultaneously, various measures of aggregate productivity increase at the expense of a higher unemployment rate. In Appendix A, we present a series of robustness checks to reinforce our findings. First, we show that alternative measures of informality exhibit a similar positive correlation with corporate income tax rates. Moreover, the same labor market dynamics persist when countries are ranked by the average profit taxes paid by their companies, rather than their statutory tax rates. Additionally, the observed cross-country patterns for each labor market outcome are robust to controlling for country-unobserved heterogeneity. Finally, we offer supplementary insights into how self-employment varies across countries and document no correlation with corporate income tax rates, which suggests that corporate tax rates correlate with informality rates across countries solely through changes in the composition of wage employment.

Our analysis sheds light on the intricate relationship between tax policies, labor markets, and economic informality. 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 that features 1) endogenous firm dynamics, 2) search frictions in the labor market, and 3) informality along the extensive and intensive margin. We focus on a stationary equilibrium, hence aggregate outcomes are time-invariant.

Time is discrete. The economy is populated by a unitary measure of workers-consumers and by an endogenous measure of firms. Workers are ex-ante homoge-

neous but differ in their employment status: they can be either wage-employed in the industrial sector, self-employed, 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 vacancies 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

$$\mathcal{U}_T = \sum_{t=T}^{\infty} \frac{c_t^{\alpha} s_t^{1-\alpha}}{(1+r)^t} \tag{1}$$

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 a variety ω. Utility maximization for a worker j with income I_j yields a demand for the homogeneous good s and for variety ω equal to

$$s = (1-\alpha) I_j \qquad \text{and} \qquad c(\omega) = \alpha \frac{I_j}{P} \left(\frac{p(\omega)}{P} \right)^{-\sigma} \quad \forall \omega \in [0, M]$$

where

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

is the exact price index for the composite good.

3.2 Production

The homogeneous good is produced by self-employed workers. Production requires labor, L_o as a unique input, homogeneous across suppliers. The self-employed produce A_o units of output per worker and face no friction in the product and labor markets. The total production of the homogeneous good is then equal to

$$y_o = A_o L_o. \tag{2}$$

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. Differences in productivity can equally well be considered differences in product varieties.

Firms also differ in the cost of setting up a formal business and 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:

$$y_i(z, \ell_i) = Az\ell_i \tag{3}$$

where A is a measure of aggregate productivity. Registered firms are allowed to combine informal and formal labor services, ℓ_i, and ℓ_f,

$$y_f(z, \ell_i, \ell_f) = Az(\ell_i + \ell_f). \tag{4}$$

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 as self-employed and their labor income 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:

$$m(U, V) = \frac{UV}{(U^\eta + V^\eta)^{\frac{1}{\eta}}} \tag{5}$$

where $\eta > 0$ governs both the efficiency of the matching function and its 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

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

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

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

where $\tilde{\phi} = \lambda(U,V)V$. Workers who get matched with a firm enter a bargaining stage to determine the wage rate, while workers who fail to match become unemployed, sustaining themselves with a benefit b. 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. Finally, wage and salary employees might lose their jobs either because of an exogenous separation shock, δ_w, or because of firm exit, which differs between unregistered and registered firms, and is equal to δ_i and δ_f respectively.

3.4 The problem of the industrial firms

Figure 3: Firms' decisions

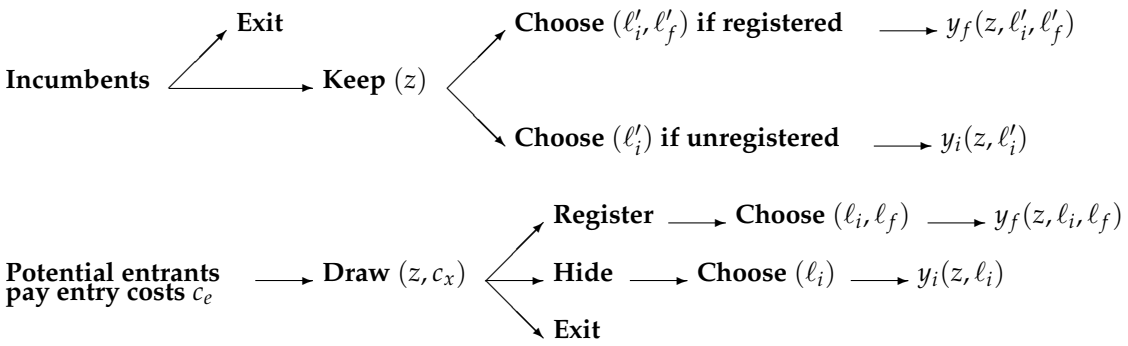


Figure 3 shows the timing of firms' decisions in the model. At the beginning of each period, potential industrial firms pay an entry cost, observe their productivity level and their cost of operating formally, and decide whether to create a new business and whether to formally register. Once incumbent, 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.

3.4.1 Revenues

Aggregating consumers' demand yields total demand for a variety ω, equal to

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

where D is an aggregate demand shifter, common to all firms, defined as

$$D = P^{\sigma-1}\alpha \int_0^1 I_j dj.$$

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:

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

Product differentiation makes the revenue function to be decreasing return to scale in the total number of employees, despite linearity in production. This ensures a well-defined firm size.

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

$$\begin{aligned} \mathcal{V}_i(z) = \max_{v_i} \quad & -c_v^i v_i + \frac{1-\delta_i}{1+r}\tilde{\mathcal{V}}_i(z, \ell_i) \\ \text{s.t.} \quad & \ell_i = \phi v_i. \end{aligned} \tag{6}$$

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:

$$\begin{aligned} \tilde{\mathcal{V}}_i(z, \ell_i) = \max_{v_i'} \quad & \pi_i(z, \ell_i) - c_v^i v_i' + \frac{1-\delta_i}{1+r}\tilde{\mathcal{V}}_i(z, \ell_i') \\ \text{s.t.} \quad & \ell_i' = (1-\delta_w)\ell_i + \phi v_i \end{aligned} \tag{7}$$

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

$$\pi_i(z, \ell_i) = R_i(z, \ell_i) - w_{ii}(z, \ell_i)\ell_i - \kappa_i(z)\ell_i$$

and $w_{ii}(z, \ell_i)$ are wages paid to informal employees. While unregistered firms do not incur any tax, they face a per-worker expected cost of informality, $\kappa_i(z)$. This cost is a reduced-form device that captures the probability of detection by the government and subsequent fines, defined as

$$\kappa_i(z) = \gamma_0 z^{\gamma_1} \quad \gamma_0 > 0, \gamma_1 > 0. \tag{8}$$

Everything else equal, more productive - hence larger - firms find it more costly to hire an extra informal worker and expand their size.[9]

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

$$\begin{aligned} \mathcal{V}_f(z) = \max_{v_i, v_f} \quad & - \sum_{j \in \{i,f\}} c_v^j v_j + \frac{1-\delta_f}{1+r} \tilde{\mathcal{V}}_f(z, \ell_i, \ell_f) \\ \text{s.t.} \quad & \ell_j = \phi v_j \qquad \forall j \in \{i, f\}. \end{aligned} \tag{9}$$

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

$$\begin{aligned} \tilde{\mathcal{V}}_f(z, \ell_i, \ell_f) = \max_{v_i', v_f'} \quad & \pi_f(z, \ell_i, \ell_f) - \sum_{j \in \{i,f\}} c_v^j v_j' + \frac{1-\delta_f}{1+r} \tilde{\mathcal{V}}_f(z, \ell_i', \ell_f') \\ \text{s.t.} \quad & \ell_j' = (1-\delta_w)\ell_j + \phi v_j' \qquad \forall j \in \{i, f\} \end{aligned} \tag{10}$$

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

$$\pi_f(z, \ell_i, \ell_f) = (1-\tau_y)[R_f(z, \ell_i, \ell_f) - (1+\tau_w) w_{ff}(z, \ell_i, \ell_f)\ell_f] - [w_{if}(z, \ell_i, \ell_f) + \kappa_f(z, \ell_i, \ell_f)]\ell_i$$

and $w_{if}(z, \ell_i, \ell_f)$ and $w_{ff}(z, \ell_i, \ell_f)$ are wages paid to informal and formal employees, respectively. 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,

[9] Given stationarity, each productivity value maps into a unique level of employment, which makes our formulation of the informality costs isomorphic to a function that depends on firm size or firm gross revenues (Dix-Carneiro et al., 2021). We choose the former to ease the numerical solution of the model.

$\kappa_f(z, \ell_i, \ell_f)$ defined as:

$$\kappa_f(z, \ell_i, \ell_f) = \gamma_2 z^{\gamma_3} \left(\frac{\ell_i}{\ell_i + \ell_f} \right)^{\gamma_4}. \tag{11}$$

Everything else equal, the cost of informality is larger for more productive firms, and it decreases with the total number of workers. The current formulation ensures a well-defined composition of formal and informal workers within registered firms: more productive firms and firms with a high share of informal employment find it more costly to hire an extra informal worker.[10]

3.4.3 Entry and formalization decision

Every period, 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

$$\mathcal{V}(z) = \int_{c_f \in \mathcal{C}} \max\{\mathcal{V}_i(z), \mathcal{V}_f(z) - c_f\} \psi_c(c_f) dc_f \tag{12}$$

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

$$\mathcal{V}^e = \int_{z \in \mathcal{Z}} \max\{\mathcal{V}(z), 0\} \psi_z(z) dz \leq c_e \tag{13}$$

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 into three groups: firms who do not enter, firms entering without registering, firms entering and registering.

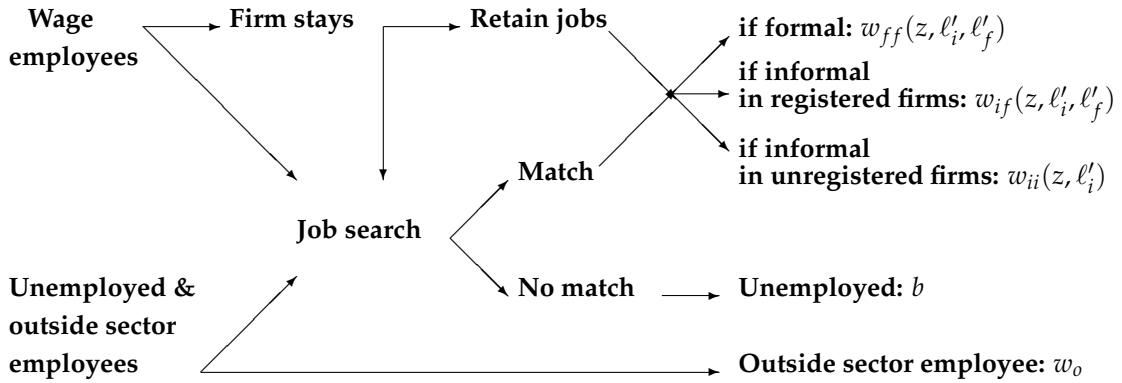
3.5 The problem of the workers

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

A worker who is not employed in a wage and salary job at the beginning of the

[10] This function microfounds 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. See also Erosa et al. (2021) for a similar formulation.

Figure 4: Workers' decisions



period chooses whether to search for it or not and solves the following problem

$$\mathcal{J}^n = \max\{\mathcal{J}^o, \mathcal{J}^s\} \tag{14}$$

where $\mathcal{J}^o$ is the value of being self-employed, equal to

$$\mathcal{J}^o = w_o + \frac{1}{1+r}\mathcal{J}^n, \tag{15}$$

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

$$\mathcal{J}^s = (1-\tilde{\phi})\mathcal{J}^u + \tilde{\phi}\mathbf{E}\mathcal{J}^e, \tag{16}$$

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

$$\mathcal{J}^u = b + \frac{1}{1+r}\mathcal{J}^n \tag{17}$$

Workers who choose to be self-employed earn w_o in the current period, and have the option of searching again next period. Workers who choose to search in the current period fail to get matched to a firm with probability $1-\tilde{\phi}$ and receive a transfer b. The expected value of matching to a firm $\mathbf{E}\mathcal{J}^e$ reads as follows:

$$\begin{aligned}\mathbf{E}\mathcal{J}^e = \Bigg[&\frac{V_{ii}}{V}\int_z\int_{\ell_i}\mathcal{J}^e_{ii}(z,\ell_i)\nu_{ii}(z,\ell_i)dzd\ell_i + \frac{V_{if}}{V}\int_z\int_{\ell_i}\int_{\ell_f}\mathcal{J}^e_{if}(z,\ell_i,\ell_f)\nu_{if}(z,\ell_i,\ell_f)dzd\ell_i d\ell_f \\ &+ \frac{V_{ff}}{V}\int_z\int_{\ell_i}\int_{\ell_f}\mathcal{J}^e_{ff}(z,\ell_i,\ell_f)\nu_{ff}(z,\ell_i,\ell_f)dzd\ell_i d\ell_f\Bigg]\end{aligned} \tag{18}$$

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 the number of employees.

Finally, it remains to specify the values of being employed, $\mathcal{J}^e_{ii}(z,\ell_i)$, $\mathcal{J}^e_{if}(z,\ell_i,\ell_f)$ and $\mathcal{J}^e_{ff}(z,\ell_i,\ell_f)$. These values are equal to

$$\mathcal{J}^e_{ii}(z,\ell_i) = w_{ii}(z,\ell_i) + \frac{1}{1+r}\left[\hat{\delta}_i\mathcal{J}^n + (1-\hat{\delta}_i)\mathcal{J}^e_{ii}(z,\ell_i)\right] \tag{19}$$

$$\mathcal{J}^e_{if}(z,\ell_i,\ell_f) = w_{if}(z,\ell_i,\ell_f) + \frac{1}{1+r}\left[\hat{\delta_f}\mathcal{J}^n + (1-\hat{\delta_f})\mathcal{J}^e_{if}(z,\ell_i,\ell_f)\right] \tag{20}$$

and

$$\mathcal{J}^e_{ff}(z,\ell_i,\ell_f) = w_{ff}(z,\ell_i,\ell_f) + \frac{1}{1+r}\left[\hat{\delta}_f\mathcal{J}^n + (1-\hat{\delta}_f)\mathcal{J}^e_{f}(z,\ell_i,\ell_f)\right] \tag{21}$$

where $\hat{\delta}_i = \delta_w + (1-\delta_w)\delta_i$ and $\hat{\delta}_f = \delta_w + (1-\delta_w)\delta_f$. Employed workers are paid $w_{ii}(z,\ell_i)$ if informal in unregistered firms, $w_{if}(z,\ell_i,\ell_f)$ if informal in registered firms and $w_{ff}(z,\ell_i,\ell_f)$ if formal. If workers lose their jobs (which happens with probabilities $\hat{\delta}_i$ and $\hat{\delta_f}$, their continuation value is $\mathcal{J}^n$, i.e. they face the option of choosing whether to search again or to move directly to the outside sector.

