Labor Market Power, Self-Employment, and Development*

Francesco Amodio† Pamela Medina‡ Monica Morlacco§

July 18, 2025

Abstract

This paper shows that self-employment shapes labor market power in low-income countries, with implications for industrial development. Using Peruvian data, we find that wage-setting power increases with employer concentration but less so where self-employment is more prevalent. A general equilibrium model shows that in oligopsonistic labor markets, self-employment raises the supply elasticity of wage labor, weakening employer market power. However, by the same mechanism, pro-competitive policies aimed at expanding wage employment and reducing reliance on self-employment may unintentionally strengthen labor market power, undermining their objectives.

Keywords: labor market power, oligopsony, self-employment, sorting, industrial development.

JEL Codes: J2, J3, J42, L10, O14, O54.

*We thank Jacob Schmidt and Paola Marino Moreno for excellent research assistance. For their insightful comments and suggestions, we thank Vittorio Bassi, Robin Burgess, Paula Bustos, Stefano Caria, Arin Dube, Rafael Dix-Carneiro, Marcela Eslava, Mayara Felix, Fernando Fernandez, Nicolas Gendron-Carrier, Kyle Herkenhoff, Jonas Hjort, Rocco Macchiavello, Isabela Manelici, Marti Mestieri, Suresh Naidu, Tommaso Porzio, Todd Schoellman, Gabriel Ulyssea, Jose Vasquez, Daniel Xu, and all participants to seminars at Bologna, BU, CEMFI, CREST, Dartmouth, Duke, EIEF, EUI, JILAEE-UChicago, Loughborough, LSE, Manchester, McGill, McMaster, Milan, Minneapolis Fed, NCID-Navarra, Nottingham, Notre Dame, Richmond Fed, SFU, Stockholm, UAB, ULosAndes, UBC, UChile, UCLA, UC Riverside, UC San Diego, UMontevideo, UofT, UPF, USC, Warwick, Western, World Bank, and many other conferences and workshops. Errors remain our own. We gratefully acknowledge financial support from the CEPR/DFID PEDL initiative (ERG 6500), McGill University and the Social Sciences and Humanities Research Council (SSHRC) of Canada - Insight Grant Program.

†francesco.amodio@mcgill.ca, McGill University and IZA, Department of Economics and ISID, 855 Sherbrooke St. West, Montreal, QC, H3A 2T7.

‡pamela.medinaquispe@rotman.utoronto.ca, University of Toronto, UTSC and Rotman School of Management, 105 St George St, Toronto, ON M5S 3E6.

§morlacco@usc.edu, University of Southern California, Department of Economics, 3620 S. Vermont Avenue, Los Angeles, CA 90089-1422.

1 Introduction

A longstanding view in economic development holds that industrialization promotes economic growth in poorer countries, with self-employment declining as modern industrial firms expand (Lewis, 1954; Rauch, 1991). As a result, increasing wage employment in the manufacturing sector has become a cornerstone of industrial development strategies, prompting significant investments and targeted policy initiatives (UN General Assembly, 2015). Despite these efforts, manufacturing self-employment remains widespread in emerging economies (La Porta and Shleifer, 2014; Gollin, 2008; Poschke, 2022), and employment at large firms has stagnated (Hsieh and Olken, 2014; McMillan and Zeufack, 2022), even as GDP per capita increases.

Labor market structure is a potentially important yet often overlooked factor influencing these outcomes. Barriers such as high entry costs, a shortage of skilled labor, and inadequate infrastructure often lead to employment being concentrated among a few dominant firms (Djankov et al., 2002; Rud and Trapeznikova, 2021; Hjort, Malmberg and Schoellman, 2022), which can reduce job opportunities and suppress wages to maximize profits. When wage jobs are scarce and unattractive, self-employment becomes a valuable alternative for workers (Blattman and Dercon, 2018; Breza, Kaur and Shamdasani, 2021), acting as a buffer against employers' market power. However, by the same mechanism, policies that seek to expand wage employment and reduce reliance on self-employment may inadvertently strengthen employers' wage-setting power, making them less effective than intended.

This paper argues that understanding the interplay between labor market power and self-employment is essential to explaining persistently high rates of self-employment and why industrial development policies often fall short of their objectives. To support these claims, we provide new evidence from Peru, an original theoretical framework, and counterfactual policy experiments. Peru is a compelling case study due to its high employer concentration, high self-employment rates, and frequent worker transitions into and out of self-employment–characteristics shared by many emerging economies.

We begin by showing that labor market power is substantial across Peruvian manufacturing local labor markets.[1] We measure it as the inverse elasticity of the labor supply curve faced by individual firms, which directly captures their ability to set wages (Manning, 2003). For estimation, we employ an instrumental variable strategy that leverages the staggered roll-out of a rural electrification program across provinces and its differential impact on firms with high versus low ex-ante constraints in accessing electricity, allowing us to isolate exogenous variation in labor demand across firms within a local labor market.

Our estimates indicate that the average firm-level inverse labor supply elasticity is positive and significant, indicating wage-setting power. The implied average wage markdown is 1.42, meaning that workers receive approximately 70 cents for each additional dollar they gener-

[1] We define a local labor market as the combination of a 2-digit industry and a commuting zone. See Section 2 for further details and motivation for this definition.

ate. Markdowns rise with market concentration, consistent with oligopsony power, though this gradient is weaker in markets with higher self-employment rates. We estimate the highest markdowns in highly concentrated markets with a low prevalence of self-employment, where workers receive only 57 cents per additional dollar they produce.

While these patterns indicate meaningful differences across markets, they do not necessarily reflect variations in underlying structural fundamentals. Our empirical strategy estimates *reduced-form* labor supply elasticities, which do not directly map to *structural* elasticities as they may reflect competitors' equilibrium responses (Berger, Herkenhoff and Mongey, 2022). Disentangling equilibrium effects from structural features requires a theoretical model and is essential to fully characterize labor market dynamics and evaluate the scope and effectiveness of policy interventions.

We thus develop a general equilibrium model of Peruvian manufacturing labor markets in which employer concentration, self-employment rates, and labor market power are jointly determined. The model is built on two key features. First, it incorporates *oligopsony power*, where a finite number of heterogeneous firms in each local labor market internalize their influence on wages and make strategic employment decisions. Second, it embeds Roy's (1951) *self-selection* framework, where heterogeneous workers choose between wage employment and self-employment based on expected earnings.

This framework captures both *demand-* and *supply-side* determinants of labor market power. This is reflected in the average wage markdown in a local labor market, which depends on two key endogenous objects. The first is the Herfindahl-Hirschman Index (HHI) for employment, a measure of employer concentration that captures oligopsony power. The second is the aggregate supply elasticity of wage labor, which counteracts the employers' wage-setting power by increasing worker responsiveness to wage changes. These mechanisms interact in equilibrium: as concentration rises, markdowns increase, depressing wages and driving more workers into self-employment. This reallocation, in turn, raises labor supply elasticity, partially offsetting labor market power.

A key takeaway from our model is that self-employment serves a *dual role* in shaping labor market power dynamics. On the one hand, it provides a fallback option for workers, constraining firms' wage-setting power when wage opportunities are limited. On the other hand, when wage employment becomes more attractive, reduced reliance on self-employment reinforces oligopsony power, making it harder for industrial policies to increase employment and wages. Through counterfactual analyses, we show that the *variable elasticity channel* is quantitatively essential in explaining the limited effectiveness of industrial development policies.

A crucial factor influencing labor supply elasticity and its higher-order moments is the distribution of worker ability. Our identification strategy relies on parametric restrictions, specifically the assumption of log-normality. Standard in empirical Roy models (French and Taber, 2011), this assumption is necessary in our setting for both estimation and tractability. It enables transparent estimation using worker-level data while providing the structure needed to embed the Roy block into our general equilibrium model and conduct counterfactual analyses. While our approach imposes a parametric structure, the log-normal distribution offers greater flexibility than the commonly used Fréchet, particularly by allowing comparative and absolute advantage to vary independently.[2]

Given the ability distribution, we employ a Method of Simulated Moments to estimate the remaining model parameters, including firm productivity, entry costs, average worker skills, and their variation across markets. We discipline them by targeting moments from the cross-sectional distributions of concentration, employment shares, earnings, and their correlations.

We validate the model by simulating the electrification program from our reduced-form analysis and assessing its ability to replicate the observed patterns of labor market power. First, we identify treated firms based on the observed correlation between the electricity wedge and productivity in the data. We then calibrate the productivity shock induced by electrification, apply it to the treated firms in the model, and simulate the resulting labor market responses. Specifically, we compute the model-implied reduced-form inverse labor supply elasticities as the ratio of the (log) wage to (log) employment responses of treated firms, aligning with the local average treatment effect (LATE) estimates from the reduced-form analysis.

The model closely matches the initial reduced-form estimates of labor market power, predicting an average inverse labor supply elasticity of 0.36, compared to 0.42 in the data, and accurately capturing the relationship between employer concentration and wage markdowns. It also replicates the mitigating effect of self-employment, revealing that the highest wage-setting power occurs in highly concentrated markets with low self-employment, with an estimated inverse elasticity of 0.62, compared to 0.75 in the data.

The model also allows for a comparison between reduced-form and structural inverse labor supply elasticities. Structural estimates tend to be more muted, as they account for the equilibrium responses of competing firms. However, the bias in reduced-form estimates does not account for the differences observed in market concentration and self-employment rates, as no consistent pattern emerges when comparing structural and reduced-form labor market power across these dimensions.

With the estimated model, we conduct two sets of counterfactual experiments. First, we assess the impact of labor market power on labor market outcomes by comparing our baseline economy to one in which employers are wage-takers. Without labor market power, the average wage employment share across markets rises from 66% to 77%. Average earnings increase by 31% and 27% in the wage and self-employment sectors, respectively, suggesting that labor market power compresses the earnings gap between wage workers and the self-employed.

Second, we evaluate how labor market power influences the effectiveness of industrial development policies to expand wage employment. We consider policies that enhance firm productivity through market integration and infrastructure improvements (Volpe Martincus, Carballo and Cusolito, 2017; Fiorini, Sanfilippo and Sundaram, 2021), lower fixed entry costs for em-

[2] See Adão (2016) and Amodio, Alvarez-Cuadrado and Poschke (2020) for a discussion.

ployers by streamlining business registration (Kaplan, Piedra and Seira, 2011; Bruhn, 2011), and improve worker skills through off- and on-the-job training programs (McKenzie, 2017; Alfonsi et al., 2020). To quantify their effects on labor market outcomes, we calibrate the policy shocks using empirical evidence from past implementations in Peru and Mexico and their estimated reduced-form impacts.

We find that the impact of industrial policies varies significantly across markets, mainly due to differences in labor market power. Changes in markdowns account for up to 99% of the variation in the policy's impact on wages and 88% of its effect on the wage employment share across markets, underscoring that a policy is only effective if it directly addresses labor market imperfections. These insights are crucial for policymakers aiming to design effective interventions for industrialization and inclusive economic growth.

Related Literature This paper contributes to several strands of the literature. First, we add to the work on informal self-employment in low-income countries. The traditional "dual" view posits that large formal firms and informal micro-enterprises operate in entirely separate economic spaces.[3] Our study challenges this perspective by emphasizing the role of worker sorting and oligopsony power in shaping employment outcomes in emerging economies while offering an explanation for the persistently high prevalence of self-employment in manufacturing. In doing so, we align with Maloney (1999), Ulyssea (2018), and Donovan, Lu and Schoellman (2023), among others, who show that formal and informal firms coexist within the same local labor markets, with frequent worker transitions between the two sectors.

Second, our paper contributes to the growing literature on labor market power by integrating self-employment into its analysis, highlighting its role in shaping labor supply elasticity to the wage employment sector and influencing policy effectiveness. This focus is particularly relevant for low- and middle-income countries, where self-employment rates are higher, and research on labor market power remains relatively limited.[4] Our theory provides a novel micro-foundation for the aggregate supply of wage labor, driven by the self-selection of heterogeneous workers, thereby explicitly incorporating supply-side determinants of wage-setting power. In doing so, we extend recent theories of monopsony power (Burdett and Mortensen, 1998; Lamadon, Mogstad and Setzler, 2019; Kroft et al., 2020) and oligopsony power (Berger, Herkenhoff and Mongey, 2022). Our findings suggest that once worker sorting is accounted for, employer concentration exhibits a *non-linear* relationship with labor market power, pro-

[3] Early contributions include Lewis (1954), Harris and Todaro (1970), and Rauch (1991). See also La Porta and Shleifer (2014) for a review and Eslava et al. (2023, 2024) for evidence on Latin America.

[4] Amodio and De Roux (2022) document substantial labor market power in Colombian manufacturing, while Felix (2022) finds similarly high labor market power in Brazil, with only limited effects of the 1990s trade liberalization. In Costa Rica, Alfaro-Ureña, Manelici and Vasquez (2021) report minor wage effects from multinational expansion, suggesting low labor market power due to high worker mobility. Beyond Latin America, studies identify significant monopsony power in India (Brooks et al., 2021), China (Muralidharan, Niehaus and Sukhtankar, 2023), and South Africa (Bassier, 2023). Finally, Armangué-Jubert, Guner and Ruggieri (2024) and Amodio et al. (2024) leverage World Bank Enterprise Survey data to analyze labor market power and its relationship with development across multiple low- and middle-income countries.

viding a potential explanation for the mixed results in the empirical literature.[5] Our theory also aligns with the findings in Felix (2022), which shows that firms in Brazilian markets with higher self-employment rates face *more elastic* labor supply curves.

Third, our analysis offers novel insights into the determinants of the earnings gap between self-employed and wage workers. In developing countries, self-employment accounts for nearly half of the workforce, yet wage employment generally pays more (Fields, 2012). We quantify the role of labor market power for the size of the gap and how it changes with policy.

Finally, our work speaks to the extensive literature on informality in low-income countries.[6] Both Dix-Carneiro and Kovak (2019) and Ponczek and Ulyssea (2021) argue that informality acts as an "unemployment buffer" by reducing trade-induced adjustment costs in the labor market. Yet, Dix-Carneiro et al. (2024) shows that in the event of a negative economic shock, welfare declines more when informality rates are high. Our analysis adopts the notion of informal self-employment as a potential outside option for workers. It demonstrates that it plays a similar role in the presence of labor market power.

The remainder of the paper is organized as follows. Section 2 introduces the data and presents the empirical facts. The model and its properties are presented in Section 3, while Section 4 discusses the model estimation procedure and results. Section 5 presents the counterfactual policy analyses. Section 6 concludes.

2 Data and Facts

The empirical analysis relies on two main datasets on firms and workers. The first dataset is the Peruvian Annual Economic Survey (*Encuesta Económica Anual*, EEA), a nationwide firm-level survey conducted annually by the national statistical agency (*Instituto Nacional de Estadística e Informática*, INEI) (Instituto Nacional de Estadística e Informática, 2004–2011*a*). This dataset includes standard balance sheet information such as revenues, input expenditures, and plant locations. The survey is mandatory for firms with net sales above a certain threshold, while smaller firms are sampled. As a result, the EEA provides comprehensive coverage of medium and large firms, along with a representative sample of smaller firms. To ensure consistency across years and account for changes in the reporting threshold, we focus on manufacturing firms with net sales exceeding 2 million Peruvian Soles (PEN) per year—approximately 700,000 USD in 2010—over the period from 2004 to 2011. Our final dataset includes 2,473 firms and 8,138 firm-year observations.

The second data source is the Peruvian National Household Survey (*Encuesta Nacional de*

[5] Several studies, mostly based on U.S. data, use employer concentration as a proxy for labor market power, showing its negative correlation with wages (Azar, Marinescu and Steinbaum, 2022; Benmelech, Bergman and Kim, 2022). However, Bassier, Dube and Naidu (2022) find no evidence that labor supply elasticities decrease with concentration, and Yeh, Macaluso and Hershbein (2022) show that wage markdowns and employer concentration in U.S. manufacturing followed different trends over recent decades.

[6] See Ulyssea (2020) for a review.

Hogares, ENAHO), conducted annually by INEI (Instituto Nacional de Estadística e Informática, 2004-2011*b*). This survey is nationally and regionally representative, covering urban and rural areas across the 24 Peruvian departments and the constitutional province of Callao. It provides information on household members' socioeconomic characteristics. Individuals aged 14 and older respond to a dedicated module with questions on employment status, pay, occupation, and industry. To align with the firm-level data, we focus on the years from 2004 to 2011 and restrict the sample to working-age individuals (25 to 65) who have completed their education and are not yet retired. ENAHO offers several panel versions where the same households are interviewed annually for five consecutive years; we use the 2007-2011 panel to track workers' transitions between employment states.

2.1 Definitions

We define a local labor market as a 2-digit ISIC industry within a specific geographical area. These areas are primarily defined by Peruvian province boundaries, which correspond to level 2 administrative divisions and are subdivisions of departments. Excluding Metropolitan Lima—the province that includes the capital city—the average province has a population of approximately 114,000. Metropolitan Lima is a significant outlier, with a population of 10 million. Following Piselli (2013), we define five distinct local labor markets within Lima province. We analyze data from 199 geographical units and 23 manufacturing industries.

Our baseline measure of concentration is the Herfindahl-Hirschman Index (HHI) for payroll, defined as $HHI^{wn}_{kt} = \sum_{i \in k} \left(s^{wn}_{ikt}\right)^2$, where $s^{wn}_{ikt} = \frac{w_{ikt} n_{ikt}}{\sum_{i \in k} w_{ikt} n_{ikt}}$ represents firm i's share of the total payroll in local labor market k in year t. Here, w_{ikt} and n_{ikt} denote the firm's wage and employment, respectively. Values near one indicate a few firms dominate the market payroll. We also consider the employment HHI, defined as $HHI^{n}_{kt} = \sum_{i \in k} \left(s^{n}_{ikt}\right)^2$, where $s^{n}_{ikt} = \frac{n_{ikt}}{\sum_{i \in k} n_{ikt}}$, along with the number of firms in the local labor market as alternative measures.

In the ENAHO survey, workers are classified into four categories: own-account workers, employers, auxiliary family workers, and employees. For our analysis, we group own-account workers and employers as *self-employed workers*, while employees are categorized as *wage workers*. We exclude auxiliary family workers from our classification as they do not report monetary compensation. Additionally, ENAHO allows us to identify informal workers. A worker is classified as informal if they meet either of the following criteria: (i) they are a wage worker without employer-provided health insurance, which is legally mandated in Peru, or (ii) they are self-employed but are not registered with the national tax authority and employ five or fewer workers.

In Online Appendix B, we perform a series of checks to validate our definitions and measures, demonstrating the robustness of the empirical findings presented below. These checks include validating the concentration measures with firm census data (Instituto Nacional de Estadística e Informática, 2007), narrowing the self-employment category to include only informal self-employment or own-account workers, and broadening the geographical boundaries of local labor markets.

2.2 Employer Concentration

Employment and wages in Peruvian local labor markets are highly concentrated among a few medium and large firms. Panel I of Table 1 shows that the average local labor market includes about six firms, with unweighted and payroll-weighted mean wage-bill HHIs of 0.65 and 0.37, respectively. Notably, 39% of these markets are dominated by just one medium-to-large firm, and these highly concentrated markets account for approximately 8% of the nationwide payroll. This highlights that, despite their smaller share, these concentrated markets exert a significant influence on the overall payroll. Importantly, location explains about 43% of the variation in wage-bill HHI across markets, while differences across 2-digit industries only account for an additional 14%.[7]

2.3 Self-Employment and Flows Into and From Wage Work

In Peruvian manufacturing, as in other low- and middle-income countries, self-employment is widespread and primarily informal (Gollin, 2008; La Porta and Shleifer, 2014). Panel II of Table 1 shows that 40% of the manufacturing workforce is self-employed, 56% are wage workers, and the remaining 4% are auxiliary family workers. Over 90% of self-employment is informal, both across all industries and within manufacturing. Of the 40% classified as self-employed, 31% are own-account workers, and 9% are employers, with only 3% of self-employed individuals operating formally as registered businesses.

Informality reduces the costs of starting and operating a business, leading to variations in self-employment rates across industries. Self-employment is more prevalent in labor-intensive industries, constituting a substantial portion of Peru's manufacturing GDP. In these industries, physical capital is less critical, credit constraints are less severe, and the potential for informality is more significant. Self-employment is lower in more capital-intensive sectors such as pharmaceuticals and metals and virtually non-existent in oil and petroleum manufacturing.

On average, earnings from self-employment are lower and more dispersed than wage work earnings. Panel II of Table 1 shows that daily earnings from self-employment are approximately 28% lower than daily wages, while their standard deviation is about 30% higher.

Transitions A defining feature of labor markets in low- and middle-income countries is the high level of worker mobility between wage work and self-employment (Maloney, 1999; Dono-

[7] Peruvian manufacturing is less geographically clustered than in the US or the UK. Using the Ellison-Glaeser index of geographic concentration over census data, we find that only 63% of industries in Peru exhibit some degree of localization, compared to 97% in the US and 94% in the UK (Ellison and Glaeser, 1997; Duranton and Overman, 2005).

Table 1: Summary Statistics

Variable	Mean	St. Dev.
Panel I. Manufacturing Local Labor Markets		
Number of Firms	6.39	10.37
Wage-bill HHI	0.65	0.33
Wage-bill HHI (Weighted by LLM payroll share)	0.37	0.03
Employment HHI	0.63	0.35
Employment HHI (Weighted by LLM empl. share)	0.31	0.02
Percent of LLMs with 1 firm	38.78	2.27
Payroll Share of LLMs with 1 firm	7.94	1.79
Employment Share of LLMs with 1 firm	7.80	1.23
Panel II. Manufacturing Workers		
Wage Worker	0.56	0.50
Daily Wage	31.84	31.85
Self-Employed	0.40	0.49
Daily Earnings from Self-Employment	23.06	41.31
W-S Transition	0.06	0.24
S-W Transition	0.04	0.21

Notes. This table reports summary statistics from EEA firm-level data across Peruvian local labor markets (Panel I) and from ENAHO worker-level data (Panel II), averaging across all years from 2004 to 2011. Transition rates are obtained using the 2007-2011 panel version of ENAHO. Worker-level statistics correspond to dummy variables indicating wage work and self-employment, wages and earnings expressed in PEN (1 PEN $\approx$ 0.35 USD in 2010), and annual transitions from the wage- to self-employment sector (W-S) and vice versa (S-W).

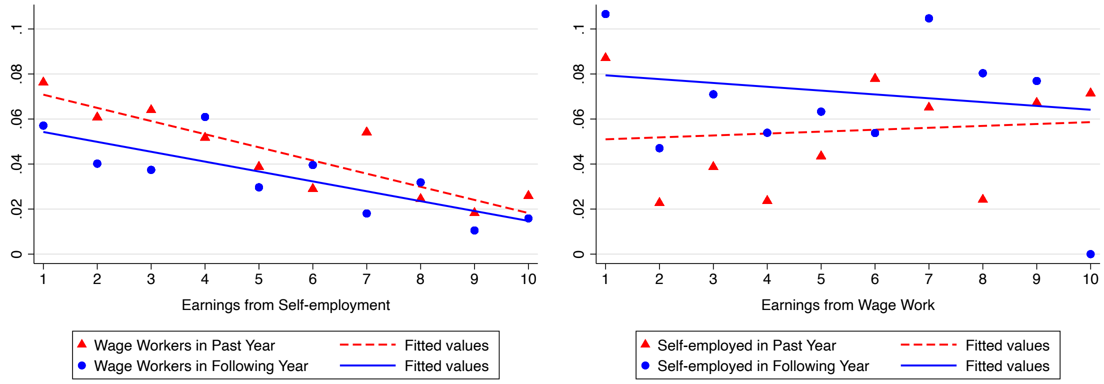
van, Lu and Schoellman, 2023). This is also evident in Peru. Panel II of Table 1 shows that approximately 4% of self-employed manufacturing workers transition to wage work the following year, while 6% of manufacturing wage workers move to self-employment. Conditional on switching either employment status or industry, transitions between wage work and self-employment are about 20% more likely than industry changes that do not involve a status shift (56% vs. 44%). Most of these transitions occur within the same geographical area. While ENAHO tracks only employment moves that do not involve a location change, the 2007 Census shows that 85% of manufacturing workers lived in the same commuting zone as in 2002, indicating limited geographical mobility. Additionally, 70% of employment transitions occur within the same 2-digit industry. This suggests that local labor markets provide a relevant unit for analyzing workers' employment transitions in the manufacturing sector.[8]

Worker transitions correlate with earnings. Figure 1 shows the likelihood of switching to or from wage and self-employment across deciles of the self-employment and wage earnings distributions. The left panel indicates that workers who have recently transitioned from wage work or are about to become wage workers are more likely to be among the lowest-earning self-employed individuals. In contrast, the right panel of Figure 1 shows that transitions to and from self-employment are not systematically correlated with earnings from wage work.

Thus, workers at the margin between self-employment and wage work consistently earn less than inframarginal self-employed workers and have similar earnings to inframarginal wage

[8] Online Appendix Table A.1 shows that workers transitioning from manufacturing self-employment to wage work (or vice versa) are about four times more likely to remain in manufacturing compared to the average worker switching employment status across all sectors.

Figure 1: Transition Probabilities Across the Earnings Distribution



Notes. The figures illustrate the relationship between the likelihood of transitioning from and into wage work and self-employment and earnings. The left panel plots average yearly transition probabilities into and from wage work across deciles of the self-employment earnings distribution. Similarly, the right panel plots average yearly transition probabilities into and from self-employment across the wage work earnings distribution deciles. The straight lines show the linear fit based on the underlying data.

workers. These findings suggest positive selection into self-employment but no selection into wage work. We will elaborate on this point later.

2.4 Concentration, Self-Employment Rates, and Earnings

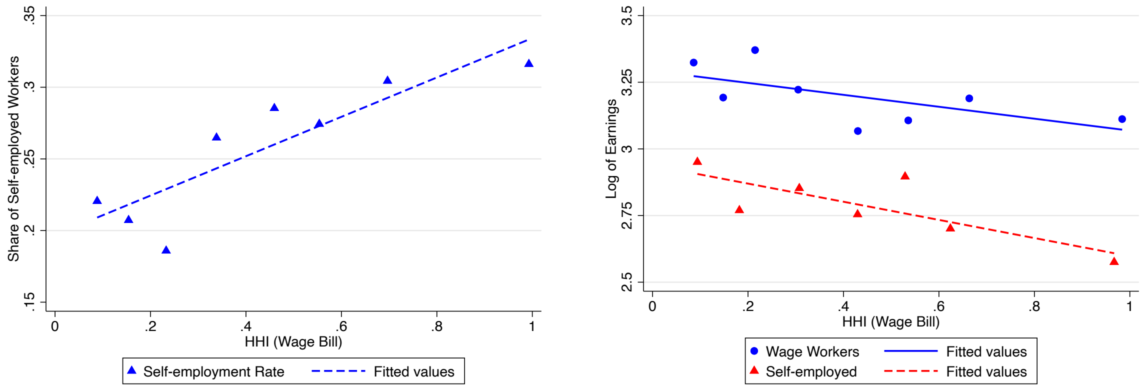
The final pattern in the data is the systematic relationship between concentration and self-employment across local labor markets. The left panel of Figure 2 shows that the average share of self-employment increases consistently across deciles of payroll HHI, indicating higher self-employment rates in more concentrated labor markets. Regression analysis further supports this correlation. We conduct a worker-level regression of a self-employment dummy on the log of wage-bill HHI in the worker's local labor market for the same year. The results, presented in Columns 1 to 3 of Online Appendix Table A.2, reveal that the relationship between concentration and self-employment is positive and significant, even after controlling for individual characteristics, industry, and location fixed effects.

The right panel of Figure 2, together with Columns 4 to 9 of Online Appendix Table A.2, examine the correlation between concentration and earnings across markets. Where concentration is higher, wages are lower, and self-employment is less lucrative.

The decline in wage employment share and wages with increasing concentration is observed among formal and informal wage workers, as shown in Online Appendix Figures A.1 and A.2. This suggests that in concentrated markets, self-employment is an alternative to formal and informal wage work. The right panels in both figures also show that, despite differences in earnings levels, the wages of both formal and informal wage workers decrease at the same rate as concentration increases.

These findings reinforce the sorting narrative proposed earlier: in highly concentrated markets, where wages are lower, more workers choose self-employment. As wage workers, these

Figure 2: Concentration, Self-Employment Rate, and Earnings



Notes. The figures illustrate the relationship between employer concentration, rate of self-employment (left), and earnings from both wage work and self-employment (right) across local labor markets. The left panel plots the share of self-employed workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for wage and self-employed workers. The straight lines show the linear fit based on the underlying data.

individuals would have earned similar wages to their peers. However, self-employed workers tend to earn less than their peers, decreasing average earnings from self-employment as market concentration increases.

2.5 Labor Market Power

The observed co-movements between concentration, self-employment, and earnings raise questions about the role of labor market power in Peruvian labor markets. While employer concentration is negatively associated with wages, this relationship alone does not prove labor market power, as concentration and wages are equilibrium outcomes. To pin down labor market power, we estimate the inverse elasticity of labor supply faced by individual firms (Manning, 2003). By examining how this varies with labor market concentration and self-employment rates, we can gain deeper insights into the role of labor market power in this context and its determinants.

Empirical Strategy We estimate the following regression model:

$$\ln w_{i(j,g)t} = \beta \ln l_{i(j,g)t} + \alpha_i + \eta_{(j,g)t} + u_{i(j,g)t}, \tag{1}$$

where $w_{i(j,g)t}$ is the wage paid by firm i in year t in its local labor market, defined by a manufacturing industry j within a province or commuting zone g, and $l_{i(j,g)t}$ is employment at the same firm. α_i is a firm fixed effect that captures differences across firms that do not change over time. $\eta_{(j,g)t}$ is a market $\times$ year fixed effect that accounts for aggregate yearly shocks at the local labor market level. This allows β to measure the firm-specific inverse supply elasticity of wage work while holding the aggregate labor supply constant.

To estimate the parameters in equation (1) consistently, we require a firm-level labor demand shifter, as OLS estimates may be biased due to the interdependence of wages and employment.

