STEG WORKING PAPER

LABOR MARKET POWER, SELF-EMPLOYMENT, AND DEVELOPMENT

Francesco Amodio, Pamela Medina,
and Monica Morlacco

AUGUST 2022
REVISED APRIL 2023
STEG WP016

Labor Market Power, Self-Employment, and Development*

Francesco Amodio[†] Pamela Medina[‡] Monica Morlacco[§]

April 15, 2023

Abstract

This paper shows that self-employment shapes labor market power in low-income countries, with implications for industrial development. Using Peruvian data, we show that wage-setting power increases with concentration, but less so where self-employment is more prevalent. We build a general equilibrium model of oligopsony with worker sorting between wage work and self-employment. Concentration depresses wages, but self-employment increases workers' sensitivity to wage changes, curbing labor market power. Policies to create salaried jobs make self-employment less attractive, reducing labor supply elasticity and increasing markdowns. Counterfactual analyses show that eliminating labor market power can boost industrial policy effectiveness by up to 60%.

Keywords: labor market power, monopsony, self-employment, sorting, 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, 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, EIEF, JILAEE-UChicago, Loughborough, LSE, Manchester, McGill, McMaster, Milan, Montevideo, NCID-Navarra, Nottingham, Notre Dame, Stockholm, UAB, UC Riverside, UofT, UPF, USC, Warwick, Western, as well as participants in the BSE Summer Forum "Market Power and the Labor Market" Workshop, the CEPR-ENS-Lyon "Firms' Behavior in Hostile Environments" Workshop, the FIMAD Seminar Series, the JADE-CEPR-TIME Conference on Economic Development, LACEA, the MEA meetings, the NYU-Columbia Trade day, PacDev, the RES Annual Conference, RIDGE, Seminario MEP, and the STEG Theme 1 Workshop. Errors remain our own. We gratefully acknowledge financial support from the CEPR/DFID PEDL initiative (ERG 6500).

†francesco.amodio@mcgill.ca, McGill University and CIREQ, 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

Millions of people in low- and middle-income countries rely on subsistence labor for their livelihoods. Still, the role of informal self-employment in economic development remains a contentious issue. The traditional view is that economic development comes from modern industrial firms and that informal self-employment will eventually disappear as the industrial sector expands (Lewis, 1954; Harris and Todaro, 1970; Rauch, 1991). As a result, the creation of manufacturing salaried jobs has been a cornerstone of industrial development policy (UN General Assembly, 2015). Despite these efforts, self-employment remains high in emerging countries (La Porta and Shleifer, 2014; Gollin, 2008; Poschke, 2022) and employment at larger firms stagnates (Hsieh and Olken, 2014; Diao et al., 2021; McMillan and Zeufack, 2022), even as GDP per capita increases. Understanding why self-employment persists and why manufacturing firms cannot absorb more workers is crucial in determining the future development trajectory of these countries and the scope for policy intervention.

Labor market structure is a potentially important but often overlooked factor influencing these outcomes. Multiple barriers to firm growth, such as high entry costs, a shortage of skilled labor, and inadequate infrastructure, result in the concentration of employment among a few firms (Djankov et al., 2002; Hsieh and Klenow, 2010; Rud and Trapeznikova, 2021; Hjort, Malmberg and Schoellman, 2022). These firms may internalize their impact on local labor market conditions, reducing job opportunities and wages to increase profits. However, self-employment represents a valuable outside option for workers. Within a local labor market, workers can easily switch between self-employment and wage work (Donovan, Lu and Schoellman, 2023), and they can opt for self-employment when posted wages are too low (Blattman and Dercon, 2018; Breza, Kaur and Shamdasani, 2021).

This paper argues that understanding the interplay between labor market power and self-employment is crucial to explain the persistently high rates of self-employment in emerging economies and why development policies aimed at boosting industrial wage employment often fall short of their objectives.[1] To support these arguments, we present new evidence from Peru, an original theoretical framework, and counterfactual policy experiments. The case of Peru is meaningful because its economy, with its high concentration and self-employment rates and extensive worker mobility, is representative of many other low- and middle-income countries.

We begin by showing that labor market power is substantial in Peru. Its extent varies across local labor markets depending on employer concentration and self-employment opportunities, with the latter acting as a constraint on employers' wage-setting power.[2] We measure labor market power as the inverse elasticity of the labor supply curve faced by individual firms, a direct measure of their ability to set wages (Manning, 2003). For estimation, we use an instrumental variable strategy that constructs firm-level labor demand shifters from the staggered

[1] See McKenzie (2017) and Bandiera et al. (2022) for a review. See also OECD (2006); World Bank (2012).

[2] We define a local labor market as the combination of a 2-digit industry and a commuting zone, similar to Berger, Herkenhoff and Mongey (2022). See Section 2 for further details and motivation for this definition.

implementation of a rural electrification program across provinces and its differential impact on firms with high vs. low ex-ante constraints in accessing electricity.