3.6 Wage determination

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 the incentive to break the match. Following Binmore et al. (1986), production delay constitutes the only credible threat in the negotiation, which makes the current-period payoffs the only relevant payoffs to split (Hall and Milgrom, 2008).

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:

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

$$\Pi^{\text{worker}}_i(z,\ell_i) = [w_{ii}(z,\ell_i) - b]\ell_i$$

Failing to reach an agreement delays production of one period and 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.[11]

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:

$$w_{ii}(z, \ell_i) = (1 - \zeta_i)b + \zeta_i \frac{R_i(z, \ell_i)}{\ell_i} \tag{22}$$

Informal workers get paid a ζ_i share of the average revenue product, $R_i(z, \ell_i)/\ell_i$, and a share $1 - \zeta_i$ of their outside option, b.

Consider now the bargaining problem between a registered firm and its employees. We assume formal and informal employees bargain separately with their employer over the average surplus they generate, net of corporate income, and payroll taxes.[12] Therefore, the surpluses accruing to registered firms and to the collective of informal employees are equal to

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

$$\Pi_f^{\text{worker}}(z, \ell_i, \ell_f) = \left[w_{if}(z, \ell_i, \ell_f) - b\right] \ell_i$$

while the surpluses shared by registered firms and the collective of formal employees are equal to

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

$$\Pi_f^{\text{worker}}(z, \ell_i, \ell_f) = \left[w_{ff}(z, \ell_i, \ell_f) - b\right] \ell_f$$

Using the same Nash splitting rule used above, the wage functions for informal and

[11] At the time of bargaining, workers do not have anymore the option of relocating to the outside sector. This makes the unemployment transfer, b, the only relevant threat point in the bargaining problem. This assumption ensures well-defined measures of workers in the outside sector versus unemployed. See Coşar et al. (2016) for a similar timing in the bargaining protocol.

[12] 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.

formal employees in registered firms are equal to

$$w_{if}(z, \ell_i, \ell_f) = (1 - \zeta_i)b + \zeta_i \frac{R_f(z, \ell_i, \ell_f)}{\ell_i + \ell_f} \tag{23}$$

and

$$w_{ff}(z, \ell_i, \ell_f) \quad = \max \left\{ \underline{w}, \frac{(1-\zeta_f)}{(1+\zeta_f[\tau_w - \tau_y(1+\tau_w)])} b + \frac{\zeta_f(1-\tau_y)}{(1+\zeta_f[\tau_w - \tau_y(1+\tau_w)])} \frac{R_f(z,\ell_i,\ell_f)}{\ell_i+\ell_f} \right\} \tag{24}$$

Both types of workers are paid a share of the average firms' revenues net of corporate income taxes, $(1 - \tau_y)R_f(z, \ell_i, \ell_f)/(\ell_i + \ell_f)$. On the other hand, wages of formal workers depend also on the payroll tax τ_w^f, which increases their bargaining power against their employers and cannot be lower than a mandated minimum wage, $\underline{w}$.

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 entrants and incumbent firms, share of unregistered firms, aggregate income, and distribution of firms across productivity values and size, such that:

- the policy functions solve the problem of workers and firms, and the value functions attain their maximum;

- workers optimally choose the sector in which they are working or seeking work, i.e. workers non-employed in a wage and salary job are indifferent between searching for a wage and salary job or not, i.e.

$$\mathcal{J}^n = \mathcal{J}^s = \mathcal{J}^o = \frac{1+r}{r} w_o, \tag{25}$$

  which implies the following value of unemployment:

$$\mathcal{J}^u = b + \frac{1}{r} w_o \tag{26}$$

- a positive mass of entrants M^e replaces exiting firms every period so that the free entry condition holds with equality:

$$\mathcal{V}^e = \int_{z \in \mathcal{Z}} \max\{\mathcal{V}(z), 0\} \psi_z(z) dz = c_e \tag{27}$$

- wages are the solution of a bargaining problem between employers and employees and determined by the equations (22), (23) and (24);
- the distributions of firms over productivity and size replicate themselves through the entry and the registration decisions, and exit shocks (equations (29) and (30));
- the product market for the outside good clears, i.e. supply matches demand (equation (32));
- the labor market for wage and salary jobs clears, i.e. the flow of workers into unemployment matches the flow of workers out of unemployment (equation (33)).

In Appendix B 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. Subsequently, shed light on several empirical facts on informal employment and, more broadly, on labor market outcomes in Peru. Following this, we 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 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 are 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 pertaining to both informal and formal workers is sourced from the Peruvian National Household Survey (ENAHO). ENAHO stands as a continuous cross-sectional representative survey encompassing the entire Peruvian population. Conducted across all regions of Peru, the survey is compartmentalized into various sec-

tions. Our focus here is directed towards the segment dedicated to independent workers and the segment that provides information on employment and income for all household members aged 14 and above. These two specific sections furnish comprehensive data on individuals, encompassing demographic attributes such as age, gender, education, and region of residency. Additionally, the survey captures crucial details including industry classification (4-digit ISIC), ownership structures, and the number of workers employed by individuals' respective employers. Notably, the surveyed jobs' characteristics enable us to categorize all employed workers within the sample into three distinct groups: informal workers along the extensive margins, informal workers along the intensive margin, and formal workers.

We label workers as informal along the "extensive margin" if they assert their employment with a company that does not maintain financial records on the online platform or software mandated by the Peruvian Tax Collection Agency (SUNAT) during the tax declaration process. Additionally, in such cases, the employer is characterized as an "unregistered employer."

Individuals professing employment in firms categorized as "registered" undergo further evaluation. We classify individuals engaged in registered family firms as intensive-informal workers if they serve as paid domestic workers. For salaried workers, we adopt the approach outlined in Cisneros-Acevedo (2022), utilizing two distinct questions for the periods 2007-2011 and 2012-2017. In the timeframe spanning 2007 to 2011, intensive-informal workers are identified as those in registered firms who claim that SUNAT did not make any deductions from their income. Conversely, between 2012 and 2014, they are individuals declaring that their employers did not contribute to their health insurance, a violation of Peruvian laws.[13] Our sample is confined to women and men aged between 25 and 60, engaged in non-military occupations, reporting positive hours worked, and functioning as wage and salary employees. Table A.5 in Appendix A describes the final sample of workers.

Information pertaining to formal firms is sourced from the World Bank Enterprise Survey (WB-ES), a cross-country survey encompassing a representative sample of private-sector firms. This survey delves into various aspects of the business environment and performance, covering general firm demographics such as age, employee count, and ownership, as well as details on sales and input purchases. Specifically focusing on Peruvian firms, the survey was conducted in the years 2006, 2010, and 2017. Our attention is directed towards formal companies, characterized by their registration with the SUNAT, and specifically those with a workforce of 5 or more employees.

Data regarding informal firms is taken from the World Bank Informal Enterprise

[13]The question regarding income deductions is applicable only between 2007 and 2011, while the question regarding health insurance is relevant only between 2012 and 2014. For further insights, refer to Cisneros-Acevedo (2022).

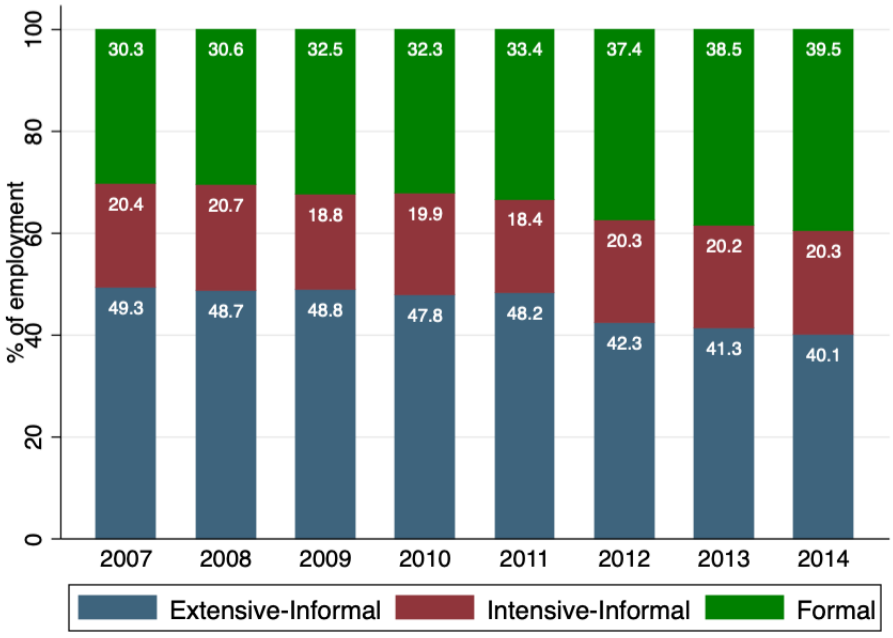
Survey (WB-IFS), a cross-country survey conducted concurrently with the WB-ES. This survey, while addressing comparable business topics, is specifically tailored to capture data on informal business activities across various countries. In its implementation, IFS equates informality with non-registration. In the context of Peru, informal firms are delineated as those not registered with the SUNAT, thus maintaining consistency with the definition employed in the ENAHO. Tables A.6 and A.7 in Appendix A describe the samples of formal and informal firms.

4.1.1 Facts on Informality in Peru

We now highlight four important facts on formal and informal workers and firms in Peru, which we 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



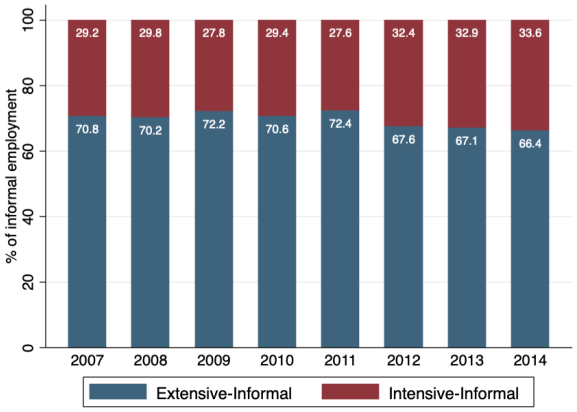
Notes: This figure reports the percentage of wage and salary employees who are informally employed in unregistered firms (blue bar), informally employed in registered firms (red bars) and formally employed (green bars). Source: ENAHO and authors' calculation.

During this period, more than 40% of workers were employed in non-registered firms, while around 20% of workers were employed off-the-books by firms that were registered with the Tax Collection Agency. Combining intensive and extensive mar-

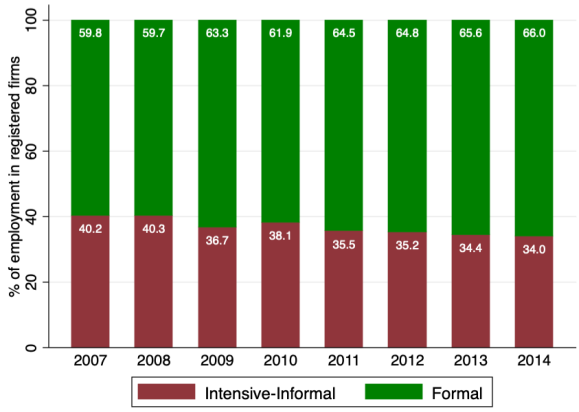
gins, between 60% and 70% of workers in Peru were employed without any safety net such as retirement, paid holidays, or sick leave.

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.

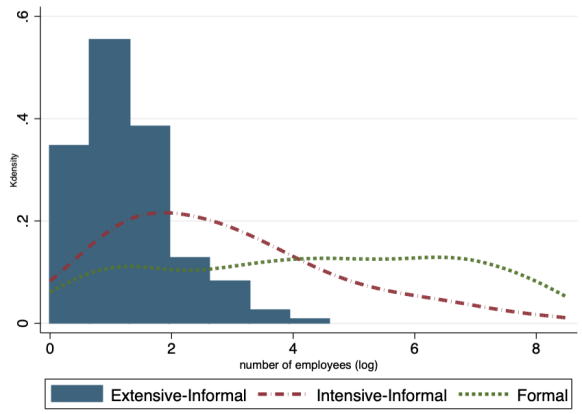
While formal employment has witnessed a rise in this decade, the proportion of informal workers along the intensive margin has remained steady. Notably, informal workers engaged in formal firms constitute 30% of the overall informal employment landscape in Peru (Figure 6A). Comparable figures of 47% in Mexico (Samaniego de la Parra, 2017) and 40% in Brazil (Ulyssea, 2018) underscore a similar trend. Furthermore, informal workers contribute to 35% and 40% of the total employment within formal firms (Figure 6B). In Appendix A, we delve into the composition of informal employment, dissecting it by education level (college and non-college), gender (male and female), and sector (manufacturing and non-manufacturing). Our analysis reveals that the prevalence of informal employment along the intensive margin persists and is substantial within each of these groups.