We address this by using the rollout of the Rural Electrification Program (*Programa de Electrificación Rural*, PER), launched by the Peruvian Ministry of Energy and Mining in 1993 to foster economic and social growth in rural areas (Dasso and Fernandez, 2015*a*). Between 1994 and 2012, the program implemented 628 projects across rural Peru, prioritizing districts with high poverty rates, low electricity coverage, and high renewable energy potential, with a total investment of USD 657.5 million (Dasso and Fernandez, 2015*b*; Dasso, Fernandez and Ñopo, 2015).

Our approach builds on the idea that electrification through the PER increased firms' marginal productivity and labor demand, especially for firms previously facing greater constraints in accessing electricity (Abeberese, Ackah and Asuming, 2019). To operationalize this approach, we first create the variable PER_{gt}, equal to the cumulative number of completed PER projects in location g up to year t. We then follow Bau and Matray (2023) to identify firms facing electricity access constraints at baseline.

For a firm i in market (j, g) producing output $y_{i(j,g)t}$ at time t and selling it in an imperfectly competitive market, the unit price $p_{i(j,g)t}$ is a markup $\mu_{i(j,g)t}$ over marginal cost. We assume that the firm uses a Cobb-Douglas production function with industry-specific input elasticities, where θ^e_j represents the output elasticity of electricity. The shadow cost of electricity, varying across firms and industries, is denoted by $\tau^e_{i(j,g)t}$. The electricity revenue share at firm i is given by $\alpha^e_{i(j,g)t} = \frac{e_{i(j,g)t}}{p_{i(j,g)t}y_{i(j,g)t}}$, where $e_{i(j,g)t}$ is total electricity bill.

Profit maximization implies $\frac{\theta^e_j}{\alpha^e_{i(j,g)t}} = \mu_{i(j,g)t}(1 + \tau^e_{i(j,g)t})$, which we can rewrite as:

$$\ln(\alpha^e_{i(j,g)t})^{-1} = \ln(\mu_{i(j,g)t}) + \ln(1 + \tau^e_{i(j,g)t}) - \ln \theta^e_j. \tag{2}$$

This shows that we can estimate the firm-level wedge $\tau^e_{i(j,g)t}$ as the residual from a regression of the log of the inverse electricity share of revenues on industry fixed effects and firm-level markups.[9] We include 4-digit ISIC Rev. 4 code fixed effects to control for industry-specific output elasticities and use second-degree polynomials of output market shares in both the local labor market and nationwide to flexibly account for firm-level markups.[10] To mitigate the impact of outliers and address measurement error, we define a dummy variable, $EC_{i(j,g)}$, which equals one for firms with an estimated wedge $\hat{\tau}^e_{i(j,g)t}$above the median at baseline, indicating tighter constraints in accessing electricity.

The interaction $PER_{gt} \times EC_{i(j,g)}$ is our instrumental variable (IV). It combines variation in program rollout across geography and over time with variation across firms within industries

[9]Intuitively, in the absence of distortions, the electricity revenue share $\alpha^e_{i(j,g)t}$ should equal the output elasticity θ^e_j. However, a firm's electricity share can fall below this optimal level if the firm either has market power, which allows it to set a higher markup $\mu_{i(j,g)t}$ or if it faces a higher shadow cost of electricity, captured by $\tau^e_{i(j,g)t}$.

[10]Output markups can be expressed as an increasing function of a firm's output market share in many macroeconomic models, such as Atkeson and Burstein (2008). This approach also aligns with our theoretical model in Section 3. The results are robust to (i) not controlling for output market shares (implicitly assuming no market power), (ii) controlling only for local labor market shares, and (iii) controlling only for national shares.

in access to electricity at baseline. The first-stage regression specification is

$$\ln l_{i(j,g)t} = \gamma PER_{gt} \times EC_{i(j,g)} + \phi_i + \delta_{(j,g)t} + v_{i(j,g)t}, \tag{3}$$

with ϕ_i and $\delta_{(j,g)t}$ capturing firm fixed effects and local labor market $\times$ year fixed effects, respectively, following the second-stage regression specification in equation (1).

The validity of this IV approach relies on three key assumptions. First, the instrument must be strongly correlated with employment, which holds if the electrification program boosts labor demand, particularly for firms with limited electricity access. Second, the instrument must be orthogonal to the wage and employment trends of electricity-constrained firms within each local labor market. This is plausible since the Ministry did not consider local firms or industries when implementing the program. Finally, the instrument must satisfy the exclusion restriction, meaning that electrification should not differentially affect labor supply to electricity-constrained firms. This ensures that changes in employment and wages reflect movements along the labor supply curve, allowing us to trace its slope. To make this assumption more plausible, we include local labor market $\times$ year fixed effects in all specifications. These fixed effects capture and control for changes in labor supply common to all firms within a market, even if these vary locally across industries. Importantly, we demonstrate below that our estimates remain robust when accounting for differences across firms over time at a more granular geographical level.

The exclusion restriction also requires that the labor demand shock does not affect wages via other channels, such as rents captured by workers. While this could be a concern, it is unlikely in Peruvian manufacturing, where workers have minimal bargaining power. Union density remained consistently low throughout the analysis period, ranging from 1.9% to 3.2%, placing Peru among the bottom 5% of countries in terms of unionization rates (International Labour Organization, 2020).

Another concern is whether the electrification program induced sufficient variation across Peruvian local labor markets. To explore this, we use districts as the geographical unit of analysis. Peru has 1,838 districts, each a subdivision of a province, with an average of about 9 districts per province. We find that 15% of districts in the firm-level IV estimation sample were affected by the program. These districts account for 41% of firm-level observations and 17% of the manufacturing workforce (based on ENAHO data). Online Appendix Figure A.3 offers additional details on the program's implementation. Initially, the targeted districts had a relatively low share of manufacturing employment. By the end of the period, however, the program had touched districts with a higher proportion of manufacturing employment, possibly due to the program spurring growth in these areas.

Results Table 2 presents the inverse elasticity IV estimates and standard errors. We report for each estimate the *F-statistic* associated with the Sanderson-Windmeijer multivariate test of excluded instruments, confirming that the instrument provides meaningful identifying variation throughout. Online Appendix Table A.3 reports the first-stage regression results.

Column 1 reports the results for the total sample of manufacturing firms. The firm-level inverse labor supply elasticity is estimated at 0.42, corresponding to a labor supply elasticity of 2.36, which is statistically significant at the 1% level. This elasticity implies a markdown of 1.42 between the marginal revenue product of labor and the wage paid. In other words, workers generate 42% more as value than what they earn at the margin, taking home 70 cents for every marginal dollar they produce.[11]

Columns 2 to 4 focus on different subsamples.[12] In Column 2, we estimate labor market power separately for markets with varying levels of labor market concentration.[13] For firms in the least concentrated labor markets ($HHI \leq 0.18$), we estimate an inverse labor supply elasticity that is statistically and economically insignificant. As concentration increases, labor market power rises. In moderately concentrated markets ($0.18 < HHI \leq 0.25$), workers take home nearly 80 cents for every marginal dollar they generate. In highly concentrated markets ($HHI > 0.25$), the wage take-home share drops to 63%.

Columns 3 and 4 further divide markets based on whether the self-employment rate is below or above the national average. In less concentrated markets, labor market power remains insignificant regardless of the self-employment rate. However, in highly concentrated markets, the degree of labor market power is influenced by the availability of self-employment opportunities. The highest level of labor market power is observed in highly concentrated markets with low self-employment rates, where the firm-level inverse labor supply elasticity is estimated at 0.75 and is statistically significant at the 1% level, corresponding to a 57% wage take-home share.[14] In contrast, in highly concentrated markets with higher self-employment rates, the estimated inverse labor supply elasticity is positive but much lower in magnitude and not statistically significant at conventional levels. Nevertheless, this estimate is statistically different at the 5% level from the corresponding estimate in less concentrated markets. Additionally, the difference between the inverse labor supply elasticity in markets with high versus low self-

[11] These figures closely align with those reported by Amodio and De Roux (2022) for Colombian manufacturing plants (inverse elasticity of 0.4) and by Deb et al. (2022) and Yeh, Macaluso and Hershbein (2022) for U.S. manufacturing (ranging from 0.37 to 0.4 and 0.53, respectively). They are slightly lower than what Felix (2022) finds for Brazil before the 1990s trade liberalization (50% wage take-home share).

[12] These estimates are derived using more flexible second- and first-stage specifications, where both the log of firm-level employment $\ln l_{i(j,g)t}$ and the instrument $PER_{gt} \times EC_{i(j,g)}$ are interacted with dummy variables that identify the different subsamples.

[13] This analysis uses contemporaneous wage-bill HHI values. Although potentially endogenous, the contemporaneous HHI determines the wage markdown size at a specific time. When we instrument contemporaneous HHIs with their lags, the resulting pattern closely mirrors the one discussed here. Similarly, when we use the previous year's self-employment rates as instruments and categorize markets by self-employment rate, the robustness of the estimates in Columns 3 and 4 is confirmed.

[14] These markets account for a significant portion of manufacturing employment, representing 24% of the manufacturing workforce across all markets for which we have firm-level data. Overall, 66% of all manufacturing workers are in provinces or commuting zones with at least one highly concentrated manufacturing labor market that features low self-employment rates. These areas tend to be less rural and have a higher share of manufacturing employment compared to agriculture.

Table 2: Estimates of Labor Market Power

			Self-Employment Rate	
			Low	High
	(1)	(2)	(3)	(4)
All Markets	0.423***			
	(0.052)			
$HHI^{wn} \in (0, 0.18]$		-0.006		
		(0.148)		
$HHI^{wn} \in (0.18, 0.25]$		0.262**		
		(0.105)		
$HHI^{wn} \in (0, 0.25]$			-0.108	-0.061
			(0.087)	(0.128)
$HHI^{wn} \in (0.25, 1]$		0.600***	0.752***	0.104
		(0.150)	(0.112)	(0.067)
SW F-statistics	178.78	222.58	215.80	130.01
		142.80	686.18	725.92
		3370.22		
Observations	6191	6191	3987	2204

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. The unit of observation is a medium to a large firm in EEA. The table reports 2SLS estimates of the firm-level inverse elasticity of supply of wage labor as captured by β in equation (1). The instrumental variable is the interaction of the cumulative number of PER projects completed in each location g up to year t (PER_{gt}) and a dummy equal to one for firms with higher than median constraints to accessing electricity at baseline ($EC_{i(j,g)}$). Estimates in Columns 2 to 4 are obtained by interacting both the log of firm-level employment $\ln l_{i(j,g)t}$ and the instrument $PER_{gt} \times EC_{i(j,g)}$ with dummy variables that identify the different subsamples as discussed in the text. Low and high self-employment rates are defined as below and above the average self-employment rate across local labor markets, respectively. We report the F-statistic associated with the Sanderson-Windmeijer multivariate test of excluded instruments for each estimate. Following equation (1), firm fixed effects and local labor market $\times$ year fixed effects are included in all specifications. Standard errors are clustered at the level of location g, i.e., province or commuting zone.

employment rates among highly concentrated markets is statistically significant at the 1% level, as is the difference-in-differences estimate between highly concentrated and less concentrated markets with varying self-employment rates.

Discussion A potential concern with our findings is that the local labor market $\times$ year fixed effects in our specification may not fully capture variations in labor supply or other general equilibrium effects of electrification that impact all firms equally within markets. For example, if firms facing tighter ex-ante constraints on accessing electricity are located in more rural areas, workers in those areas might respond differently to electrification, potentially violating the exclusion restriction. To address this, we redefine local labor markets as 2-digit industries j within districts d, substantially increasing the granularity of the local labor market $\times$ year fixed effects. This refinement enables us to control for changes in labor supply and general equilibrium effects at a geographical level approximately 10 times finer than in the baseline analysis. Results are presented in Online Appendix Table A.4, and all first-stage regression results are reported in Online Appendix Table A.5. The point estimates of the inverse elasticity of labor supply remain broadly consistent with those in Table 2, exhibiting similar patterns of market heterogeneity. Although the estimates are somewhat higher, the standard errors also increase, reflecting the reduced identifying variation caused by the more granular fixed effects.

Finally, a critical consideration is that the labor supply to a given firm is not independent of the one faced by other firms in the same market. By affecting more those firms facing tighter constraints on accessing electricity at baseline, electrification leads to labor reallocation across firms due to both supply and demand effects. In other words, the *reduced-form* inverse elasticity obtained through our identification strategy does not directly translate into the *structural* inverse elasticity, which measures the elasticity of firm-level wages to employment changes while holding competitors' wages and employment constant. Berger, Herkenhoff and Mongey (2022) study this issue in detail, showing that in granular labor markets, there is no closed-form mapping between the two elasticities; thus, a model is needed to determine the structural, welfare-relevant elasticity. Furthermore, the bias could be either negative or positive depending on the number of treated firms, their market shares, and those of their competitors. Our approach involves replicating the electrification quasi-experiment within the estimated model and comparing the implied reduced-form inverse elasticities with those in Table 2 as a means of testing the model's validity. We then derive the corresponding structural elasticities and rule out the possibility that the observed heterogeneity across markets, related to concentration and self-employment rates, is driven by bias rather than by the underlying structural features of the economy.

The evidence in this section reveals significant interactions between concentration, self-employment opportunities, and wage-setting power. Workers shift between wage employment and self-employment based on earnings. High concentration levels lead to increased oligopsony power, which results in fewer wage jobs and lower wages, thus pushing more workers towards self-employment. As wages fall, more workers are displaced from wage employment, making wage workers more responsive to wage changes and, in turn, reducing employers' labor market power. The following section provides a theoretical characterization of these dynamics.

3 Model

This section introduces a general equilibrium model of Peruvian manufacturing, where employer concentration, self-employment rates, and labor market power are jointly determined. The model aims to reconcile the empirical evidence discussed in the previous section and to conduct counterfactual analyses, specifically concerning industrial development policies.

3.1 Environment

The economy consists of a continuum of local labor markets, indexed by $k \in [0, 1]$, each characterized by M_k heterogeneous firms and a fixed measure of workers, L_k. Workers choose employment in one of two sectors: wage employment (sector F) or self-employment (sector

S). Workers are immobile across local labor markets and make employment decisions within their respective markets. Although workers differ in their sector-specific abilities, they share preferences over consumption goods and incur no disutility from labor. In addition to their labor earnings, workers hold equity in firms and receive income from profit distributions.

Preferences Workers' preferences over the consumption of market-level goods $\{C_k\}_{k\in[0,1]}$ are of the Cobb-Douglas form:

$$C = \exp\left\{\int_0^1 \alpha_k \log C_k \, dk\right\},$$

where C_k is market k's output and $\{\alpha_k\}_{k\in[0,1]}$ are nonnegative parameters determining the shares of income spent on each market, such that $\int_0^1 \alpha_k \, dk = 1$.

Each good C_k is an aggregate of two sectoral goods: $C_{F,k}$, produced by firms in sector F, and $C_{S,k}$, produced by self-employed workers in sector S. Within sector F, each firm $i = \{1, .., M_k\}$ produces a unique variety of $C_{F,k}$, denoted $c_{iF,k}$. The corresponding CES aggregators are given by:

$$C_k = \left[\zeta C_{F,k}^{\frac{\rho-1}{\rho}} + C_{S,k}^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \qquad \text{where} \qquad C_{F,k} = \left(\sum_{i=1}^{M_k} c_{iF,k}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}. \tag{4}$$

Consumers substitute between $C_{F,k}$ and $C_{S,k}$ with a constant elasticity ρ, and among firm-level varieties within sector F with a constant elasticity η. We impose $\eta > \rho > 1$, implying that consumers are more willing to substitute within than across sectors. Additionally, we allow for a preference shifter for good F, denoted as $\zeta > 0$.

Given this demand structure, consumer expenditure on market k's goods is:

$$P_{F,k}C_{F,k} = \gamma_{F,k}\alpha_k Y \qquad \text{and} \qquad P_{S,k}C_{S,k} = (1-\gamma_{F,k})\alpha_k Y, \tag{5}$$

where $\gamma_{F,k} = \zeta^\rho \left(\frac{P_{F,k}}{P_k}\right)^{1-\rho}$ represents the share of expenditure on variety F of good k relative to total expenditure in market k, with the sector- and market-level price indices defined as $P_{F,k} = \left(\sum_{i=1}^{M_k} p_{iF,k}^{1-\eta}\right)^{\frac{1}{1-\eta}}$ and $P_k = \left(\zeta^\rho P_{F,k}^{1-\rho} + P_{S,k}^{1-\rho}\right)^{\frac{1}{1-\rho}}$, respectively.

Similarly, expenditure on firm-level varieties within sector F is given by:

$$p_{iF,k}c_{iF,k} = s_{iF,k}\gamma_{F,k}\alpha_k Y, \tag{6}$$

where $s_{iF,k} = \left(\frac{p_{iF,k}}{P_{F,k}}\right)^{1-\eta}$ is the share of expenditure on variety i of good k in sector F over the total expenditure on sector F in market k.

Labor Supply Each worker h is endowed with sector-specific efficiency labor units, $\mathbf{a}^h \equiv (a_F^h, a_S^h)$, drawn from a joint distribution G_k, whose parameters may vary across markets.

Workers self-select into wage employment or self-employment to maximize their earnings. In sector F, workers perceive all M_k firms as identical despite underlying productivity differences. As a result, the wage per efficiency unit, $W_{F,k}$, is uniform across all firms. In sector S, workers earn a distinct wage per efficiency unit, $W_{S,k}$. Given their abilities, worker h's earnings in sector $J \in \{F, S\}$ are given by $W_{J,k}a_J^h$. Worker h chooses sector F if and only if:

$$\hat{W}_k \geq \left(\frac{a_F^h}{a_S^h}\right)^{-1}, \tag{7}$$

where $\hat{W}_k \equiv \frac{W_{F,k}}{W_{S,k}}$ denotes the relative wage in market k.

The sorting rule in equation (7) determines how workers allocate themselves across sectors based on their relative abilities and the relative wage $\hat{W}_k$. It implies that the aggregate supply of efficiency labor units in sector F is:

$$N_{F,k} \equiv N_{F,k}(\hat{W}_k) = L_k \int_0^\infty \int_0^{a_F \hat{W}_k} a_F\, g_k(a_F, a_S)\, da_F\, da_S, \tag{8}$$

where L_k is the total labor force in market k, and $g_k(a_F, a_S)$ represents the joint density of sector-specific abilities. The labor supply $N_{F,k}$ is an increasing function of the relative wage, i.e., $N'_{F,k} > 0$, since a higher $\hat{W}_k$ leads more workers to choose wage employment.

The sensitivity of labor supply to changes in $\hat{W}_k$ is captured by the wage labor elasticity:

$$\epsilon_{F,k} \equiv \epsilon_{F,k}(\hat{W}_k) = \frac{\hat{W}_k \int_0^\infty a_F^2\, g_k(a_F, a_F\hat{W}_k)\, da_F}{\int_0^\infty \int_0^{a_F\hat{W}_k} a_F\, g_k(a_F, a_S)\, da_F\, da_S} > 0, \tag{9}$$

which depends on the relative wage $\hat{W}_k$ and is thus determined in equilibrium. In what follows, we will show how adjustments in $\epsilon_{F,k}$ play a central role in shaping the dynamics of labor market power in this economy.

Technology Production technology in sectors S and F is linear in efficiency units of labor. Total output in sector S equals the efficiency labor units allocated to it:

$$Y_{S,k} = N_{S,k}, \tag{10}$$

where $N_{S,k}$ denotes labor supply in sector S.

In sector F, each firm i produces output according to $y_{iF,k} = z_{iF,k}n_{iF,k}$, where $z_{iF,k}$ is a productivity term, and $n_{iF,k}$ is labor demand. Aggregating across firms according to (4) leads to the following expression for total output in sector F:

$$Y_{F,k} = Z_{F,k}N_{F,k}, \tag{11}$$

where $N_{F,k} \equiv \sum_{i=1}^{M_k} n_{iF,k}$, and $Z_{F,k} \left(\sum_{i=1}^{M_k} s_{iF,k}^{\frac{\eta}{\eta-1}} z_{iF,k}^{-1}\right)^{-1}$ is a productivity index, with $s_{iF,k}$ defined above. Labor market clearing ensures that $N_{F,k}$ in (11) equals aggregate labor supply to sector F from equation (8). $N_{S,k}$ in (10) can be derived analogously.

Market Structure Sector S operates under perfect competition, while firms in sector F engage in Nash-Cournot competition in both the product and labor markets.

Firms in sector F produce differentiated varieties of the final good, leading to heterogeneous markups and prices in the output market. However, because workers perceive these firms as perfect substitutes, the labor market features oligopsonistic competition among effectively homogeneous employers, resulting in a uniform wage.

Firm's Problem Each firm i chooses its labor demand to maximize profits, taking as given the aggregate prices P_k and $P_{S,k}$, the self-employment wage $W_{S,k}$, and the employment decisions of its competitors, $N_{F,k/\{i\}}$. Firms internalize the impact of their labor demand on the sector F price index $P_{F,k}$, labor supply $N_{F,k}$, and the aggregate wage $W_{F,k}$.

The first order necessary condition of the firm problem determines its labor demand:

$$W_{F,k} = \frac{MRPL_{iF,k}}{\psi_{iF,k}}, \tag{12}$$

where $MRPL_{iF,k} \equiv \frac{\partial R_{F,k}}{\partial n_{iF,k}} = \frac{p_{iF,k} z_{iF,k}}{\mu_{iF,k}}$ is the firm's marginal revenue product of labor. The term $\psi_{iF,k}$ captures the wage markdown and is expressed as:

$$\psi_{iF,k} = 1 + \frac{s_{iF,k}^N}{\epsilon_{F,k}} \geq 1, \tag{13}$$

where $s_{iF,k}^N \equiv \frac{n_{iF,k}}{N_{F,k}}$ denotes the firm's employment share in market k, and $\epsilon_{F,k}$ is the wage labor supply elasticity defined in (9). A markdown above one indicates that the equilibrium marginal revenue product of labor exceeds the wage, reflecting wage-setting power. This outcome requires two conditions: (i) a positive employment share for the firm ($s_{iF,k}^N > 0$) and (ii) a finite labor supply elasticity ($\epsilon_{F,k} < \infty$).[15]

The optimality condition in (12) can be rearranged to express the firm i's output price as:

$$p_{iF,k} = \mu_{iF,k} \psi_{iF,k} \frac{W_{F,k}}{z_{iF,k}}, \tag{14}$$

[15]The markdown can also be expressed as $\psi_{iF,k} \equiv 1 + \epsilon_{iF,k}^{-1}$, where $\epsilon_{iF,k} \equiv \frac{\partial \ln n_{iF,k}}{\partial \ln W_{F,k}}$ is the firm *residual* labor supply elasticity, *holding other firms' employment fixed.* This is found as:

$$\epsilon_{iF,k} = \frac{\partial \ln n_{iF,k}}{\partial \ln N_{F,k}} \cdot \frac{\partial \ln N_{F,k}}{\partial \ln W_{F,k}} = \frac{\epsilon_{F,k}}{s_{iF,k}^N},$$

where we derived $\frac{\partial \ln n_{iF,k}}{\partial \ln N_{F,k}} = \frac{1}{s_{iF,k}^N}$, from $N_{F,k} = \sum_{i=1}^{M_k} n_{iF,k}$.

where $\mu_{iF,k}$ represents the firm's markup over marginal cost. This is defined as:

$$\mu_{iF,k} = \frac{\varepsilon_{iF,k}}{\varepsilon_{iF,k} - 1}, \quad \text{where} \quad \varepsilon_{iF,k} = \varepsilon(s_{iF,k}) = \left[\frac{1}{\eta}(1 - s_{iF,k}) + \frac{1}{\rho} s_{iF,k}\right]^{-1}, \tag{15}$$

is the firm's price elasticity of demand. This elasticity depends on the firm's output market share $s_{iF,k}$, as is standard in models of oligopolistic competition and product differentiation (e.g., Atkeson and Burstein, 2008).

3.2 Equilibrium

With segmented labor markets and Cobb-Douglas preferences across market-level goods, interactions across markets occur solely through changes in expenditures $Y_k \equiv \alpha_k Y$, where $\{\alpha_k\}_{k\in(0,1)}$ are the constant expenditure shares. This means that given Y, each market's equilibrium can be determined independently of the others.

This setup simplifies the model's solution into two distinct components: market equilibrium and general equilibrium. Below, we briefly outline these components while detailing the algorithm and numerical implementation in Online Appendix C.1.

Market Equilibrium The market equilibrium refers to the equilibrium in each local labor market, given aggregate income Y and the model's fundamentals. It is characterized by a vector $\hat{\mathbf{K}}_k \equiv \left\{M_k, \hat{W}_k, \mathbf{\Lambda}_k\right\}$ for each market k, where M_k is the number of active firms, $\hat{W}_k$ is the relative wage, and $\mathbf{\Lambda}_k = \{s_{iF,k}, s^N_{iF,k}, \mu_{iF,k}, \psi_{iF,k}\}_{i=1}^{M_k}$ is the vector of output and employment shares, markups, and markdowns for each firm.

To solve for $\hat{\mathbf{K}}_k$, we first assume that the set of employers M_k and their productivity are known for each k. Given M_k and a guess for $\hat{W}_k$, equations (6), (13), (14) and (15) define a fixed-point problem that determines the vector $\mathbf{\Lambda}_k$. In turn, given $\mathbf{\Lambda}_k$, the relative wage $\hat{W}_k$ is updated using equations (5), (10), and (11). The resulting fixed point in each market k constitutes the vector of market equilibria for a given guess of M_k.

Solving for the Number of Entrants We model firm entry as a sequential game in which, upon entry, firms incur a fixed cost $f^e_{i,k}$, measured in units of the final good, with more productive firms entering first. This process uses a fixed-point algorithm to solve for equilibrium wages and market shares iteratively. The equilibrium is reached when the marginal entrant's profits are non-negative, and any additional entrants would incur losses, leading to a unique and stable cutoff equilibrium where only firms above a certain productivity threshold enter.

To reduce the computational load of solving the fixed-point algorithm for each candidate M_k, we employ a simplified entry model for baseline calibration, similar to Gaubert and Itskhoki (2021). In this model, firms at the entry stage are assumed to behave "naively," anticipating atomistic markups and markdowns, which considerably streamlines the computation of market shares and equilibrium conditions. In Section 4.5.3, we examine the robustness of our findings to a more comprehensive entry game, which produces quantitatively similar results.

General Equilibrium The general equilibrium of the economy is given by a vector of expenditures and prices, $\mathbf{X} = (Y, P)$, such that aggregate income equals aggregate expenditure and that product markets clear. The first condition can be formulated as:

$$Y = W + F, \tag{16}$$

where W denotes total workers' income, equal to the sum of labor income and profits:

$$W = \int_{k\in(0,1)} \left(W_{S,k}N_{S,k} + W_{F,k}N_{F,k} + \sum_{i\in[1,M_k]} \pi_{iF,k} \right) dk, \tag{17}$$

and $F = \int_{k\in(0,1)} \left(\sum_{i=1}^{M_k} f_i^e\right) dk$ are total entry costs.[16]

To ensure that product markets clear, the real aggregate expenditure must also be equal to aggregate output, which leads to the second equilibrium condition:

$$\frac{Y}{P} = C. \tag{18}$$

With P normalized to 1, the conditions (16)-(17) and (18) yield the general equilibrium vector $\mathbf{X}$, given a market equilibrium $\mathbf{K}$.[17] In turn, given $\mathbf{X}$, the algorithm for the market equilibrium described above yields $\mathbf{K}$. The fixed point $(\mathbf{X}; \mathbf{K})$ is the economy equilibrium.

3.3 Characterization

Table 2 indicates that labor market power increases with employer concentration but less so in areas with higher self-employment rates. As noted in Section 2, reduced-form evidence alone makes it challenging to disentangle these patterns, as the estimates in Table 2 capture both direct and equilibrium effects. We now turn to our structural model for deeper insights.

Define $\bar{\psi}_{F,k} \equiv \sum_{i\in M_k} s_i^N \psi_{iF,k}$ as the employment-weighted average markdown in local labor market k. Substituting for $\psi_{iF,k}$ using equation (13), we obtain:

$$\bar{\psi}_{F,k} = 1 + \frac{HHI_{F,k}^n}{\epsilon_{F,k}}, \tag{19}$$

where $HHI_{F,k}^n$ is the employment-based HHI in sector F of market k, which also coincides with the wage-bill HHI in our model. This equation reveals that while employer concentration

[16] Here, W also captures aggregate welfare. In our counterfactual analysis, we will focus on local labor income $W_{S,k}N_{S,k} + W_{F,k}N_{F,k}$, which serves as a direct and relevant indicator of welfare that varies significantly across local labor markets.

[17] Walras' law implies that one of the two general equilibrium conditions is redundant.

increases average labor market power, reflecting more substantial oligopsony power, the aggregate labor supply elasticity $\epsilon_{F,k}$ weakens this effect. It also shows that this elasticity depends on the relative wage $\hat{W}_k$ and the joint distribution of workers' abilities G_k.

To further examine these relationships, we impose structure on the distribution by assuming a log-normal functional form–a common choice in empirical Roy models due to its desirable identification properties, as discussed in Section 4.

Let the random variable $Z \equiv a_F - a_S$ represent workers' comparative advantage, with mean $\hat{\mu}$ and standard deviation σ^*, both assumed constant across markets for simplicity. Under the log-normal assumption, and using parameter values consistent with our data and prior empirical studies, the aggregate wage labor supply elasticity $\epsilon_{F,k}$ can be approximated as:

$$\epsilon_{F,k} \approx \frac{\lambda(c_{F,k})}{\sigma^*}, \quad \text{where} \quad c_{F,k} \equiv \frac{\ln \hat{W}_k + \hat{\mu}}{\sigma^*}. \tag{20}$$

Here, $\lambda(x) = \frac{\phi(x)}{\Phi(x)}$ is the inverse Mills ratio, where $\phi(x)$ and $\Phi(x)$ denote the standard normal probability density and cumulative distribution functions, respectively. Notably, since $\lambda'(x) < 0$, the labor supply elasticity decreases as the relative wage $\hat{W}_k$ rises.[18]

Equation (20) shows that under the log-normal assumption, the elasticity function $\epsilon_F(\cdot)$ depends only on a few parameters of the ability distribution. In Section 4.2, we demonstrate how these parameters can be identified from cross-sectional earnings data, allowing us to estimate $\epsilon_F(\cdot)$ using worker-level data alone.