We find that the average firm-level inverse labor supply elasticity across local labor markets is positive and significant, indicating substantial wage-setting power. The implied average wage markdown is 1.43, meaning that manufacturing workers receive about 70 cents as a wage for every additional dollar they produce. The markdown widens with market concentration, suggesting oligopsony power among employers (Berger, Herkenhoff and Mongey, 2022). However, this positive relationship weakens in markets where self-employment is more prevalent. We find the widest markdowns in markets with high concentration and low self-employment rates, where workers receive only 57 cents for the marginal dollar they produce. Conversely, the labor supply is relatively more elastic in markets with a large self-employment sector.[3]

Motivated by the empirical evidence, we build a general equilibrium model where Peru is treated as a collection of segmented local labor markets. Within each labor market, workers can work for a firm or be self-employed. The model's first key feature is *oligopsony*. There exists an endogenous finite number of heterogeneous firms in each local labor market, which internalize their impact on market-level wages and make strategic decisions accordingly. The model's second key feature is Roy's (1951) structure of *self-selection of heterogeneous workers* across wage work and self-employment based on comparative advantage and unit wages. Within-market worker heterogeneity is defined in terms of sector-specific skills or efficiency labor units; it is governed by a joint ability distribution, which we allow to be labor market-specific.

The theory sheds light on the structural determinants of labor market power. In equilibrium, the (payroll-weighted) average markdown in a local labor market is an exact function of two endogenous variables. The first one is the payroll Herfindahl-Hirschman Index (HHI), a standard measure of employer concentration. Concentration positively affects the markdown and captures the *demand-side* determinants of labor market power, specifically the employers' oligopsony power (Berger, Herkenhoff and Mongey, 2022; Felix, 2022). The second variable is the aggregate supply elasticity of wage work, which has a negative effect. This term reflects *supply-side* forces, notably how wage changes affect the sorting of workers across wage work and self-employment. As the relative unit wage falls, more workers choose self-employment. Falling wages make it easier to push workers out of wage employment, resulting in an increase in the overall supply elasticity of wage work and a decline in the average markdown.[4] Similar forces can also increase labor market power when wage employment becomes more attractive, for instance as a result of policy interventions.

Our framework captures the dual role played by self-employment in the presence of labor market power. It can shield workers from the wage-setting power of firms by providing a livelihood when wage opportunities are scarce. However, it can also increase labor market power

[3] In Brazil, Felix (2022) also found that firms in local labor markets where self-employment is more prevalent face more elastic labor supply curves.

[4] In Appendix B, we discuss conditions on the workers' ability distribution for the supply elasticity of wage employment to always *decrease* as relative unit earnings increase.

when wage employment becomes more attractive, making it difficult for industrial policies to boost wage employment and wages and potentially hindering the growth prospect of countries. Through counterfactual analysis, we show that the variable elasticity channel is quantitatively essential to understand the limited impact of industrial policies in emerging economies.

We use the model to decompose the response of sectoral average earnings to economic shocks. The average wage response reflects two components: a direct effect of the shock on the efficiency unit wage and a compositional effect on the average worker ability. The change in the unit wage can be further decomposed into the change in workers' marginal revenue product (MRPL) and the change in markdown. In turn, the first effect is the sum of a *misallocation* channel and a *general equilibrium* channel. Misallocation arises because the larger and more productive firms have wider markdowns. A shock that reduces the average markdown reallocates market shares towards the most productive firms, increasing market-level MRPL and unit wages. Moreover, markdowns and misallocation induce aggregate output losses that negatively affect all economic outcomes through general equilibrium forces.[5]

We refer to the compositional effect on workers' ability as the *selection* channel. This effect reflects the difference in efficiency between sector-switchers and sector-stayers. The strength and direction of this channel depends on the schedules of workers' comparative and absolute advantage in the two sectors and their correlation (Adão, 2015; Alvarez-Cuadrado, Amodio and Poschke, 2020). We estimate that the workers' abilities in the two sectors are positively correlated and slightly more dispersed in the self-employment sector. For estimation, we impose that ability endowments are jointly log-normally distributed, allowing us to estimate the relevant ability parameters from cross-sectional data on earnings and employment shares (Heckman and Sedlacek, 1985; Heckman and Honoré, 1990). Our results imply a weak negative correlation between absolute and comparative advantage in wage work, resulting in positive selection in self-employment and weak negative selection in wage work. As a consequence, when the size of the wage employment sector decreases (increases), average labor productivity decreases (increases) in both sectors, reducing (increasing) average wages.

Given the ability distribution parameters, we rely on a Method of Simulated Moments (MSM) strategy to estimate the remaining model's parameters. We discipline the model by matching moments derived from our stylized facts on cross-sectional dispersion of concentration and self-employment. We test the model by showing that it quantitatively replicates the reduced-form patterns of labor market power across local labor markets, which were not targeted for estimation.