Informal workers are more likely to be employed in smaller firms. The share of informal workers in registered firms declines with size. Figure 7A illustrates the distribution of formal and informal workers based on the size of their employing firm, with size measured by 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. Conversely, informal workers associated with registered firms are more prevalent in medium-sized companies.[14]

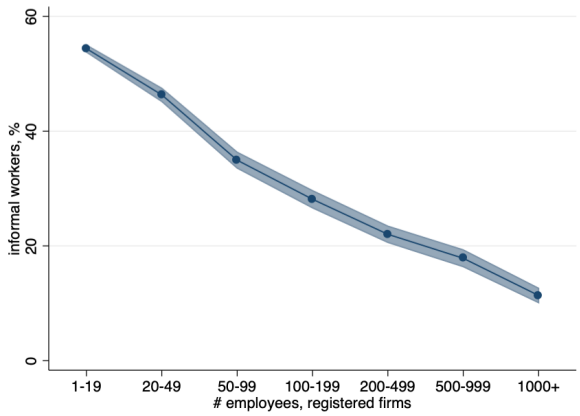
[14]In the Appendix we show that the distribution of firm size remains qualitatively the same when we restrict the sample to workers in the manufacturing and non-manufacturing sectors, male and female, college and non-college educated workers. See Figure A.5.

Figure 7: Firm size across formal and informal workers

(A) Distribution of firm-size



(B) Informal workers in registered firms

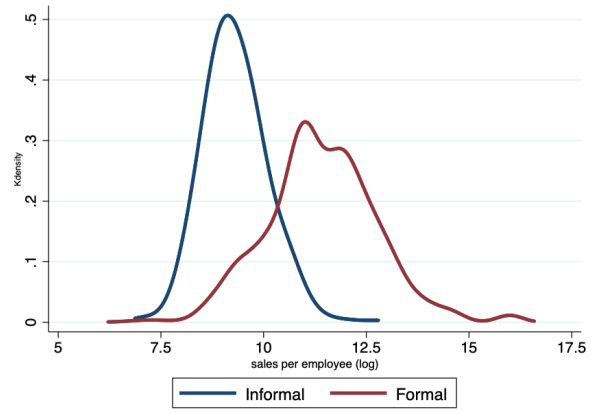


Notes: Panel A reports the distribution of firm size for workers who are informally employed in unregistered firms (blue bars), informally employed in registered firms (red dashed line), and formally employed (green dotted line). Panel B reports the average share of informal workers employed in registered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

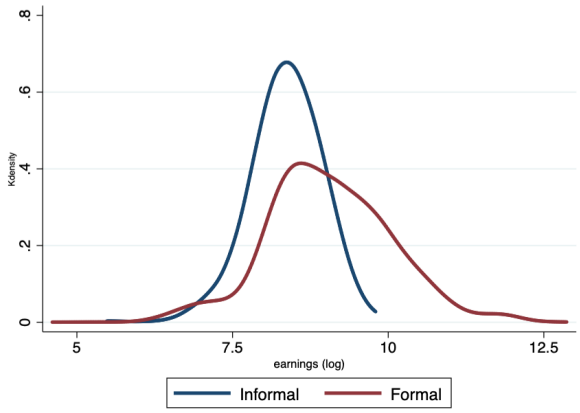
While formal employment can be found in firms of almost any size, the largest share of formal workers has a job in big companies. Figure 7B reports the percent of informal workers in registered firms for different employers ranked by their firm size. Larger firms are composed of a significantly higher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

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 the 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.

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 type of firm. 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:

$$\log w_{it} = \alpha \mathbf{1}[\text{Formal}]_{it} + \beta \mathbf{1}[\text{Intensive Informal}]_{it} + \mu_t + \gamma X_{it} + \epsilon_{it} \tag{28}$$

where w_{it} is the real monthly earnings of worker i at time t, the variables $\mathbf{1}[\text{Formal}]_{it}$ and $\mathbf{1}[\text{Intensive Informal}]_{it}$ are indicators denoting whether the worker is employed formally and informally 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)
$\mathbf{1}[\text{Intensive Informal}]_{it}$		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. $\mathbf{1}[\text{Intensive Informal}]_{it}$ 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 (28). 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, the earnings of informal workers depend on whether workers are employed in a registered firm or not. Informal workers in registered firms 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 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 A.8 in Appendix A).

4.2 Estimation

The estimation strategy proceeds 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, as for tax rates, set to their statutory values. 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%
A	Aggregate productivity	1	normalization
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_w	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	Transfer to the unemployed	0	OECD (2019)
$\underline{w}$	Minimum wage, % of median	95.0	CEDLAS (2010-2015)

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). We normalize aggregate productivity A to 1. 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_w is chosen to have a monthly E-U rate of 7.6% (Reynaga and Ramírez-Rondán, 2021). The 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, following the discussion in OECD (2019), we set the transfer to the unemployed, b, to 0, and we fix the minimum wage to 95% of the median wage in the economy (Soares, 2018).

Firms productivity is drawn from a log-normal distribution, $z \sim \log\mathcal{N}(0, \varphi_z)$, with $\varphi_z > 0$, 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 squared inverse of each empirical moments.[15]

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 the number of employees. The last two groups include labor market outcomes such as the rate of wage employment, the shares of wage employment that is informal along the extensive and the intensive margins, measured overall and separately by the number of employees, and finally the overall job-finding rate and the finding rate for informal jobs.

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 measures of wage

[15] 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.

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, $\mathbf{E}[\log R_i]$	7.385	8.146	Wage employment rate	0.450	0.444
Average log-size, $\mathbf{E}[\log \ell_i]$	0.204	0.186	Extensive-informal wage employment, share	0.436	0.395
Log-size dispersion, $\mathbf{std}[\log \ell_i]$, %	0.364	0.295	Intensive-informal wage employment, share	0.221	0.189
			Share intensive informal, 1-19 employees	0.544	0.429
Formal firms			Share intensive informal, 20-49 employees	0.461	0.379
Average log-revenues, $\mathbf{E}[\log R_f]$	11.97	11.76	Share intensive informal, 50-99 employees	0.351	0.349
Average log-size, $\mathbf{E}[\log(\ell_i + \ell_f)]$	3.227	3.186	Share intensive informal, 100-199 employees	0.281	0.317
Log-size dispersion, $\mathbf{std}[\log(\ell_i + \ell_f)]$, %	1.303	1.187	Share intensive informal, 200+ employees	0.166	0.268
Log-size, 20th cutoff	2.079	2.257			
Log-size, 40th cutoff	2.639	2.678	*Aggregate outcomes*		
Log-size, 60th cutoff	3.296	3.256	Job finding rate (overall)	0.437	0.437
Log-size, 80th cutoff	4.249	4.173	Job finding rate (informal)	0.283	0.260

inequality, i.e. the wage gap between formal and informal workers in registered firms, and the gap between informal workers in registered versus unregistered firms.

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 role in identifying a particular parameter. Entry cost c_e and formality costs $\bar{c}_f$ 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.

The parameters governing the expected costs of informality for informal and formal firms, γ_0, γ_1, γ_2, γ_3, and γ_4 are identified by the distribution of both types of firms, and by the share of informal workers in formal firms of different size.

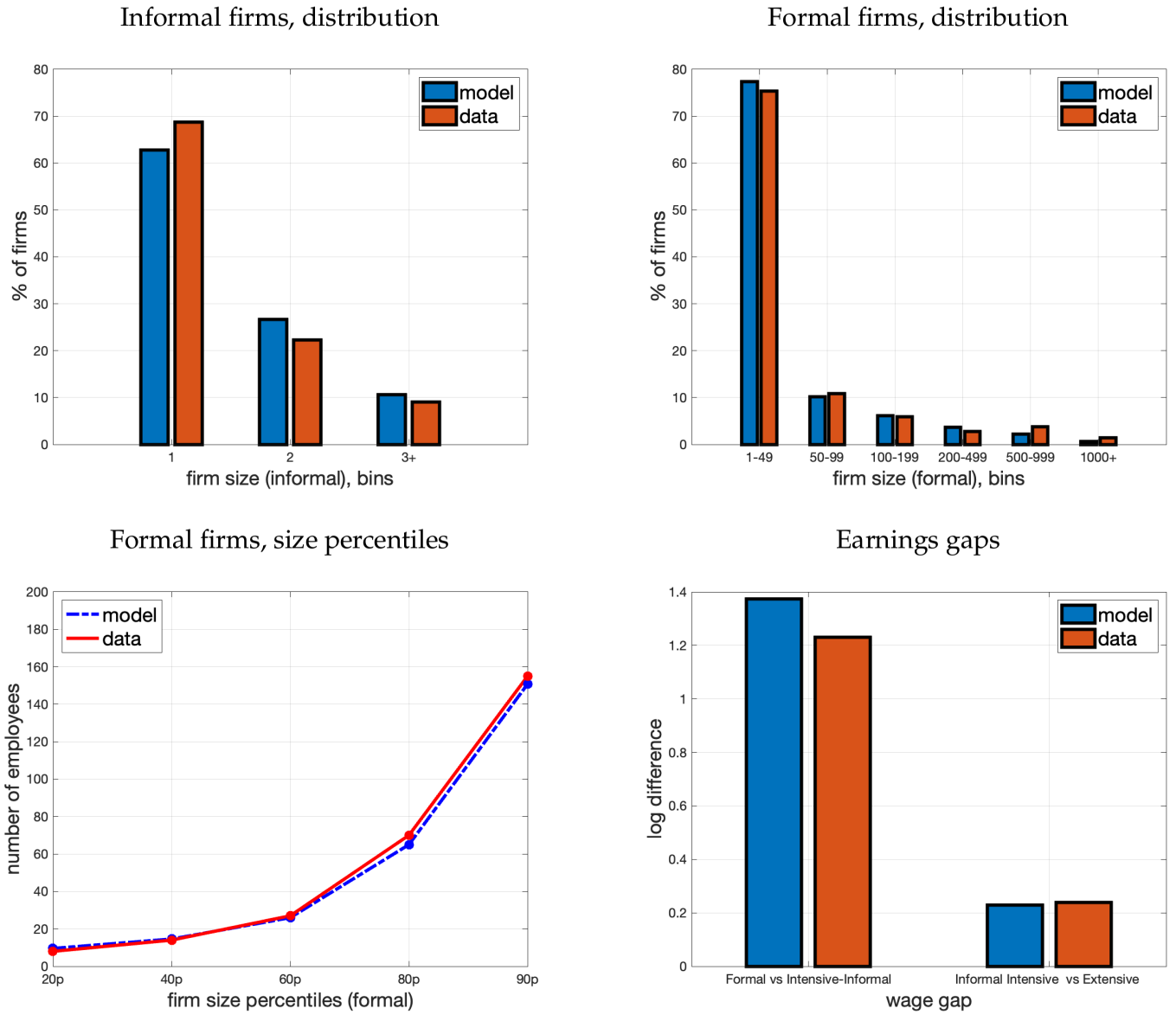
We interpret employment in the outside sector as composed of those who are not wage employed, i.e. those who report to be self-employed.[16] Therefore the wage-employment rate identifies the consumption share of differentiated varieties, α. The job-finding rates for formal and informal jobs in the industry 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 the dispersion of log size and wage gaps, respectively.

4.3 Estimates and model fit

Overall, the model is able to 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 generates 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 signifi-

[16] Following the definition from the ILO, the self-employed are defined as workers who report to be persons working on their own account, including unpaid family workers.

Figure 9: Selected targeted moments



Notes: This figure reports the model-based and the empirical distributions of informal and formal firms, size percentiles for formal firms, and earnings gaps between formal and informal workers.

cantly smaller, the majority being composed of one or two workers. Formal firms are larger, and more than 10 percent of those have more than 100 employees.

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

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 between 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 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.[17] In equilibrium, the earnings of those employed in the outside sector, w_o equals the productivity of the outside good, A_o. We calculate this to be $S/1{,}051.92 \times 0.353 = \$371.33$ per month, \$4,455.93 per year. This estimate implies the earnings of those employed outside the industry are on average 89% of the earnings of those employed in the industry.

[17] 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.

Expressed in dollars, the sunk cost of creating a new firm is estimated to be $S/3{,}832.66 \times 0.353$ = \$1352,92, while the costs of operating formally vary uniformly between 0 and $S/98{,}010.8 \times 0.353$ = \$34,597.81. The estimates imply an average entry cost 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.[18] Fajgelbaum (2016) uses official tax records of the manufacturing sector of Argentina and estimates the entry cost net of operating costs to be \$25,000.

Combining formal and informal firms, the average entry cost 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 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 parameter, $\eta = 2.11$ 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.

4.4 Non-targeted statistics

Table 7 compares data and model-based measures of wage inequality. Despite being non-targeted, the model 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 in 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.

[18] 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.