Letting $\chi_{S,k}$ denote the self-employment share, we derive the following relationship:

$$c_{F,k} = \Phi^{-1}(1 - \chi_{S,k}), \tag{21}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function with $\Phi' > 0$. This result implies that, in the log-normal case, the wage labor supply elasticity is directly linked to the self-employment share, which serves as a sufficient statistic for labor supply responsiveness to wage changes.

Equations (19)-(21) illustrate a key takeaway from our model: greater employer concentration increases average markdowns by strengthening oligopsony power, which suppresses wages. As wages fall, more workers shift into self-employment, altering the sectoral workforce composition. This, in turn, raises labor supply elasticity, mitigating labor market power.

The Pass-Through of Shocks to Average Earnings We now use our theory to characterize the effects of shocks to the economic environment on labor market outcomes. For illustration, we focus on average sectoral earnings, a central policy outcome in the counterfactuals below.

[18] See Online Appendix C.2.1 for a detailed discussion.

In sector F, we can decompose the change in (log) average earnings, $\ln \bar{E}_{F,k}$, as:

$$d\ln \bar{E}_{F,k} = -\underbrace{d\ln \bar{\psi}_{F,k}}_{\text{labor market power}} + \underbrace{d\ln Z_{F,k} + d\ln P_{F,k} - d\ln \bar{\mu}_{F,k}}_{\text{labor revenue productivity } (\overline{MRPL}_{F,k})} + \underbrace{d\ln \bar{A}_{F,k}}_{\text{selection}}, \tag{22}$$

where $P_{F,k}$ is the sectoral price index and $\bar{\mu}_{F,k}$ is the average markup, defined as $\bar{\mu}_{F,k} \equiv \left(\sum_i s_{iF,k} \cdot (\mu_{iF,k})^{-1}\right)^{-1}$.

Shocks influence average earnings in sector F through two channels: a *direct* effect on the wage per efficiency unit, captured by the first two bracketed terms, and a *selection* effect, which reflects how changes in worker composition alter average ability in response to the shock. The sign and size of the selection effect depend on the distribution of worker abilities.

The direct effect itself consists of two components. The *labor market power* channel captures how the average markdown responds to the shock, while the *labor revenue productivity* channel reflects how the average marginal revenue product of labor changes through adjustments in aggregate productivity ($Z_{F,k}$), prices ($P_{F,k}$), and average markups ($\bar{\mu}_{F,k}$).

Changes in earnings in the self-employment sector can be decomposed in a similar way, but with perfect competition, the expression simplifies to:

$$d\ln \bar{E}_{S,k} = \underbrace{d\ln P_{S,k}}_{\overline{MRPL}_{S,k}} + \underbrace{d\ln \bar{A}_{S,k}}_{\text{selection}}. \tag{23}$$

The sorting of heterogeneous workers across sectors influences how shocks affect average earnings by altering the average ability of workers in each sector, i.e., through the *selection* channel. These effects depend on the correlation between workers' abilities in wage work and self-employment and their relative dispersion (Adão, 2016; Amodio, Alvarez-Cuadrado and Poschke, 2020). In Online Appendix C.2.2, we show that, when the two abilities are strongly positively correlated and more dispersed in self-employment than in wage work, the mean ability of workers in both sectors increases when the relative wage $\hat{W}_k$ rises in response to shocks. This is because absolute and comparative advantage are negatively correlated in wage work but positively correlated in self-employment: as more workers select into wage employment, average ability increases everywhere. Vice versa, if abilities are negatively or weakly positively correlated, the mean ability of wage workers decreases, and the one of self-employed workers increases when more workers choose wage employment.

3.4 Model Discussion

We conclude with a discussion of four critical assumptions in our model. First, we assume that workers perceive all firms in sector F as homogeneous, enabling us to model the labor market as a standard Cournot oligopsony. Firms strategically choose labor demand but offer a uniform wage despite productivity differences. This simplifies our model compared to recent oligopsony theories where firms are imperfect substitutes, leading to wage variation.[19] This choice is motivated by two key factors: (i) it ensures analytical tractability, allowing us to examine the sources of market-level differences in labor market power–an essential aspect of both the theoretical and empirical analysis; and (ii) it is crucial for model estimation, as discussed in the next section. However, a caveat is that in our framework, only Cournot competition generates markdowns, whereas Bertrand competition would lead to efficient outcomes.

The second key assumption is that we restrict worker mobility across markets, effectively segmenting labor markets. This assumption is supported by the evidence in Section 2 showing that transitions between wage work and self-employment are more frequent than movements between local labor markets within wage work, a common focus in the literature. The main advantage of this assumption is that it simplifies worker sorting into a binary decision, as in the classic Roy (1951) model. Together with our assumptions about labor market structure and the log-normal parameterization of the ability distribution, this facilitates mapping the ability distribution to cross-sectional earnings data for identification. As discussed in Section 3.3, an important implication is that we can use worker-level data to discipline the labor supply determinants of labor market power in the model.

Finally, a notable feature of our model is that it incorporates oligopoly power in the output market, oligopsony power in the labor markets, and endogenous entry. This sets our approach apart from much of the existing literature, which often assumes fixed entry due to the computational challenges of modeling entry games with oligopsony. Notable difficulties in modeling endogenous entry include accounting for existing competitors, which can lead to multiple equilibria, and determining entry patterns across interdependent markets (MacKenzie, 2021).

Two assumptions make the entry problem tractable in our model. First, the boundaries of the product and labor markets align, with product varieties and worker decisions being determined within both industry and location.[20] Second, we assume Cobb-Douglas preferences across market-level goods. Combined with segmented labor markets, these assumptions ensure that a firm's market share depends only on local competitors, allowing firms to make independent entry decisions across different markets.

4 Model Identification and Estimation

This section explains how the model is identified using Peruvian firm- and worker-level data. First, we set up the model for quantitative analysis by imposing parametric assumptions. Next, we outline how the parameters are identified using direct and indirect methods. Finally, we present the estimation results.

[19] See, e.g., MacKenzie (2021); Berger, Herkenhoff and Mongey (2022); Felix (2022); Gutiérrez (2023).

[20] While this is common for labor markets, it is less typical for product markets, which are often defined at the national level. Berger, Herkenhoff and Mongey (2022) assume perfect competition in the output market. Gutiérrez (2023) discusses the challenges that arise when product and labor market boundaries do not align in a theory of oligopoly and oligopsony with fixed entry.

4.1 Parameterization

We parameterize the model to capture heterogeneity across local labor markets in several factors affecting firms' and workers' decisions, such as firm productivity, entry costs, and worker skills. Since our model's structure does not permit a straightforward inversion–where each labor market or firm in the data can be directly matched to a counterpart in the model–we treat each local labor market as a multi-dimensional draw from the structural data-generating process defined by the model, focusing on the estimation of common parameters shared across markets.

Firm Productivity Firm productivity is influenced by both idiosyncratic factors and broader market characteristics. We adopt a parameterization that incorporates both types of variation.

Let M_k^* denote the potential (shadow) entrants in the wage sector of local labor market k, assumed to follow a Poisson distribution with $\mathbb{E}(M_k^*) = \bar{M}_k^*$. Each potential entrant independently draws a productivity level from a Pareto distribution with a lower bound $\underline{z}_k$ and shape parameter θ, where a smaller θ implies a more dispersed and skewed productivity distribution. Under this Poisson-Pareto framework, the parameter $T_k \equiv \bar{M}_k^* \cdot \underline{z}_k^\theta$ acts as a sufficient statistic for expected productivity in a local labor market, provided the least efficient firm remains inactive (Gaubert and Itskhoki, 2021).[21]

To capture heterogeneity in productivity across local labor markets, we assume that the T_k values are drawn from a log-normal distribution with parameters μ_T and σ_T:

$$T_k \sim \log \mathcal{N}(\mu_T, \sigma_T).$$

This parsimonious productivity structure generates a realistic cross-sectional distribution of sales. Online Appendix Figure A.5 illustrates this, showing that a log-normal distribution closely approximates the distribution of log sales across local labor markets.

Firm Entry Cost In many models, entry costs are assumed to be constant across firms. We make two critical adjustments in our baseline specification. First, we set the entry cost for the first firm to zero, $f^e(1) = 0$, ensuring that at least one firm operates in every local labor market. This is essential to maintain a fixed number of active local labor markets. Second, we allow entry costs to increase with the number of active firms, specified as:

$$f^e(n) = f_0 + f_1\sqrt{n}, \quad \text{for } n \geq 2,$$

where n denotes the rank of an entrant. While not critical for the main results, this adjustment prevents the least productive shadow firm from becoming active. As discussed earlier, this is

[21]Specifically, the number of shadow firms with productivity exceeding $z > \underline{z}_k$ follows a Poisson distribution with mean $T_k z^{-\theta}$ (Eaton, Kortum and Sotelo, 2012).

necessary to ensure that the parameter T_k accurately reflects average market productivity.

Section 4.5 demonstrates that our results remain robust under an alternative parameterization in which the fixed cost f_k^e is a constant that may vary across labor markets.

Worker Ability For workers' ability, we assume that each worker's ability vector $\mathbf{a} = (a_F, a_S)$ is drawn from a joint log-normal distribution:

$$\log \mathbf{a} \sim \mathcal{N}\left(\boldsymbol{\mu_k}, \boldsymbol{\Sigma_k}\right), \quad \text{where} \quad \boldsymbol{\mu_k} = \begin{pmatrix} \mu_{F,k} \\ \mu_{S,k} \end{pmatrix}, \quad \boldsymbol{\Sigma_k} = \begin{pmatrix} \sigma_{F,k}^2 & \varrho_k \sigma_{F,k} \sigma_{S,k} \\ \varrho_k \sigma_{F,k} \sigma_{S,k} & \sigma_{S,k}^2 \end{pmatrix}, \tag{24}$$

where, $\boldsymbol{\mu_k}$ represents the mean efficiency labor units in wage work and self-employment, while $\boldsymbol{\Sigma_k}$ is the variance-covariance matrix.

The log-normal assumption for G_k is common in empirical Roy models as it facilitates the identification of the underlying parameters (French and Taber, 2011). In Section 4.2.1, we explain how, within our framework, the variance-covariance parameters $\boldsymbol{\Sigma_k}$ and the relative comparative advantage $\hat{\mu}_k \equiv \mu_{F,k} - \mu_{S,k}$ can be identified on a market-by-market basis using cross-sectional earnings data from both wage and self-employment sectors. However, the absolute advantage parameters $\mu_{F,k}$ and $\mu_{S,k}$ cannot be directly identified. To address this, we employ an indirect approach by assuming that $\mu_{S,k}$ is drawn from a normal distribution:[22]

$$\mu_{S,k} \sim \mathcal{N}\left(\mu_{\mu_S}, \sigma_{\mu_S}\right),$$

where the parameters μ_{μ_S} and σ_{μ_S} are estimated using the procedure described in Section 4.2.2. Given a draw for $\mu_{S,k}$ and an estimate for $\hat{\mu}_k$, we can then recover $\mu_{F,k}$ as $\mu_{F,k} = \mu_{S,k} + \hat{\mu}_k$. Later, we will show that this parameterization of the ability distribution yields a plausible distribution of workers' abilities and earnings.

4.2 Identification

We now discuss the identification of all model parameters, beginning with the parameters governing the distribution of workers' abilities, $\boldsymbol{\Sigma_k}$ and $\hat{\mu}_k$. For brevity, we provide only a high-level overview here, while Online Appendix D.1 offers a detailed explanation of our approach. Section 4.2.2 then discusses identification via indirect inference.

4.2.1 Workers' Ability Distribution

Identification of the variance-covariance parameters $\boldsymbol{\Sigma_k} = \{\sigma_{F,k}, \sigma_{S,k}, \varrho_k\}$ builds on the approach in Heckman and Sedlacek (1985) and relies on the relationship between observed labor market outcomes and the underlying ability distribution of workers. Assuming log-normality,

[22] The assumption of normality for $\mu_{S,k}$ is supported by the evidence in Online Appendix Figure A.5 showing that a normal distribution well approximates the distribution of years of education across local labor markets.

these relationships can be expressed in a tractable form, yielding a system of three equations that map the three parameters of interest to observed employment shares and sectoral earnings data on a market-by-market basis.

Mean Comparative Advantage The mean comparative advantage $\hat{\mu}_k$ governs the relationship between the average ability gap across sectors, sectoral shares, and the parameters $\boldsymbol{\Sigma_k}$. While the log-normal assumption provides an analytically tractable way to express this relationship, the lack of data on worker abilities still presents an identification challenge for $\hat{\mu}_k$.

To address this, we impose an additional restriction by assuming that mean log ability across sectors can be mapped to the ratio of log years of education across sectors, with this mapping governed by an unknown elasticity β, which we assume to be constant across markets. This parameter captures how differences in average (log) education levels between wage and self-employed workers translate into differences in average (log) abilities across sectors.

We then use the model equations to derive an expression that maps differences in log earnings across sectors to differences in abilities. Using education as a proxy for ability, we can infer the value of β. To address concerns about measurement error in earnings data or potential violations of the exclusion restriction, we also experiment with alternative regressors and different values of β. With estimates for β and $\boldsymbol{\Sigma_k}$, we infer $\hat{\mu}_k$ for each market using the expression for the mean log ability gap under the log-normal assumption.

This approach allows us to integrate the education-based proxy into the structural model and estimate the ability gap. While Heckman and Sedlacek (1985) address the challenge of identification of mean absolute advantages using instrumental variables with time-series wage data, our method combines direct and indirect estimation techniques, leveraging the specific structure of our model for identification.

Setting Parameters to Constant Despite our ability to recover market-specific parameters, we maintain a constant variance-covariance matrix and mean comparative advantage across markets in our baseline model for two reasons. First, this parsimonious approach enhances transparency by minimizing market heterogeneity and the impact of measurement error in earnings data. Second, it aligns with our method of matching the model to the data, where the mapping between markets in the data and the model is indirect. In Section 4.5, we discuss an extension of the model that allows for heterogeneity in these parameters.

Discussion Our approach to the identification of the workers' ability distribution relies on parametric restrictions, a common but not fundamental feature of empirical Roy models. In our setting, these restrictions allow for transparent estimation using worker-level data while providing the necessary structure to embed the Roy block into our general equilibrium model and conduct counterfactual analysis.

In principle, one could relax parametric assumptions if valid instruments that shift skill distributions without directly affecting wages were available (Heckman and Honoré, 1990; French and Taber, 2011). A recent application by Adão (2016) demonstrates how non-parametric identification can be obtained using cross-regional variation in sectoral responses to industry-specific price shocks. While this approach avoids strong parametric assumptions, it relies on exogenous variation in labor demand shifters and additional model restrictions, such as price-taking behavior, which are inconsistent with our setting. Given these challenges, implementing a similar approach is beyond the scope of this paper.

Crucially, while our assumption of log-normality imposes parametric structure, it avoids the strict selection patterns that would arise under the more widely used Fréchet distribution (Hsieh et al., 2019; Burstein, Morales and Vogel, 2019). In contrast, as explained in Section 3.3, the log-normal specification allows comparative and absolute advantage to vary independently, offering sufficient flexibility to capture the core mechanism of the Roy model while maintaining tractability within a general equilibrium framework.

In sum, our parametric approach strikes a balance between flexibility and tractability, making it both empirically reasonable and methodologically appropriate for our analysis. Nevertheless, we acknowledge its limitations. As Heckman and Honoré (1990) emphasizes, relying solely on normality for identification can introduce biases if the assumption does not hold.

4.2.2 Targeted Moments and Identification of Remaining Parameters

We now describe the procedure used to estimate the remaining parameter vector, given by $\Phi = (\mu_T, \sigma_T, \theta, f_0, f_1, \mu_{\mu_S}, \sigma_{\mu_S}, \eta, \rho, \zeta)$. We target 27 empirical moments that capture key local labor market outcomes, informed by the empirical findings in Section 2. These moments characterize the cross-sectional patterns of concentration, self-employment, and their co-movements. While all parameters influence every moment to some extent, certain parameters are more directly linked to specific moments, which plays a crucial role in identification, as discussed below.

Targeted Moments First, we focus on capturing the prevalence of concentration across local labor markets. From Table 1, we target the mean and standard deviation of the (log) number of firms in each market and the employment-based HHI, both weighted and unweighted. We also target the share of local labor markets with a single employer and the corresponding percentage of total wage employment.[23]

These moments are essential for identifying productivity parameters and the fixed cost distribution. Intuitively, fixed cost parameters (f_0 and f_1), along with productivity parameters for firms and workers (θ and μ_{μ_S}), determine the incidence of concentration across markets. A market can feature fewer firms due to high fixed costs or high firm or worker productivity.

[23] We specifically target the moments presented in Online Appendix Table A.6, which replicates Table 1 for the merged sample of local labor markets where both firm-level and worker-level data are available across sectors.

We also target moments of the distribution of log total sales across markets, specifically the interquartile ratio and the 90-10 ratio, along with the mean and standard deviation of the sales concentration ratio (CR1 and CR4) within markets. These moments help identify the average market productivity parameters (μ_T and σ_T) and within-market productivity variation (θ).

Next, we aim to capture self-employment patterns across local labor markets. We target the mean and standard deviation of the wage employment share and the relative (log) worker earnings between wage and self-employed workers, as reported in Table 1. Under our parameterization of the ability distribution, with fixed $\boldsymbol{\Sigma}$ and $\hat{\mu}$, the wage employment share is a monotonic function of the relative wage $\hat{W}$, which depends on the elasticity parameter ρ and the preference shifter ζ. These two parameters also influence the mean (log) relative earnings.

Although most ability distribution parameters are estimated externally, we target moments of the relative average schooling between wage and self-employed workers—mapped to relative average ability in the model—focusing on the interquartile and 90-10 ratios. These moments help identify the mean absolute advantage parameters $(\mu_{\mu_S}, \sigma_{\mu_S})$, which determine the dispersion in average ability across markets.

Finally, we target correlations between the employment-based HHI and several labor market outcomes, specifically the wage-employment share and (log) earnings in both sectors. The sensitivity of these variables to changes in employer concentration mainly depends on the across-sector elasticity (ρ) and the absolute advantage parameters $(\mu_{\mu_S}, \sigma_{\mu_S})$.

Normalization To improve estimation precision, we apply the following normalizations. First, since the elasticity of substitution η is weakly separately identified from the productivity parameter θ, both being linked to the Pareto tail of the sales distribution across firms, we fix $\eta = 6$ and estimate θ in the MSM routine. Second, as the parameter f_1 shows minimal sensitivity to the targeted moments, we set it externally to 5×10^{-7}. This reduces the parameter vector to $\Phi = (\mu_T, \sigma_T, \theta, f_0, \mu_{\mu_S}, \sigma_{\mu_S}, \rho, \zeta)$, thereby improving estimation precision.

Identification To support our identification strategy, we formally examine the connection between parameters and moments by computing the elasticity of each model-generated moment to each parameter, following standard practices in the literature (e.g., Kaboski and Townsend 2011). The full Jacobian matrix is provided in Online Appendix Figure A.7. Below, we offer some insights based on the results of this analysis.

The parameters θ, ζ, and ρ are the most influential for most targeted moments. This is expected, as these parameters play key roles in determining the equilibrium relative wage $\hat{W}_k$ in each local labor market, as shown in equation (C.3) in the Online Appendix C.1. By targeting multiple moments directly linked to the relative wage, more than the number of unknown parameters, we ensure sufficient identifying variation for these parameters. Additionally, only the moments related to relative ability show sensitivity to changes in the dispersion parameter for workers' absolute advantages. This is reassuring, as the market equilibrium—which determines most (other) moments—depends solely on $\hat{\mu}$, not on the absolute advantage parameters.

The Jacobian matrix confirms the identification argument, demonstrating that the model provides enough variation to effectively identify the remaining parameters.

Table 3: Summary of Model Parameters

Parameter		Value
Panel I. Externally Fixed		
η	Substitution elasticity within sector F	6
f_1	Fixed cost slope parameter	5×10^{-7}
Panel II. Externally Estimated		
σ_F	St. dev. of log ability as a wage worker	0.81
σ_S	St. dev. of log ability as a self-employed	0.91
ϱ	Correlation of log abilities	0.89
$\hat{\mu}$	Mean comparative advantage	-0.12
Panel III. Estimated via MSM		
μ_T	Mean of market-level productivity	0.82
σ_T	St. dev. of market-level productivity	0.96
f_0	Fixed cost intercept parameter	1.97×10^{-4}
θ	Firm-level productivity dispersion parameter	2.34
μ_μ	Mean of market-level mean absolute advantage	1.68
σ_μ	St. dev. of market-level mean absolute advantage	0.10
ρ	Substitution elasticity across F and S	2.72
ζ	Sector F preference shifter	1.88

Notes. This table reports the parameter values for the quantitative model. See Section 4 for details on parameter estimation.

4.3 Estimation Results

We estimate the model in three steps. First, we calibrate the Cobb-Douglas expenditure shares $\{\alpha_k\}_{k\in[0,1]}$ and population shares $\{L_k\}_{k\in[0,1]}$ from the data, equating them to the income share and the share of the workforce in each local labor market.[24]

We then use our matched firm-worker level data to estimate the parameters of the variance-covariance matrix of the workers' ability distribution $(\boldsymbol{\Sigma_k})$ as well as the mean comparative advantage $(\hat{\mu}_k)$ using the procedure described in 4.2.1. Lastly, we implement a MSM procedure to estimate all the remaining parameters using the procedure detailed in Section 4.2.2.

Table 3 summarizes the estimated parameter vector. Panel II provides the median estimates of the key Roy model parameters, based on the direct inference approach outlined in Sec-

[24] For the model calibration, we use a merged sample comprising local labor markets where both firm-level and worker-level data across both sectors are available. Summary statistics for this sample are provided in Online Appendix Table C.1. The sample includes 1,040 local labor market-year observations. For computational efficiency, we reduce the sample size to 234 for the baseline calibration. A histogram of the resulting vector $\{\alpha_k, L_k\}$ is shown in Online Appendix Figure A.4.

tion 4.2.1. Online Appendix Figure A.6 displays the histograms of the estimated variance-covariance parameters and mean comparative advantage across markets.

The two abilities are highly correlated, with an estimated correlation coefficient of $\hat{\varrho} = 0.89$, and the ability for self-employment is more dispersed than the ability for wage work, i.e., $\hat{\sigma}_S > \hat{\sigma}_F$. These parameters are precisely estimated, with bootstrap standard errors ranging from 0.02 to 0.07. The estimates suggest no correlation between workers' comparative and absolute advantage in wage employment, but a positive correlation in self-employment advantages. This is consistent with the evidence in Figure 1, which shows that transitions into and out of wage employment are more common among lower-earning self-employed workers, while transitions to self-employment are unrelated to wage earnings.[25] We also estimate $\hat{\mu} = -0.12$, indicating that the average worker in the population has a comparative advantage in self-employment. These findings align with experimental evidence suggesting that workers in low-income countries, given their current skill levels and wage rates, often choose self-employment over industrial jobs (Blattman and Dercon, 2018).

Panel III presents the estimated parameter vector from the MSM procedure, with the corresponding model moments summarized in Table 4. The model demonstrates a strong fit to the data, which is noteworthy given that only 8 parameters were used to target 27 moments.

The model effectively captures various measures of concentration across local labor markets. It closely replicates the high share of monopsonistic labor markets, with 39% observed in the data and 36% predicted by the model, as well as the corresponding payroll share–8% in the data and 7% in the model. Additionally, the model predicts that approximately 66% of workers are wage employees, compared to 71% in the data. Wage workers in the model earn about 0.4 log points more than self-employed workers, in line with the empirical evidence.

The model also successfully replicates the negative cross-sectional correlations between earnings in both sectors and wage-employment rates with the employment-based HHI. All coefficients are both economically and statistically significant. However, the model falls short of accurately matching the correlations between concentration and mean log earnings in the two sectors. Although the signs of the coefficients are correct, their magnitudes deviate from the observed data. We find that this discrepancy is mostly driven by markets with one firm, to which the model assigns lower average earnings than in the data.

Overall, the model's ability to replicate the core patterns documented in Section 2 builds confidence in our estimated parameters, which are broadly consistent with findings from other studies in the literature. For instance, we estimate a Pareto shape parameter of $\theta = 2.34$, which is higher but close to the 1.5 in Huang et al. (2024) for Chilean importers.We also find

[25]These transitions can be explained by shocks to relative unit earnings $\hat{W} = W_F/W_S$ combined with the estimates of the variance-covariance matrix of the joint ability distribution. Positive selection in self-employment implies that transitions are more frequent among lower-earning self-employed workers. Meanwhile, the lack of selection in wage work suggests that sector switchers earn wages comparable to those of inframarginal wage workers. Changes in $\hat{W}$ and their heterogeneity in terms of sign, size, and frequency across markets, combined with the estimated sign and strength of selection, are therefore sufficient to generate the patterns in Figure 1.

Table 4: Targeted Moments and Model Fit

Moment	Model	Data	Moment	Model	Data
Panel I. Distribution Moments					
Log Number of Firms			Wage-bill Share of		
Mean	0.97	1.22	Markets with 1 firm	0.07	0.08
Standard Deviation	0.95	1.17	Markets with <10 firms	0.89	0.84
			Markets with <50 firms	1.00	0.99
Log of Sales			Share of Wage Employment		
Ratio p75/p25	2.94	2.92	Mean	0.66	0.71
Ratio p90/p10	5.29	5.30	Standard Deviation	0.12	0.32
CR_1, Mean	0.66	0.69			
CR_1, Standard Deviation	0.29	0.29	Log of Earnings$_F$/Earnings$_S$		
CR_4, Mean	0.94	0.91	Mean	0.41	0.40
CR_4, Standard Deviation	0.11	0.15	Standard Deviation	0.58	0.93
Employment HHI					
Mean, Unweighted	0.57	0.59	Log of Schooling$_F$/Schooling$_S$ (Ability$_F$/Ability$_S$)		
Standard Deviation	0.34	0.35	Ratio p75/p25	1.42	1.28
Mean, Weighted	0.30	0.33	Ratio p90/p10	1.18	1.04
Percent of Markets with 1 firm	0.36	0.39			
Panel II. Regression Coefficients					
% Wage Employment on (Log) HHI^n			(Log) Earnings$_F$ on (Log) HHI^n		
Point Estimate	-0.04	-0.07	Point Estimate	-1.36	-0.14
Standard Error	0.01	0.01	Standard Error	0.13	0.02
(Log) Earnings$_S$ on (Log) HHI^n					
Point Estimate	-1.17	-0.11			
Standard Error	0.12	0.03			

Notes. This table reports the moments used in the estimation and compares them with those calculated from the estimated model. The data moments are computed in the sample of local labor markets where at least one formal firm is active and the share of self-employed workers and wage workers is strictly between 0 and 1. See Section 4 for more details on the moments' construction.

limited variation in absolute advantage across markets ($\sigma_\mu = 0.10$), which aligns with our observation of remarkably stable estimates for the parameters of the ability distribution across markets. Lastly, we estimate a substitution elasticity between sector goods within a market of $\rho = 2.72$, lower than the substitution elasticity within sector F but higher than the substitution elasticity across product markets. The latter is implicitly set to 1 in our case due to the Cobb-Douglas assumption and estimated at 1.5 in Gutiérrez (2023). This is consistent with gradually decreasing substitution elasticity as we move to upper utility nests.

4.4 Model Fit

We begin by evaluating how effectively the model replicates the distributions of key variables across local labor markets. Online Appendix Figure A.9 illustrates the distribution of normalized (log) sales, the (log) earnings gap between wage and self-employed workers, and the number of firms. The red bars represent the data, while the blue bars show the model's predicted distributions. Despite only targeting a few key moments in the estimation, the model's distributions closely match those observed in the data, indicating a good overall fit.

Model-Implied Reduced-Form Elasticities We now evaluate how well the model captures labor market power in the Peruvian economy. In Section 2.5, we provided evidence of significant labor market power across markets. However, those estimates derived from the data are reduced-form inverse labor supply elasticity estimates, which incorporate competitors' employment responses and do not directly map to the structural elasticity (Berger, Herkenhoff and Mongey, 2022). To assess the model's fit, we replicate Table 2 using model simulations by applying the same shock to firm productivity and labor demand used in Section 2.5. This is challenging for several reasons, including the model's static nature, the wage homogeneity within markets, and the absence of electricity as a production input for firms.

We overcome these challenges by following a three-step procedure, detailed in Online Appendix D.4. First, we identify treated firms in the model by assigning each firm an electricity wedge τ. This assignment relies on the empirical relationship between τ and firm productivity, inferred from the data. A firm is classified as treated if its τ exceeds the economy-wide median, following the approach in Section 2.5. Second, we determine the magnitude of the productivity shock induced by electrification by estimating its effect on firm productivity in the data. Lastly, starting from the baseline model equilibrium, we simulate a 2.3% productivity shock to the treated firms, corresponding to the estimated average effect. The model's average reduced-form inverse labor supply elasticities are then calculated by taking the ratio of the (log) wage to (log) employment responses of treated firms in markets with more than one firm, consistent with the local average treatment effect (LATE) within-market estimates reported in Table 2.

Table 5 presents the model-implied reduced-form estimates of labor market power across markets, along with bootstrap standard errors, averaged across all markets as well as within different subsamples. The results closely align with those in Table 2. The model estimates an average inverse elasticity of 0.36 across local labor markets, compared to 0.42 in the data. The model effectively replicates the relationship between average inverse elasticity and market concentration, capturing its gradient with precision. Additionally, it reflects the mitigating effect of self-employment, showing that labor market power is highest in markets with high employer concentration and low self-employment rates. In these markets, the model predicts an inverse elasticity of 0.62, compared to 0.75 in the data. Conversely, labor market power is lower in concentrated markets with a higher-than-average self-employment share, though the model underestimates the difference here, overshooting the inverse elasticity estimate in this latter group of markets. We attribute this discrepancy to the model's inability to fully capture the broad variation in self-employment rates across markets, with a standard deviation of 0.09 in the model compared to 0.32 in the data. However, the difference-in-differences between highly concentrated and less concentrated markets with varying self-employment rates is preserved.