Armed with the estimated model, we perform two counterfactual experiments. First, we evaluate the effect of labor market power on labor market outcomes in Peru by comparing our baseline economy with one where employers act as wage-takers. Our estimates imply that eliminating labor market power would raise the average share of wage employment across markets

[5] In the self-employment sector, the markdown is always constant and equal to one, while the MRPL is only affected by general equilibrium forces.

by over ten percentage points, from 60% to 71%. Furthermore, average earnings would increase by 31% in the wage employment sector and 27% in the self-employment sector. These effects occur through selection, markdown, misallocation, and general equilibrium channels. Worker self-selection stands out as a crucial margin through which labor market power decreases worker earnings in the self-employment sector.

The second objective is to investigate the role of labor market power in limiting the impact of industrial development policies. We examine three categories of policies aimed at promoting manufacturing wage employment: (i) enhancing firm productivity through market integration or infrastructure improvement policies (Volpe Martincus, Carballo and Cusolito, 2017; McCaig and Pavcnik, 2018; Fiorini, Sanfilippo and Sundaram, 2021); (ii) reducing fixed entry costs for employers by simplifying business registration regulations (Kaplan, Piedra and Seira, 2011; Bruhn, 2011) or alleviating financial constraints for firms (Buera, Kaboski and Shin, 2011); (iii) improving workers' skills through off and on-the-job training programs (McKenzie, 2017; Alfonsi et al., 2020). We estimate the impact of these policies on labor market outcomes in both our baseline economy and in the counterfactual scenario where firms are wage-takers, and identify the effect of labor market power using a *difference-in-differences* approach. To inform the size of the policy shocks, we analyze actual policies implemented in Peru and Mexico and their reduced-form estimated effects.

Our findings indicate that industrial policies can boost labor income only when they substantially decrease wage markdowns, which is rare. In most cases, a policy-induced expansion of the wage employment sector reduces the wage work supply elasticity. This may increase labor market power, even if concentration decreases. Through the variable elasticity channel, labor market power significantly reduces the impact of policies. Our estimates show that policies aimed at raising firm productivity or improving worker skills could increase labor income by 64% and 40% more without labor market power, respectively.

Related Literature This paper contributes to several strands of the literature. Firstly, we contribute to the literature on informal self-employment in low-income countries. The traditional "dual" view suggests that medium and large formal firms and informal micro-enterprises are fundamentally different and operate in entirely different economic spheres.[6] This view has been challenged by Maloney (1999), Ulyssea (2018), and Donovan, Lu and Schoellman (2023), among others, who demonstrate that formal and informal firms coexist in the same local labor markets, characterized by frequent worker transitions between the two sectors. Our analysis belongs to this second view in that it emphasizes the importance of worker sorting across the two sectors for understanding labor market outcomes in emerging countries in the presence of labor market power, as well as the high persistence of self-employment in these contexts, even as GDP per capita increases.

[6] Early contributors to this literature include Lewis (1954), Harris and Todaro (1970), and Rauch (1991). See also La Porta and Shleifer (2014) for a review article.

Second, our paper contributes to the growing literature on labor market power. Recent evidence shows that U.S. employers enjoy some degree of market power in the labor market (Azar, Berry and Marinescu, 2022; Berger, Herkenhoff and Mongey, 2022). Several studies use employer concentration as a proxy for labor market power showing that it correlates negatively with wages (Azar, Marinescu and Steinbaum, 2022; Benmelech, Bergman and Kim, 2022). Yet, using matched employer-employee data from Oregon, Bassier, Dube and Naidu (2022) find no evidence that labor supply elasticities decrease with concentration. Similarly, in U.S. manufacturing data, Yeh, Macaluso and Hershbein (2022) find that aggregate wage markdowns and employer concentration moved along different trends over the last two decades. We introduce an original micro-foundation for the firm-level labor supply curve based on the self-selection of heterogeneous workers between wage work and self-employment.[7] We therefore consider both demand- and supply-side determinants of labor market power to show that, with sorting, employer concentration can have a non-linear relationship with labor market power, providing a rationale for the mixed findings in the literature.

The literature on labor market power in lower-income countries is more limited. Amodio and De Roux (2022) use plant and customs data from Colombia to estimate firms' wage-setting power, concluding that workers produce 40% more than their wage level. Felix (2022) studies the impact of trade liberalization on concentration and wages in Brazil, estimating high levels of labor market power before the 1990s liberalization, but minor labor market power effects of trade.[8] Méndez-Chacón and Van Patten (2022) document the critical role of labor mobility and workers' outside option on determining the degree of monopsony power of private companies and their investment in local amenities in Costa Rica. Still in Costa Rica, Alfaro-Ureña, Manelici and Vasquez (2021) find minor wage effects following multinational companies' expansion, indicating low labor market power.[9] We add to this literature by presenting new evidence on the interplay between labor market power, concentration, and self-employment rates. We propose and estimate a novel general equilibrium model to demonstrate that self-employment acts as a check on employers' market power while, at the same time, undermining development policies in low-income countries.