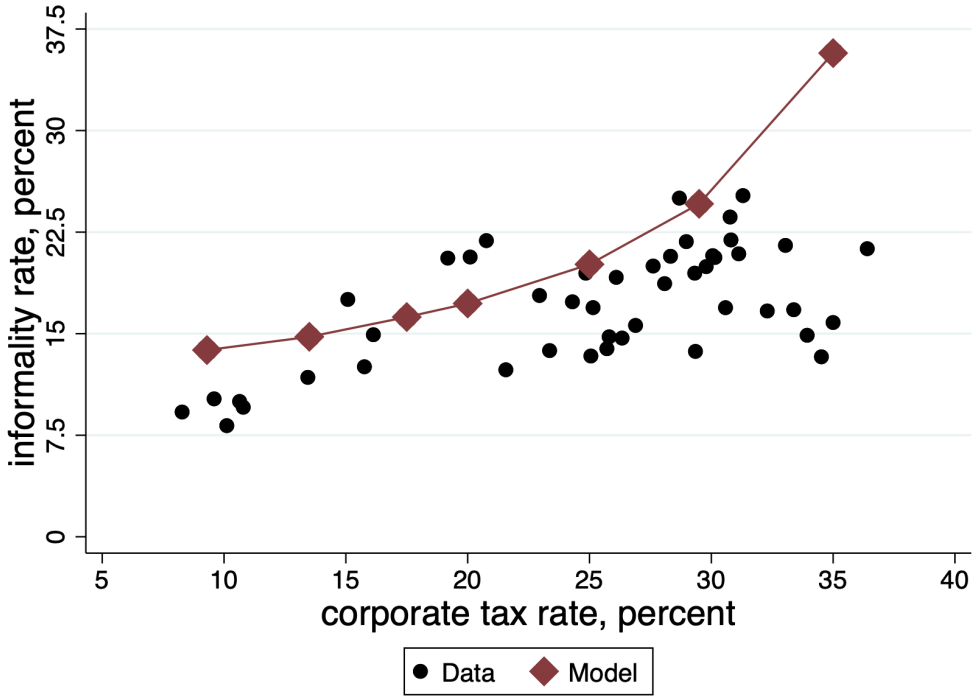
Table 7: Additional Statistics

Moment	Data	Model
Wage dispersion std$[\log w]$	0.875	0.517
Unemployment rate	0.037	0.041

5 Corporate income taxes around the world, revisited

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.[19] 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: Taxes and Informality: Model vs. Data

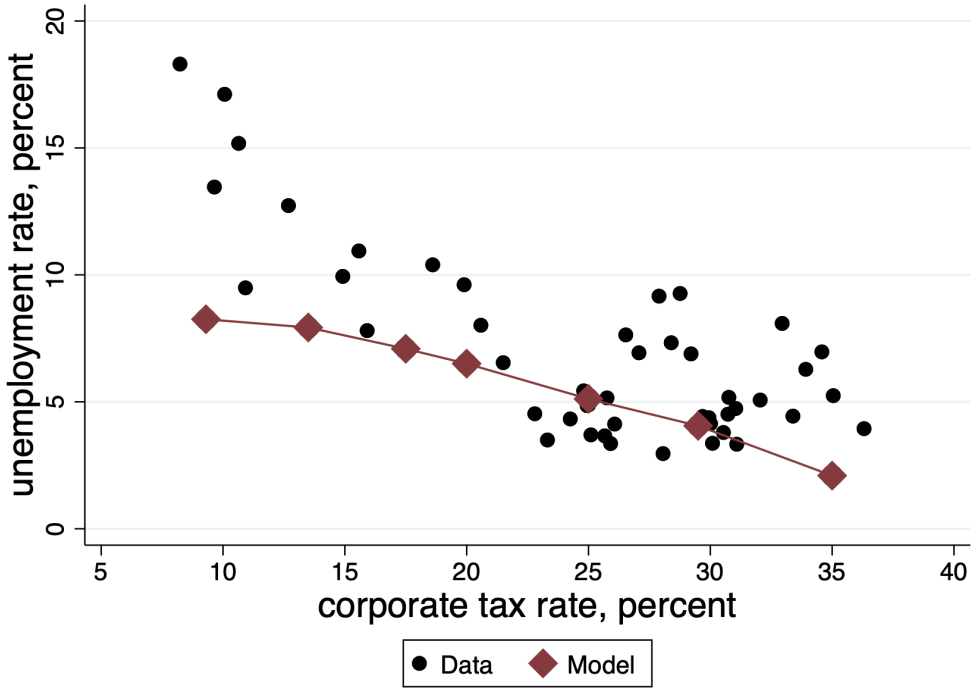


Notes: This figure shows the rate of informal wage employment rate for countries with different corporate income tax rates. The black dots represent the data and the red diamonds the model.

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 line with the data. As we move from low- to high-corporate tax environments, the share of informal wage and salary workers increases from 35% to 15% of aggregate

[19] In equilibrium, tax revenues are not rebated to workers. See Section 6.2 for the analysis of a revenue-neutral reform.

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 the red diamonds the model.

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).

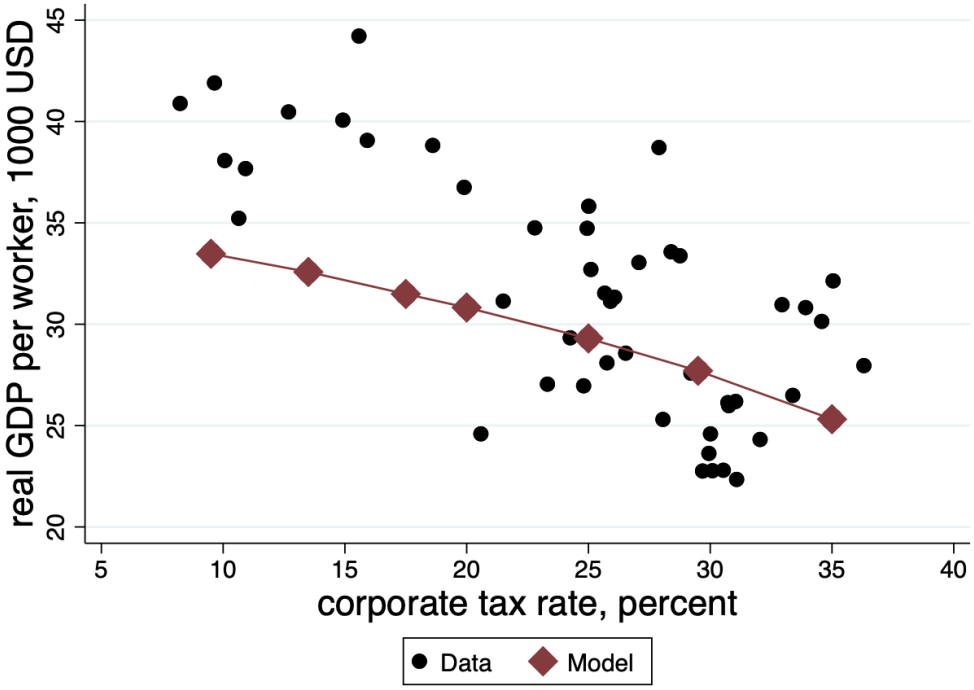
Figures 11 and 12 report data and model predictions for the unemployment rate and real GDP per worker. The model reproduces the pattern for the unemployment rate that we observe across countries: as we lower the corporate income tax rate, the unemployment rate increases by 5 p.p., from 3% to 8%. The 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.

Finally, consistent with the data, a reduction in corporate income tax rates increases GDP per worker in the model. This is achieved through job reallocation from low-productivity informal firms to high-productivity formal firms (see next section). Everything else equal, a model reduction of 25 p.p. 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.[20] For the informality rate, the model generates a slope of 0.789, which overpredicts the relation observed in the data (≈ 0.37). This arises be-

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

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 the red diamonds the model.

cause of the simulated outcome in the highest tax rate country. Ignoring the latter, the model is able to quantitatively mimic the negative correlation between the informal wage employment rate and the corporate tax rate.

Table 8: Slope Coefficients: Data vs. Model

Moment	Data	Model
Informality rate	0.371	0.789
	(0.088)	(0.183)
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.

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

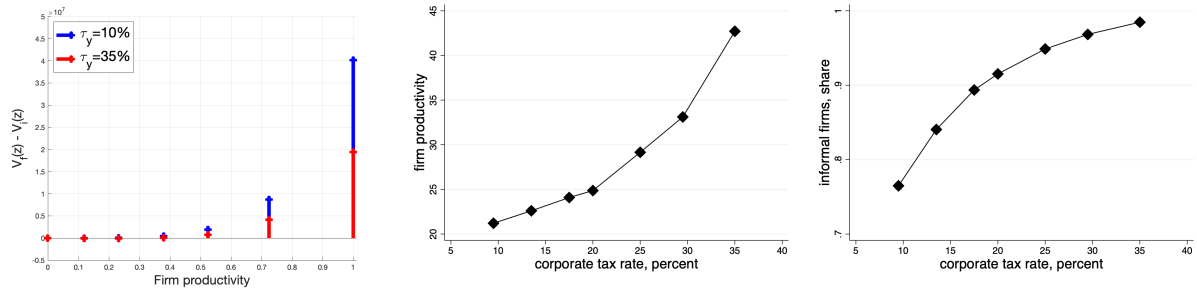
generates a slope of -0.262 for GDP per worker, which is about 45 percent of what is predicted by the data (-0.564).

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 the reallocation of workers across jobs. To describe it, Figure 13A 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

(A) Value of registering **(B)** Formality threshold **(C)** Informal firms



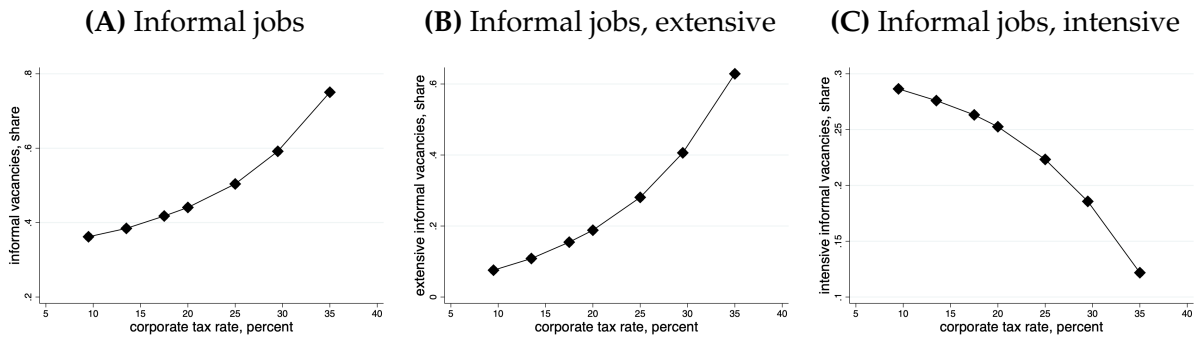
Notes: Panel A shows the relative value of forming a formal business ($V_f(z) - V_i(z)$) for firms with different productivity levels when the corporate income tax rate is equal to 10% (blue line) and 35% (red line). Panel B reports the average productivity thresholds for formal firms. Panel C reports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

Corporate tax rates act 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 13A). This happens across the board for all firms and, although high-productivity formal businesses gain the most, a tax relief allows low-productivity to cover the cost of formalization and push them out of informality. As a 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 13B), and the share of informal firms declines by more than 20 p.p. (from 98% to 77%, Figure 13C)

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 p.p. (from

75% to 36%, Figure 14.A). This is achieved through the formalization of jobs along the extensive margin, as opposed to the intensive margin, whose share increases with lower corporate tax rates, but not enough to overturn the trend. A higher share of informal vacancies translates into a 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 rates.

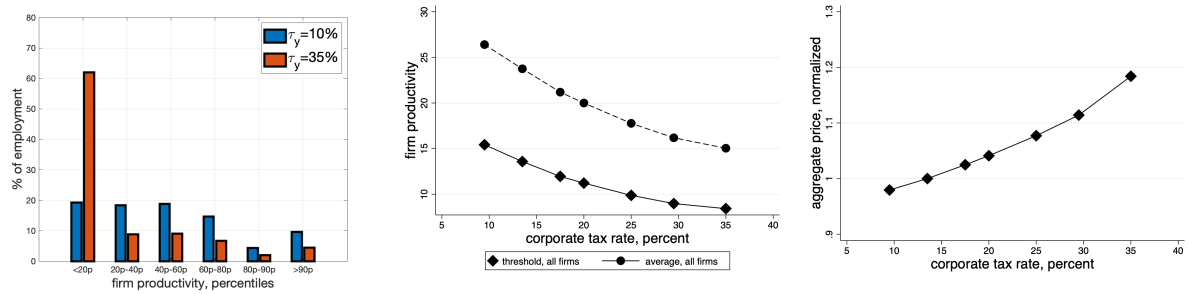
Second, the reallocation of firms increases also aggregate efficiency. Figure 15A reports the employment distribution across firm productivity percentiles in two selected counterfactual economies: one with a relatively low corporate income tax ($\tau_y = 10\%$, blue bars) versus one with a relatively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

High-productivity firms take advantage of lower taxes to charge a lower price and expand. This drives low-productivity firms out of the industry and lets workers reallocate to high-productivity firms (Figure 15A). The distribution of employment with respect to firms' productivity becomes left-skewed, and the share of workers in the top 10 percentile of the productivity distribution doubles from 5% to 10%. Due to higher firm selection, the productivity threshold for incumbent firms rises, which makes aggregate productivity increase (Figure 15B). Efficiency gains are expressed in the form of lower aggregate prices (Figure 15C). This effect resembles the one in 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. As documented in Section 2, increased efficiency and reduced prices result in a higher real GDP per employed worker.

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

Figure 15: Efficiency and prices

(A) Employment distribution **(B)** Firm productivity **(C)** Aggregate price index

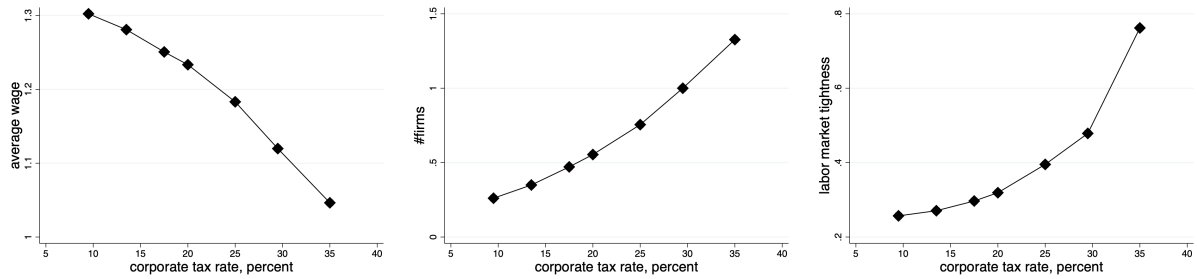


Notes: Panel A shows the distribution of employment across firms with different productivity levels when the corporate income tax rate is equal to 10% (blue bar) and 35% (red bar). Panel B and C report average productivity and productivity thresholds, and aggregate price index, for various counterfactual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

in the economy (panel B), and the labor market tightness (panel C).

Figure 16: Scale effect and 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 rates. Panel B and C report the mass of firms, and labor market tightness, for various counterfactual economies with different corporate income tax rates.

The 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 (25). The reduction in corporate taxes amplifies the average labor earnings for individuals with wage and salary positions compared to the value of earnings in the external sector (Figure 16A). Which enhances the attractiveness of seeking wage and salary jobs.