The estimated model also allows for comparison between reduced-form inverse elasticities and their structural counterparts, reported in Online Appendix Table A.7. The structural inverse

Table 5: Reduced-Form Inverse Supply Elasticity – Model Estimates

			Self-Empl. Rate	
			Low	High
	(1)	(2)	(3)	(4)
All Markets	0.364			
	(0.070)			
$HHI^{wn} \in (0, 0.18]$		0.242		
		(0.071)		
$HHI^{wn} \in (0.18, 0.25]$		0.319		
		(0.046)		
$HHI^{wn} \in (0, 0.25]$			0.214	0.284
			(0.033)	(0.134)
$HHI^{wn} \in (0.25, 1]$		0.560	0.621	0.479
		(0.160)	(0.177)	(0.294)

Notes. This table presents the estimates of the average inverse labor supply elasticity of treated firms, i.e., $\hat{\epsilon}_{iF,k} \equiv \Delta \ln W_{F,k}/\Delta \ln n_{iF,k}$, across all markets with more than one firm (Column 1) and within different market subsets (Columns 2 to 4) in the estimated model. These estimates are compared to the reduced-form markdown estimates provided in Table 2. The procedure to obtain these estimates is detailed in Online Appendix D.4. Low and high self-employment rates are defined as being below or above the average self-employment rate across local labor markets, respectively. Bootstrap standard errors are in parentheses. These are obtained by redrawing the iid shock associated with the assignment of τ values 1,000 times.

elasticities are lower than the reduced-form estimates in Table 5, underscoring the importance of equilibrium responses from competitors. However, as discussed in Section 3.3, the bias in reduced-form estimates does not account for the heterogeneity observed across markets in terms of concentration or self-employment prevalence. Online Appendix Table A.7 supports this conclusion quantitatively. Additionally, Online Appendix Figure A.8 illustrates that no discernible pattern emerges when comparing structural and reduced-form labor market power against employer concentration and self-employment rates.

4.5 Robustness

Our structural approach to identifying labor market power relies on parametric assumptions. To validate the robustness of our findings, we explore several alternative parameterizations and their effect on labor market power and its variation across markets.

4.5.1 Heterogeneity in Roy Parameters

In our baseline specification, we assume a constant variance-covariance matrix and mean comparative advantage for workers' abilities across local labor markets. However, Online Appendix Figure A.6 shows that markets vary in their relative skill endowments, raising concerns that overlooking this heterogeneity may lead to biases. In particular, inspection of equations (19) and (20) reveals that variation in labor supply elasticity may be partially driven by these additional sources of heterogeneity, rather than self-employment shares alone.

To address this concern, we modify the model to allow for heterogeneity in both the variance-covariance matrix, $\boldsymbol{\Sigma_k}$, and the mean comparative advantage, $\hat{\mu}_k$. Specifically, we group markets into I clusters based on population quantiles. For each cluster $i = 1, \ldots, I$, we obtain the group-specific parameters $(\sigma_{F,i}, \sigma_{S,i}, \varrho_i, \hat{\mu}_i)$ as the within-group median. Markets in the model are then assigned to groups according to their population quantile. In the baseline model, we implicitly set $I = 1$. In the robustness exercise, we set $I = 3$.

Online Appendix Table A.8 presents the estimated Roy parameters at baseline and under group heterogeneity, showing minimal variation across groups. Column 2 of Online Appendix Table A.9 reports the sensitivity of our labor market power estimates to this robustness check, demonstrating that the overall incidence of labor market power and its relationship with market concentration and self-employment remains essentially unchanged. These findings support our decision to use constant parameters for the baseline calibration.

4.5.2 Heterogeneity in Fixed Costs

Firm entry costs into the wage sector represent barriers to starting a formal firm, including regulatory procedures, high licensing fees, limited access to credit, and inadequate infrastructure. These barriers are likely heterogeneous across local labor markets. In our baseline calibration, we assumed a constant cost structure across markets, implying that entry—and therefore market concentration—was fully proportional to variable profits and productivity. This assumption may influence labor market power estimates.

To address this, we recalibrate the model, allowing for an alternative parameterization of fixed costs, assuming each market draws f_k^e from a Weibull distribution with shape parameter f_κ and scale parameter f_λ. Column 3 of Online Appendix Table A.9 shows that our labor market power estimates remain robust under this assumption. Online Appendix Table A.10 demonstrates that this adjustment improves the model's ability to replicate the correlation between market concentration and earnings across local labor markets, particularly in highly concentrated markets where the baseline model underestimates earnings.

However, this adjustment has little impact on the model's ability to capture labor market power dynamics or its broader implications. While it resolves one specific issue, it also introduces additional assumptions about the relationship between fixed costs and productivity. Moreover, it worsens the model's fit in other dimensions–particularly in matching the number of firms across markets–making it a less appealing solution overall.

4.5.3 Full Entry Game

In the baseline model, we simplify the entry process to avoid the computational complexity of solving for the exact equilibrium values of M_k, instead using an approximation where firms are treated as infinitesimally small at the entry stage. To assess robustness, we recalibrate the model under the full entry game. Column 4 of Online Appendix Table A.9 shows that even with this more complex entry framework, our labor market power estimates–and their relationship with market concentration and self-employment–remain unchanged. This is reassuring but expected, as discussed in Section 4.1 since the entry assumption primarily affects equilibrium through the number of firms and market concentration directly targeted in calibration without altering the model's core outcomes.

4.5.4 Census Data

In our baseline model estimation, we target moments derived from the EEA firm-level data. Online Appendix B shows that the empirical facts presented in Section 2 are robust to using the 2007 Economic Census data to measure concentration and capture its variation across markets. We use those same data also to validate the model estimation. Specifically, we recalculate all moments using firm-level census data and the same local labor markets as in our baseline calibration. We exclude one-person firms from the census to omit self-employment. We then follow the procedure outlined in Section 4.2.2 and show the results in Column 5 of Online Appendix Table A.9. Using census data reduces average concentration across markets, lowering the average markdown from 1.45 to 1.31. Most importantly, the relationship between labor market power, market concentration, and self-employment remains robust.

5 Counterfactual Policy Analyses

Armed with the estimated model, we conduct two sets of counterfactual experiments to address our key research questions. First, we quantitatively assess the role of labor market power in shaping labor market outcomes in Peru. Second, we simulate three industrial policies aimed at promoting industrialization and increasing wage employment by targeting firms or workers. We measure their aggregate and distributional impacts and investigate how labor market power affects the success of these policies.

5.1 Impact of Labor Market Power

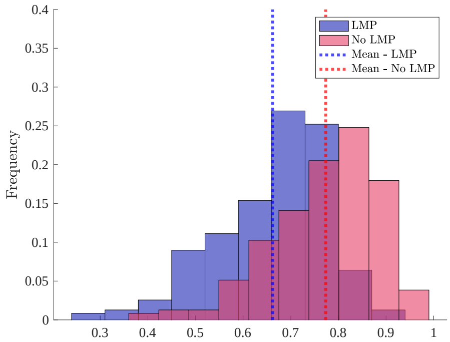
To investigate the impact of labor market power on labor market outcomes, we introduce a conduct parameter, $\iota \equiv \frac{dN_{F,k}}{dn_{iF,k}} = \{0, 1\}$, which captures the firm's perceived effect of its labor demand on aggregate market variables. When $\iota = 1$, firms are strategic and fully internalize the impact of their labor demand on aggregate labor demand and wages, as in our baseline model. Conversely, when $\iota = 0$, firms behave as wage-takers, irrespective of market concentration.

Under this generalization of market conduct, the firm i's markdown in market k becomes:

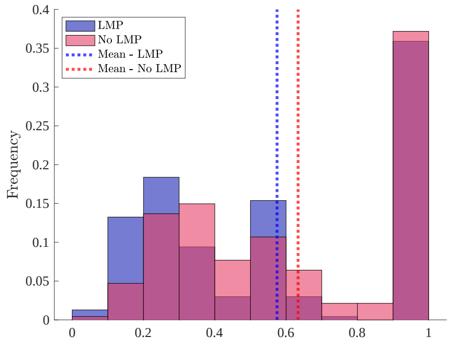
$$\psi_{iF,k} = 1 + \iota \frac{s_{iF,n}}{\epsilon(\hat{W}_k)}, \tag{25}$$

which highlights that the markdown equals one whenever firms act as wage takers.[26] We quantify the role of labor market power in the economy by comparing the baseline economy with one where $\iota = 0$, holding other things equal.

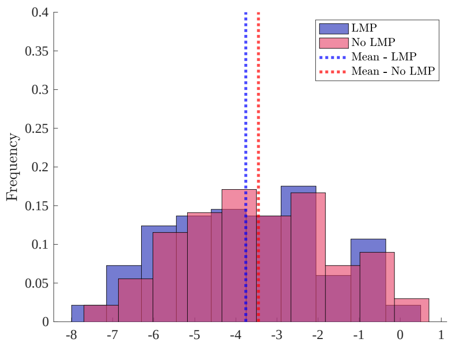
Figure 3: Effects of Labor Market Power



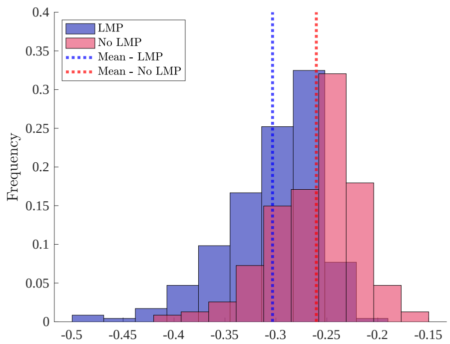
(a) Wage Share



(b) Payroll HHI



(c) Average Wages (Log)



(d) Earnings Gap (Log)

Notes. The four panels show the distribution of key labor market outcomes across markets in the baseline (blue) and in the counterfactual economy (red) with no labor market power. Online Appendix Table A.12 complements these figures by reporting the average of selected outcomes across markets in the two economies together with the difference between the two.

Figure 3 presents some key distributions of interest, while Online Appendix Table A.12 shows the average outcomes in the two economies. Panel (a) of Figure 3 compares the distribution of wage employment across markets, showing that labor market power significantly limits wage employment in Peru. With perfectly competitive labor markets, the share of wage employment would increase by over ten percentage points, from 66 to 77%.

Panel (b) shows that despite increased competition, concentration persists and even rises in the absence of labor market power due to market share reallocation. Without labor market

[26]The wage markdown of firm i in market k can be written as:

$$\psi_{iF,k} = 1 + \frac{d \ln W_{F,k}}{d \ln n_{iF,k}} = 1 + \frac{d \ln W_{F,k}}{d \ln N_{F,k}} \cdot \frac{dN_{F,k}}{dn_{iF,k}} \cdot \frac{n_{iF,k}}{N_{F,k}} = 1 + \iota \frac{s_{iF,n}}{\epsilon(\hat{W}_k)}.$$

power, the most productive firms gain market share, leading to higher concentration.[27]

Panel (c) shows that labor market power depresses average wages, while panel (d) reveals that it narrows the earnings gap between wage workers and the self-employed. Without labor market power, average earnings would rise by 31% in the wage employment sector and by 27% in the self-employment sector.

Drawing on insights from Section 3.3, we can decompose the changes in average earnings in both sectors into their key components.[28] The increase in average wages in the no-labor-market-power economy is entirely driven by higher earnings per efficiency unit, with the elimination of markdowns accounting for a 35% wage increase. This is partially offset by a 4% reduction in labor revenue productivity (MRPL). As market share shifts toward more productive, high-markup firms, output prices fall, accounting for most of the decline in MRPL. The selection channel has no impact, as expected from our parameter estimates.

In the self-employment sector, average earnings also rise, and about two-thirds of the increase is driven by higher unit earnings, fully attributable to higher output prices. Unlike in the wage sector, the selection channel plays a crucial role here. In the counterfactual economy, the average ability of self-employed workers is 11% higher, as more workers transition to wage employment, leaving the remaining self-employed workers increasingly positively selected.

Overall, the results in this section demonstrate that labor market power has a strong hold on the Peruvian economy. It contributes to the scarcity of wage jobs, lowers firm size, and reduces wages and self-employment earnings through markdowns, selection, and revenue productivity effects. Our findings also show that worker self-selection is a key factor through which labor market power in wage employment reduces earnings in the self-employment sector, thereby influencing the earnings gap between wage workers and the self-employed.

5.2 Industrial Policy

Despite long-standing efforts to increase wage employment as a means of promoting inclusive growth, policy interventions have often had limited impact (McKenzie, 2017; Bandiera et al., 2022). This section examines the extent to which the interaction between labor market power and self-employment contributes to this outcome. We simulate three industrial policy interventions—targeting firm productivity, worker productivity, and entry costs—and use our model to evaluate how labor market power affects their overall effectiveness, both qualitatively and quantitatively.

5.2.1 Firm Productivity

Policy efforts to boost firm productivity have often focused on market integration, primarily through infrastructure improvements (Fiorini, Sanfilippo and Sundaram, 2021). The goal is to

[27] See Eslava, Haltiwanger and Urdaneta (2024) for empirical evidence of this channel.

[28] These results are detailed in Online Appendix Table A.12.

expand market access, enhance productivity, and reduce both information frictions and shipping costs for inputs and outputs. To assess such interventions, we examine a road infrastructure project in Peru. Between 2003 and 2010, the country added over 5,000 kilometers of new roads, expanding the network by more than 10%. Volpe Martincus, Carballo and Cusolito (2017) evaluate this intervention and find that firm exports increased by an average of 3.7% as a result. Building on this evidence, we calibrate a shock in our model that shifts expected market-level productivity (T_k) to achieve a comparable increase in firm sales across markets.[29]

Although the shock is applied uniformly across markets, its effects vary significantly. Panel (a) of Figure 4 shows the estimated average impact across markets, segmented by quintiles of the baseline wage markdown distribution.[30] Wage employment shares and wages rise on average across all market groups, as intended by the policy. However, the rise in productivity does not fully translate into higher wages. The incomplete pass-through stems from an increase in wage markdowns, resulting from the policy's effects on its two key determinants: employer concentration and labor supply elasticity. As firm productivity rises, market entry increases, which reduces concentration. Yet, the labor supply elasticity also decreases as more workers opt for wage employment. The latter effect dominates in most markets, as shown by Online Appendix Figure A.10, leading to higher wage markdowns.[31]

Second, despite higher markdowns and stronger positive selection into self-employment, the earnings gap between wage and self-employed workers increases moderately. This occurs because the rising productivity of formal firms, combined with general equilibrium effects, more than offsets the opposing effects of markdowns and selection.

Third, and most importantly, the figure demonstrates that labor market power and its determinants strongly influence the policy's impact. Increasing firm productivity is significantly more effective in markets with lower baseline wage markdowns. This is true despite the higher baseline levels of wage employment shares and wages in these markets, as discussed in Section 5.1. The average increase in wage employment share and wages is 57% and 43% higher, respectively, in markets in the bottom quintile of the baseline wage markdown distribution compared to those in the top quintile.[32] Most notably, rising markdowns account for 98% of the variation in the policy's impact on wages and 85% of the variation in its effect on the wage employment share across markets.[33]

[29] We estimate that T_k must increase by 8% to reach the targeted sales growth.

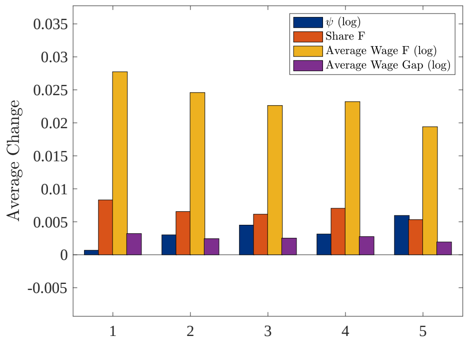
[30] Online Appendix Table A.13 provides average changes in selected outcomes across all markets.

[31] We also compare the effects of this policy in general equilibrium (GE) versus a partial equilibrium (PE) setting, where aggregate income is fixed. The results shown in Online Appendix Table A.11 reveal significant differences: in the PE model, a productivity shock leads to declining unit wages due to fixed demand and price reductions, causing lower wages in both sectors as workers reallocate. Conversely, the GE model shows that falling prices boost aggregate demand, increasing the marginal revenue product of labor and raising wages across all sectors. This underscores the importance of GE structure in accurately capturing wage responses, as the GE model significantly mitigates the negative wage impacts observed in the PE scenario.

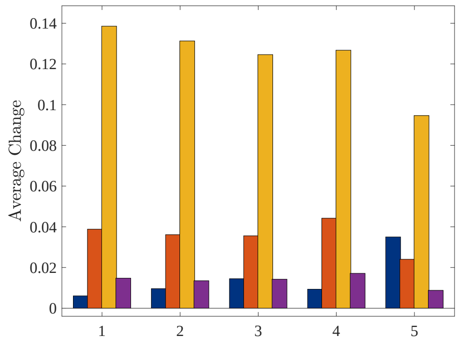
[32] Wages, for instance, increase by 1.94% on average in the least competitive labor markets and by 2.77% in the most competitive ones, so $(2.77 - 1.94)/1.94 = 0.43$.

[33] These figures are derived from the R^2 of a simple regression implemented on simulated data, where the local labor market is the unit of observation. The dependent variable is the change in average wage or wage employment

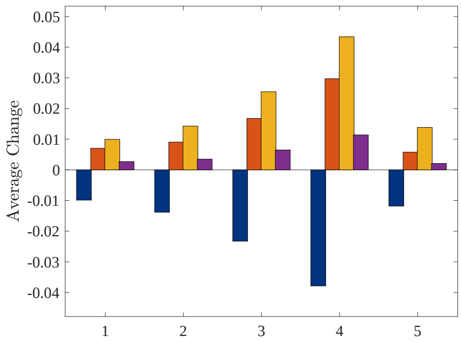
Figure 4: Effects of Policy Shocks Across Markets



(a) Firm Productivity



(b) Worker Skills



(c) Fixed Cost

Notes. The three panels illustrate the estimated change in wage markdown, wage employment share, average wage, and average earnings gap between wage workers and self-employed workers across local labor markets resulting from the three policy experiments. It does so for separate bins determined by the size of the wage markdown at baseline. Online Appendix Figure A.10 complements these figures by showing the change in the wage markdown and its determinants, i.e., employer concentration and wage labor supply elasticity.

5.2.2 Worker Skills

Next, we examine policies aimed at boosting the supply of wage labor by enhancing worker skills through targeted training programs. These initiatives operate on the belief that unemployment stems from a lack of specific technical skills, which can be addressed through short-term training (McKenzie, 2017). Several such programs have been implemented across Latin America, including Peru's Job Youth Training Program, known as *Projoven*. Operating from 1996 to 2010, the program aimed to equip young people from low-income backgrounds with training and labor market experience aligned with the needs of the productive sector, catering directly to employer demands. An experimental evaluation of Projoven by Díaz and Rosas-Shady (2016) found that, two years after completing the program, participants had a 3.6 percentage point higher probability of securing wage employment compared to the control group, although the

share following the policy shock, and the independent variable is the change in the wage markdown.

result was not statistically significant.[34] To simulate a similar training program in our model, we introduce a shock to workers' mean comparative advantage in wage employment, $\hat{\mu}$.[35]

Panel (b) of Figure 4 shows that, by improving worker skills, the program raises productivity in the formal sector, similar to a productivity shock to firms. This reduces concentration and increases wage employment. Additionally, as worker skills improve, the mean comparative advantage in wage employment increases, making self-employment less attractive and decreasing the supply elasticity of wage labor, as shown in Online Appendix Figure A.10. This reduction in elasticity outweighs the decline in concentration, resulting in higher wage markdowns. Despite this, wages still increase as skills improve. Notably, the model closely replicates the 13.4% rise in monthly earnings reported by Díaz and Rosas-Shady (2016).

Labor market power and its changes are crucial in shaping the policy's impact. The average effects on wage employment and wages are 62% and 47% higher, respectively, in the most competitive labor markets compared to the least competitive ones. As with policies that boost firm productivity, rising markdowns account for 98% of the variation in the policy's effect on average wages and 85% of the variation in its impact on the wage employment share.

5.2.3 Firm Entry Cost

Policies to reduce entry costs typically involve government programs that simplify entry regulations. A prime example is the Mexican Rapid Business Opening System (SARE), which aimed to streamline local business registration procedures across various municipalities starting in May 2002. Both Kaplan, Piedra and Seira (2011) and Bruhn (2011) evaluate the impact of this reform on several economic outcomes at the municipality level.[36] For our policy experiment, we target the 2.2% increase in the fraction of wage earners documented in Bruhn (2011).[37]

Panel (c) of Figure 4 shows the policy impact across markets. The policy encourages firm entry and reduces concentration. Competition intensifies in the output and labor market, leading to lower markups and prices.[38] Wage employment increases, but the effect is modest compared to the large reduction in concentration. This is because the drop in concentration stems from the entry of relatively unproductive firms, which have little impact on aggregate labor demand. As a result, the policy reduces concentration more than it affects labor supply elasticity, and wage markdowns decrease, as shown in Online Appendix Figure A.10.

The policy's impact is highly heterogeneous across markets. However, unlike the firm pro-

[34] However, significant positive effects were found in the likelihood of obtaining formal employment, such as jobs with health insurance and pensions.

[35] To increase average wage employment by 3.6 percentage points, $\hat{\mu}$ must rise from -0.12 to 0.04 across all markets.

[36] The main difference between the two studies is that Bruhn (2011) uses household data from labor market surveys, whereas Kaplan, Piedra and Seira (2011) uses social security data.

[37] We find that a 40% reduction in fixed costs across local labor markets is required to achieve an average 2.2% increase in wage employment in our model.

[38] Remarkably, the model replicates the 1% decrease in prices found by Bruhn (2011). Specifically, this refers to the average change across markets in the price index in the wage employment sector, $P_{F,k}$, as reported in Online Appendix Table A.13.

ductivity shock, reducing entry costs is the least effective in the most competitive labor markets. It is also relatively ineffective in the most concentrated and least competitive markets. In markets with moderately high markdowns, reducing entry costs is more effective because concentration is on the margin more responsive to the negative cost shift.

Yet, as in the previous two cases, changes in labor market power are crucial to understanding the effects of reducing fixed costs. Changes in markdowns explain 99% of the variation in the policy's impact on wages and 88% of the variation in its effect on the wage employment share.

5.2.4 Discussion

These counterfactual exercises show that our framework provides a valuable perspective for understanding the impact of industrialization policies. The effectiveness of these policies is closely tied to labor market power and its key drivers—employer concentration and labor supply elasticity. While these policies help create more wage jobs, they also make self-employment less attractive, reducing labor supply elasticity. In some cases, this means that pro-competitive policies can have the unintended effect of making the labor market less competitive. For productivity-enhancing policies, the drop in elasticity outweighs the reduction in concentration, leading to higher wage markdowns and weakening the policy's overall impact.

In contrast, policies that lower entry costs reduce both concentration and markdowns, making them more effective in markets with moderate labor market power. In all cases, markdown changes largely explain the variation in policy outcomes across markets. These insights are crucial for researchers and policymakers seeking to better understand labor market dynamics and develop effective strategies for industrialization and inclusive growth.

6 Conclusions

Addressing the scarcity of good jobs in developing countries remains a pressing challenge for policymakers. This paper draws on new evidence from Peru, a general equilibrium model, and counterfactual analyses to examine the crucial role of labor market power and its interaction with self-employment in shaping labor market outcomes, informing effective policy design.

Our findings show that self-employment in manufacturing plays a dual role in the presence of labor market power. On the one hand, it provides workers with an outside option when wage employment is scarce or unattractive, limiting employers' ability to suppress wages. On the other hand, policies aimed at expanding wage employment and reducing reliance on self-employment may inadvertently strengthen firms' wage-setting power, ultimately undermining their intended objectives.

More broadly, our results challenge the conventional view of self-employment as a perfectly elastic labor reservoir for industrialization (Lewis, 1954; Rauch, 1991). Instead, we demonstrate that labor market power can impede the efficient movement of workers into the modern industrial sector, leading to an over-reliance on self-employment that may slow economic development. These findings highlight the need for industrial policies that explicitly account for labor market frictions to be effective (Donovan and Schoellman, 2023).

An important caveat is that our analysis focuses on manufacturing, which accounts for a relatively small share of employment and value added in most economies. However, manufacturing has historically played a central role in economic development and remains a key target of industrial policy. In many low-income countries, manufacturing value added has grown without a corresponding increase in large-firm employment, while self-employment remains persistently high (Alfaro et al., 2023; Huneeus and Rogerson, 2023). This divergence, despite targeted industrial policies (Juhász, Lane and Rodrik, 2023), underscores the need to better understand the barriers to manufacturing wage employment expansion.

At the same time, our findings—particularly on the dual role of self-employment under labor market power—are relevant beyond manufacturing and low-income countries. The digital economy has transformed labor markets in wealthier nations, increasing self-employment and flexible work arrangements while reducing reliance on traditional jobs (Mas and Pallais, 2017; Katz and Krueger, 2019), yet firms continue to exert significant labor market power. Understanding how these trends interact is an important avenue for future research with meaningful policy implications. Additionally, extending our analysis to sectors such as agriculture and services, where self-employment is more prevalent, would provide further insights into labor markets in developing countries.

References

Abeberese, Ama Baafra, Charles Godfred Ackah, and Patrick Opoku Asuming. 2019. "Productivity Losses and Firm Responses to Electricity Shortages: Evidence from Ghana." *The World Bank Economic Review*, 35(1): 1–18.

Adão, Rodrigo. 2016. "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from brazil." Mimeo.

Alfaro, Laura, Marcela Eslava, Martí Mestieri, and Luis Felipe Saenz. 2023. "Jobless Industrialization." Mimeo.

Alfaro-Ureña, Alonso, Isabela Manelici, and Jose P. Vasquez. 2021. "The Effects of Multinationals on Workers: Evidence from Costa Rican Microdata." Mimeo.

Alfonsi, Livia, Oriana Bandiera, Vittorio Bassi, Robin Burgess, Imran Rasul, Munshi Sulaiman, and Anna Vitali. 2020. "Tackling youth unemployment: Evidence from a labor market experiment in Uganda." *Econometrica*, 88(6): 2369–2414.

Amodio, Francesco, and Nicolas De Roux. 2022. "Measuring Labor Market Power in Developing Countries: Evidence from Colombian Plants." *Journal of Labor Economics*. Forthcoming.

Amodio, Francesco, Emanuele Brancati, Peter Brummond, Nicolas de Roux, and Michele Di Maio. 2024. "Global Labor Market Power." CEPR Discussion Papers 18828.

Amodio, Francesco, Francisco Alvarez-Cuadrado, and Markus Poschke. 2020. "Selection and Absolute Advantage in Farming and Entrepreneurship." CEPR Discussion Papers 14269.

Armangué-Jubert, Tristany, Nezih Guner, and Alessandro Ruggieri. 2024. "Labor Market Power and Development." *American Economic Review: Insights*. Forthcoming.

Atencio-De-Leon, Andrea, Munseob Lee, and Claudia Macaluso. 2023. "Does Turnover Inhibit Specialization? Evidence from a Skill Survey in Peru." Institute of Labor Economics (IZA) IZA Discussion Papers 16671.

Atkeson, Andrew, and Ariel Burstein. 2008. "Pricing-to-Market, Trade Costs, and International Relative Prices." *American Economic Review*, 98(5): 1998–2031.

Azar, José, Ioana Marinescu, and Marshall Steinbaum. 2022. "Labor Market Concentration." *Journal of Human Resources*, 57(S): 167–199.

Bandiera, Oriana, Ahmed Elsayed, Andrea Smurra, and Céline Zipfel. 2022. "Young Adults and Labor Markets in Africa." *Journal of Economic Perspectives*, 36(1): 81–100.

Bassier, Ihsaan. 2023. "Firms and inequality when unemployment is high." *Journal of Development Economics*, 161: 103029.

Bassier, Ihsaan, Arindrajit Dube, and Suresh Naidu. 2022. "Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies." *Journal of Human Resources*, 57(S): 50–86.

Bassi, Vittorio, Jung Hyuk Lee, Alessandra Peter, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2023. "Self-Employment Within the Firm." National Bureau of Economic Research, Inc NBER Working Papers 31740.

Bau, Natalie, and Adrien Matray. 2023. "Misallocation and Capital Market Integration: Evidence

From India." *Econometrica*, 91(1): 67–106.

Benmelech, Efraim, Nittai K. Bergman, and Hyunseob Kim. 2022. "Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages?" *Journal of Human Resources*, 57(S): 200–250.

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor Market Power." *American Economic Review*, 112(4): 1147–93.

Blattman, Christopher, and Stefan Dercon. 2018. "The Impacts of Industrial and Entrepreneurial Work on Income and Health: Experimental Evidence from Ethiopia." *American Economic Journal: Applied Economics*, 10(3): 1–38.

Breza, Emily, Supreet Kaur, and Yogita Shamdasani. 2021. "Labor Rationing." *American Economic Review*, 111(10): 3184–3224.

Brooks, Wyatt J., Joseph P. Kaboski, Yao Amber Li, and Wei Qian. 2021. "Exploitation of labor? Classical monopsony power and labor's share." *Journal of Development Economics*, 150: 102627.

Bruhn, Miriam. 2011. "License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico." *The Review of Economics and Statistics*, 93(1): 382–386.

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

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel. 2019. "Changes in Between-Group Inequality: Computers, Occupations, and International Trade." *American Economic Journal: Macroeconomics*, 11(2): 348–400.

Dasso, Rosamaría, and Fernando Fernandez. 2015*a*. "Dataset for: The effects of electrification on employment in rural Peru." https://izajold.springeropen.com/articles/10.1186/s40175-015-0028-4, IZA Journal of Labor & Development.

Dasso, Rosamaría, and Fernando Fernandez. 2015*b*. "The effects of electrification on employment in rural Peru." *IZA Journal of Labor & Development*, 4(6).

Dasso, Rosamaría, Fernando Fernandez, and Hugo R. Ñopo. 2015. "Electrification and Educational Outcomes in Rural Peru." Institute of Labor Economics (IZA) IZA Discussion Papers 8928.

Deb, Shubhdeep, Jan Eeckhout, Aseem Patel, and Lawrence Warren. 2022. "What drives Wage Stagnation: Monopsony or Monopoly?" *Journal of the European Economic Association*. jvac060.

Díaz, Juan José, and David Rosas-Shady. 2016. "Impact Evaluation of the Job Youth Training Program Projoven." Inter-American Development Bank Working Paper IDB-WP-693.

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

Dix-Carneiro, Rafael, Pinelopi Koujianou Goldberg, Costas Meghir, and Gabriel Ulyssea. 2024. "Trade and Domestic Distortions: The Case of Informality." Cowles Foundation for Research in Economics, Yale University Cowles Foundation Discussion Papers 2384.