Finally, our work speaks to the extensive literature on informality in low-income countries (Ulyssea, 2020). 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. (2021) show that unemployment buffer does not

[7] Kahn and Tracy (2019) study how local monopsony power affects the cross-sectional spatial distribution of wages and rents across cities incorporating as an extension worker sorting across sectors *à la* Roy.

[8] See also Pham (2019) and MacKenzie (2019) on the interactions between trade and labor market distortions in China and India, respectively.

[9] Outside Latin America, Brooks et al. (2021*b*,*a*) combine theory and data to show evidence of labor market power in China and India. Muralidharan, Niehaus and Sukhtankar (2017) show through experimental evidence that the labor market effects of public employment programs in rural India are consistent with monopsony power in private-sector employment. In South Africa, Bassier (2023) uses a variety of worker separation designs to estimate firm-level labor supply elasticities and finds high levels of monopsony.

necessarily imply "welfare buffer" meaning that, in the event of a negative economic shock, welfare declines by less when informality rates are modest.[10] Our analysis adopts the notion of informal self-employment as a potential outside option for workers, and shows it has a similar dual role in the presence of labor market power: it shields workers against the wage-setting power of employers when wages are too low, but also makes it more difficult for policies that seek to boost wage employment and wages to succeed.

The remainder of the paper is organized as follows. Section 2 introduces the data sources and definitions. Section 3 presents the empirical facts. The model and its properties are presented in Section 4, while Section 5 discusses the model estimation procedure and results. Section 6 presents the counterfactual policy analyses. Section 7 concludes.

2 Data and Definitions

The empirical analysis draws from two main datasets with information on firms and workers. The first data source is the Peruvian Annual Economic Survey (*Encuesta Económica Anual*, EEA), a national firm-level survey administered yearly by the national statistical agency (*Instituto Nacional de Estadística e Informática*, INEI) to characterize the structural composition of the economy at the national and sub-national level.

The dataset includes standard balance sheet information, such as revenues, labor and material expenditures, and plant location. The survey questionnaire is filed electronically and required for all firms with net sales above a given known threshold. As a result, the EEA provides information on the universe of medium and large firms. To ensure consistency across years, we focus on all firms surveyed from 2004 to 2011 operating in the manufacturing industry and reporting net sales per year above 2 million Peruvian Soles (PEN) – equal to around 700K USD in 2010. Our final dataset counts 2,473 firms and 8,138 firms $\times$ year observations. As explained below, we validate our EEA sample by comparing it with the 2007 Peruvian Economic Census of all establishments in the manufacturing sector. We derive summary statistics at the local labor market level and find them remarkably close in the two cases.

The second data source is the Peruvian National Household Survey (*Encuesta Nacional de Hogares*, ENAHO), which is carried out by the INEI every year to measure households' living conditions and the impact of social programs. The survey covers urban and rural areas across the 24 Peruvian departments and the constitutional province of Callao; it is representative at the national and regional levels. The data provide information on all household members' demographics, education, and other individual characteristics. Respondents aged 14 or older fill out a specific module that includes questions on employment status, pay, occupation, and industry of employment. To be consistent with the firm-level data, we focus on 2004 to 2011

[10] However, Meghir, Narita and Robin (2015) show that, with search frictions and firms posting wages in both the formal and informal sector, increasing enforcement against informality does not increase unemployment and increases wages and welfare.

and restrict the sample to working-age individuals aged 25 to 65 who have completed their education and are not yet retired. ENAHO has several panel versions available where the same subset of households is interviewed every year for five consecutive years. We use the 2007-2011 panel to document workers' transitions across employment states.

The survey classifies workers as own-account workers, employers, auxiliary family workers, or employees. We label the first two categories as *self-employed* and the latter as *wage workers*. We exclude auxiliary family workers from our classification as they do not report monetary compensation. The information contained in ENAHO also allows identifying informal workers. A worker is labeled as informal if s/he (i) is a paid worker but reports not having health insurance,[11] or (ii) is self-employed, but s/he is not registered or follows the procedures demanded by the national tax authority and has five or fewer employees.

We define a local labor market as a 2-digit industry within a geographical area.[12] Our focus on industries rather than occupations is similar to Berger, Herkenhoff and Mongey (2022) and motivated by the fact that 86% of (self-employed or wage) workers are observed in the same 2-digit industry for two consecutive years. This probability declines at about 60% when looking at 3-digit occupational codes. Moreover, information on the worker's industry of employment is available much more frequently than information on occupation in ENAHO. This increases the size of our worker-level sample by 17%.