To restore equilibrium, jobs concentrate within a smaller pool of firms (Figure 16B), diminishing market tightness (Figure 16C) and making it more challenging for workers to secure wage and salary positions. As discussed in Section 2, this decline in job-finding rates leads to increased unemployment. This mechanism shares similarities with the findings in Feng et al. (2018), where shifts in the productivity of a modern sector elevate the value of job search away from a traditional sector, raising the risk of unemployment and subsequently elevating labor earnings inequality. In Appendix

C, we delve into the extent to which disparities in aggregate productivity versus corporate taxes can elucidate the labor market outcomes observed across countries, as detailed in Section 2. We demonstrate that a significant portion of cross-country patterns can still be attributed to the latter force.

5.2 Informality as a buffer

How does informality interact with corporate taxes? We answer this question by implementing a corporate income tax reform to a version of the model where we prevent firms from creating informal jobs. Table 9 reports the outcome of this exercise. Each entry corresponds to percentage points changes in selected outcomes following a reduction in corporate tax rate from 35% to 10%. Column (1) refers to the baseline model. Column (2) refers to a model where registered firms are not allowed to hire workers off-the-books and informality arises only along the extensive margin. Column (3) refers to a version of the model in column (2) where firms are also not allowed to operate informally.

Table 9: Corporate tax reform with and without informality

	Baseline (1)	Only extensive informality (2)	No informality (3)
Informality rate	-21.9	-31.7	-
Unemployment rate	+6.158	+10.72	+13.89
Real GDP per worker	+1.322	+1.443	+1.271

Notes: Each entry denotes a percentage point change following a reduction in corporate income tax from 35% to 10%. Column (1) refers to the baseline model. Column (2) refers to a version of the baseline model where informality arises only along the extensive margins. Column (3) refers to a version of the baseline model with no informality.

Abstracting from informality alters the aggregate and the distributional consequences of firm-level tax reforms. Changes in corporate tax rates induce a much larger decline in the informality rate when only the extensive margin is considered (column 2) compared to the baseline model (column 1). This is due to informal firms deciding to switch their status to registered. As such, they can no longer hire off-the-books. Consequently, informal vacancies in registered firms are relatively less scarce when corporate taxes are low (Figure 14.C).

Importantly, we find that the response of the unemployment rate is amplified when informal jobs are not explicitly modeled. The unemployment rate increases by 6 p.p. in the baseline model, by about 11 p.p. when only extensive margin informality is modeled, and by almost 14 p.p. when no informality is allowed. These results suggest that abstracting from informality would lead to significantly overestimating the

distributional cost of corporate tax reforms.

Finally, when the two margins of informality are present, the gains in output per worker from reducing corporate taxes are about 32.2%. These gains are lower than the ones in the economy with only the extensive margin, equal to 44.3%. This highlights that the ability to hire workers off-the-books reduces the combined negative effects of search frictions and taxes. On the other hand, the gains are higher than the ones in the economy with no informality at all, equal to 27.1%. This suggests that the presence of informal businesses amplifies the misallocation costs of corporate taxation.

6 Firm-level policy interventions

We now use the estimated model to study the labor market outcomes of alternative firm-policy interventions and compare them to corporate income tax reforms. The first policy is a change in payroll taxes for formal workers in registered firms, τ_w. The second policy is a change in the expected cost of hiding for informal firms, captured by the parameter γ_0 in the cost equation (8). The last policy is a change in expected fine from hiring workers off-the-books for formal firms, which we implement by changing the parameter γ_2 in the cost equation (11).[21]

6.1 Welfare gains

To assess the welfare properties and the efficiency-equity trade-off generated by each of these policies, 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.[22]

To study how inclusive the welfare gains are, we define and study a policy possibility frontier.[23] The frontier confronts two feasible outcomes for firm-level policies.

[21] We report the full set of counterfactual outcomes in Appendices D to F. In addition, we also explore the effects of labor market policy intervention as lump-sum transfers to the unemployed and minimum wages. See Tables G.1 and G.2 in Appendix G.

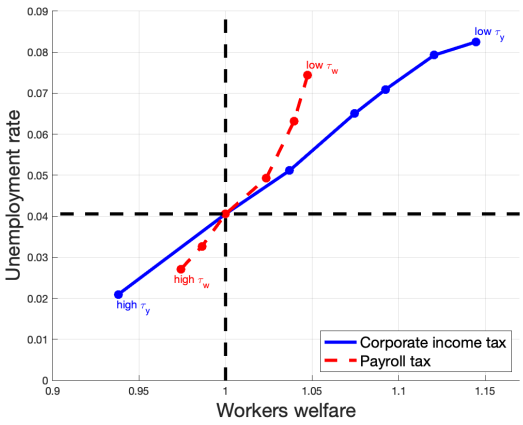
[22] Because both $\mathcal{J}_o$ and $\mathcal{J}_u$ are linked to r through the no-arbitrage condition, their values are invariant across counterfactuals. Changes in average welfare only occur because of changes in $\mathbf{E}\mathcal{J}^e$ - through workers' reallocation across firms in the industrial sector - or because of changes in employment shares across sectors.

[23] As an alternative, we could pose the existence of a social welfare function that maps the vector of

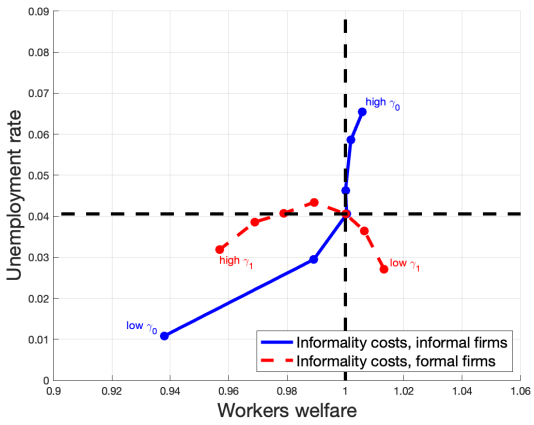
Specifically, Figure 17 scatters the unemployment rate against workers' average welfare for different levels of corporate income and payroll taxes (panel A), and for different regulatory costs of informal and formal firms (panel B). Each dot in the figure corresponds to a different counterfactual economy and welfare in the estimated economy is normalized to one.

Figure 17: Efficiency-equity trade-off

(A) Tax rates



(B) Informality costs



We start by focusing on tax policies, i.e. corporate income tax and payroll taxes (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 the unemployment rate is lower for the latter. Fixing a change of 1 p.p. in the unemployment rate, changes in corporate tax rates are associated with almost 2 times higher changes in welfare gains. This implies that neither policy is dominated by the other. While low values of corporate tax rates ensure higher welfare for the same level of the unemployment rate (right-upper quadrant of Figure 17A), low payroll taxes generate a lower unemployment rate for the same level of welfare (lower-left quadrant of Figure 17A).

Figure 17B replicates the same plot for expected costs of informality, i.e. informality costs for unregistered firms (blue line) and informality costs for hiring workers off-the-books in registered firms (red line). In this case, the latter unambiguously dominates the former. By ensuring relatively higher welfare and a relatively lower unemployment rate than any other combination of policies, lower informality costs move the possibility frontier to the lower-right quadrant 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 unemploy-

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.

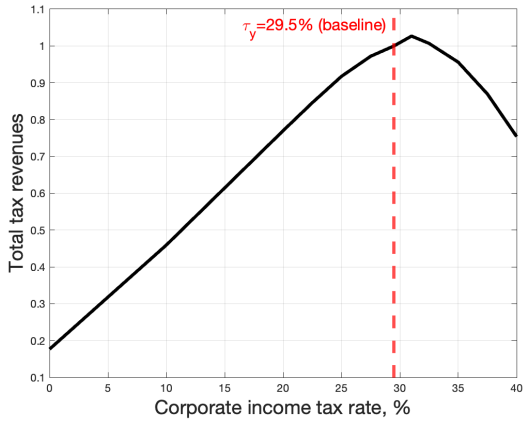
ment reducing rather than expanding. Overall, economies with a larger share of registered firms and a larger share of informal workers hired by formal firms are welfare dominants in the efficiency-equity space. [24]

6.2 Laffer curve and optimality

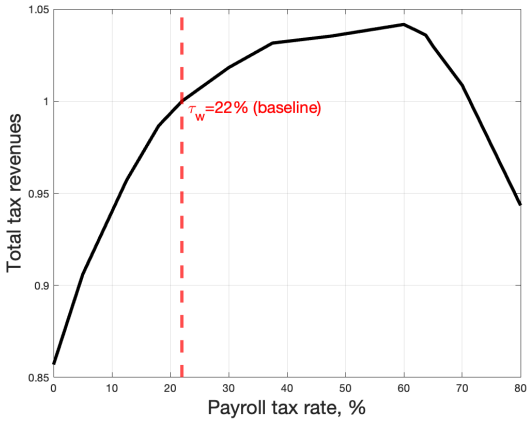
How much would it cost to implement such fiscal reforms? To answer this question, in Figure 18 we plot total tax revenues collected for different values of the corporate income tax rate (panel A) and payroll tax rates (panel B). Aggregate tax revenues are normalized by the amount collected in the baseline economy (which equals one). The dashed line refers to corporate income and payroll tax rates in the baseline economy, equal to 29.5% and 22% respectively.

Figure 18: Laffer curves

(A) Corporate income tax



(B) Payroll taxes



Notes: The laffer curves are computed keeping everything else equal, i.e. each parameter at its baseline value. Total tax revenues are normalized relative to the baseline economy.

A reduction in either corporate income or payroll tax rates would bring significantly fewer tax revenues to the government. Despite the gains in GDP per worker due to better allocative efficiency, and despite the increase in the share of formal companies that would increase tax capacity, a reduction in the corporate income tax rate of 10 p.p. would bring 55% fewer aggregate revenues to the government. The gradient of the Laffer curve is less steep for payroll taxes: a reduction of 10 p.p. would reduce aggregate revenues collected by slightly more than 5%. Both results suggest that tax reforms aimed at reducing firms' tax burden cannot simultaneously achieve output gains, accomplish employment formalization, and increase tax revenues. On the other

[24] Changes in welfare admits a decomposition between the fraction coming from the changing allocation of workers across employment states and the fraction coming from within-state changes. Counterfactual calculations suggest that, quantitatively, the latter plays only a minor role. Results are available upon request.

hand, keeping aggregate tax revenues constant, a reduction in payroll tax happens to be much less effective than changes in corporate tax rates.

Table 10 illustrates this point. We compare labor market outcomes in the baseline economy (column 1) against two alternative counterfactual economies (columns 2 to 3). In column 2 we consider an economy with no payroll tax and with a corporate income tax rate as in the baseline. In column 3 we leave the payroll tax unchanged and lower the corporate income tax rate from 29.5% to 22.5%, a value selected such that the aggregate tax revenues would drop by 15.8%, the same amount as in column 2. These two counterfactual economies are only equivalent in terms of foregone tax revenues: corporate tax rate cuts are in fact more effective in reducing informality rate (by 0.8 p.p.), they increase real GDP per worker by 1 p.p. more and do not generate as much unemployment as cuts in payroll taxes (0.059 versus 0.074, against a baseline of 0.041).

Table 10: Corporate income versus payroll taxes

	Baseline	Counterfactuals				
	(1)	(2)	(3)	(4)	(5)	(6)
Corporate tax rate	0.295	0.295	0.225	0.375	0	0.225
Payroll tax rate	0.220	0	0.220	0	2.250	0.421
Aggregate tax revenues	1	0.842	0.842	1	1	1
Informality rate	0.246	0.189	0.184	0.271	0.277	0.224
Unemployment rate	0.041	0.074	0.059	0.042	0.023	0.043
Real GDP per worker	1	1.078	1.087	0.968	0.984	1.024
Welfare	1	1.048	1.056	0.982	0.998	1.021

Notes: This table reports labor market outcomes for baseline (column 1) and counterfactual economies (columns 2 to 6). Column (2) refers to an economy with the same corporate tax rate as the baseline and no payroll taxes. Column (3) refers to an economy with the same payroll tax rate as the baseline and a corporate tax rate that keeps the aggregate tax revenues as in (2). Columns (4) to (6) refer to an economy with a mix of payroll tax and a corporate tax rate that keeps the government budget as the baseline. Aggregate tax revenues and real GDP per worker are normalized as a fraction of their values in the baseline economies.

In columns (4) to (6) of Table 10 we simulate a revenue-neutral change of labor and corporate income taxes. These economies are equivalent to the baseline only in terms of aggregate tax revenues collected by the government: a shift of the tax burden from labor to corporate income taxes increases informality and unemployment by 2.5 and 0.1 p.p. (relative to the baseline) and reduces real GDP per worker and welfare by 3.2% and 1.8% respectively (column 4). Shifting the tax burden from labor to corporate income taxes has comparable effects: while reducing unemployment by 1.8 p.p., it increases informality rates by 3.1 p.p. and it costs 1.6% or real GDP per capita (column 5)

We find the pair of corporate and payroll income tax rates that keep the aggregate

tax revenue constant and maximize aggregate welfare to be 22.5% and 42.1% (column 6). A lower corporate income tax rate fosters formalization and improves allocative efficiency: the informality rate declines by 2.2 p.p., while real GDP per capita and aggregate welfare increase by 2.4% and 2.1% respectively. At the same time a higher payroll rate allows the government to avoid the negative distributional effects of the reform — the unemployment rate increases by only 0.2 p.p., while keeping the budget balanced. Hence the policy mix in column (6) provides a Pareto improvement relative to the status quo in column (1).

Finally, Figure 18 shows that under the current policies, the Peruvian government budget is placed to the left of the Laffer Curve, especially regarding payroll taxes. If the objective function of the Peruvian government were to maximize aggregate revenues, it would have room to raise total revenues by increasing corporate income tax rates to 31%, or by raising payroll tax rates to 60%, values at which the Laffer curve peaks. However, this would cost a significant reduction in total output and a large increase in the overall informality rate.