Djankov, Simeon, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. 2002. "The regulation of entry." *The Quarterly Journal of Economics*, 117(1): 1–37.

Donovan, Kevin, and Todd Schoellman. 2023. "The role of labor market frictions in structural transformation." *Oxford Development Studies*, 51(4): 362–374.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman. 2023. "Labor Market Dynamics and Development." *The Quarterly Journal of Economics*, 138(4): 2287–2325.

Duranton, Gilles, and Henry G. Overman. 2005. "Testing for Localization Using Micro-Geographic Data." *The Review of Economic Studies*, 72(4): 1077–1106.

Eaton, Jonathan, Samuel S. Kortum, and Sebastian Sotelo. 2012. "International Trade: Linking Micro and Macro." NBER Working Papers 17864.

Ellison, Glenn, and Edward L. Glaeser. 1997. "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach." *Journal of Political Economy*, 105(5): 889–927.

Eslava, Marcela, John Haltiwanger, and Nicolas Urdaneta. 2024. "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand, and Wedges." *The Review of Economic Studies*, 91(1): 259–300.

Eslava, Marcela, Marcela Melendez Arjona, Laura Daniela Tenjo Galvis, and Nicolás Urdaneta. 2023. "Business Size, Development, and Inequality in Latin America : A Tale of One Tail." The World Bank Policy Research Working Paper Series 10584.

Eslava, Marcela, Marcela Meléndez, Gabriel Ulyssea, Nicolás Urdaneta, and Ignacio Flores. 2024. "Firms and Inequality in Latin America." Inter-American Development Bank IDB Publications (Working Papers) 13480.

Felix, Mayara. 2022. "Trade, Labor Market Concentration, and Wages." Mimeo.

Fields, Gary. 2012. *Working Hard, Working Poor: A Global Journey.* Oxford University Press.

Fiorini, Matteo, Marco Sanfilippo, and Asha Sundaram. 2021. "Trade liberalization, roads and firm productivity." *Journal of Development Economics*, 153(C).

French, Eric, and Christopher Taber. 2011. "Identification of models of the labor market." *Handbook of labor economics*, 4: 537–617.

Gaubert, Cecile, and Oleg Itskhoki. 2021. "Granular comparative advantage." *Journal of Political Economy*, 129(3): 871–939.

Gollin, Douglas. 2008. "Nobody's business but my own: Self-employment and small enterprise in economic development." *Journal of Monetary Economics*, 55(2): 219 – 233.

Gutiérrez, Agustín. 2023. "Labor Market Power and the Pro-competitive Gains from Trade." Mimeo.

Harris, John R., and Michael P. Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *The American Economic Review*, 60(1): 126–142.

Heckman, James J., and Bo E. Honoré. 1990. "The Empirical Content of the Roy Model." *Econometrica*, 58(5): 1121–1149.

Heckman, James J, and Guilherme Sedlacek. 1985. "Heterogeneity, aggregation, and market wage functions: an empirical model of self-selection in the labor market." *Journal of Political Economy*, 93(6): 1077–1125.

Hjort, Jonas, Hannes Malmberg, and Todd Schoellman. 2022. "The Missing Middle Managers: Labor Costs, Firm Structure, and Development." NBER Working Papers 30592.

Hsieh, Chang-Tai, and Benjamin A Olken. 2014. "The missing "missing middle"." *Journal of Economic Perspectives*, 28(3): 89–108.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow. 2019. "The Allocation of Talent and U.S. Economic Growth." *Econometrica*, 87(5): 1439–1474.

Huang, Hanwei, Kalina Manova, Oscar Perello, and Frank Pisch. 2024. "Firm Heterogeneity and Imperfect Competition in Global Production Networks." Mimeo.

Huneeus, Federico, and Richard Rogerson. 2023. "Heterogeneous Paths of Industrialization." *The Review of Economic Studies*, 91(3): 1746–1774.

Instituto Nacional de Estadística e Informática. 2004–2011*a*. "Encuesta Económica Anual (EEA)." https://proyectos.inei.gob.pe/microdatos/, Lima, Peru. Downloaded on 2018.

Instituto Nacional de Estadística e Informática. 2004-2011*b*. "Encuesta Nacional de Hogares (ENAHO)." https://proyectos.inei.gob.pe/microdatos/, Lima, Peru. Downloaded on 2018.

Instituto Nacional de Estadística e Informática. 2007. "Censo Económico (CENEC)." Lima, Peru.

International Labour Organization. 2020. "ILO modelled estimates database."

Juhász, Réka, Nathan J. Lane, and Dani Rodrik. 2023. "The New Economics of Industrial Policy." NBER Working Papers 31538.

Kaboski, Joseph P., and Robert M. Townsend. 2011. "A structural evaluation of a large-scale quasi-experimental microfinance initiative." *Econometrica*, 79(5): 1357–1406.

Kaplan, David S., Eduardo Piedra, and Enrique Seira. 2011. "Entry regulation and business start-ups: Evidence from Mexico." *Journal of Public Economics*, 95(11): 1501–1515.

Katz, Lawrence F, and Alan B Krueger. 2019. "The rise and nature of alternative work arrangements in the United States, 1995–2015." *ILR review*, 72(2): 382–416.

Kroft, Kory, Fabian Lange, Matthew J. Notowidigdo, and Matthew Tudball. 2020. "Long Time Out: Unemployment and Joblessness in Canada and the United States." *Journal of Labor Economics*, 38(S1): S255–S297.

Lamadon, Thibaut, Magne Mogstad, and Bradley Setzler. 2019. "Imperfect Competition, Compensating Differentials, and Rent Sharing in the U.S. Labor Market." National Bureau of Economic Research NBER Working Paper 25954.

La Porta, Rafael, and Andrei Shleifer. 2014. "Informality and Development." *Journal of Economic Perspectives*, 28(3): 109–26.

Lewis, W. Arthur. 1954. "Economic Development with Unlimited Supplies of Labour." *The Manchester School*, 22(2): 139–191.

MacKenzie, Gaelan. 2021. "Trade and Market Power in Product and Labor Markets." Bank of Canada Staff Working Papers 21-17.

Maloney, William F. 1999. “Does informality imply segmentation in urban labor markets? Evidence from sectoral transitions in Mexico.” *The World Bank Economic Review*, 13(2): 275–302.

Manning, Alan. 2003. *Monopsony in Motion: Imperfect Competition in Labor Markets.* Princeton University Press.

Mas, Alexandre, and Amanda Pallais. 2017. “Valuing alternative work arrangements.” *American Economic Review*, 107(12): 3722–59.

McKenzie, David. 2017. “How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence.” *The World Bank Research Observer*, 32(2): 127–154.

McMillan, Margaret, and Albert Zeufack. 2022. “Labor Productivity Growth and Industrialization in Africa.” *Journal of Economic Perspectives*, 36(1): 3–32.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2023. “General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence From India.” *Econometrica*, 91(4): 1261–1295.

Piselli, Roberto. 2013. “Patrones de transporte en Lima Metropolitana: adónde, cuánto y por qué viajan los limenos,.” *Revista Argumentos*, 25–28.

Ponczek, Vladimir, and Gabriel Ulyssea. 2021. “Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil.” *The Economic Journal*. ueab052.

Poschke, Markus. 2022. “Wage employment, unemployment and self-employment across countries.” Mimeo.

Rauch, James. 1991. “Modelling the informal sector formally.” *Journal of Development Economics*, 35(1): 33–47.

Roy, Andrew Donald. 1951. “Some thoughts on the distribution of earnings.” *Oxford economic papers*, 3(2): 135–146.

Rud, Juan Pablo, and Ija Trapeznikova. 2021. “Job creation and wages in least developed countries: Evidence from sub-saharan africa.” *The Economic Journal*, 131(635): 1331–1364.

Ulyssea, Gabriel. 2018. “Firms, Informality, and Development: Theory and Evidence from Brazil.” *American Economic Review*, 108(8): 2015–47.

Ulyssea, Gabriel. 2020. “Informality: Causes and Consequences for Development.” *Annual Review of Economics*, 12(1): 525–546.

UN General Assembly. 2015. “Transforming our world: the 2030 Agenda for Sustainable Development.” *United Nations: New York, NY, USA*.

Volpe Martincus, Christian, Jerónimo Carballo, and Ana Cusolito. 2017. “Roads, exports and employment: Evidence from a developing country.” *Journal of Development Economics*, 125(C): 21–39.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein. 2022. “Monopsony in the US Labor Market.” *American Economic Review*, 112(7): 2099–2138.

Online Appendix

Labor Market Power, Self-Employment, and Development

Francesco Amodio, Pamela Medina, Monica Morlacco*

This Appendix is organized as follows. Section A contains the additional tables and figures discussed in the text. Section B shows that the empirical facts shown in Section 2 are robust to the use of alternative concentration measures and variable definitions. Section C integrates Section 3 by showing additional theoretical results. Section D complements Section 4 and provides further details on model estimation.

*Amodio: francesco.amodio@mcgill.ca, McGill University and IZA, Department of Economics and ISID, 855 Sherbrooke St. West, Montreal, QC, H3A 2T7; Medina: pamela.medinaquispe@rotman.utoronto.ca, University of Toronto, UTSC and Rotman School of Management, 105 St George St, Toronto, ON M5S 3E6; Morlacco: morlacco@usc.edu, University of Southern California, Department of Economics, 3620 S. Vermont Avenue, Los Angeles, CA 90089-1422.

A Additional Tables and Figures

Table A.1: Employment Distribution Across Sectors and Transitions Across Wage Work and Self-Employment

	All Workers	Wage Workers			Self-Employed		
		All	Self-Employed at $t-1$	Self-Employed Manuf. at $t-1$	All	Wage Workers at $t-1$	Wage Workers Manuf. at $t-1$
Agriculture	0.36	0.15	0.25	0.12	0.42	0.37	0.12
Mining	0.01	0.03	0.03	0.02	0.01	0.01	0.01
Manufacturing	0.08	0.10	0.08	0.31	0.08	0.07	0.28
Utilities	0.00	0.01	0.01	0.02	0.00	0.00	0.00
Construction	0.05	0.09	0.14	0.12	0.03	0.08	0.03
Retails	0.14	0.05	0.04	0.05	0.22	0.14	0.20
Transportation	0.05	0.03	0.04	0.02	0.07	0.10	0.11
Other Services	0.30	0.56	0.42	0.35	0.17	0.24	0.25
Observations	29970	11008	1736	113	15709	1746	162

Notes. This table shows the distribution of employment across sectors amongst all workers, wage workers, and self-employed workers in the 2007-2011 panel version of ENAHO. It reports these distributions at the end of each two-year period. It also shows the distribution of wage work (and self-employment) among workers who were self-employed (wage workers) in the previous year, across all sectors and in manufacturing only.

Table A.2: Concentration, Self-Employment, and Earnings

	Self-Employed {0,1}			Log of Earnings					
				Wage Workers			Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	0.050***	0.049***	0.062***	-0.085***	-0.100***	-0.052**	-0.124**	-0.158***	-0.051
(Log of Wage-bill HHI)	(0.012)	(0.014)	(0.015)	(0.020)	(0.021)	(0.021)	(0.057)	(0.048)	(0.052)
Female	0.122***	0.121***	0.111***	-0.426***	-0.381***	-0.382***	-1.313***	-1.228***	-1.211***
	(0.023)	(0.023)	(0.023)	(0.041)	(0.036)	(0.035)	(0.072)	(0.075)	(0.078)
Age	0.018***	0.017***	0.016***	0.023**	0.027***	0.029***	0.116***	0.115***	0.110***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)	(0.018)	(0.017)	(0.018)
Age sq.	-0.000	-0.000	-0.000	-0.000	-0.000**	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.006	-0.001	0.000	0.178***	0.161***	0.156***	0.111***	0.109***	0.101***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.007)	(0.018)	(0.017)	(0.019)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7637	7637	7634	4707	4706	4698	2054	2054	2047
R^2	0.102	0.132	0.156	0.308	0.363	0.395	0.327	0.383	0.399

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. Unit of observation is a working-age individual surveyed in ENAHO. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates and their standard errors obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln HHI^{wn}_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln HHI^{wn}_{(j,g)t}$ is the log of wage-bill HHI in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Table A.3: Estimates of Labor Market Power – First Stage Regression Results

$HHI^{wn} \in$	$(0,1]$	$(0,0.18]$	$(0.18,0.25]$	$(0.25,1]$	$(0,0.25]$	$(0,0.25]$	$(0.25,1]$	$(0.25,1]$
Self-employment Rate	All	All	All	All	Low	High	Low	High
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Log of Employment							
$PER_{gt} \times EC_{i(j,g)}$	0.005***							
	(0.000)							
$\times \mathbb{I}\{HHI^{wn} \in (0,0.18]\}$		-0.001	0.002***	0.004***				
		(0.002)	(0.001)	(0.001)				
$\times \mathbb{I}\{HHI^{wn} \in (0.18,0.25]\}$		0.019***	-0.026***	0.011***				
		(0.001)	(0.001)	(0.002)				
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\}$		0.012***	-0.003**	-0.006***				
		(0.002)	(0.001)	(0.002)				
$\times \mathbb{I}\{HHI^{wn} \in (0,0.25]\} \times \mathbb{I}\{\text{Low SE}\}$					-0.010***	0.009***	0.003***	0.002**
					(0.001)	(0.000)	(0.000)	(0.001)
$\times \mathbb{I}\{HHI^{wn} \in (0,0.25]\} \times \mathbb{I}\{\text{High SE}\}$					0.012***	-0.010***	0.003***	0.002**
					(0.001)	(0.001)	(0.000)	(0.001)
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\} \times \mathbb{I}\{\text{Low SE}\}$					0.001	0.003	-0.009***	0.010***
					(0.001)	(0.003)	(0.002)	(0.003)
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\} \times \mathbb{I}\{\text{High SE}\}$					0.010***	0.011***	0.006***	-0.026***
					(0.002)	(0.002)	(0.002)	(0.003)
Observations	6191	6191	6191	6191	6191	6191	6191	6191
R^2	0.952	0.974	0.959	0.979	0.966	0.957	0.973	0.971

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. Unit of observation is a medium to large firm in EEA. The table reports the first stage estimates corresponding to the second stage estimates reported in Table 2. The dependent variable is the log of firm-level employment $\ln l_{i(j,g)t}$. The instrumental variable is the interaction of the cumulative number of PER projects completed in each location g up to year t (PER_{gt}) and a dummy equal to one for firms with higher than median constraints to access electricity at baseline ($EC_{i(j,g)}$). Column 1 reports the first-stage estimates associated with column 1 of Table 2. Columns 2 to 4 report the first-stage results from the three first-stage regressions associated with column 2 of Table 2. Columns 5 to 8 report those from the four first-stage regressions associated with column 3 and 4 of Table 2. Following equation 3, firm fixed effects and local labor market × year fixed effects are included in all specifications. Standard errors are clustered at the level of location g, i.e. province or commuting zone.

Table A.4: Estimates of Labor Market Power
Robustness to District $\times$ Industry $\times$ Year Fixed Effects

			Self-Employment Rate	
			Low	High
	(1)	(2)	(3)	(4)
All Markets	0.565***			
	(0.105)			
$HHI^{wn} \in (0, 0.18]$		-0.062		
		(0.064)		
$HHI^{wn} \in (0.18, 0.25]$		0.450***		
		(0.142)		
$HHI^{wn} \in (0, 0.25]$			-0.127	-0.041
			(0.309)	(0.279)
$HHI^{wn} \in (0.25, 1]$		0.861***	1.725***	-0.502
		(0.050)	(0.365)	(0.643)
SW F-statistics	376.08	848.96	282.72	1065.85
		1545.52	446.12	1211.87
		14295.91		
Observations	4954	4954	3257	1697

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is a medium to a large firm in EEA. The table reports 2SLS estimates of the firm-level inverse elasticity of supply of wage labor as captured by β in equation (1). The instrumental variable is the interaction of the cumulative number of PER projects completed in each location g up to year t (PER_{gt}) and a dummy equal to one for firms with higher than median constraints to accessing electricity at baseline ($EC_{i(j,g)}$). Estimates in Columns 2 to 4 are obtained by interacting both the log of firm-level employment $\ln l_{i(j,g)t}$ and the instrument $PER_{gt} \times EC_{i(j,g)}$ with dummy variables that identify the different subsamples as discussed in the text. Low and high self-employment rates are defined as below and above the average self-employment rate across local labor markets, respectively. We report the F-statistic associated with the Sanderson-Windmeijer multivariate test of excluded instruments for each estimate. Firm fixed effects and district (instead of province or commuting zone as in baseline) $\times$ industry $\times$ year fixed effects are included in all specifications. Standard errors are clustered at the level of location g, i.e., province or commuting zone.

Table A.5: Estimates of Labor Market Power – First Stage Regression Results
Robustness to District $\times$ Industry $\times$ Year Fixed Effects

$HHI^{wn} \in$ Self-employment Rate	$(0,1]$ All (1)	$(0,0.18]$ All (2)	$(0.18,0.25]$ All (3)	$(0.25,1]$ All (4)	$(0,0.25]$ Low (5)	$(0,0.25]$ High (6)	$(0.25,1]$ Low (7)	$(0.25,1]$ High (8)
	Log of Employment							
$PER_{gt} \times EC_{i(j,g)}$	0.006***							
	(0.000)							
$\times \mathbb{I}\{HHI^{wn} \in (0,0.18]\}$		-0.000	0.002***	0.004***				
		(0.001)	(0.000)	(0.000)				
$\times \mathbb{I}\{HHI^{wn} \in (0.18,0.25]\}$		0.024***	-0.032***	0.014***				
		(0.003)	(0.004)	(0.001)				
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\}$		0.014***	0.001	-0.005***				
		(0.001)	(0.002)	(0.001)				
$\times \mathbb{I}\{HHI^{wn} \in (0,0.25]\} \times \mathbb{I}\{\text{Low SE}\}$					-0.008***	0.009***	0.003***	0.002***
					(0.000)	(0.000)	(0.000)	(0.000)
$\times \mathbb{I}\{HHI^{wn} \in (0,0.25]\} \times \mathbb{I}\{\text{High SE}\}$					0.013***	-0.012***	0.003***	0.003***
					(0.001)	(0.000)	(0.000)	(0.000)
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\} \times \mathbb{I}\{\text{Low SE}\}$					0.003*	0.008***	-0.002	0.003
					(0.001)	(0.001)	(0.002)	(0.003)
$\times \mathbb{I}\{HHI^{wn} \in (0.25,1]\} \times \mathbb{I}\{\text{High SE}\}$					0.013***	0.016***	-0.002	-0.019***
					(0.001)	(0.001)	(0.002)	(0.003)
Observations	4954	4954	4954	4954	4954	4954	4954	4954
R^2	0.954	0.980	0.970	0.984	0.973	0.967	0.980	0.979

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. Unit of observation is a medium to large firm in EEA. The table reports the first stage estimates corresponding to the second stage estimates reported in Table 2. The dependent variable is the log of firm-level employment $\ln l_{i(j,g)t}$. The instrumental variable is the interaction of the cumulative number of PER projects completed in each location g up to year t (PER_{gt}) and a dummy equal to one for firms with higher than median constraints to access electricity at baseline ($EC_{i(j,g)}$). Column 1 reports the first-stage estimates associated with column 1 of Online Appendix Table A.4. Columns 2 to 4 report the first-stage results from the three first-stage regressions associated with column 2 of Online Appendix Table A.4. Columns 5 to 8 report those from the four first-stage regressions associated with column 3 and 4 of Online Appendix Table A.4. Firm fixed effects and district (instead of province or commuting zone as in baseline) $\times$ industry $\times$ year fixed effects are included in all specifications. Standard errors are clustered at the level of location g, i.e. province or commuting zone.

Table A.6: Employer Concentration Across Local Labor Markets
Alternative Samples and Definitions

	Full Sample		*Merged Sample*	
	Mean	St. Dev.	Mean	St. Dev.
Number of Firms	6.39	10.37	7.25	11.19
Wage-bill HHI	0.65	0.33	0.61	0.34
Wage-bill HHI (Payroll Weighted)	0.37	0.03	0.37	0.03
Wage-bill HHI (Employment Weighted)	0.34	0.03	0.34	0.03
Employment HHI	0.63	0.35	0.59	0.35
Employment HHI (Payroll Weighted)	0.33	0.03	0.33	0.03
Employment HHI (Employment Weighted)	0.31	0.02	0.31	0.02
Percent of LLMs with 1 firm	38.78	2.27	38.78	2.29
Payroll Share of LLMs with 1 firm	7.94	1.79	7.96	1.81
Employment Share of LLMs with 1 firm	7.80	1.23	7.81	1.25
Number of Local Labor Markets	280		228	
Number of Locations	61		48	
Industries	23		22	

Notes. This table presents summary statistics and employer concentration measures derived from EEA firm-level data across Peruvian local labor markets, averaged over the years 2004 to 2011. The data are shown separately for the entire sample and the subset merged with worker-level data from ENAHO. In Local labor markets are defined by 2-digit industries within locations, with locations corresponding to Peruvian provinces or commuting zones.

Table A.7: Structural Inverse Supply Elasticity – Model Estimates

			Self-Empl. Rate	
			Low	High
	(1)	(2)	(3)	(4)
All Markets	0.298			
$HHI^{wn} \in (0, 0.18]$		0.125		
$HHI^{wn} \in (0.18, 0.25]$		0.191		
$HHI^{wn} \in (0, 0.25]$			0.182	0.129
$HHI^{wn} \in (0.25, 1]$		0.373	0.461	0.267

Notes. This table presents the estimates of the average structural inverse labor supply elasticity of treated firms, obtained as $\epsilon_{iF,k} \equiv \psi_{iF,k} - 1$, across all markets (Column 1) and within different market subsets (Columns 2 to 4) in the estimated model. Low and high self-employment rates are defined as being below or above the average self-employment rate across local labor markets, respectively.

Table A.8: Median Estimates of Roy Parameters

	σ_F	σ_S	ϱ	$\hat{\mu}$ $\ln \frac{\text{Earnings}_F}{\text{Earnings}_S}$	$\ln \frac{\text{Revenues}_F}{\text{Earnings}_S}$	$\beta = 1$
Panel I. Baseline Estimates						
	0.81	0.91	0.89	-0.16	-0.1	-0.09
Panel II. Group Heterogeneity						
Group 1	0.91	0.93	0.87	-0.14	-0.02	-0.03
Group 2	0.83	0.91	0.96	-0.12	-0.11	-0.10
Group 3	0.69	0.91	0.88	-0.27	-0.16	-0.19

Notes. This table presents the median estimates of the Roy model parameters, including the variance-covariance matrix parameters (σ_F, σ_S, ϱ) and the relative mean ability ($\hat{\mu}$) across different market groups. Panel I displays the baseline estimates, where these parameters are held constant across markets to simplify the estimation process and reduce the potential impact of measurement errors. Panel II shows estimates allowing for heterogeneity across three market groups, clustered based on population terciles. The three columns under relative mean ability refer to robustness tests, as described in Appendix D.1.2, with $\ln \frac{\text{Earnings}_F}{\text{Earnings}_S}$ as our baseline.

Table A.9: Labor Market Power – Robustness

	(1) Baseline	(2) Roy Estimates	(3) Fixed Costs	(4) Entry	(5) Census Data
$\bar{\psi}_k$	1.45	1.43	1.49	1.44	1.31
— High HHI	1.71	1.71	1.78	1.67	1.56
— High self-employment	1.40	1.45	1.35	1.40	1.28

Notes. This table presents the estimates of the average markdown across different model calibrations and in different subgroups of markets. $\bar{\psi}_k$ represents the average labor market power across all markets. The second and third rows show the average labor market power in markets with high concentration and high self-employment rates, respectively.

Table A.10: Targeted Moments and Model Fit – Robustness

Moment	Baseline	Fixed Costs	Entry
Panel I. Distribution Moments			
Log Number of Firms			
Mean	0.97	2.29	1.00
Standard Deviation	0.95	1.71	0.85
Log of Sales			
Ratio p75/p25	2.94	2.92	2.93
Ratio p90/p10	5.29	5.16	5.30
CR_1, Mean	0.66	0.70	0.67
CR_1, Standard Deviation	0.29	0.25	0.26
CR_4, Mean	0.94	0.96	0.95
CR_4, Standard Deviation	0.11	0.08	0.10
Employment HHI			
Mean, Unweighted	0.57	0.55	0.56
Standard Deviation	0.34	0.30	0.31
Mean, Weighted	0.30	0.45	0.29
Percent of Markets with 1 firm	0.36	0.20	0.29
Wage-bill Share of			
Markets with 1 firm	0.07	0.08	0.01
Markets with $<$10 firms	0.89	0.49	0.91
Markets with $<$50 firms	1.00	0.72	1.00
Share of Wage Employment			
Mean	0.66	0.68	0.67
Standard Deviation	0.12	0.10	0.12
Log of Earnings$_F$/Earnings$_S$			
Mean	0.41	0.52	0.43
Standard Deviation	0.58	0.53	0.59
Log of Schooling$_F$/Schooling$_S$ (Ability$_F$/Ability$_S$)			
Ratio p75/p25	1.42	1.63	1.34
Ratio p90/p10	1.18	1.28	1.15
Panel II. Regression Coefficients			
% Wage Employment on (Log) HHI^n			
Point Estimate	-0.04	-0.01	-0.06
Standard Error	0.01	0.01	0.01
(Log) Earnings$_S$ on (Log) HHI^n			
Point Estimate	-1.17	-0.27	-1.61
Standard Error	0.12	0.16	0.12
(Log) Earnings$_F$ on (Log) HHI^n			
Point Estimate	-1.36	-0.31	-1.90
Standard Error	0.13	0.15	0.13

Notes. This table reports the moments calculated from the baseline estimated model and compares them those estimated when relaxing some parametric assumptions, as described in Section 4.5.

Table A.11: Effect of Productivity Shock in Partial and General Equilibrium

	$\Delta\bar{Y}$ (%)	
	PE	GE
Log Avg. Wage $\bar{a}_F W_F$	-0.46	2.35
Log Avg. Ability $\bar{a}_F$	0.00	0.00
Log Unit Wage W_F	-0.46	2.35
Log Markdown $\bar{\psi}_{F,k}$	0.44	0.35
Log $\overline{MRPL}_{F,k}$	-0.01	2.70
Log Price Index $P_{F,k}$	-3.30	-0.61
Log Productivity Index $Z_{F,k}$	3.28	3.24
Log $\overline{\mu}_{F,k}$	-0.01	-0.06
Log Avg. Self-Empl. earnings $\bar{a}_S W_S$	-0.70	2.09
Log Avg. Ability $\bar{a}_S$	0.50	0.54
Log Unit Earnings W_S	-1.20	1.55

Notes. This table reports the percentage change in the average of selected outcomes across markets following the same productivity shock in partial equilibrium (PE) and general equilibrium (GE). In the PE exercise, we keep aggregate income Y (the only GE variable in our model) constant at its baseline level, focusing solely on the market responses to the productivity shock. In the GE exercise, we allow aggregate income to adjust by solving for the full model.

Table A.12: Impact of Labor Market Power Across Markets

	$\bar{Y}_{\iota=1}$	$\bar{Y}_{\iota=0}$	$\bar{Y}_{\iota=1} - \bar{Y}_{\iota=0}$
Wage Employment Share	0.66	0.77	0.11
Wage-bill Concentration HHI_k^{wb}	0.57	0.63	0.06
Log Avg. Wage $\bar{a}_F W_F$	-3.76	-3.45	0.31
Log Avg. Ability $\bar{a}_F$	1.89	1.89	0.00
Log Unit Wage W_F	-5.64	-5.34	0.31
Log Markdown $\bar{\psi}_{F,k}$	0.35	0	-0.35
Log $\overline{MRPL}_{F,k}$	-5.29	-5.34	-0.04
Log Price Index $P_{F,k}$	-5.59	-5.64	-0.05
Log Productivity Index $Z_{F,k}$	0.64	0.66	0.02
Log $\overline{\mu}_{F,k}$	0.34	0.35	0.01
Log Avg. Self-Empl. earnings $\bar{a}_S W_S$	-3.46	-3.19	0.27
Log Avg. Ability $\bar{a}_S$	2.49	2.6	0.11
Log Unit Earnings W_S	-5.95	-5.79	0.15
Log Labor Income	-3.72	-3.46	0.26
Wage Labor Supply Elasticity $\epsilon(\hat{W}_k)$	0.24	-0.18	-0.42

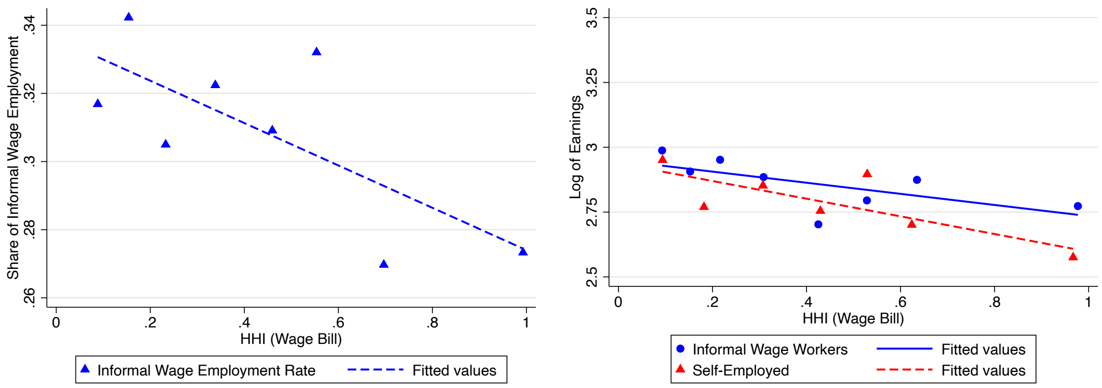
Notes. This table reports the average of selected outcomes across markets in the baseline economy ($\bar{Y}_{\iota=1}$) and in the counterfactual economy with no labor market power ($\bar{Y}_{\iota=0}$) together with the difference between the two.