The geographical areas are primarily informed by Peruvian province boundaries. These are level 2 administrative units (the homolog of counties in the US), and sub-divisions of departments. Excluding Metropolitan Lima – the province that includes the capital city of Lima – the average province has a population of approximately 114,000. Metropolitan Lima is an outlier, with a population of 10 million. We thus follow Piselli (2013) and define separate local labor markets within the Lima province. The Survey of Transport, Labor, and Technology Use assembled by the Peruvian Studies Institute (IEP) identifies five distinct zones in which people do most of their activities: Lima Center, Lima North, Lima South, Lima East, and Lima West. In total, we have information on 199 geographical units and 23 manufacturing industries.

3 Facts

We begin by documenting several facts about Peruvian manufacturing labor markets. We find a systematic correlation across local labor markets between concentration and self-employment prevalence and between concentration and earnings from both wage work and self-employment. Workers frequently switch between sectors, and transitions are correlated with earnings. We then estimate that labor market power is widespread and positively associated with employer

[11] Employers in Peru are required by law to provide health insurance to their employees.

[12] We consider 2-digit CIIU Rev. 4 code industries in the manufacturing sector. Examples of adjacent 2-digit industry codes are 10 - Manufacture of Food Products, 11 - Manufacture of Beverages 12 - Manufacture of Tobacco Products and 13 - Manufacture of Textile Products.

concentration, but less so in markets with high self-employment rates.

3.1 Employer Concentration

Our baseline measure of concentration is the Herfindahl-Herschman index (HHI) for payroll. Let w_{ikt} and n_{ikt} denote wage and employment, respectively, of firm i in local labor market k in year t. The payroll or wage-bill HHI is defined as $HHI_{kt}^{wn} = \sum_{i\in k}(s_{ikt}^{wn})^2$, where $s_{ikt}^{wn} = w_{ikt}n_{ikt}/\sum_{i\in k} w_{ikt}n_{ikt}$ is firm i's payroll share in the local labor market. Values of the HHI close to one indicate that a few firms account for a large share of payroll in the market.[13]

Across Peruvian local labor markets, employment and wages are concentrated in a few medium and large firms.[14] As reported in Panel A of Table 1, the average local labor market counts about six firms. The unweighted and payroll-weighted average HHI for wages are 0.65 and 0.37, respectively, revealing that more concentrated markets account for a smaller share of nationwide payroll.[15] Still, the 39% of local labor markets with only one medium-to-large firm account for around 8% of total employment and payroll.

Location is more important than industry in explaining concentration. 43% of the variation in wage-bill HHI across markets is between locations. Variation between 2-digit industries accounts for another 14% of the total.

3.2 Self-Employment and Flows Into and From Wage Work

In Peruvian manufacturing, like in other low- and middle-income countries, self-employment is widespread and predominantly informal (Gollin, 2008; La Porta and Shleifer, 2014). Panel B of Table 1 reports that wage workers account for 56% of the manufacturing workforce, while 40% are self-employed, and the remaining 4% are auxiliary family workers. Over 90% of self-employment is informal across all industries and within manufacturing. In contrast, about half the wage workers are informal, with numbers declining over time.[16]

Informality is a crucial factor determining the prevalence of self-employment as it lowers the costs associated with starting and operating a business by reducing the need to comply

[13] As alternative measures of concentration, we consider (i) employment HHI, $HHI_{kt}^{n} = \sum_{i\in k}(s_{ikt}^{n})^2$ with $s_{ikt}^{n} = n_{ikt}/\sum_{i\in k} n_{ikt}$, and (ii) the number of firms in the local labor market. Online Appendix Figure A.1 plots the distributions of the three concentration measures. Figure A.2 shows the three concentration measures correlate strongly. Our findings are robust to using these alternative measures.

[14] The medium and large firms in the EEA account for the vast majority of wage employment in their local labor markets. In Online Appendix Figure A.3, we show that the concentration measures obtained using the EEA data align closely with those obtained from the 2007 Peruvian Economic Census.

[15] The numbers for Peru are more significant than the corresponding ones for the United States, for which Berger, Herkenhoff and Mongey (2022) report 0.48 and 0.17, respectively, as unweighted and payroll-weighted average wage-bill HHI across local labor markets in 2014.

[16] Informal workers account for 73% of the workforce in our data. This number is close to that reported by the INEI in 2007 (80%), which estimates that the informal sector accounts for about a fifth of aggregate GDP. The high rate of informal self-employment is in stark contrast with the tiny unemployment numbers. In our data, the national unemployment rate is about 3%, very close to the 3.2 to 3.6% reported by the International Labor Organization for Peru over the same period (International Labour Organization, 2020).

Table 1: Summary Statistics

	Mean	St. Dev.
A. 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
B. Manufacturing Workers		
Wage Worker	0.56	0.50
Self-Employed	0.40	0.49
W-S Transition	0.08	0.27
S-W Transition	0.09	0.28

Notes. This table reports summary statistics from EEA firm-level data across Peruvian local labor markets (Panel A) and from ENAHO worker-level data (Panel B), averaging across all years from 2004 to 2011. Local labor markets are defined as 2-digit industries within locations, corresponding to Peruvian provinces or commuting zones. Worker-level statistics are for dummy variables indicating wage work, self-employment, and annual transitions from the wage- to self-employment sector (W-S) and vice versa (S-W).