7 Conclusion

In this paper, we study the distributional consequence of firm-level taxes and regulations in developing countries. Our initial findings reveal a correlation between higher corporate income tax rates and elevated informality rates, coupled with diminished aggregate productivity and lower unemployment rates. Based on these empirical insights, we construct a general equilibrium model encompassing firm dynamics and a frictional labor market. Our model unveils that a reduction in corporate tax rates underpins the observed cross-country patterns through two mechanisms: i.e. 1) a reallocation of jobs from low- to high-productivity firms, and 2) a augmented concentration of employers in the labor market.

While our framework incorporates features from macroeconomic and and development literature, such as corporate taxation, state-dependent distortions, entry barriers, and informal employment, we advance the state of the art by merging these elements within a search framework. This innovation empowers us to study the distributional consequences of diverse firm- and labor-market policy interventions, particularly those targeting informality along both extensive and 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 by shedding light on the inherent trade-offs between welfare gains and equity across workers concerning alternative firm-level policies. Our work emphasizes the importance of considering both margins of job informality in policy evaluation and underscores the nuanced trade-offs between efficiency gains

and government tax revenues.

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.

Ahn, H. J., Hobijn, B., and Şahin, A. (2023). The dual us labor market uncovered. Technical report, National Bureau of Economic Research.

Amodio, F. and De Roux, N. (2021). Labor market power in developing countries: Evidence from colombian plants. *CEPR Discussion Paper*, (DP16180).

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.

Bentolila, S., Cahuc, P., Dolado, J. J., and Le Barbanchon, T. (2010). Two-tier labor markets in the great recession: France vs. spain.

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.

Caselli, F. (2016). *Technology Differences over Space and Time*. Princeton University Press.

Cisneros-Acevedo, C. (2022). Unfolding trade effect in two margins of informality. the peruvian case. *The World Bank Economic Review*, 36(1):141 – 170.

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

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.

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.

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.

Feenstra, R. C., Inklaar, R., and Timmer, M. P. (2015). The next generation of the penn world table. *American economic review*, 105(10):3150–3182.

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.

Guner, N. and Ruggieri, A. (2022). Misallocation and inequality. *CEPR Discussion Paper No. DP17113*.

Hall, R. E. and Milgrom, P. R. (2008). The limited influence of unemployment on the wage bargain. *American economic review*, 98(4):1653–1674.

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.

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.

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

OECD (2019). *Investing in Youth: Peru*.

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

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

Pijoan-Mas, J. and Roldan-Blanco, P. (2022). Dual labor markets and the equilibrium distribution of firms. *Centre for Economic Policy Research*.

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.

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.

Soares, S. (2018). Minimum wage: global challenges and perspectives. *Policy in Focus*.

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.

Online Appendix

A Data and supplementary evidence

A.1 Cross-country data

Figure A.1 reports the list of countries and years covered by our dataset.

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

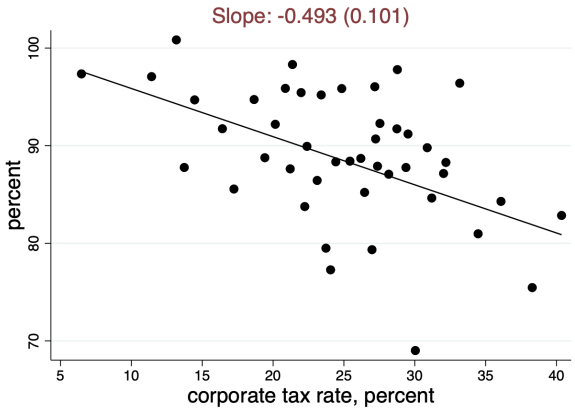
Notes: This table reports countries and years covered by our cross-country dataset.

A.2 Supplementary cross-country evidence

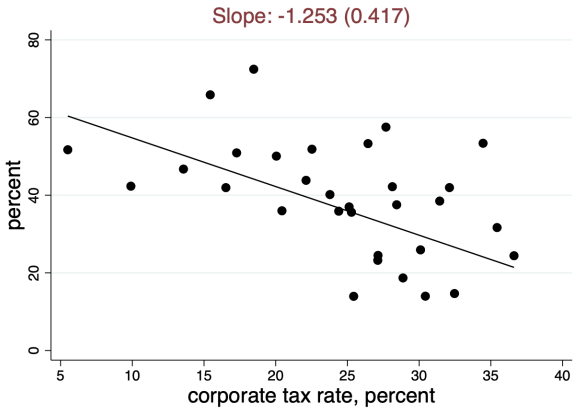
In this section of the Online Appendix, we show that the cross-country labor market outcomes patterns are robust. Figure A.1 scatters two alternative measures of informality over corporate income tax rates across countries.

Figure A.1: Informality and corporate income taxes

(A) Firms registered when started operations



(B) Workers covered by social security



Notes: Corporate tax rates refer to the standard statutory corporate income tax rates levied on domestic businesses. Source: ILO-stat, World Bank Enterprise Survey, Tax Foundation, and authors' calculation.

The first measure refers to the percentage of formally registered firms that started their operation as formal (Figure A.1.A). The second refers to the percentage of employed workers covered by social security in case of an injury at work (Figure A.1.B).

Both measures are taken from the World Bank Enterprise Survey Database. As we move from low- to high-corporate income tax economies the share of formal firms that were formal at the start of their operation significantly decline. The same pattern is followed by the share of employed workers covered by social security.

As a further robustness check, we show that cross-country labor market patterns are not accounted for by country-specific unobserved heterogeneity. To do so we estimate the following cross-country regression equation:

$$y_{it} = \alpha + \mu_t + \mu_{c(i)} + t \times \mu_{c(i)} + \beta\tau_{it} + \epsilon_{it}$$

where the dependent variable, y_{it}, denotes either a measure of informality, unemployment rate, or aggregate productivity for country i at time t, μ_t are time fixed-effects, and $\mu_{c(i)}$ denotes country unobserved-heterogeneity, t denotes time, τ_{it} denotes corporate income tax rate and ϵ_{it} is an error term.

We control for country-specific unobserved heterogeneity by clustering countries into selected groups, based on the average GDP per capita of each country in our sample. We identify again 10 clusters using a *k-mean* clustering algorithm.

Table A.2: Labor market outcome across countries

	Informal wage employment			Unemployment rate		
	(1)	(2)	(3)	(1)	(2)	(3)
Corporate tax rate, τ_{it}	0.377***	0.394***	0.383***	-0.198***	-0.203***	-0.202***
	(0.111)	(0.112)	(0.115)	(0.0294)	(0.0299)	(0.0299)
Observations	326	326	326	326	326	326
R-squared	0.372	0.390	0.417	0.229	0.240	0.251
GDP p.c. cluster FE	✓	✓	✓	✓	✓	✓
Time FE		✓	✓		✓	✓
GDP p.c. cluster trend			✓			✓

	Real GDP p.w. (1000 USD)			Real TFP (US=100)		
	(1)	(2)	(3)	(1)	(2)	(3)
Corporate tax rate, τ_{it}	-0.985***	-0.985***	-0.984***	-1.052***	-1.055***	-1.064***
	(0.0798)	(0.0799)	(0.0804)	(0.111)	(0.112)	(0.109)
Observations	326	326	326	326	326	326
R-squared	0.173	0.173	0.176	0.187	0.191	0.197
GDP p.c. cluster FE	✓	✓	✓	✓	✓	✓
Time FE		✓	✓		✓	✓
GDP p.c. cluster trend			✓			✓

Notes: Robust standard errors (in parenthesis) are clustered at country level. *p≤0.10, **p≤0.05, ***p≤0.01.
Source: ILO-stat, Tax Foundation, World Bank Enterprise Survey, and authors' calculation.

Tables A.2 and A.3 report the estimation outcomes of this exercise.

Table A.3: Labor market outcome across countries

	Firms registered when starting operation, %			Workers covered by social security, %		
	(1)	(2)	(3)	(1)	(2)	(3)
Corporate tax rate, τ_{it}	-0.483***	-0.546***	-0.569***	-0.448***	-0.363***	-0.372***
	(0.211)	(0.233)	(0.259)	(0.0850)	(0.101)	(0.111)
Observations	140	140	140	202	202	202
R-squared	0.667	0.700	0.710	0.262	0.315	0.327
GDP p.c. cluster FE	✓	✓	✓	✓	✓	✓
Time FE		✓	✓		✓	✓
GDP p.c. cluster trend			✓			✓

Notes: Robust standard errors (in parenthesis) are clustered at country level. *p≤0.10, **p≤0.05, ***p≤0.01.
Source: ILO-stat, Tax Foundation, World Bank Enterprise Survey, and authors' calculation.

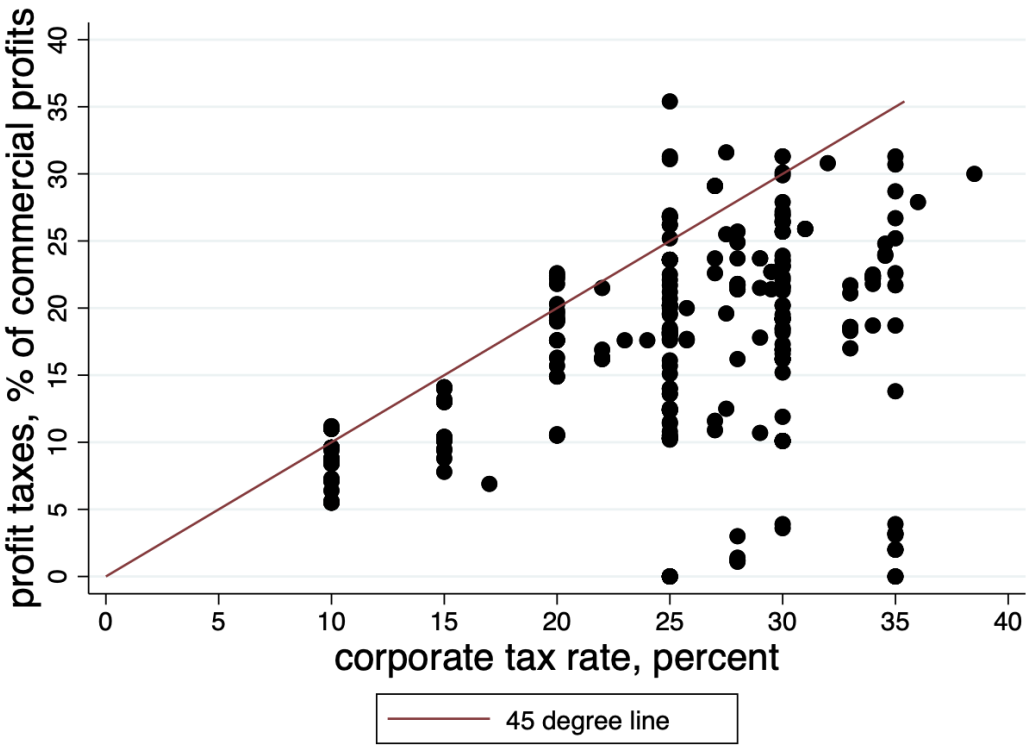
Specifically, Table A.2 reports the estimates for informal wage employment, unemployment rate, real GDP per worker, and TFP. Table A.3 reports the estimates for the alternative two measures of informality described above. Controlling for unobserved heterogeneity across countries in the form of GDP per capita cluster fixed effect, time fixed effects, and cluster-specific trend does not alter how informality rate, unemployment rate, and aggregate productivity change with corporate income tax rates across

countries.

Finally, we check the robustness of our findings by using a measure of average profit tax paid by registered companies instead of the statutory tax rate on corporate income. We source the former variable from the World Bank Doing Business project dataset, where it is reported as the total tax rate payable by limited liability businesses divided by their total commercial profits.

The average profit tax might differ from the statutory tax rate, which is the factor applied to the tax base because it measures all taxes and contributions that are government-mandated at any level - federal, state, or local. Figure A.2 scatters the average profit tax against the statutory corporate income tax rate for each country in the sample.

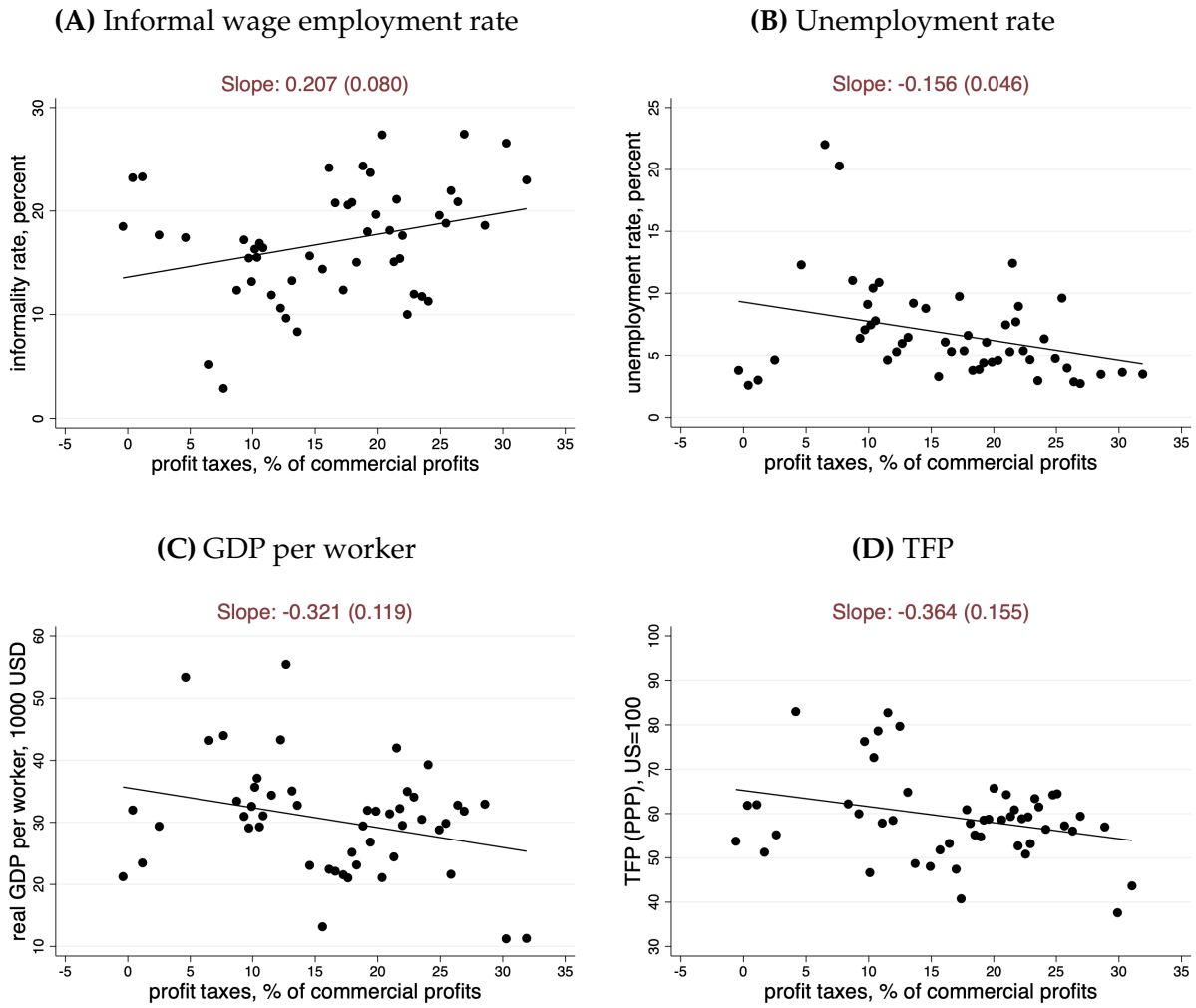
Figure A.2: Statutory corporate tax rate VS average profit tax



Notes: Corporate tax rates refer to the standard statutory corporate income tax rates levied on domestic businesses. Source: World Bank Doing Business, Tax Foundation, and authors' calculation.