Table A.13: Average Policy Impact Across Markets

		$\Delta\bar{Y}$	
	Firm Productivity ΔT_k	Fixed Cost Δf_k^e	Worker Skills $\Delta\hat{\mu}_k$
Wage Employment Share	0.67	1.37	3.57
Wage-bill Concentration HHI_k^{wb}	-0.30	-5.54	-2.49
Log Avg. Wage $\bar{a}_F W_F$	2.35	2.14	12.32
Log Avg. Ability $\bar{a}_F$	0	0	16.5
Log Unit Wage W_F	2.35	2.14	-4.18
Log Markdown $\bar{\psi}_{F,k}$	0.35	-1.93	1.49
Log $\overline{MRPL}_{F,k}$	2.70	0.21	-2.69
Log Price Index $\overline{P}_{F,k}$	-0.61	-1.43	-3.28
Log Productivity Index $\overline{Z}_{F,k}$	3.24	0.20	-0.05
Log Markup $\overline{\mu}_{F,k}$	-0.06	-1.44	-0.64
Log Avg. Self-Empl. earnings $\bar{a}_S W_S$	2.09	1.61	10.95
Log Avg. Ability $\bar{a}_S$	0.54	1.06	2.94
Log Unit Earnings W_S	1.55	0.55	8.01
Log Labor Income	2.03	1.47	10.62
Wage Labor Supply Elasticity $\epsilon(\hat{W}_k)$	-1.91	-3.69	-10.42

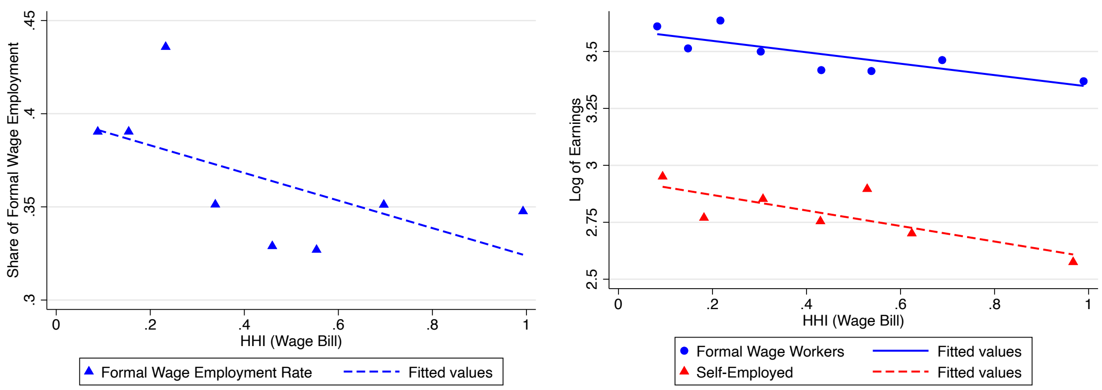
Notes. This table reports the percentage change in the average of selected outcomes across markets following the policy shocks discussed in Section 5.2.

Figure A.1: Concentration, Informal Wage Employment Rate, and Earnings



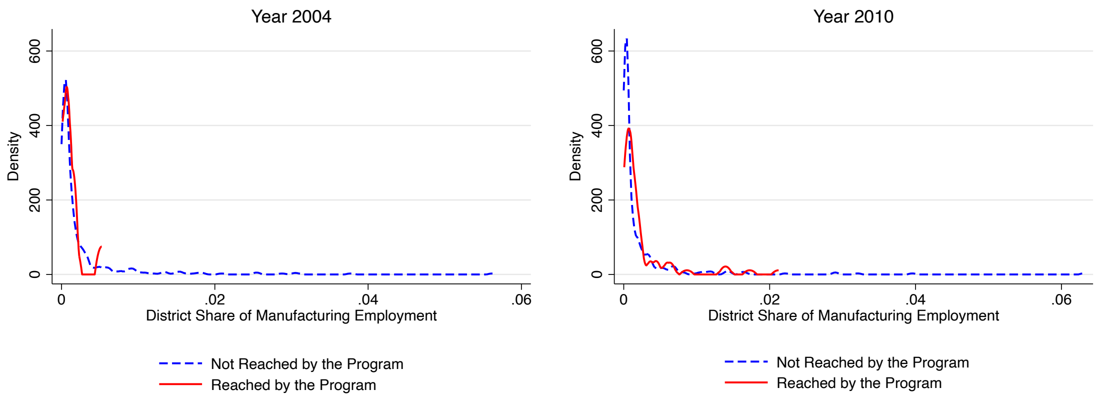
Notes. The figures illustrate the relationship between employer concentration, rate of informal wage employment (left), and earnings from both informal wage work and informal self-employment (right) across local labor markets. The left panel plots the share of informal wage workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for informal wage and self-employed workers. The straight lines show the linear fit based on the underlying data.

Figure A.2: Concentration, Formal Wage Employment Rate, and Earnings



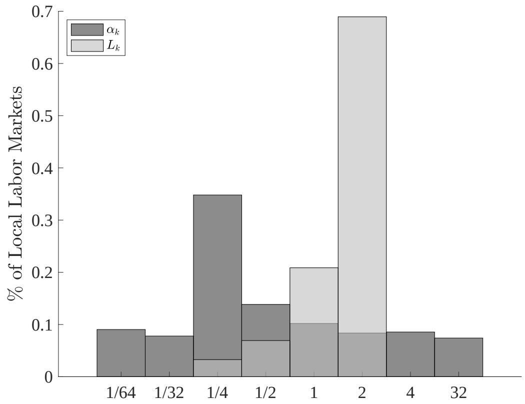
Notes. The figures illustrate the relationship between employer concentration, rate of formal wage employment (left), and earnings from both formal wage work and informal self-employment (right) across local labor markets. The left panel plots the share of formal wage workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for formal wage and self-employed workers. The straight lines show the linear fit based on the underlying data.

Figure A.3: Electrification Program Implementation Across Districts



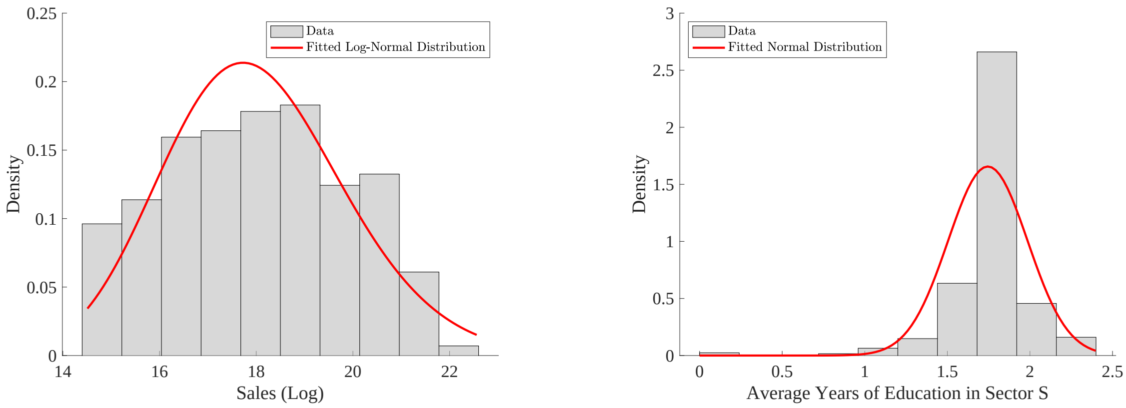
Notes. The figures shows the distribution of manufacturing employment share across districts reached vs. not reached by the electrification program in the first and last year of the IV estimation sample, i.e. 2004 and 2010.

Figure A.4: Cobb-Douglas and Population Shares in the Data



Notes. The figure displays histograms of aggregate sales and population shares across local labor markets, with shares normalized to average 1. The summary statistics are as follows: For the Cobb-Douglas shares $\tilde{\alpha}$, the mean is 0.09%. The largest Cobb-Douglas share is 2.1%, the 90th percentile is 0.33%, the median is 0.02%, and the 10th percentile is 0.002%. For the population shares $\tilde{L}_k$, the mean is 0.09%. The largest population share is 14.6%, the 90th percentile is 12.3%, the median is 0.10%, and the 10th percentile is 0.04%. The correlation coefficient from a regression of expenditure shares on population shares is 0.79, with a standard error of 0.20.

Figure A.5: Sales and Education in the Data – Fitted Distributions

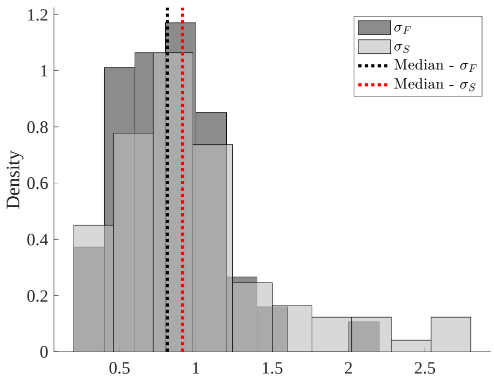


(a) Variance

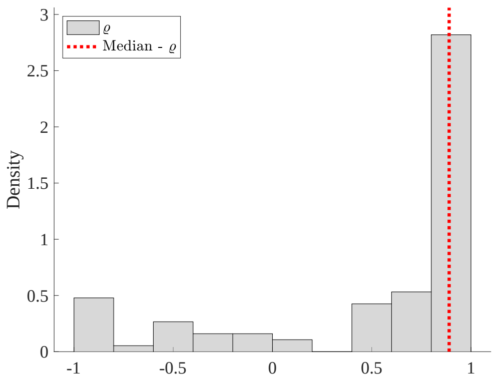
(b) Correlation

Notes. The figure displays the distributions of log sales and average years of education among self-employed workers across local labor markets, along with the fitted log normal and normal distribution, respectively.

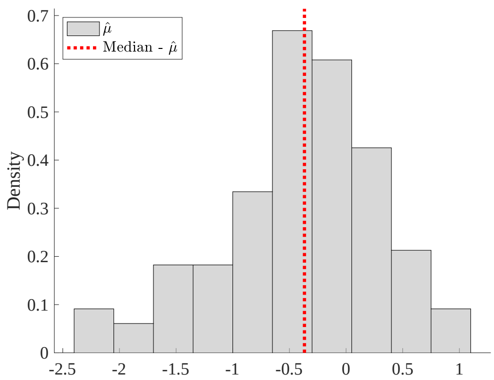
Figure A.6: Parameter Estimates for Joint Ability Distribution



(a) Variance



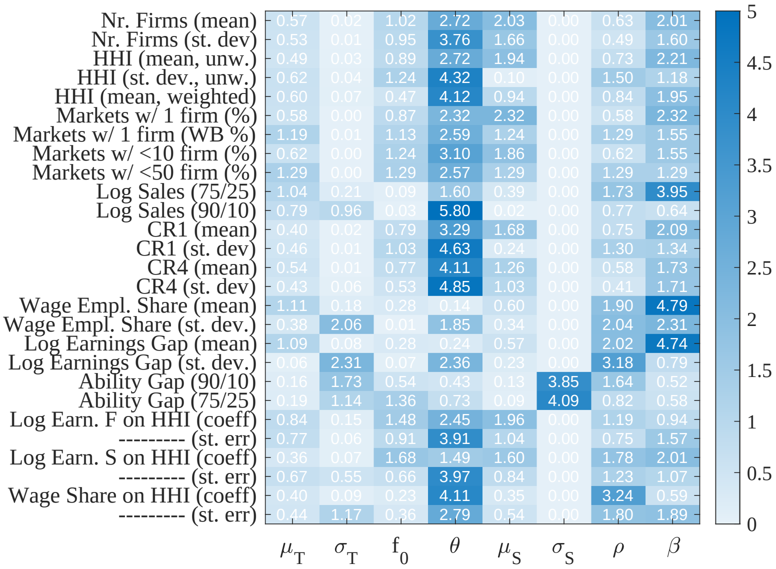
(b) Correlation



(c) Mean Comparative Advantage

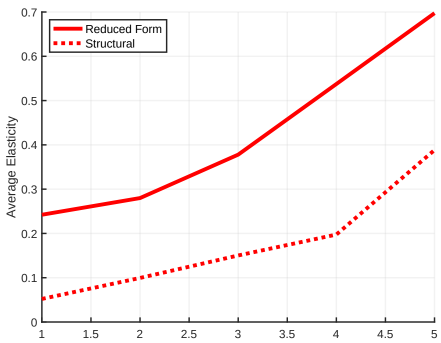
Notes. The figure shows histograms of the parameter estimates for the ability distribution, as discussed in Section 4.3. Panel (a) presents histograms for the estimates of $\sigma_{F,k}$ and $\sigma_{S,k}$ across local labor markets. Panel (b) displays the histograms for the parameter ϱ_k, while panel (c) presents the estimates for $\hat{\mu}_k$. Each panel also includes the median estimate, which is used in the calibration of our baseline model.

Figure A.7: Normalized Partial Derivatives of Moments with Respect to Parameters

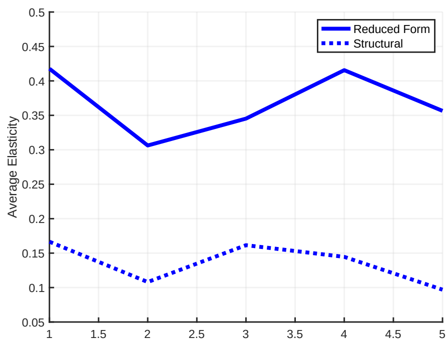


Notes. The Jacobian matrix includes the normalized values of the elasticity of each moment i with respect to a 10% change in parameter j around its estimated value while keeping all the other parameters constant. Each row is a moment and each column is a parameter.

Figure A.8: Inverse Elasticity, Concentration, and Self-Employment Shares



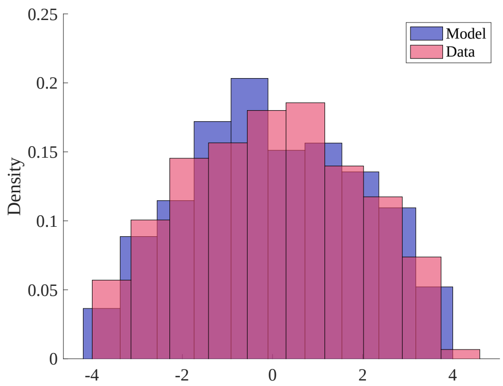
(a) HHI Quantiles



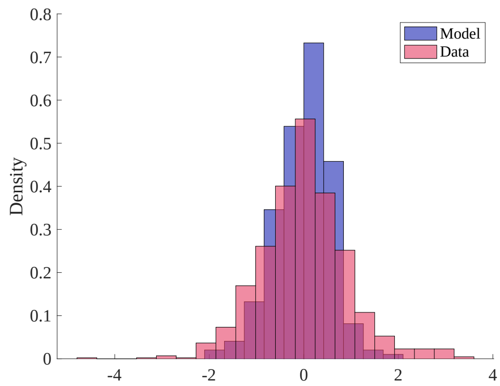
(b) Self-Employment Quantiles

Notes. The figure presents the average reduced-form and structural inverse elasticity of wage labor across different quantiles of the wage-bill HHI and self-employment shares. Subfigure (a) shows the elasticity variation across HHI quantiles, while subfigure (b) shows the elasticity variation across self-employment quantiles. The solid red lines indicate reduced form elasticities, and the dotted red lines represent structural elasticities.

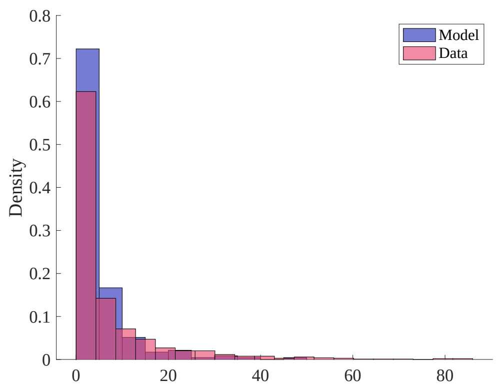
Figure A.9: Model Fit – Sales, Earnings Gap, and Number of Firms



(a) Sales (Log)



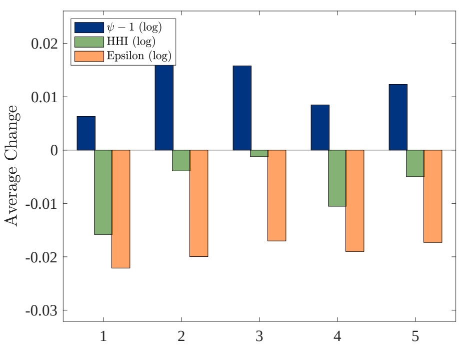
(b) Earnings Gap (Log)



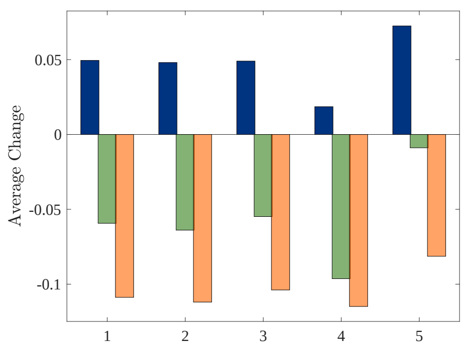
(c) Number of Firms

Notes. The figure shows histograms of log sales, log earnings gap, and number of firms, in the model and in the data.

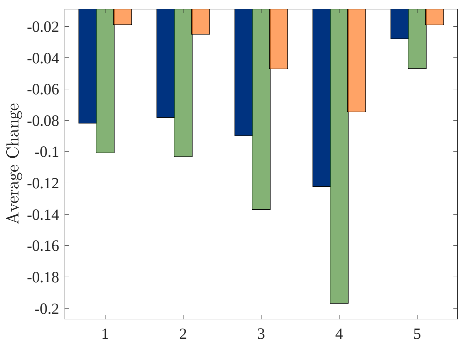
Figure A.10: Effect of Policy Shocks on Wage Markdown and Its Determinants



(a) Firm Productivity



(b) Worker Skills



(c) Fixed Cost

Notes. The three panels complement those in Figure 4 by illustrating the estimated change in wage markdown, concentration, and wage labor supply elasticity across local labor markets resulting from the three policy experiments. It does so for separate bins determined by the size of the wage markdown at baseline.

B Robustness of Stylized Facts

In this section, we present and discuss a series of checks to demonstrate that the empirical facts shown in Section 2 remain robust when using alternative measures of concentration, redefining self-employment, and redrawing local labor market boundaries.

Concentration. A primary concern is whether the *Encuesta Económica Anual* (EEA) firm-level dataset accurately reflects the population of firms that employ wage workers. The EEA is mandatory for—and is therefore a census of—firms with net sales above a certain threshold. To ensure consistency across years and account for changes in such threshold over time, we focus on manufacturing firms with yearly net sales exceeding 2 million Peruvian Soles—approximately 700,000 USD in 2010. Importantly, the EEA is a longitudinal survey, and the panel dimension is crucial for implementing the identification strategy used to estimate labor market power. This feature is lacking in alternative sources of firm-level data like the Economic Census.

We regard the use of EEA data and the decision to adopt a single sales threshold as inconsequential for the analysis for several reasons. First, according to 2007 Economic Census data, medium and large firms constitute the overwhelming majority of wage employment and sales in their respective local labor markets. Within manufacturing, firms with sales above the threshold account for 95% of total sales and 76% of total employment. Second, these choices do not introduce a systematic bias in the extent to which concentration varies across markets. Figure B.1 shows the distribution of wage-bill HHIs in the 2007 Economic Census and in our EEA data. The average unweighted HHI is lower in the Census data, due to the exclusion of smaller firms in the EEA. Yet, the average weighted HHI is very similar across the two. Most importantly, Figure B.2 and Table B.1 demonstrate that the two measures are strongly correlated, even after accounting for industry and location fixed effects.

A separate concern is about the use of wage-bill HHI as concentration measure. We also consider the employment HHI and the number of firms alternative measures, all defined in Section 2. Figure B.3 shows that the three measures are strongly correlated. Importantly, Tables B.2 and B.3 report the coefficient estimates that we obtain when using these measures instead of wage-bill HHI to investigate the correlation between concentration, self-employment rates and earnings. The results are very similar to the baseline ones in Table A.2.

Self-Employment. A possible concern regarding our definition of self-employment is that it includes both own-account workers and employers. First, note that employers are a small fraction of workers in the ENAHO data: out of the 40% of workers that we classify as self-employed in manufacturing, 31% are *own-account* workers while the remaining 9% are *employers*. Among these employers, only 14%—equivalent to about 1.2% of all workers in our data—have their businesses registered as legal entities and thus operate formally. Note that registration is a necessary condition for them to be included in the EEA, which is our source of data on medium to large firms.

Our decision to pool own-account workers and employers into a single self-employment category stems from the dual nature of manufacturing labor markets in poor countries, with wage employment at medium and large manufacturing firms on the one hand and self-employment on the other hand. The businesses operated by employers in this second category differ significantly from the larger firms that employ wage workers. Both workers and owners of these micro-enterprises earn low incomes, with minimal specialization between them, making these businesses resemble a group of self-employed individuals sharing a workspace rather than functioning as modern firms (Bassi et al., 2023; Atencio-De-Leon, Lee and Macaluso, 2023). These enterprises, captured in the ENAHO survey, represent the readily accessible, labor- and product-market price-taking self-employment opportunities that our model conceptualizes. For this reason, we chose to pool them with own-account self-employed workers.

We nonetheless investigate whether excluding employers from the sample affects our results. Table B.5 shows the summary statistics for manufacturing workers that we obtain when we exclude employers. Both average self-employment earnings and their dispersion are lower compared to Table 1, which is to be expected. Despite these differences, all the stylized facts hold when focusing solely on *own-account self-employment* only as the relevant alternative to wage work. This is illustrated in Figures B.6 and B.7, which closely resemble Figures 1 and 2. The same is true for the results in Table B.6, which mirror those in Table A.2.

All stylized facts are also robust to focusing exclusively on informal self-employment as the relevant alternative to wage work, as opposed to overall self-employment. This is not surprising considering that over 90% of self-employment is informal, also within manufacturing. The results are shown in Figure B.4, Figure B.5 and Table B.4.

Local Labor Markets. Another possible concern is with our definition of local labor markets, and specifically with whether the province or commuting zone boundaries are the appropriate geographical unit to consider. This is relevant considering that the model does not feature mobility of workers across markets. To address this concern, we consider a broader definition using 2-digit industries within departments as units of analysis, as opposed to the baseline definition of 2-digit industries within provinces or commuting zones. Departments are Level 1 administrative units, while provinces are Level 2. On average, the 24 departments (plus the Callao constitutional province) in the country contain around 9 provinces.

In Table B.7, we present summary statistics across local labor markets using the alternative definition. The number of local labor markets decreases from 280 to 179, with an average of about 10 medium to large firms in each, compared to 6 using the baseline definition. The unweighted average wage-bill HHI slightly increases to 0.68, and the percentage of markets with only one firm increases from 30% to 45%. However, the weighted concentration measures, either for payroll or employment, decrease by around one-third compared to the baseline (0.26

vs. 0.37 for wage-bill HHI and 0.21 vs. 0.31 for employment HHI). Figure B.8 displays the distributions of wage-bill HHI using both the baseline and broadened definition.

Figure B.9 mirrors Figure 1, plotting the average likelihood of transitioning into and from wage work and self-employment across the earnings distribution in the two sectors. In Figure 1, we group workers by deciles of the earning distribution within each local labor market as defined by a 2-digit industry within a province or commuting zone. In Figure B.9, we do the same but using the new broadened definition of a 2-digit industry within a department. Substantial variation exists in earnings within such large geographical areas. Despite the change in definition, the patterns in Figure B.9 are very similar to those in Figure 1.

Next, Figure B.10 mirrors Figure 2 and illustrates the relationship between employer concentration, the rate of self-employment (left), and earnings from both wage work and self-employment (right) across local labor markets, now defined as 2-digit industries within departments. Once again, the patterns are very similar to those found using the baseline definition. Higher concentration is consistently associated with higher self-employment rates and lower earnings from both wage work and self-employment. The results in Table B.8 show that these relationships still stand when controlling for individual characteristics, industry and department fixed effects—mirroring the baseline results in Table A.2.

Table B.1: Correlation Between Concentration Measures in Census and EEA Data

	EEA Wage-bill HHI			
	(1)	(2)	(3)	(4)
Census Wage-bill HHI	0.804***	0.845***	0.712***	0.715***
	(0.093)	(0.069)	(0.109)	(0.104)
	[8.60]	[12.19]	[6.56]	[6.85]
2-digit Industry FE	No	Yes	No	Yes
Location FE	No	No	Yes	Yes
Observations	194	194	169	169
R^2	0.500	0.618	0.634	0.735

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. The unit of observation is a local labor market as defined by a 2-digit industry within location, the latter corresponding to Peruvian provinces or commuting zones. The table reports the coefficient estimates and their standard errors obtained when regressing wage-bill HHI as obtained from the 2007 EEA data over the same variable obtained from the 2007 economic census, focusing on manufacturing firms. The standard errors and *t-statistics* associated with each estimate are reported in round and square brackets, respectively. Standard errors are clustered at the level of location in all specifications.

Table B.2: Concentration, Self-employment, and Earnings – Employment HHI

	Self-Employed {0,1}			Log of Earnings					
				Wage Workers			Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	0.053***	0.050***	0.063***	-0.084***	-0.096***	-0.045**	-0.122**	-0.146***	-0.032
(Log of Employment HHI)	(0.012)	(0.013)	(0.014)	(0.019)	(0.019)	(0.019)	(0.052)	(0.045)	(0.050)
Female	0.121***	0.121***	0.111***	-0.424***	-0.381***	-0.382***	-1.312***	-1.229***	-1.211***
	(0.023)	(0.023)	(0.023)	(0.041)	(0.036)	(0.035)	(0.072)	(0.075)	(0.078)
Age	0.018***	0.017***	0.016***	0.023**	0.027***	0.029***	0.116***	0.115***	0.109***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)	(0.018)	(0.017)	(0.018)
Age sq.	-0.000	-0.000	-0.000	-0.000	-0.000**	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.005	-0.001	-0.000	0.178***	0.162***	0.156***	0.111***	0.109***	0.102***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.007)	(0.018)	(0.017)	(0.019)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7637	7637	7634	4707	4706	4698	2054	2054	2047
R^2	0.104	0.133	0.156	0.308	0.363	0.395	0.327	0.382	0.399

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. Unit of observation is a working-age individual surveyed in ENAHO. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates and their standard errors obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln HHI^n_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln HHI^n_{(j,g)t}$ is the log of employment HHI in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Table B.3: Concentration, Self-Employment, and Earnings – Number of Firms

	Self-Employed {0,1}			Log of Earnings					
				Wage Workers			Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	-0.035***	-0.033***	-0.041***	0.060***	0.068***	0.036***	0.083**	0.109***	0.039
(Log of Number of Firms)	(0.008)	(0.009)	(0.011)	(0.015)	(0.014)	(0.013)	(0.035)	(0.030)	(0.037)
Female	0.121***	0.120***	0.111***	-0.425***	-0.379***	-0.381***	-1.311***	-1.224***	-1.209***
	(0.023)	(0.023)	(0.023)	(0.041)	(0.036)	(0.035)	(0.072)	(0.075)	(0.079)
Age	0.018***	0.017***	0.016***	0.023**	0.028***	0.029***	0.114***	0.113***	0.109***
	(0.005)	(0.005)	(0.005)	(0.009)	(0.009)	(0.009)	(0.018)	(0.017)	(0.018)
Age sq.	-0.000	-0.000	-0.000	-0.000	-0.000**	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.006	-0.001	-0.000	0.178***	0.162***	0.156***	0.110***	0.108***	0.102***
	(0.004)	(0.004)	(0.004)	(0.008)	(0.008)	(0.007)	(0.018)	(0.017)	(0.019)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7637	7637	7634	4707	4706	4698	2054	2054	2047
R^2	0.102	0.132	0.156	0.308	0.363	0.395	0.327	0.383	0.399

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. Unit of observation is a working-age individual surveyed in ENAHO. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates and their standard errors obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln M_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln M_{(j,g)t}$ is the log of the number of firms in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Table B.4: Concentration, Informal Self-Employment, and Earnings

	Informal Self-Employed {0,1}			Log of Earnings: Wage Workers			Log of Earnings: Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	0.049***	0.048***	0.056***	-0.085***	-0.100***	-0.052**	-0.101*	-0.143***	-0.025
(Log of Wage-bill HHI)	(0.012)	(0.014)	(0.015)	(0.020)	(0.021)	(0.021)	(0.053)	(0.045)	(0.050)
Female	0.145***	0.138***	0.129***	-0.426***	-0.381***	-0.382***	-1.233***	-1.169***	-1.161***
	(0.023)	(0.023)	(0.023)	(0.041)	(0.036)	(0.035)	(0.074)	(0.075)	(0.076)
Age	0.012***	0.011**	0.010**	0.023**	0.027***	0.029***	0.093***	0.095***	0.088***
	(0.005)	(0.004)	(0.005)	(0.009)	(0.009)	(0.009)	(0.019)	(0.019)	(0.020)
Age sq.	-0.000	-0.000	0.000	-0.000	-0.000**	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.013***	-0.009**	-0.007**	0.178***	0.161***	0.156***	0.065***	0.063***	0.057***
	(0.004)	(0.004)	(0.003)	(0.008)	(0.008)	(0.007)	(0.018)	(0.018)	(0.019)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7309	7309	7305	4707	4706	4698	1726	1725	1718
R^2	0.105	0.132	0.155	0.308	0.363	0.395	0.296	0.362	0.381

Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. Unit of observation is a working-age individual surveyed in ENAHO, excluding all formal self-employed workers. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln HHI^{wn}_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln HHI^{wn}_{(j,g)t}$ is the log of wage-bill HHI in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Table B.5: Worker Summary Statistics – Excluding Employers

Variable	Mean	St. Dev.
Manufacturing Workers		
Wage Worker	0.61	0.49
Daily Wage	31.84	31.85
Own-Account Self-Employed	0.35	0.48
Daily Earnings from Own-Account Self-Employment	13.89	19.49
W-OAS Transition	0.05	0.21
OAS-W Transition	0.04	0.20

Notes. This table reports summary statistics from ENAHO worker-level data (Panel II), averaging across all years from 2004 to 2011, and excluding employers. Transition rates are obtained using the 2007-2011 panel version of ENAHO. Worker-level statistics are for dummy variables indicating wage work, self-employment, earnings (in PEN, 1 PEN $\approx$ 0.35 USD in 2010), and annual transitions from the wage- to self-employment sector (W-S) and vice versa (S-W).