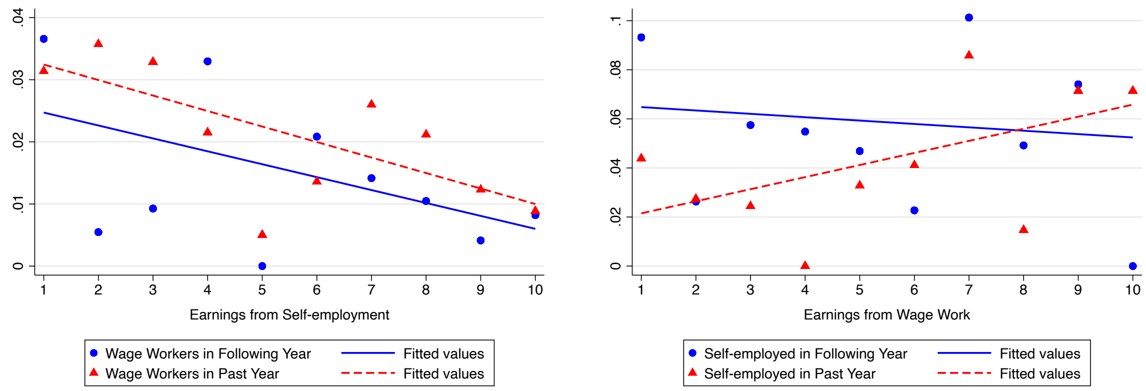
with tax and labor regulations. The variation in the rate of self-employment across different industries reflects these costs. In labor-intensive industries, the initial investment required for physical capital is low, credit constraints are less severe, and the scope for informality is larger. As a result, self-employment rates are higher in these industries, which account for most of the Peruvian manufacturing GDP. For instance, clothing and furniture manufacturing have self-employment rates that exceeds 50%. On the contrary, self-employment is comparatively lower in more capital-intensive sectors such as pharmaceuticals or metals and essentially absent in oil and petroleum manufacturing.

A crucial aspect of the Peruvian labor market is the ease of moving from wage work to self-employment and vice versa. Panel B of Table 1 shows that around 8% of self-employed manufacturing workers in a given year become wage workers the following year. Similarly, 9% of all manufacturing wage workers in a given year transition to self-employment in the next year. These numbers resonate with the evidence in Maloney (1999) and Donovan, Lu and Schoellman (2023) for Mexico, Peru and other low- and middle-income countries.[17]

Worker transitions within a local labor market are systematically related to earnings. The left panel of Figure 1 plots the average yearly probability of transitioning to or from wage work

[17] As explained above, the vast majority of worker transitions within manufacturing (86%) occur within the same 2-digit industry. Additionally, there is limited evidence of substantial annual migration flows, indicating that worker mobility appears confined within local labor markets.

Figure 1: Transitions Probabilities Across the Earnings Distribution



Notes. The figures illustrate the relationship between the likelihood of transitioning across sectors and earnings. The left panel plots average yearly transition probabilities into and from self-employment across the wage work earnings distribution deciles. Similarly, the right panel plots average yearly transition probabilities into and from wage work across deciles of the self-employment earnings distribution.

across deciles of the self-employment earnings distribution. These moves are concentrated at the bottom of the distribution, meaning that workers who have just transitioned from wage work or will become wage workers in the next period are more likely to be among the lowest-earning self-employed. Similarly, the right panel of Figure 1 plots the likelihood of transitions into and from self-employment across deciles of wage work earnings distribution. Future transitions to self-employment are evenly spread across the distribution. In contrast, those who have just transitioned to wage work from self-employment are among the highest-earning wage workers.

These patterns show that sector movers earn systematically less than self-employment sector stayers but (weakly) more than the wage employment sector stayers. Interpreted through the lens of a model of worker sorting, they suggest positive selection in self-employment and (mild) negative selection in wage work. We will return to this point below.

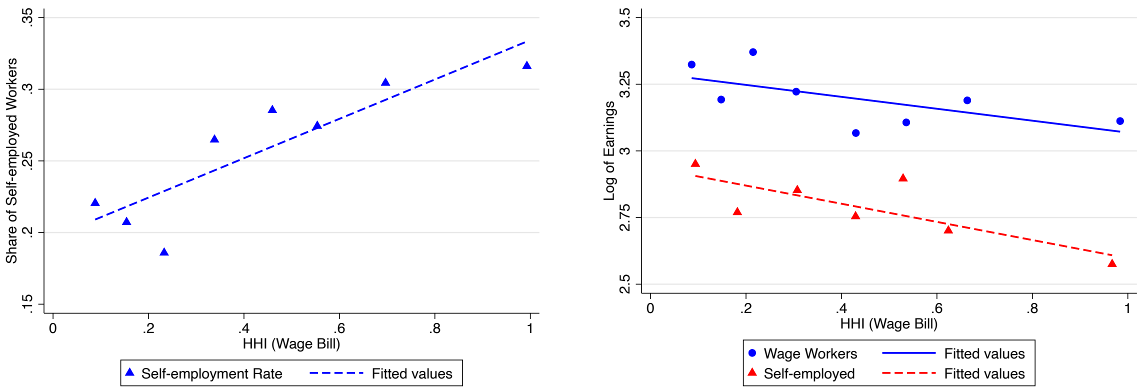
3.3 Concentration, Self-Employment Rates, and Earnings