Three main considerations are in order. First, all the dots in the scatter lie below the 45-degree line, which implies that the statutory tax rate is higher than the average profit taxes paid by companies. This is because tax deductions available to companies - like for instance those for recorded depreciation of tangible fixed assets and investment property, are not accounted for by the statutory rates. Second, there is a large dispersion in average profit taxes for a given statutory rate, reflecting heterogeneity in tax deductions across countries with the same corporate tax rate. Finally, the correla-

Figure A.3: Labor market outcomes and average profit taxes



Notes: Informal employment is expressed as a percent of total employment and comprises persons who in their main or secondary jobs were 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 coverage by the social security system, entitlement to paid annual or sick leave, or written employment contracts. The unemployment rate is reported in percent of the labor force. GDP per worker is measured in 2017 USD and expressed in 1000 USD. Profit taxes refer to the average amount of taxes on profits paid by the business as a percent of total commercial profits. Source: ILO-stat, World Bank Doing Business, and authors' calculation.

tion between the average and statutory rate is positive and significant. One p.p. higher statutory rate is associated with 0.43 p.p. higher average profit tax paid by companies.

In Figure A.3 we scatter the cross-country informality rate (panel A), unemployment rate (panel B), real GDP per worker (panel C), and total factor productivity (panel D), against the average profit taxes. In red we report the estimated slope of each relationship and robust standard errors, clustered at the country level.

Using the average profit tax paid does not alter the main findings. Countries with more burdensome corporate taxation have higher informality rates, lower unemployment rates, and lower aggregate productivity, both in terms of real GDP per worker

and aggregate TFP.

A.2.1 Additional results

Table A.4 reports the estimation outcomes from regressing self-employment rate on corporate income tax rates (columns 1 to 3), controlling for unobserved heterogeneity across countries, in the form of GDP per capita cluster fixed effects, time fixed effects and cluster-specific time trend. Data for self-employment is taken from the World-Bank database and refers to the share of own-account workers out of total employment.

Table A.4: Self-employment across countries

	Self-employment rate, %		
	(1)	(2)	(3)
Corporate tax rate, τ_{it}	0.131	0.106	0.100
	(0.105)	(0.106)	(0.108)
Observations	326	326	326
R-squared	0.741	0.746	0.752
GDP p.c. cluster FE	✓	✓	✓
Time FE		✓	✓
GDP p.c. cluster trend			✓

Notes: Robust standard errors (in parenthesis) are clustered at country level. *p≤0.10, **p≤0.05, ***p≤0.01. Source: ILO-stat, Tax Foundation, World Bank, and authors' calculation.

Differences in corporate income tax rates across countries are not associated with different rates of self-employment. The estimate coefficients are small in magnitude and not significant at 10 percent level. This result, together with the evidence provided in Section 2, suggests that changes in corporate tax rates are correlated to informality rates across countries only through differences in the composition of wage employment.

A.3 Workers and Firms in Peru

Tables A.5, A.6, and A.7 report summary statistics for workers, formal and informal firms in Peru, respectively.

Table A.5: 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 A.6: 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

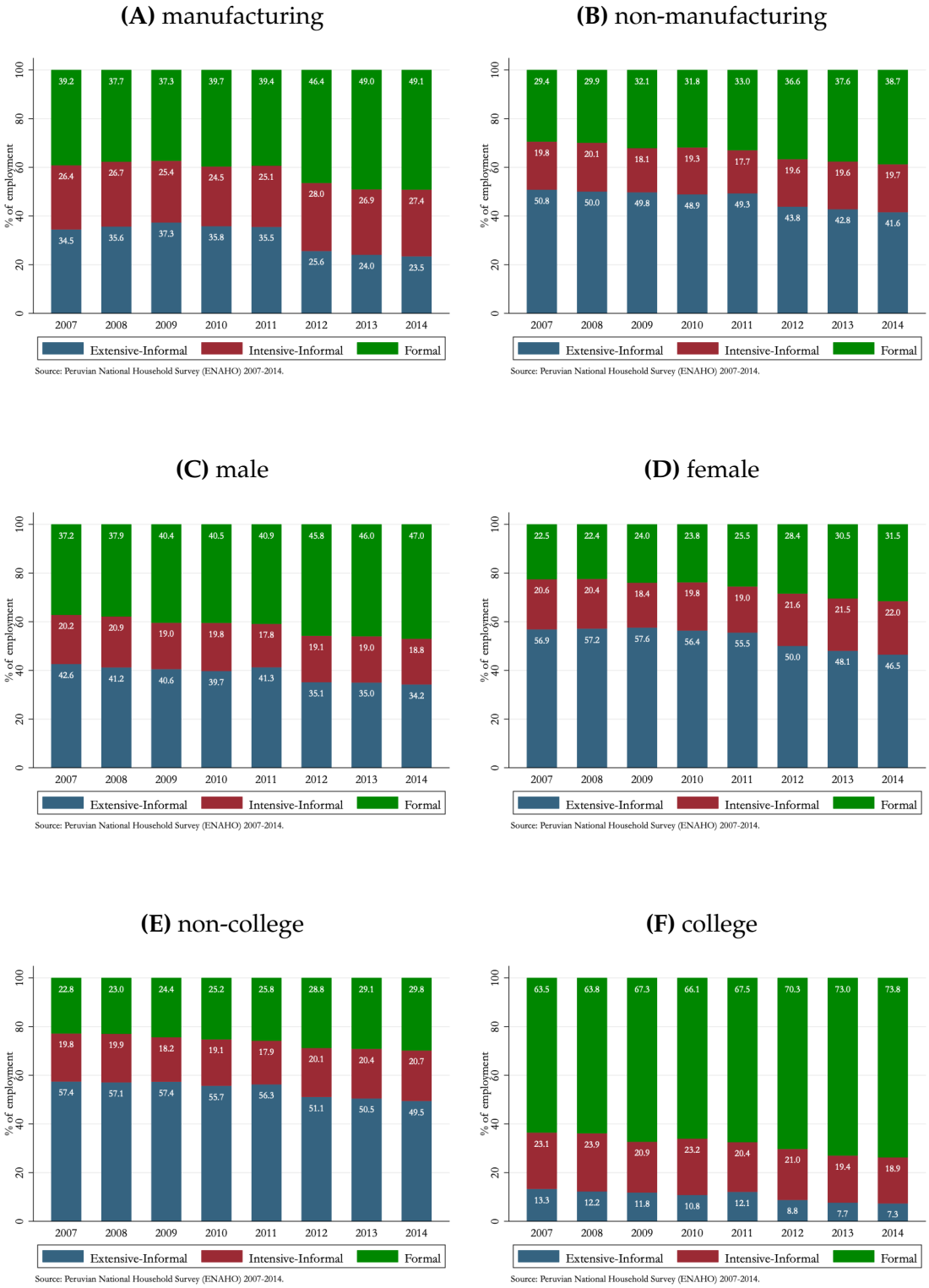
Notes: This table reports selected statistics for formal firms in Peru. Sales and average payroll are reported in 2010 Peruvian local currency (Nuevo sol). Source: WB-ES, 2006, 2010, 2017

Table A.7: 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

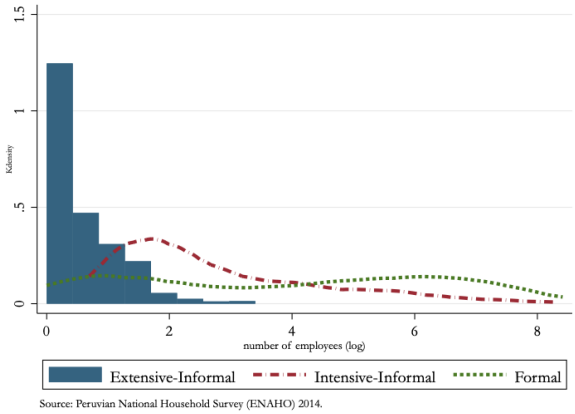
Figure A.4: Employment composition



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

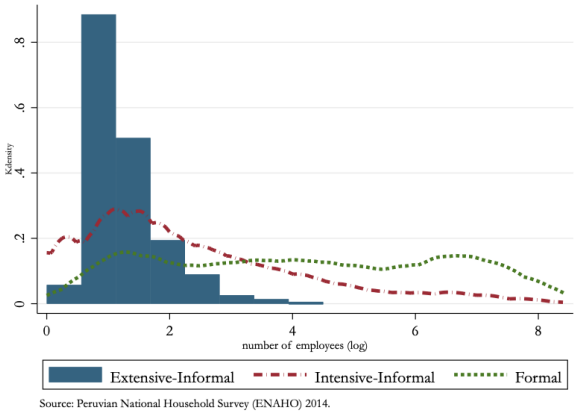
Figure A.5: Firm size distribution across workers

(A) manufacturing



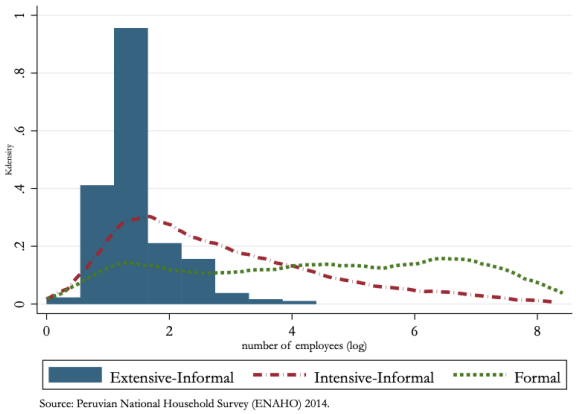
Source: Peruvian National Household Survey (ENAHO) 2014.

(B) non-manufacturing



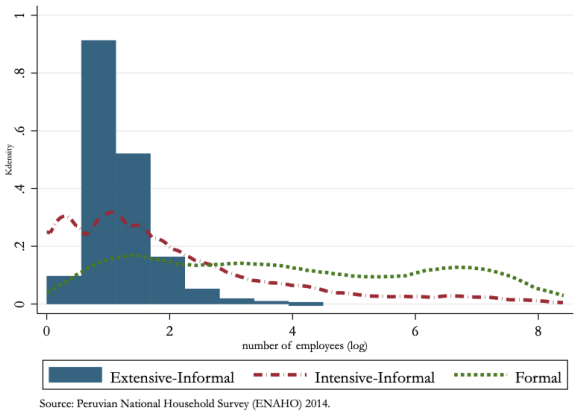
Source: Peruvian National Household Survey (ENAHO) 2014.

(C) male



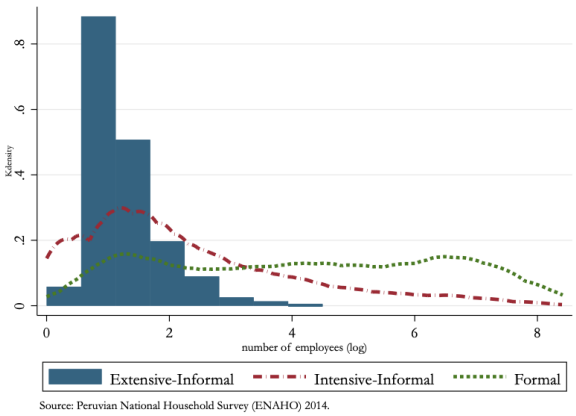
Source: Peruvian National Household Survey (ENAHO) 2014.

(D) female



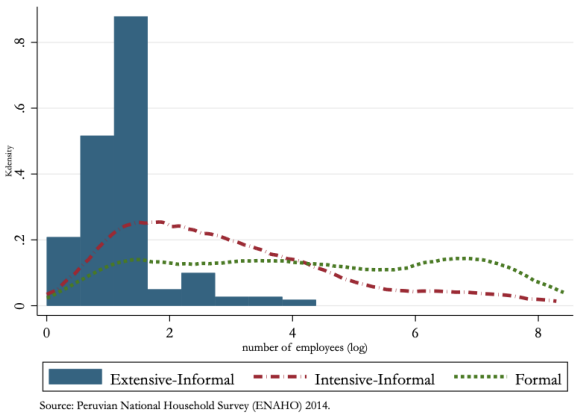
Source: Peruvian National Household Survey (ENAHO) 2014.