Table B.6: Concentration, Own-Account Self-Employment, and Earnings

	Own-Account Self-Employed {0,1}			Log of Earnings					
				Wage Workers			Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	0.043***	0.045***	0.052***	-0.085***	-0.100***	-0.052**	-0.093*	-0.123***	0.036
(Log of Wage-bill HHI)	(0.012)	(0.013)	(0.014)	(0.020)	(0.021)	(0.021)	(0.051)	(0.042)	(0.052)
Female	0.167***	0.154***	0.143***	-0.426***	-0.381***	-0.382***	-1.217***	-1.178***	-1.176***
	(0.022)	(0.022)	(0.022)	(0.041)	(0.036)	(0.035)	(0.070)	(0.081)	(0.084)
Age	0.009**	0.008**	0.007	0.023**	0.027***	0.029***	0.095***	0.099***	0.094***
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)	(0.023)	(0.023)	(0.024)
Age sq.	0.000	0.000	0.000	-0.000	-0.000**	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.012***	-0.008**	-0.008**	0.178***	0.161***	0.156***	0.060***	0.052***	0.039**
	(0.004)	(0.004)	(0.003)	(0.008)	(0.008)	(0.007)	(0.018)	(0.018)	(0.018)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	7003	7003	6999	4707	4706	4698	1470	1469	1463
R^2	0.122	0.147	0.170	0.308	0.363	0.395	0.294	0.349	0.369

Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. Unit of observation is a working-age individual surveyed in ENAHO, excluding all employers. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln HHI^{wn}_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln HHI^{wn}_{(j,g)t}$ is the log of wage-bill HHI in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Table B.7: Employer Concentration Across Broadened Local Labor Markets

	Full Sample		*Broadened LLMs*	
	Mean	St. Dev.	Mean	St. Dev.
Number of Firms	6.39	10.37	10.24	22.83
Wage-bill HHI	0.65	0.33	0.68	0.35
Wage-bill HHI (Payroll Weighted)	0.37	0.03	0.26	0.03
Wage-bill HHI (Employment Weighted)	0.34	0.03	0.34	0.03
Employment HHI	0.63	0.35	0.66	0.36
Employment HHI (Payroll Weighted)	0.33	0.03	0.21	0.02
Employment HHI (Employment Weighted)	0.31	0.02	0.21	0.02
Percent of LLMs with 1 firm	38.78	2.27	44.65	2.58
Payroll Share of LLMs with 1 firm	7.94	1.79	6.65	1.72
Employment Share of LLMs with 1 firm	7.80	1.23	6.38	1.10
Number of Local Labor Markets	280		179	
Number of Locations	61		23	
Industries	23		23	

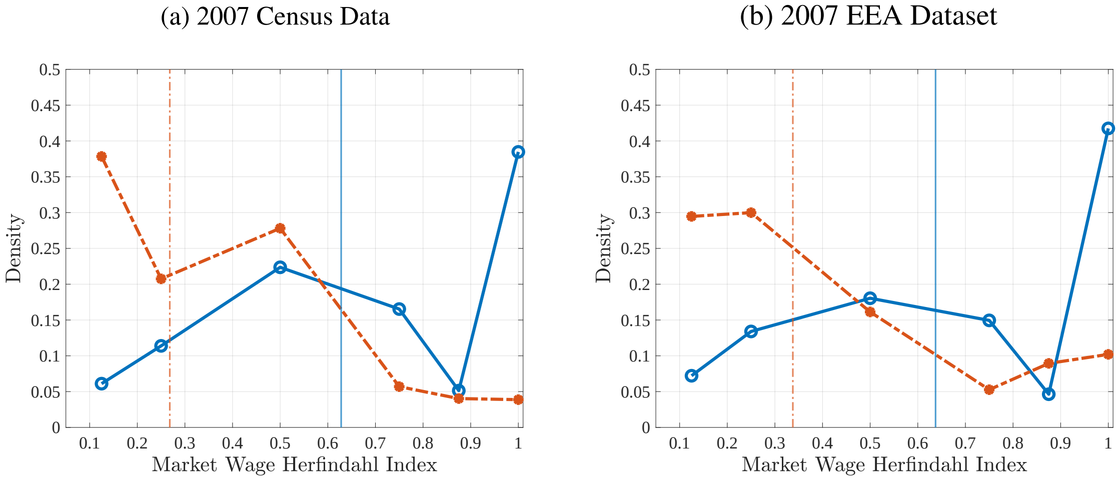
Notes. This table presents summary statistics and employer concentration measures derived from EEA firm-level data across Peruvian local labor markets, averaged over the years 2004 to 2011. The data are shown separately for the entire sample and the full sample where local labor markets are more broadly defined as 2-digit industries within Peruvian departments instead of provinces or commuting zones.

Table B.8: Concentration, Self-Employment, and Earnings – Broadened Local Labor Markets

	Self-Employed {0,1}			Log of Earnings					
				Wage Workers			Self-employed		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Employer Concentration	0.044***	0.050***	0.058***	-0.057***	-0.075***	-0.048**	-0.103**	-0.151***	-0.024
(Log of Wage-bill HHI)	(0.010)	(0.010)	(0.017)	(0.019)	(0.018)	(0.020)	(0.044)	(0.026)	(0.055)
Female	0.130***	0.130***	0.106***	-0.427***	-0.397***	-0.396***	-1.345***	-1.252***	-1.216***
	(0.026)	(0.026)	(0.027)	(0.045)	(0.039)	(0.037)	(0.074)	(0.073)	(0.074)
Age	0.023***	0.023***	0.019***	0.021**	0.025***	0.027***	0.098***	0.088***	0.094***
	(0.004)	(0.004)	(0.004)	(0.009)	(0.009)	(0.009)	(0.015)	(0.014)	(0.014)
Age sq.	-0.000***	-0.000***	-0.000*	-0.000	-0.000*	-0.000**	-0.001***	-0.001***	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Schooling	-0.015***	-0.011***	-0.002	0.180***	0.165***	0.156***	0.115***	0.119***	0.092***
	(0.004)	(0.004)	(0.003)	(0.009)	(0.007)	(0.007)	(0.019)	(0.019)	(0.017)
Year FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Industry FE	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
Location FE	No	No	Yes	No	No	Yes	No	No	Yes
Observations	9613	9613	9596	5573	5573	5541	2937	2937	2904
R^2	0.119	0.150	0.206	0.300	0.349	0.409	0.343	0.401	0.458

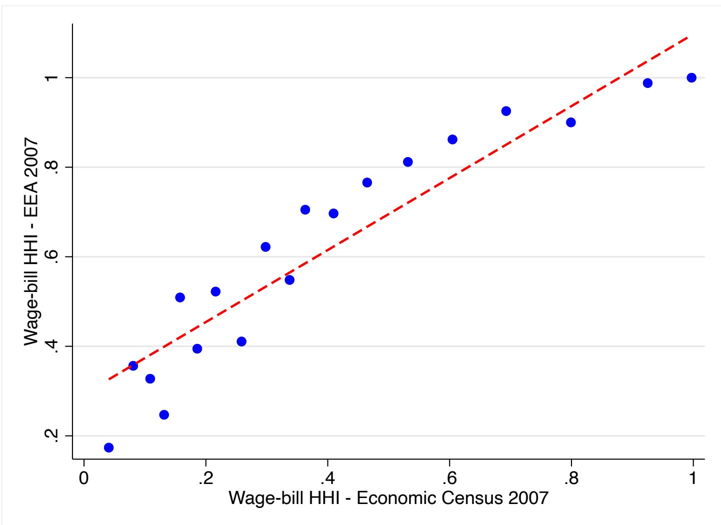
Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. Unit of observation is a working-age individual surveyed in ENAHO. A local labor market k is defined by a 2-digit industry j within a department g. This table reports the coefficient estimates and their standard errors obtained when estimating the regression specification: $y_{i(j,g)t} = \beta \ln HHI^{wn}_{(j,g)t} + \mathbf{X}'_{i(j,g)t}\theta + \gamma_j + \lambda_g + \delta_t + u_{i(j,g)t}$, where $y_{i(j,g)t}$ is the labor market outcome of worker i in local labor market k as defined by a manufacturing industry j within a province or commuting zone g in year t. The first regressor $\ln HHI^{wn}_{(j,g)t}$ is the log of wage-bill HHI in the market in the same year. $\mathbf{X}_{i(j,g)t}$ is a vector of individual characteristics, while γ_j, λ_g and δ_t stand for industry, location, and year fixed effects respectively. Standard errors are clustered at the local labor market level.

Figure B.1: Employer Concentration Across Local Labor Markets in Census and EEA Data



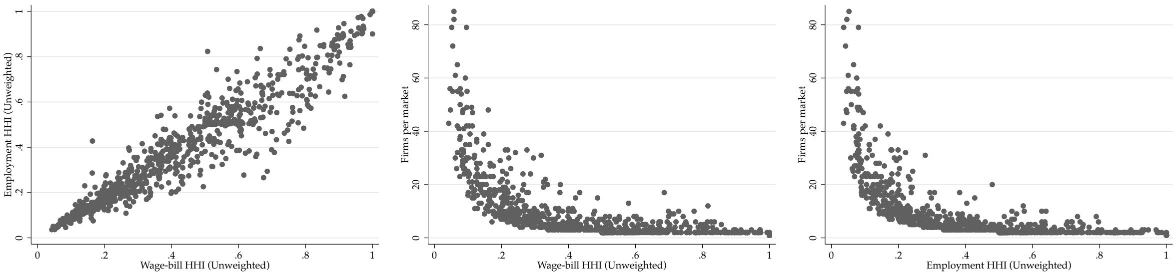
Notes. The figure plots the distribution of wage-bill HHI computed from the 2007 Peruvian Economic Census (left panel) and the same distribution computed from the 2007 EEA dataset (right panel) across local labor markets in the manufacturing sector. The blue solid line in both panels corresponds to the unweighted average, while the dashed line corresponds to the weighted average, where weights are given by the local labor market's share of nation-wide payroll.

Figure B.2: Correlation Between Concentration Measures in Census and EEA Data



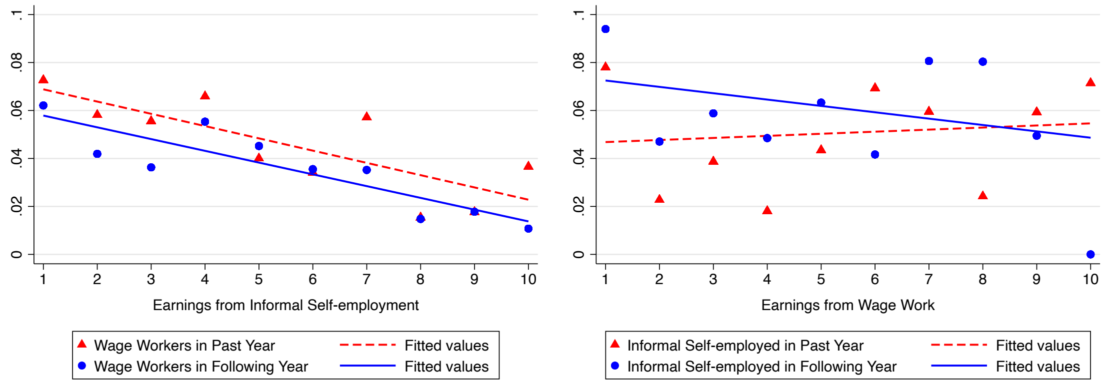
Notes. The figure illustrates the correlation between wage-bill HHI across local labor markets computed from the 2007 Peruvian Economic Census and the 2007 EEA dataset. Both variables are grouped into equal-sized bins, each point showing the average within bins. The dashed line shows the linear fit based on the underlying data, its slope equal to the corresponding parameter in the first column of Online Appendix Table B.1.

Figure B.3: Correlation Between Employer Concentration Measures



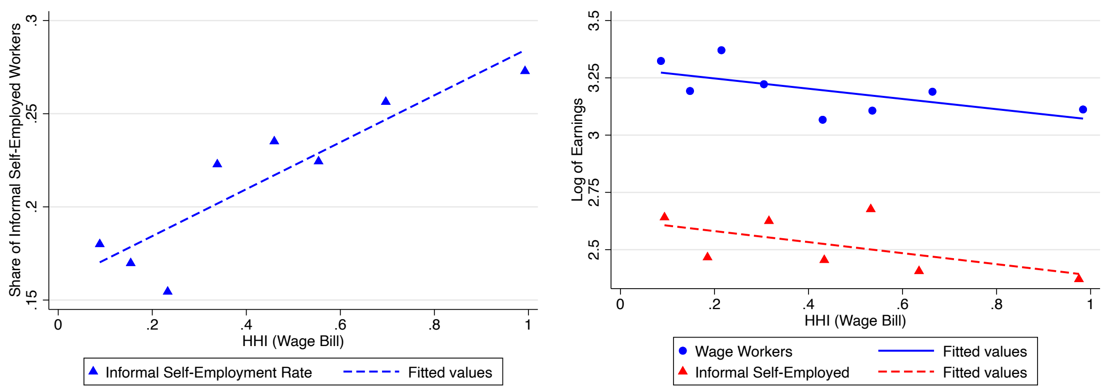
Notes. The figure plots the raw correlation of the three employer concentration measures – wage-bill HHI, employment HHI, and number of firms (bottom center panel) – one against the other across all local labor market-level observations. Wage-bill and employment HHI are strongly positively correlated and they are both strongly negatively correlated to the number of firms.

Figure B.4: Transitions Probabilities and Earnings – Informal Self-Employment



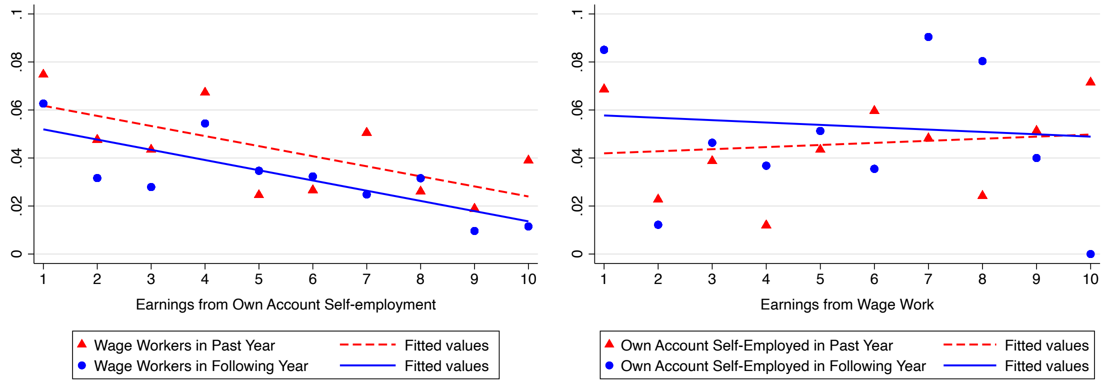
Notes. The figures illustrate the relationship between the likelihood of transitioning from and into wage work and informal self-employment, and earnings. The left panel plots average yearly transition probabilities into and from wage work across deciles of the informal self-employment earnings distribution. Similarly, the right plots average yearly transition probabilities into and from informal self-employment across the wage work earnings distribution deciles. The straight lines show the linear fit based on the underlying data.

Figure B.5: Concentration, Informal Self-Employment Rate, and Earnings



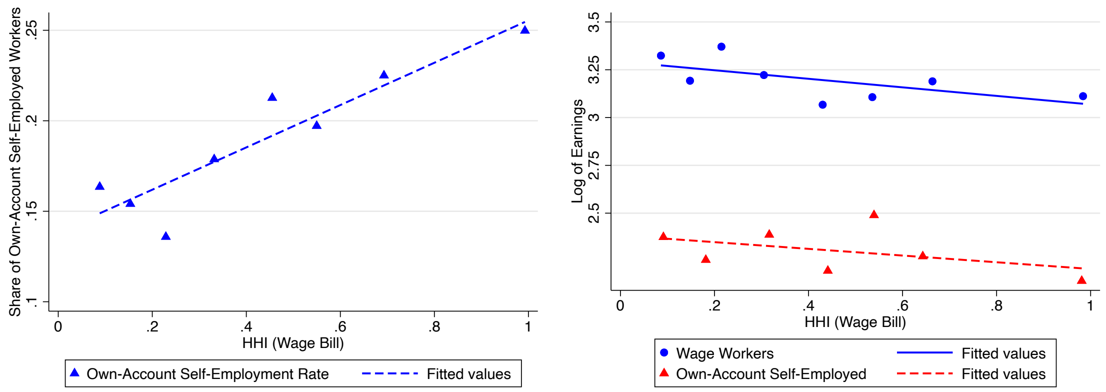
Notes. The figures illustrate the relationship between employer concentration, rate of informal self-employment (left), and earnings from both wage work and informal self-employment (right) across local labor markets. The left panel plots the share of informal self-employed workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for wage and informal self-employed workers. The straight lines show the linear fit based on the underlying data.

Figure B.6: Transitions Probabilities and Earnings – Own-Account Self-Employment



Notes. The figures illustrate the relationship between the likelihood of transitioning from and into wage work and own-account self-employment (thus excluding employers), and earnings. The left panel plots average yearly transition probabilities into and from wage work across deciles of the own-account self-employment earnings distribution. Similarly, the right panel plots average yearly transition probabilities into and from own-account self-employment across the wage work earnings distribution deciles. The straight lines show the linear fit based on the underlying data.

Figure B.7: Concentration, Own-Account Self-Employment Rate, and Earnings

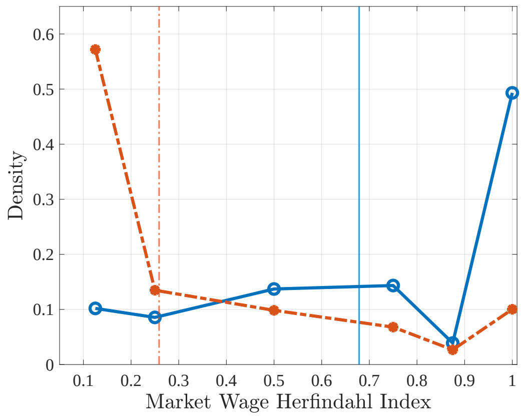


Notes. The figures illustrate the relationship between employer concentration, rate of informal self-employment (left), and earnings from both wage work and informal self-employment (right) across local labor markets. The left panel plots the share of informal self-employed workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for wage and informal self-employed workers. The straight lines show the linear fit based on the underlying data.

Figure B.8: Employer Concentration Across Local Labor Markets – Baseline vs. Broadened

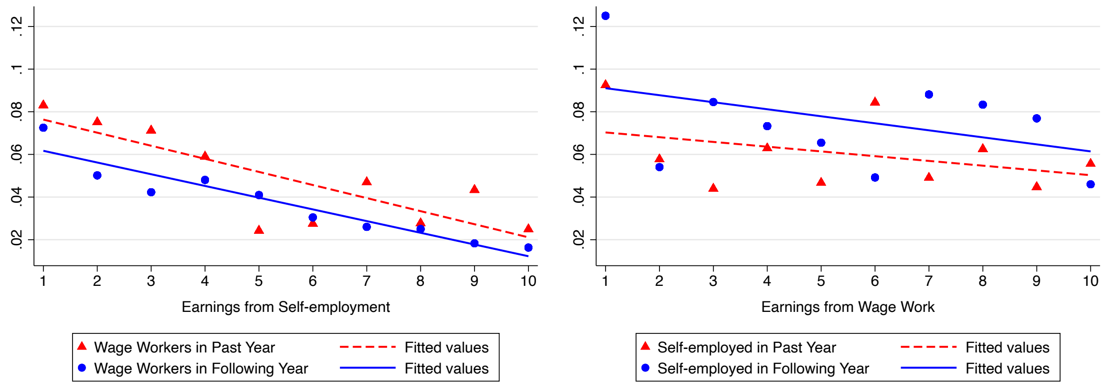
(a) Baseline

(b) Broadened



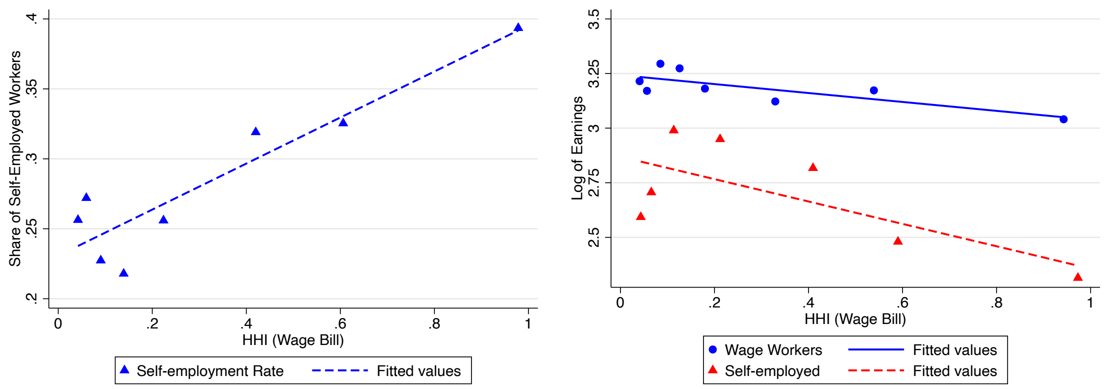
Notes. The figure plots the distribution of wage-bill HHI computed from the EEA dataset using the baseline (left panel) and broadened (right panel) definition of local labor markets. At baseline, local labor markets are defined by a 2-digit industry within a province or commuting zone. In the broadened definition, they are 2-digit industries within a department. The blue solid line in both panels corresponds to the unweighted average, while the dashed line corresponds to the weighted average, where weights are given by the local labor market's share of nation-wide payroll.

Figure B.9: Transition Probabilities and Earnings – Broadened Local Labor Markets



Notes. The figures illustrate the relationship between the likelihood of transitioning from and into wage work and self-employment, and earnings, where the latter are residuals from a regression of daily earnings over the full set of province or commuting zone fixed effects. The left panel plots average yearly transition probabilities into and from wage work across deciles of the self-employment earnings distribution within local labor markets as defined by 2-digit industries within departments. Similarly, the right panel plots average yearly transition probabilities into and from self-employment across the wage work earnings distribution deciles derived at the same level. The straight lines show the linear fit based on the underlying data.

Figure B.10: Concentration, S.-E. Rate, and Earnings – Broadened Local Labor Markets



Notes. The figures illustrate the relationship between employer concentration, rate of self-employment (left), and earnings from both wage work and self-employment (right) across local labor markets as defined by 2-digit industries within departments. The left panel plots the share of self-employed workers in each decile of the wage-bill HHI distribution across local labor markets. The right panel plots the average log of daily earnings in each decile and separately for wage and self-employed workers. The straight lines show the linear fit based on the underlying data.

C Theory Appendix

This section provides further details on the theory. Section C.1 describes the approach to solving the model's general equilibrium (GE). Section C.2 explores the implications of assuming a log-normal distribution for workers' ability, a restriction applied in our empirical analysis.

C.1 Model Solution

With segmented labor markets, interactions across markets occur solely through changes in expenditures $Y_k = \alpha_k Y$, where $\{\alpha_k\}_{k\in(0,1)}$ are the constant expenditure shares. Consequently, given Y, the equilibrium in each market can be determined independently of the others.

This feature of the model allows decomposing its solution into a market equilibrium component and a general equilibrium component. The market equilibrium refers to the process of solving for equilibrium in each local labor market given Y. Each market equilibrium, in turn, provides a value for Y based on market-clearing conditions. The final step involves determining the general equilibrium Y by solving the corresponding fixed-point algorithm.

C.1.1 Market Equilibrium

Let $\boldsymbol{\Lambda}_k \equiv \{s_{iF,k}, s^N_{iF,k}, \mu_{iF,k}, \psi_{iF,k}\}_{i=1}^{M_k}$ represent the vector of output and employment shares, markups, and markdowns for each active firm in market k. Let $\mathbf{K}$ denote the vector of number of entrants, relative wages, and $\boldsymbol{\Lambda}_k = \left\{M_k, \hat{W}_k, \boldsymbol{\Lambda}_k\right\}_{k\in(0,1)}$ in each k. A market equilibrium consists of the vector $\hat{\mathbf{K}}$ such that, given a value for Y, expenditure shares $\{\alpha_k\}_{k\in(0,1)}$, and model primitives $\left\{\{z_{iF,k}\}_{i=1}^{\bar{M}_k^*}, f_k^e, G_k\right\}$, all model equations (4)-(15) are satisfied.

Equilibrium given M_k Let's first assume that the number of entrants M_k is known in each k. From equations (10) and (11), we have:

$$\frac{Y_{F,k}}{Y_{S,k}} = Z_{F,k}\hat{N}_k(\hat{W}_k), \tag{C.1}$$

where $\hat{N}_k(\hat{W}_k) \equiv \frac{N_{F,k}}{N_{S,k}}$ is the relative labor supply, which is a known function of $\hat{W}_k$, given $G_k(\cdot)$, and $Z_{F,k} \equiv \left(\sum_{i=1}^{M_k} s_{iF,k}^{\frac{\eta}{\eta-1}} z_{iF,k}^{-1}\right)^{-1}$ is the productivity index for sector F in market k. With CES preferences, the left-hand side of (C.1) is equal to:

$$\frac{Y_{F,k}}{Y_{S,k}} = \zeta^\rho \left(\frac{P_{F,k}}{P_{S,k}}\right)^{-\rho}. \tag{C.2}$$

Using the pricing rule from (14) and aggregating across firms, we find:

$$P_{F,k} = \left(\sum_{i=1}^{M_k} (p_{iF,k})^{1-\eta} \right)^{\frac{1}{1-\eta}} = \Phi_{F,k} W_{F,k},$$

where $\Phi_{F,k} \equiv \left(\sum_{i=1}^{M_k} \left(\frac{\mu_{iF,k}\psi_{iF,k}}{z_{iF,k}} \right)^{1-\eta} \right)^{\frac{1}{1-\eta}}$ is a market-level index reflecting the aggregate effects of productivity, markups, and markdowns. The self-employment sector is competitive, so aggregate prices reflect marginal cost: $P_{S,k} = W_{S,k}$. Combining these, we obtain:

$$\hat{N}_k(\hat{W}_k) \left(\hat{W}_k \right)^{\rho} = \zeta^{\rho} \left(\Phi_{F,k} \right)^{-\rho} Z_{F,k}^{-1}. \tag{C.3}$$

Equation (C.3) represents the first equilibrium block. The unknowns in this equation are the relative wage $\hat{W}_k$ and the market-level indices $Z_{F,k}$ and $\Phi_{F,k}$, which depend on the vector $\mathbf{\Lambda}_k \equiv \{s_{iF,k}, s_{iF,k}^N, \mu_{iF,k}, \psi_{iF,k}\}_{i=1}^{M_k}$ of shares, markups and markdowns.

The second equilibrium block is defined by the following expressions for firms' market shares, markups, and markdowns, which together form a fixed-point problem given $\hat{W}_k$:

$$s_{iF,k} = \left(\frac{p_{iF,k}}{P_{F,k}} \right)^{1-\eta} = \frac{\left(\frac{\mu_{iF,k}\psi_{iF,k}}{z_{iF,k}} \right)^{1-\eta}}{\sum_{i=1}^{M_k} \left(\frac{\mu_{iF,k}\psi_{iF,k}}{z_{iF,k}} \right)^{1-\eta}}, \tag{C.4}$$

$$\mu_{iF,k} = \frac{\varepsilon_{iF,k}}{\varepsilon_{iF,k} - 1}, \quad \text{where} \quad \varepsilon_{iF,k} = \left[\frac{1}{\eta}(1 - s_{iF,k}) + \frac{1}{\rho} s_{iF,k} \right]^{-1}, \tag{C.5}$$

$$\psi_{iF,k} = \left(\frac{s_{iF,k}^N}{\epsilon_{F,k}(\hat{W}_k)} + 1 \right), \quad \text{with} \quad s_{iF,k}^N = \frac{s_{iF,k}^{\frac{\eta}{\eta-1}} (z_{iF,k})^{-1}}{\sum_{i=1}^{M_k} s_{iF,k}^{\frac{\eta}{\eta-1}} (z_{iF,k})^{-1}}. \tag{C.6}$$

Given these expressions, we can now outline an algebraic algorithm to solve for the market equilibrium. Specifically, given M_k and market-level draws $\{z_{iF,k}\}_{i \in [1, M_k]}$, the equilibrium in market k consists of a relative wage $\hat{W}_k$ and a vector $\mathbf{\Lambda}_k$ such that:

1. Given $\hat{W}_k$, $\mathbf{\Lambda}_k$ solves the fixed-point problem defined by equations (C.4)-(C.6).
2. Given $\mathbf{\Lambda}_k$, the wage $\hat{W}_k$ solves equation (C.3).

The market equilibrium can be found by applying the implied iterative fixed-point procedure.

Solving for M_k As is standard in the literature, we solve the entry problem by considering a sequential entry game where shadow firms with higher productivity draws move first. The equilibrium number of entrants in each market can be determined using the following iterative procedure. For each candidate number of entrants, $M_k = 1, .., \bar{M}_k^*$, we find the equilibrium $(\hat{W}_k, \mathbf{\Lambda}_k)$ using the procedure outlined above. We then compute the profits of the marginal entrant $\underline{i}$ in market k as:

$$\pi_{\underline{i}F,k} = s_{\underline{i}F,k}\gamma_{F,k}Y_k\left(1 - \frac{1}{\mu_{\underline{i}F,k}\psi_{\underline{i}F,k}}\right) - f^e_{\underline{i},k},$$

where $f^e_{\underline{i},k}$ is the entry cost for firm $\underline{i}$, which depends on the firm's ranking of entry, and $\gamma_{F,k}$ is the expenditure share on sector F goods, which solves:

$$\left(\frac{\gamma_{F,k}}{1-\gamma_{F,k}}\right) = \zeta^\rho\left(\hat{W}_k\right)^{1-\rho}\Phi^{\rho-1}_{F,k}.$$

An equilibrium of the entry game is achieved when the equilibrium profits for the marginal entrant $\underline{i}$ are positive, while those for any additional entrant would be negative. With sequential entry, this entry game has a unique cutoff equilibrium, meaning that only firms with productivity above a certain cutoff enter the market.