Across local labor markets, concentration correlates strongly with self-employment rates. The left panel of Figure 2 presents the proportion of self-employed workers in each local labor market across the wage-bill HHI distribution. Higher concentration is consistently associated with higher self-employment and lower wage employment shares. Regression analysis confirms this relationship. We implement a worker-level regression where a dummy indicating self-employment is regressed over the log of wage-bill HHI in the local labor market in the same year. The results, presented in Columns 1 to 3 of Online Appendix Table A.2, demonstrate that the relationship between concentration and self-employment is positive and significant, even after controlling for individual factors, industry, and location fixed effects.[18]

The right panel of Figure 2 shows the correlation between concentration and earnings from

[18]Online Appendix Tables A.3 and A.4 report the coefficient estimates obtained using employment HHI and the number of firms, respectively, as alternative concentration measures, showing similar results.

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 local labor market across the wage-bill HHI distribution deciles. The right panel does the same but plots the average log of daily earnings separately for wage and self-employed workers.

the two sectors.[19] In local labor markets where concentration is high, wages are lower, and self-employment—while more common—is less lucrative. The regression results, presented in Columns 4 to 9 of Online Appendix Table A.2, lend additional support to these patterns.

These findings support the sorting story proposed earlier: where concentration is high, wages are lower and more workers choose self-employment. These workers would have been among the highest earners as wage workers, but as self-employed individuals, they earn less than their peers. This ultimately leads to lower average earnings in both sectors.

3.4 Labor Market Power

The correlation between concentration, self-employment, and earnings in the data raises the question about labor market power in this association. While employer concentration is strongly negatively associated with wages, this correlation alone is not evidence of labor market power as both concentration and wages are endogenously determined by supply and demand factors in the market.

To estimate labor market power, we leverage the rollout of a nationwide electrification program to estimate the inverse elasticity of labor supply faced by individual firms. This inverse elasticity is a direct measure of wage-setting power.[20] By examining how this elasticity varies across local labor markets with different levels of concentration and self-employment rates, we

[19]On average, self-employed workers earn considerably less than wage workers. Across all industries and within manufacturing, the average wage worker earns about 50% more than the average self-employed worker. These gaps are even higher ($\sim$150%) when comparing formal wage workers with informal self-employed workers.

[20]Following Manning (2003), consider a firm that produces output Y using labor N as input, i.e. $Y = Y(N)$ with $Y'(N) > 0$ and $Y''(N) < 0$. The firm chooses the amount of labor that maximizes $Y(N) - w(N)N$ which leads to the first order condition $\frac{\partial Y(N)}{\partial N} = w(N)\left[1 + \frac{\partial w(N)}{\partial N}\frac{N}{w(N)}\right]$. The wage is below the marginal revenue product of labor, and the markdown is exactly equal to 1 plus the inverse elasticity of firm-level labor supply $\varepsilon^{-1} = \frac{\partial w(N)}{\partial N}\frac{N}{w(N)}$.

can better understand the role of labor market power in this relationship.

3.4.1 Empirical Strategy

To estimate the inverse elasticity of the firm-level labor supply curve, we implement the following regression specification:

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

where $w_{i(j,g)t}$ is the wage paid by firm i in year t in its local labor market as 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-specific constant that captures differences across firms that do not change over time. $\eta_{(j,g)t}$ is a local labor 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 labor supply elasticity of wage work while holding the aggregate labor supply constant.

Estimating the parameters in equation (1) using Ordinary Least Squares (OLS) can yield biased and inconsistent results due to the interdependence of wages and employment. To determine the slope of the firm-level labor supply curve, we need a labor demand shifter at the firm level. To this end, we rely on an electrification program for identification. The Rural Electrification Program (*Programa de Electrificación Rural*, PER) was introduced by the Peruvian Ministry of Energy and Mining in 1993 to promote economic and social growth in rural areas (Dasso and Fernandez, 2015). The program was implemented from 1994 to 2012 and involved 628 electrification projects across rural Peru, with a total cost of USD 657.5 million. The Ministry prioritized districts with high poverty rates, low electricity coverage, low cost per connection, and a high potential for renewable energy use (Dasso, Fernandez and Nopo, 2015).