(E) non-college



Source: Peruvian National Household Survey (ENAHO) 2014.

(F) college

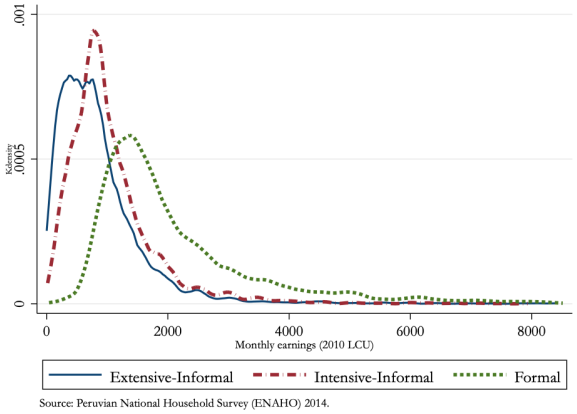


Source: Peruvian National Household Survey (ENAHO) 2014.

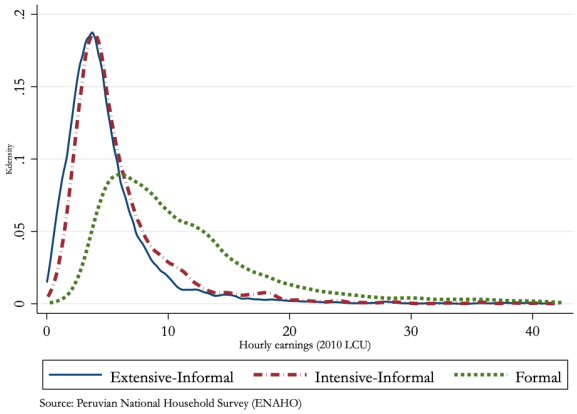
Notes: This figure reports the distribution of formal and informal workers over the size of their employers, separately for manufacturing and non-manufacturing workers, males and females, college and non-college workers. Source: ENAHO and authors' calculation

Figure A.6: Earnings distribution

(A) monthly earnings



(B) hourly earnings



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

Table A.8: 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)
$\mathbf{1}[\text{Intensive Informal}]_{it}$		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. $\mathbf{1}[\text{Intensive Informal}]_{it}$ 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

B Model

B.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, $\tilde{\phi}$ and the job filling probability, ϕ, 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, the share of unregistered firms s_i, distributions of firms across productivity values and size, $\psi_i(z, \ell_i)$ and $\psi_f(z, \ell_i, \ell_f)$, 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 the entry and registration decisions and exit shocks. Since all entering firms start the interim period with a productivity draw from $\psi_z(z)$ and with a registration cost drawn from $\psi_c(c_f)$, we can measure formal and informal firms in their respective states as

$$\psi_i(z, \ell_i) = \delta_i \int_{z,c_f} [\mathbf{1}^e(z)][1 - \mathbf{1}^f(z, c_f)]\psi_z(z)\psi_c(c_f)dc_f + (1 - \delta_i)\psi_i(z, \ell_i) \quad (29)$$

and

$$\psi_f(z, \ell_i, \ell_f) = \delta_f \int_{z,c_f} [\mathbf{1}^e(z)][\mathbf{1}^f(z, c_f)]\psi_z(z)\psi_c(c_f)dc_f + (1 - \delta_f)\psi_f(z, \ell_i, \ell_f) \quad (30)$$

2. **Firm dynamics.** In equilibrium, the measure of entrants M^e is equal to the fraction of formal and informal firms that turn over every period, i.e.

$$\begin{aligned} \frac{M^e}{M} = {} & \delta_i \int_{z,c_f} [\mathbf{1}^e(z)][1 - \mathbf{1}^f(z, c_f)]\psi_z(z)\psi_c(c_f)dc_f \\ & + \delta_f \int_{z,c_f} [\mathbf{1}^e(z)][\mathbf{1}^f(z, c_f)]\psi_z(z)\psi_c(c_f)dc_f \end{aligned} \quad (31)$$

3. **Equilibrium in the outside sector**. Demand for the outside sector goods 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, earnings of workers in the outside sector, and transfer to the unem-

ployed i.e.

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

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

$$A_o L_o = (1 - \alpha) I + M^e (c_e + \bar{c}_f) + M \bar{c}_v \tag{32}$$

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.

4. **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

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

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

$$\tilde{\phi} U = \mu_d L_e \tag{33}$$

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:

$$1 = L_o + L_u + L_e$$

5. **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.

$$\mathcal{J}^n = \mathcal{J}^s = \mathcal{J}^o = \frac{1+r}{r} w_o \tag{34}$$

B.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 the workers' probability of finding 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 the policy functions for hiring 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.

B.3 Estimation algorithm

In the calibration algorithm, we exploit the worker's no-arbitrage condition between the value of searching and the value of being employed in the outside sector

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

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 parameter 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 the workers' probability of finding 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 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.

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

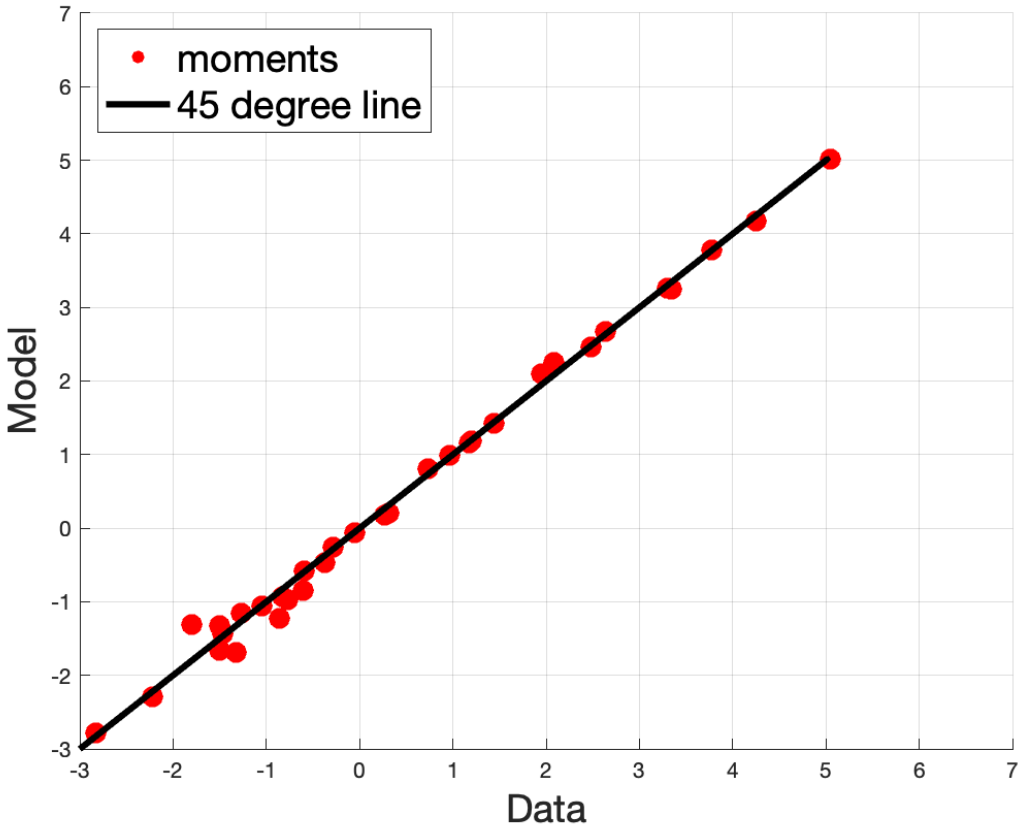
- Use ϕ^0 and $\tilde{\phi}^0$, the definition of matching function, and the market clearing for the outside sector, to obtain a solution for the endogenous measure of incumbent firms $M(\vartheta^0)$ and a measure of workers searching for wage and salary jobs, $U(\vartheta^0)$.
- Use the labor market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



C The Role of Aggregate Productivity

Differences in technologies are an alternative potential candidate to explain how informality and unemployment vary across countries. Richer countries adopt more productive technologies and have higher production efficiency than poorer countries (Caselli, 2016). As robustness, in this section, we explore to which extent this force vis-à-vis corporate income taxes can account for the cross-country patterns documented in Section 2.

The experiment we conduct goes as follows. We use aggregate productivity A to proxy for differences in technologies across countries and re-calibrate it to match the average unemployment rate observed in low- and high-tax economies within our cross-country dataset. Low- (high-) tax economies correspond to economies with a 10% (35%) corporate income tax. While doing so, we adjust the entry cost, c_e, and the productivity of the outside sector, A_o by the same factor of A.

Table C.1: Corporate taxes versus productivity changes

	Low-tax high-productivity (1)	High-tax low-productivity (2)	Low-tax low-productivity (3)	% explained by productivity (4)
Corporate income tax rate, τ_y	10%	35%	10%	-
Aggregate productivity, A	1.202	0.997	0.997	-
Productivity in outside sector, A_o	1264.20	1048.76	1048.76	-
Entry cost, c_e	4606.09	3821.16	3821.16	-
Informality rate	0.153	0.313	0.138	-10.7%
Unemployment rate	0.189	0.033	0.055	85.9%
Real GDP per worker	1.443	0.916	1.205	45.2%
Self-employment rate	0.463	0.589	0.542	-

Notes: Columns (1) and (2) report the calibrated values of aggregate productivity, A, productivity in the outside sector A_o, and entry cost c_e, and selected labor market outcomes, for counterfactual low- and high-tax economies, respectively. Column (3) reports labor market outcomes in a counterfactual economy with low corporate income tax and low aggregate productivity. Entries in column (4) refer to the percent change in labor market outcome explained by changes in tax rates and are computed as the difference between columns (1) and (3) over the difference between columns (1) and (2). Real GDP per capita is expressed as a ratio of the baseline economy (Peru).

Columns (1) and (2) in Table C.1 report the outcomes of this calibration. Aggregate productivity is recalibrated to a value of 1.202 in low-tax economies (column 1), and to 0.995 in high-tax economies (column 2), ensuring the unemployment rate is equal to the averages observed in the data, 18.9%, and 3.3% respectively. We then construct a third counterfactual economy by applying a low aggregate productivity level to a low-tax environment. This allows us to separately identify the impact that aggregate productivity has on labor market outcomes. Column 3 in Table C.1 reports the results of this experiment. A comparison between columns (1) and (3) versus columns (1) and (2) suggests that changes in aggregate productivity can explain up to 86% of differences in the unemployment rate, up to 45% of differences in real GDP per worker, and cannot explain differences in informality rate. This confirms that a large fraction of cross-country patterns can still be attributed to differences in corporate taxes.

D Expected cost for informal firms

Table D.1 reports firm-level and aggregate outcomes for different counterfactual values of γ_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 informal wage employment (as a share of total wage employment), the measure of firms, the labor market tightness, the unemployment rate, the real average wage, and the real GDP per employee.

Table D.1: Expected informality cost for informal firms

Informality cost, γ_0	13%	16%	18%*	22%	26%
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					
Informal wage employment	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. The average wage is expressed as a function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline. The informality cost per employee, γ_0, is reported as a percent of the average earnings for a firm with an average productivity $\bar{z}$.

The counterfactual outcomes following an increase in the expected informality cost for informal firms mirror 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 informal wage employment 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 a higher unemployment rate. Quantitatively, doubling the expected cost of informality (from 33.41 to 66.83, columns 1 and 5 of Table D.1) 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).

E Expected cost of hiring informal workers

Table E.1 reports the firm-level and aggregate counterfactual outcomes for different values of γ_2. The implications of this policy are qualitatively different from changing the expected cost to informal firms.

Table E.1: Expected cost of hiring informal workers for registered firms

Informality cost, γ_2	18%	27%	54%	107%	143%
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					
Informal wage employment	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 a function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline. The informality cost per employee, γ_2, is reported as a percent of the average earnings for a firm with an average productivity, $\bar{z}$, and 50% informal workforce.

First, the 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 expand 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 effect 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 start 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 rate, as a result of higher informal jobs in registered business.

F Payroll tax reform

Table F.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 informal wage employment (as a share of total wage employment), the measure of firms, the labor market tightness, the unemployment rate, the real average wage, and the real GDP per employee.

Table F.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					
Informal wage employment	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 a function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

By targeting informality along the intensive margin, changes in payroll taxes could lead to qualitatively similar dynamics in the labor markets as those obtained by changing the expected cost of hiring informal workers for formal firms. On the other hand, in our quantitative exercise, a reform that eliminates payroll taxes is not strong enough to make job formalization non-monotonic.

G Labor market policy interventions

Transfer to the unemployed. Table G.1 reports a set of counterfactual outcomes for different transfers to the unemployed, 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 informal wage employment (as a share of wage employment), the measure of firms, the labor market tightness, the unemployment rate, the real average wage, and the real GDP per employee.

Table G.1: Transfers to the unemployed

Transfers, 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					
Informal wage employment	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. The average wage is expressed as a function of the earnings in the outside sector. Real GDP per worker is expressed in terms of baseline.

We express the transfer 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 the amount transferred to the unemployed has the same qualitative implications as a reduction in corporate income taxes, although firm-level and aggregate outcomes change by a much smaller magnitude.

Minimum wage. Table G.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:

$$w_{ff}(z, \ell_i, \ell_f) = \max\left\{\underline{w}, \frac{(1-\zeta_f)}{(1+\zeta_f[\tau_w-\tau_y(1+\tau_w)])}b + \frac{\zeta_f(1-\tau_y)}{(1+\zeta_f[\tau_w-\tau_y(1+\tau_w)])}\frac{R_f(z,\ell_i,\ell_f)}{\ell_i+\ell_f}\right\}$$

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

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

Minimum wage, $\underline{w}$	0	$w_o{}^*$	$1.5w_o$	$2w_o$	$2.5w_o$	$3w_o$
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						
Informal wage employment	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. The average wage is expressed as a 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, increases the number of employers per working population, and raises the probability of finding a wage and salary job for workers.