Simplified Entry Game Solving for the exact equilibrium values of M_k can be computationally intensive, as it requires solving for $(\hat{W}_k, \mathbf{\Lambda}_k)$ for each candidate value, which involves a fixed-point algorithm. To mitigate this complexity, we adopt a simplified entry game approach, inspired by Gaubert and Itskhoki (2021), where firms are treated as 'naive' at the entry stage. Specifically, we assume that upon entry, firms expect to charge the minimum markup and markdown as if they were infinitesimal. For markups, this means setting $\mu_i = \frac{\eta}{\eta-1}$; for markdowns, it implies $\psi_{iF,k} = 1$ for all i. Moreover, under the assumption that all firms behave atomistically, the market shares simplify to:

$$s_{iF,k} = \frac{(z_{iF,k})^{\eta-1}}{\sum_i (z_{iF,k})^{\eta-1}},$$

and the market index $\Phi_{F,k}$ becomes:

$$\Phi_{F,k} = \left(\frac{\eta-1}{\eta}\right)Z_{F,k}, \quad \text{where} \quad Z_{F,k} = \left[\sum_i (z_{iF,k})^{\eta-1}\right]^{\frac{1}{\eta-1}}.$$

As a result, profits simplify to:

$$\pi_{\underline{i}F,k} = s_{\underline{i}F,k}\gamma_{F,k}\frac{\alpha_k Y}{\eta} - f^e_{\underline{i}}, \tag{C.7}$$

where $\gamma_{F,k}$ is given by:

$$\frac{\gamma_{F,k}}{1-\gamma_{F,k}} = \zeta^\rho\left(\frac{\eta}{\eta-1}\right)^{1-\rho}Z^{\rho-1}_{F,k}\left(\hat{W}_k\right)^{1-\rho}, \tag{C.8}$$

and where $\hat{W}_k$ can be found by simple inversion, solving the 'simplified' equilibrium condition:

$$\hat{W}_k^{\rho}\hat{N}_k = \left(\frac{\eta}{\eta-1}\right)^{-\rho}\zeta^{\rho}Z_{F,k}^{\rho-1}. \tag{C.9}$$

The number of entrants M_k can be determined using the iterative procedure described above with these simplified expressions.

C.2 Sorting and Log Normality

This section examines the properties of the model when the joint ability distribution $G_k(a_F, a_S)$ is log-normally distributed, as specified in equation (24). For simplicity, we focus on a single market k and omit the market-level subscript unless needed.

C.2.1 Aggregate Wage Labor Supply Elasticity

We begin by providing sufficient conditions for the wage work supply elasticity to be a decreasing function of the relative wage $\hat{W}$. Without loss of generality, let $L = 1$. The aggregate supply of wage work $N_F(\hat{W})$ and its log can be expressed as:

$$\begin{aligned} N_F(\hat{W}) &= \Pr(h \in \text{wage sector}) \times \mathbb{E}\left(a_F \mid a_F\hat{W} \geqslant a_S\right), \\ \ln N_F(\hat{W}) &= \ln\Pr(h \in \text{wage sector}) + \ln\mathbb{E}\left(a_F \mid a_F\hat{W} \geqslant a_S\right). \end{aligned} \tag{C.10}$$

We want to find conditions for $\epsilon_F(\hat{W}) \equiv \frac{\partial \ln N_F}{\partial \ln \hat{W}}$ to decrease with $\hat{W}$, which is equivalent to show that $\ln N_F$ is a concave function of $\ln \hat{W}$, i.e. $\frac{\partial^2 \ln N_F}{\partial \ln \hat{W}^2} < 0$.

Under log-normality we have

$$\Pr(h \in \text{wage sector}) = \Phi\left(\frac{\ln\hat{W} + \hat{\mu}}{\sigma^*}\right) = \Phi(c_F), \tag{C.11}$$

with $\sigma^* \equiv \sqrt{\sigma_F^2 + \sigma_S^2 - 2\varrho\sigma_F\sigma_S}$ and $\hat{\mu} \equiv \mu_F - \mu_S$.

Following Heckman and Sedlacek (1985) and knowing that $\ln\mathbb{E}(x) \approx \mathbb{E}(\ln x) + \frac{1}{2}\text{Var}(\ln x)$, we also get:

$$\begin{aligned} \ln\mathbb{E}\left(a_F \mid a_F\hat{W} \geqslant a_S\right) &\approx \mathbb{E}\left(\ln a_F | a_F\hat{W} \geqslant a_S\right) + \frac{1}{2}\text{Var}\left(\ln a_F | a_F\hat{W} \geqslant a_S\right) \\ \ln\mathbb{E}\left(a_F \mid a_F\hat{W} \geqslant a_S\right) &\approx \mu_F + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F) + \frac{1}{2}\left\{\sigma_F^2 + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2\lambda'(c_F)\right\} \end{aligned} \tag{C.12}$$

where $\lambda(x) \equiv \phi(x)/\Phi(x)$ is a decreasing and convex function of x.

Let $\alpha = \frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}$. Substituting (C.11) and (C.12) into equation (C.10) and taking the derivative with respect to $\ln \hat{W}$ we get the wage labor supply elasticity:

$$\epsilon_F(\hat{W}) \equiv \frac{\partial \ln N_F}{\partial \ln \hat{W}} = \frac{1}{\sigma^*}\left[\lambda(c_F) + \alpha\lambda'(c_F) + \frac{1}{2}\alpha^2\lambda''(c_F)\right] > 0, \tag{C.13}$$

This is positive for any given α. When $\alpha \leq 0$, since $\lambda(\cdot), \lambda''(\cdot) > 0$, and $\lambda'(\cdot) \in (-1, 0)$, the expression is always positive. When $\alpha > 0$, for any given c_F the expression has a minimum at $\alpha = -\frac{\lambda'(c_F)}{\lambda''(c_F)}$, and is always positive when evaluated at that minimum.

Finally, taking the second derivative we get:

$$\frac{\partial^2 \ln N_F(\hat{W})}{\partial \ln \hat{W}^2} = \frac{1}{(\sigma^*)^2}\left[\lambda'(c_F) + \alpha\lambda''(c_F) + \frac{1}{2}\alpha^2\lambda'''(c_F)\right]. \tag{C.14}$$

which can be shown to be negative if $\alpha \in (-1, 1)$, thus ruling out cases of particularly extreme negative or positive selection (see Section C.2.2). This is because it is negative for $\alpha = -1$ and $\alpha = 1$ at any given value of c_F. It is also monotone increasing in α for any $c_F \leq 1$ since $-\frac{\lambda''(\cdot)}{\lambda'''(\cdot)} \leq -1 < \alpha$. For all $c_F > 1$, the term in squared brackets is bounded between $\lambda'(1) + \alpha\lambda''(1)$ and zero and always negative for $\alpha < -\frac{\lambda'(1)}{\lambda''(1)} \approx 1$. Hence the following proposition.

Proposition C.1. *When the joint ability distribution is log-normal and $\alpha \in (-1, 1)$, the aggregate supply elasticity of wage labor $\epsilon_F(\hat{W})$ is positive and decreases with the relative efficiency unit wage $\hat{W}$.*

Note that our estimates of $(\sigma_F, \sigma_S, \varrho, \hat{\mu}) = (0.81, 0.91, 0.89, -0.12)$ imply $\alpha \approx 0$. Equations (C.10)-(C.13) also imply a one-to-one negative relationship between $\epsilon_F(\hat{W})$ and the equilibrium self-employment share. Denoting the self-employment share as χ_S, we have:

$$1 - \chi_S = \Phi\left(c_F\right) \Leftrightarrow c_F = \Phi^{-1}(1 - \chi_S).$$

Since the function $\Phi(x)$ is monotone increasing, its inverse function $\Phi^{-1}(x)$ is also monotone increasing. As the self-employment share χ_S increases, $\Phi^{-1}(1 - \chi_S)$ decreases, implying that $\epsilon_F(\hat{W})$ increases. Hence the following proposition.

Proposition C.2. *When the joint ability distribution is log-normal and $\alpha \in (-1, 1)$, the aggregate supply elasticity of wage labor $\epsilon_F(\hat{W})$ is a one-to-one increasing function of the self-employment share.*

C.2.2 Ability Distribution and Sectoral Earnings

In Section 3.3, we argued that the scope and sign of the selection channel depend on the parameters of the workers' ability distribution, particularly those governing absolute and comparative advantage. Here, we illustrate this point using the log-normal case.

Following Heckman and Sedlacek (1985), the mean average ability in each sector, defined as the log endowment of efficiency units of labor, can be written as:

$$\begin{aligned} A_F &\equiv \mathbb{E}\left(\ln a_F \mid a_F \hat{W} \geqslant a_S\right) = \mu_F + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F), \\ A_S &\equiv \mathbb{E}\left(\ln a_S \mid a_F \hat{W} < a_S\right) = \mu_S + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_S), \end{aligned} \tag{C.15}$$

where $c_S = -c_F$ and all other parameters have been defined in Section C.2.1.

Now, consider a shock ϑ to the economic environment that lowers the relative wage per efficiency unit $\hat{W}$, thereby shrinking the wage employment sector. Given the system in (C.15), we can express the response of average ability in the two sectors as:

$$\begin{aligned} \frac{dA_F}{d\vartheta} &= \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\cdot\frac{d\lambda(c_F)}{dc_F}\cdot\frac{dc_F}{d\vartheta}, \\ \frac{dA_S}{d\vartheta} &= \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\cdot\frac{d\lambda(c_S)}{dc_S}\cdot\frac{dc_S}{d\vartheta}. \end{aligned} \tag{C.16}$$

By construction, $\frac{dc_F}{d\vartheta} < 0$ and $\frac{dc_S}{d\vartheta} > 0$, implying that $\frac{d\lambda(c_F)}{dc_F} < 0$ and $\frac{d\lambda(c_S)}{dc_S} < 0$. This indicates that the signs of $\frac{dA_F}{d\vartheta}$ and $\frac{dA_S}{d\vartheta}$ depend solely on $\left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)$ and $\left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)$, respectively.

If the two abilities are uncorrelated ($\varrho = 0$) or negatively correlated ($\varrho < 0$), it follows that $\frac{dA_F}{d\vartheta} > 0$ and $\frac{dA_S}{d\vartheta} < 0$. In this case, average ability will decrease in the wage employment sector and increase in self-employment as the relative wage $\hat{W}$ increases.

Conversely, if $\varrho > 0$ and $\sigma_F^2 < \varrho\sigma_F\sigma_S < \sigma_S^2$, then $\left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right) < 0$ and $\left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right) > 0$, implying $\frac{dA_F}{d\vartheta} < 0$ and $\frac{dA_S}{d\vartheta} < 0$. This means that if abilities are more dispersed in self-employment compared to the wage employment sector, and provided that ϱ is positive and sufficiently high, average ability will increase in both the wage employment and self-employment sectors as the relative wage $\hat{W}$ increases.

D Estimation Appendix

This section outlines the model's estimation strategy. Section D.1 covers the identification and direct estimation of the Roy parameters, focusing on the parameters of the variance-covariance matrix Σ_k and the relative mean comparative advantage $\hat{\mu}_k$. Section D.2 provides details on the identifying variation of remaining parameters using the Method of Simulated Moments. Section D.3 describes the computational algorithm used to solve the model given estimated parameters. Section D.4 provides details on non-targeted model fit.

D.1 Identification of Ability Distribution Parameters

D.1.1 Variance-Covariance Matrix

We focus on a single market k and omit the market-level subscript unless needed. We denote the share of workers in the wage sector as c_1, and write it as:

$$c_1 \equiv \Pr(h \in \text{wage sector}) = \Phi\left(\frac{\ln \hat{W} + \hat{\mu}}{\sigma^*}\right). \tag{D.1}$$

with $\sigma^* \equiv \sqrt{\sigma_F^2 + \sigma_S^2 - 2\varrho\sigma_F\sigma_S}$. Following Heckman and Sedlacek (1985), the mean log earnings in the wage employment and self-employment sectors, which we denote as c_3 and c_4, can be expressed as:

$$\begin{aligned} c_3 &\equiv \mathbb{E}\left(\ln a_F W_F \mid a_F\hat{W} \geq a_S\right) = \ln W_F + \mu_F + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F), \\ c_4 &\equiv \mathbb{E}\left(\ln a_S W_S \mid a_F\hat{W} \leq a_S\right) = \ln W_S + \mu_S + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_S), \end{aligned} \tag{D.2}$$

and the corresponding variances, denoted as c_5 and c_6, as:

$$\begin{aligned} c_5 &\equiv \text{Var}\left(\ln a_F W_F \mid a_F\hat{W} \geq a_S\right) = \sigma_F^2 + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2\left[-\lambda(c_F)c_F - \lambda^2(c_F)\right], \\ c_6 &\equiv \text{Var}\left(\ln a_S W_S \mid a_F\hat{W} \leq a_S\right) = \sigma_S^2 + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2\left[-\lambda(c_S)c_S - \lambda^2(c_S)\right], \end{aligned} \tag{D.3}$$

where c_F is as defined in equation (C.11), and $c_S = -c_F$.

The variables $c_1, .., c_6$ can be observed for each local labor markets where a cross section of workers' earnings across the two sectors is available. Similarly, equation (C.11) shows that we can easily recover the terms c_F (and c_S) from simple inversion of the observed employment shares in the two sectors, from which we can also get $\lambda(c_F), \lambda(c_S)$.

From simple algebra, one can derive the following system of equations holding for each market:

$$
\begin{cases}
c_1 & = \Phi(c_F), \\
c_3 - c_4 & = \sigma^* c_F + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F) - \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(-c_F), \\
c_5 & = \sigma_F^2 + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[\lambda(c_F)c_F - \lambda^2(c_F)\right], \\
c_6 & = \sigma_S^2 + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[-\lambda(-c_F)c_F - \lambda^2(-c_F)\right].
\end{cases}
\tag{D.4}
$$

The one above is a system of $4 \times \bar{K}$ equations in $4 \times \bar{K}$ unknowns $(c_{F,k}, \sigma_{F,k}, \sigma_{S,k}, \varrho_k)$, where $\bar{K}$ is the number of local labor markets in our data where earnings data are available for both sectors. It follows that observing multiple individuals in each sector in a given market k will suffice for the identification of the parameter vector $\mathbf{\Theta}_k = (\sigma_{F,k}, \sigma_{S,k}, \varrho_k)$ in that market. We recover the vector $\mathbf{\Theta}_k = (\sigma_{F,k}, \sigma_{S,k}, \varrho_k)$ from a constrained Minimum Distance Estimation (MDE) procedure, where we restrict the variance coefficients to be non-negative and the correlation parameter to be $\varrho_k \in [-1, 1]$.

D.1.2 Mean Comparative Advantage

We now address the identification of the mean comparative advantage $\hat{\mu}$. From Equations (C.15), we derive the following expression for the relative mean log abilities across sectors:

$$
\begin{aligned}
\mathbb{E}\left(\ln \frac{a_F}{a_S} \mid a_F \hat{W} \geq a_S\right) &\equiv \mathbb{E}\left(\ln a_F \mid a_F \hat{W} \geq a_S\right) - \mathbb{E}\left(\ln a_S \mid a_F \hat{W} \leq a_S\right) \\
&= \hat{\mu} + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F) - \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_S).
\end{aligned}
\tag{D.5}
$$

While the terms involving the variance-covariance parameters are known, the left-hand side of (D.5) is unobserved, hindering the identification of $\hat{\mu}$. To overcome this, we use years of education as a proxy for abilities and assume the average log ability in sector I relates to average log education as $\mathbb{E}\left(\ln a_I \mid \cdot\right) = \delta + \beta\mathbb{E}(\ln \text{edu}_I) + \varepsilon_I$, where δ is market-specific. This gives:

$$
\mathbb{E}\left(\ln \frac{a_F}{a_S} \mid a_F \hat{W} \geq a_S\right) = \beta\mathbb{E}\left(\ln \frac{\text{edu}_F}{\text{edu}_S}\right) + \varepsilon. \tag{D.6}
$$

Here, β represents the elasticity of ability with respect to education, assumed constant across markets and sectors, while ε is a zero-mean i.i.d. error term. By examining sectoral differences, we control for market-level factors affecting the education-ability mapping.

The left-hand-side of equation D.6 can be decomposed as:

$$
\mathbb{E}\left(\ln \frac{a_F}{a_S} \mid a_F \hat{W} \geq a_S\right) = \mathbb{E}\left(\ln \frac{N_F}{N_S} \mid a_F \hat{W} \geq a_S\right) - \mathbb{E}\left(\ln \frac{Em_F}{Em_S} \mid a_F \hat{W} \geq a_S\right), \tag{D.7}
$$

which expresses the mean (log) ability gap as the difference between the mean (log) effective labor supply gap and relative employment in the two sectors. For log-normal distributions, the mean (log) effective labor supply gap can be expressed as:

$$\mathbb{E}\left(\ln \frac{N_F}{N_S} \mid a_F \hat{W} \geq a_S\right) = \ln \frac{N_F}{N_S} - \frac{1}{2}\left(c_5 - c_6\right), \tag{D.8}$$

where c_5 and c_6 are the observed variances of log earnings in sectors F and S.

Combining equations (D.6)-(D.8), we obtain:

$$\ln \frac{N_F}{N_S} = \alpha + \beta^* \mathbb{E}\left(\ln \frac{\bar{\text{edu}}_F}{\bar{\text{edu}}_S}\right) + \gamma_1 \left(\ln \frac{c_1}{1-c_1}\right) + \gamma_2 \left(c_5 - c_6\right) + u, \tag{D.9}$$

where $\beta^* = \beta \frac{\rho-1}{\rho}$ and u represents measurement error.

While the left-hand-side is unobserved, we can use our model's equation to write it as:[1]

$$\ln \frac{N_F}{N_S} \propto \frac{\rho}{\rho-1} \ln \left(\frac{\text{Revenues}_F}{\text{Earnings}_S}\right) - \ln Z_F. \tag{D.10}$$

We also consider an alternative specification where we measure the left hand side as:[2]

$$\ln \frac{N_F}{N_S} \propto \frac{\rho}{\rho-1} \ln \left(\frac{\text{Earnings}_F}{\text{Earnings}_S} + \ln \Phi_F\right) - \ln Z_F. \tag{D.11}$$

Assuming u is uncorrelated with education, this provides a consistent estimate of $\hat{\beta} = \hat{\beta}^* \frac{\rho}{\rho-1}$ for a given choice of ρ.[3] With $\hat{\beta}$ estimated, we impute the mean log ability gap from equation (D.6) and finally derive $\hat{\mu}$ using equation (D.5) market by market.

Robustness A potential concern with our approach is that the estimate of β may be biased if (i) there is measurement error in the earnings data or (ii) the orthogonality assumption is violated. While considering two alternative methods to measure $\ln \frac{N_F}{N_S}$ partially addresses the

[1]Our model implies:

$$\begin{aligned} \frac{N_F}{N_S} &= \frac{Y_F}{Y_S}\left(Z_F\right)^{-1} \\ &= \zeta^{-\frac{\rho}{\rho-1}}\left(\frac{R_F}{R_S}\right)^{\frac{\rho}{\rho-1}}\left(Z_F\right)^{-1}, \end{aligned}$$

where $\frac{Y_F}{Y_S}$ is relative output across sectors, which under the CES structure of demand, can be expressed as a function of revenues. In the self-employment sector, total revenues coincide with self-employment earnings ($R_S = \text{Earnings}_S$). Taking logs yields equation (D.10).

[2]The rationale for this approach is that since firm revenues are measured from the firm survey and self-employment earnings from the worker survey, the variable $\ln \frac{\text{Revenues}_F}{\text{Earnings}_S}$ could be measured with error. We therefore express $\text{Revenues}_F = \text{Earnings}_F \cdot \Phi_F$ using the model's structure, leading to equation (D.11). This approach reduces measurement error by measuring relative earnings only with worker-level data. The disadvantage is that the wedge Φ_F is not observed, so that we need to rely on additional proxy controls to mitigate its potential confounding effects.

[3]For implementation, we use a $\rho = 2.7$, equal to the estimated value using the MSM method.

first concern, we also consider the simpler case of setting $\beta = 1$, imposing

$$\mathbb{E}\left(\ln \frac{a_F}{a_S} \mid a_F \hat{W} \geq a_S\right) = \mathbb{E}\left(\ln \frac{\text{edu}_F}{\text{edu}_S}\right).$$

While more restrictive, this approach is less susceptible to measurement error and omitted variable bias.

D.1.3 Results

Figure A.6 presents histograms of the estimated variance-covariance parameters and mean comparative advantage. Panel (a) shows the distributions of $\hat{\sigma}_F$ and $\hat{\sigma}_S$ across local labor markets, panel (b) depicts the histogram of the correlation parameter $\hat{\varrho}$, and panel (c) illustrates the distribution of the estimated mean comparative advantage $\hat{\mu}$. While there is some heterogeneity across markets, these estimates are generally well-behaved and centered around their means.

As explained in the main text, we set the $\boldsymbol{\Sigma_k}$ parameters and $\hat{\mu}$ constant across markets. Panel I of Table A.8 reports their median values. When computing the median of the $\boldsymbol{\Sigma_k}$ parameters, we exclude markets where the constrained minimum-distance procedure failed to find an interior solution to limit the influence of outliers. We verify that this exclusion does not significantly affect the correlation patterns between absolute and comparative advantage.

For $\hat{\mu}$, the table provides estimates from the three alternative cases discussed earlier. We calibrate the baseline model using the average of their median values.

D.2 Sensitivity of Model Moments to Parameters

To examine the relationship between model parameters and the generated moments, we follow a method similar to that of Kaboski and Townsend (2011), computing the sensitivity of each moment to each model parameter. The process involves the following steps:

1. We begin with the estimated vector of parameters, denoted by Φ^*, and create 18 alternative parameter vectors. For each parameter j, we generate two variations: one where Φ_j is reduced by 5%, $\Phi^- = \{\Phi^*_{-j}, 0.95 \cdot \Phi^*_j\}$, and one where it is increased by 5%, $\Phi^+ = \{\Phi^*_{-j}, 1.05 \cdot \Phi^*_j\}$. In both cases, all other parameters remain unchanged.
2. Using these adjusted parameter vectors, we then simulate the model to obtain the corresponding moment vectors. For each parameter change, we calculate the difference in each moment r, denoted as $\Delta_{jr} = m_r(\Phi^+) - m_r(\Phi^-)$. This difference quantifies how moment r changes when parameter j is altered by 10%, while keeping the other parameters fixed.

To facilitate comparison across moments, we normalize Δ_{jr} for each parameter such that, after rounding, the sum of the values across all parameters equals 9. This normalization creates an interpretable scale: if all parameters have an equal influence on a particular moment, the corresponding row in the matrix will show a value of 1 for each parameter. Alternatively, if only three parameters significantly affect a moment and their impacts are equal, each will show a value of 3, with the rest being 0, and so on.

The resulting Jacobian matrix, displayed in Figure A.7, clearly highlights the parameters that most strongly affect each moment. This intuitive mapping links specific parameters to the moments we would expect, as detailed in Section 4.2.2.

D.3 Full Estimation Algorithm

This section outlines the algorithm used to solve the model given the estimated parameters.

1. Using the estimated Roy parameters, Σ and $\hat{\mu}$, derive the functional expressions for labor supply and labor supply elasticity for a benchmark '0'-market, where we set $\mu_S = 0$.
2. For given parameter values $(\mu_T, \sigma_T, \theta)$, (f_0, f_1), (μ_μ, σ_μ), and (ρ, η, ζ), draw K local labor market productivities T_k from a log-normal distribution with parameters (μ_T, σ_T). Similarly, draw K mean absolute advantage parameters μ_k from a log-normal distribution with parameters (μ_μ, σ_μ). The seed for all random draws remains constant during estimation.[4]
3. For given values of parameter θ and realization of T_k in each market $k = 1, .., K$, we draw productivities of potential entrants $\{z_{iF,k}\}_{i=1}^{\bar{M}}$ as follows. We follow Eaton, Kortum and Sotelo (2012) and draw the productivity of the most productive firm and each firm thereafter, with spacings following an exponential distribution. Specifically, denote $U_k^{(n)} \equiv T_k z_{F,k}^{(n)-Z_\theta}$, where n is the rank of the firm in market k. Then $U_k^{(1)}$, $(U_k^{(2)} - U_k^{(1)})$, $(U_k^{(3)} - U_k^{(2)})$, etc., are i.i.d. exponential with cdf $G_U(u) = 1 - e^{-u}$ (Eaton, Kortum and Sotelo, 2012). We use the transformation to convert the exponential draws into productivity draws $\{z_{iF,k}\}_{i=1}^{\bar{M}}$. We cap the number of shadow firms $\bar{M}$ at 85, which is the maximum number of firms observed in the data.
4. With the calibrated value of local labor market shares and populations $\{\alpha_k, L_k\}_{k=1}^{K}$, the normalization $P = 1$, and given the functional forms for $\hat{N}(\cdot)$ and $\epsilon_F(\cdot)$, draws of $\{T_k, \mu_{\mu_k}, \{z_{iF,k}\}_{i=1}^{\bar{M}}\}_{k=1}^{K}$, and the remaining model parameters, we implement the following fixed point procedure:
   (a) Take an initial guess for aggregate income Y_0, which completes the general equilibrium vector $\mathbf{X} = (Y, 1)$.
   (b) Given $\mathbf{X}$, we solve for the market equilibrium $\mathbf{K} = \{\mathbf{M}, \hat{\mathbf{W}}, \mathbf{\Lambda}\}$, as detailed in Appendix C.1.1.
   (c) Given $\mathbf{K}$, use the parameters $\mu_{F,k}$ and $\mu_{S,k}$ to obtain the corresponding labor supply functions for each market by affinity with the functions of the benchmark market. Note that the market equilibrium is not affected by the specific value of these parameters, as it only depends on $\hat{N}$ and ϵ_F, which only depend on Σ and $\hat{\mu}$.

[4] To avoid mechanical correlations between the different distributions, we use separate random seed values for each distribution and verify that the correlations are close to zero.

(d) Given $\mathbf{K}$, use the general equilibrium conditions to solve for the new values of Y.

(e) Update the initial values of Y_0 taking the midpoint between the initial vector from step (a) and the new vector from step (c), and loop over until convergence.

(f) Upon convergence of the equilibrium vector $(\mathbf{X}, \mathbf{K})$, simulate the model and calculate the moment vector $\{m_k(\Phi)\}_{k=1}^{K}$ for all markets $k = 1, .., K$, corresponding to parameter vector Φ.

5. On a grid for parameters Φ with 100,000 points, evaluate the moment function $m_k(\Phi)$, with moments described in Table 4, and the associated MSM loss function:

$$\mathcal{L}(\Phi) = \hat{\mathbf{f}}(\Phi)'\mathbb{W}\hat{\mathbf{f}}(\Phi),$$
$$\text{where} \quad \hat{\mathbf{f}}(\Phi) \equiv f(m_k(\Phi)) - f(\tilde{m}_k)$$

and where $\tilde{m}_k$ are the values of the moments in our empirical dataset, the function $f(\cdot)$ is the simple average: $f(x_k) = K^{-1}\sum_k x_k$, and $\mathbb{W}$ is the weighting matrix, which we chose to be diagonal and inversely proportional to $\tilde{\mathbf{m}}$.[5] We use a Halton sequence to define the grid points, so that it covers the whole parameter space more efficiently than if points were regularly spaced.

6. With the results from the first Halton grid, we recompute a second finer Halton grid of 50,000 points. We restrict this grid to be wide enough to encompass the 50 best fitting parameter values of the previous grid, but exclude the regions with the highest loss function. We iterate this procedure several times, until convergence to a narrow region of the parameter space.

7. We take as our estimate (the global minimizer) the point of local convergence with the lowest loss function, $\hat{\Phi} = \arg\min_{\Phi} \mathcal{L}(\Phi)$.

D.4 Model Fit

To assess the model's fit, we replicate the reduced-form elasticities from Table 2 by simulating the shock to firm productivity and labor demand used in Section 2.5.

We proceed in three steps. First, we identify the counterpart of treated firms in the model. We extend the model by assigning an electricity wedge (τ) to all firms, as estimated from the data. Specifically, we restrict the sample to the baseline year and use OLS to estimate the parameters from the following specification:

$$\tau_{i(j,g)t_0} = \beta_\tau z_{i(j,g)t_0} + e_{i(j,g)t_0},$$

where $\tau_{i(j,g)t_0}$ represents the electricity wedge estimated at the firm level, as detailed in Section 2.5, $z_{i(j,g)t_0}$ denotes value added per worker, and $e_{i(j,g)t_0}$ is a normally distributed i.i.d. shock.

[5]Because of the poor fit in matching the correlation coefficient mgintudes between concentration and log earnings, we downweight these moment sby a factor of 30 to improve overall precision.

We estimate $\hat{\beta}_\tau = 0.08$, significant at the 1% level.

In the model, we then calculate the associated $\hat{\tau}_{ik}$ for each firm as

$$\hat{\tau}_{ik} = \hat{\beta}_\tau z_{ik} + e_{ik},$$

where z_{ik} is the firm-level productivity draw in local labor market k and $e_{ik} \sim \mathcal{N}(0,1)$. Once each firm is assigned a wedge, we consider a firm treated if their $\hat{\tau}_{ik}$ is greater than the economywide median, mirroring the strategy described in Section 2.5.

Second, from the data, we back up the size of the electrification shock in terms of productivity improvements—specifically, the average productivity increase at the firm level induced by the treatment. We do this by estimating using OLS the parameters from the following specification:

$$\ln z_{i(j,g)t} = \gamma PER_{gt} \times EC_{i(j,g)} + \phi_i + \delta_{(j,g)t} + v_{i(j,g)t}, \tag{D.12}$$

which is akin to the specification in equation (3) but has the log of value added per worker as dependent variable. We estimate $\gamma = 0.0019$, significant at the 5% level. Given that the mean of the treatment ($PER_{gt} \times EC_{i(j,g)}$) is approximately 12, the average treatment effect is an increase of 2.3% in firm-level productivity.

Third, starting from the baseline model equilibrium, we simulate a 2.3% productivity shock for the treated firms. We then obtain the firm-level inverse elasticity by taking the ratio between the (log) wage and the (log) employment responses $\hat{\epsilon}_{iF,k} \equiv \Delta \ln W_{F,k} / \Delta \ln n_{iF,k}$. We do this for treated firms only and in markets with more than one firm, consistent with the within-market identification strategy and the LATE nature of the estimates in Table 2.