Our approach is based on the idea that electrification through the program had a positive effect on firms' marginal revenue product and labor demand. This was especially true for firms previously facing more severe constraints in accessing electricity (Abeberese, Ackah and Asuming, 2019). To operationalize this, we create a variable, PER_{gt}, counting the total number of completed PER projects in location g up to year t. We then follow Bau and Matray (2023) to identify firms in each industry that face constraints in accessing electricity.

Consider firm i in local labor market (j, g) that produces output $y_{i(j,g)t}$ at time t. The output market is imperfectly competitive, and output is sold at a unit price $p_{i(j,g)t}$, which is a markup $\mu_{i(j,g)t}$ over marginal cost. The firm produces using a Cobb-Douglas production function with industry j-specific input elasticities. Inputs include labor, electricity, and others. Electricity's output elasticity is denoted by θ_j^e. Its shadow cost varies across firms and industries and is captured by $\tau_{i(j,g)t}^e$. Finally, we denote by $e_{i(j,g)t}$ the firm's total electricity bill. Profit maximization implies

$$\theta_j^e \frac{p_{i(j,g)t} y_{i(j,g)t}}{e_{i(j,g)t}} = \mu_{i(j,g)t}(1 + \tau^e_{i(j,g)t}),$$
$$\ln\left(\frac{p_{i(j,g)t} y_{i(j,g)t}}{e_{i(j,g)t}}\right) = \ln(\mu_{i(j,g)t}) + \ln(1 + \tau^e_{i(j,g)t}) - \ln \theta_j^e. \tag{2}$$

This equation shows that one can retrieve firm-level estimates of the wedge $\tau^e_{i(j,g)t}$ from the residuals of a regression of the (log of) inverse electricity share on industry fixed effects and firm-level markups. To account for the standard technology within industries, we include fixed effects at the 4-digit CIIU Rev. 4 code level, effectively comparing electricity expenditure across firms within the same narrowly defined industry. To account flexibly for firm-level markups, we include second-degree polynomials of output market shares in both the local labor market and economywide.[21] Finally, to minimize the impact of outliers and partially address possible measurement error, we create a dummy variable, $EC_{i(j,g)}$, equal to one for firms with an estimated wedge $\hat{\tau}^e_{i(j,g)t}$ above the median value at baseline (i.e., in the first year they are observed in the data), indicating tighter constraints to accessing electricity.[22]

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 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 rests upon three assumptions. First, the instrument must be strongly positively correlated with employment. This is achieved if electrification through the program raises labor demand and employment, more so for firms that had limited access to electricity at baseline. Second, the instrument must be independent of firm-level outcomes. Specifically, it must be orthogonal to the wage and employment trajectories of electricity-constrained firms within each local labor market. This is likely the case since the Ministry did not consider local firms or industries when rolling out the program. Finally, the instrument must satisfy the exclusion restriction, meaning that electrification must not differentially affect labor supply to electricity-constrained firms, aside from increasing wages. This concern is mostly addressed by the inclusion of local labor market $\times$ year fixed effects, which can capture the overall impact of electrification on aggregate labor supply, even if it varies across industries.[23]

[21] This approach is motivated by our theoretical model in Section 4, where the firm-level markup is an exact function of the firm's output market share. The results are robust to (i) not controlling for output market shares (thus implicitly assuming that the firm has no output market power), (ii), controlling for local labor market shares only, and (iii) controlling for national shares only.

[22] Online Appendix Figure A.4 shows the distribution of these wedges and the median value used as a cutoff.

[23] Berger, Herkenhoff and Mongey (2022) show that the reduced-form firm-level elasticity estimated using firm-level shocks overestimates the structural elasticity so that our IV estimate is a lower bound of the inverse structural elasticity.

One potential concern with the exclusion restriction is that if workers had bargaining power, they would receive a share of any rent generated at the firm level due to increased access to electricity.[24] In this case, the labor demand shock could affect wages also through this alternative channel. However, this scenario is unlikely in the Peruvian manufacturing industry since workers have very little bargaining power. Peru has consistently low levels of union density, ranging between 1.9% and 3.2% during the analysis period, placing it among the bottom 5% of the 121 countries with available data from International Labour Organization (2020).[25]

3.4.2 Results

Table 2 shows 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, showing that the instrument carries meaningful identifying variation throughout.[26] The results from using the total sample of manufacturing firms are reported in Column 1. The firm-level inverse labor supply elasticity is estimated to be 0.43, which corresponds to a labor supply elasticity of 2.33, statistically significant at the 1% level. This implies that the difference between wages and the marginal revenue product of labor is 1.43, meaning that workers generate 43% more than what they earn in wages, taking home approximately 70 cents for every marginal dollar produced. These numbers are close to 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 US manufacturing (0.37 to 0.4 and 0.53, respectively). They are slightly lower than those found by Felix (2022) in Brazil prior to the 1990s trade liberalization (wage take-home share of 50%).

Columns 2 to 4 focus on different subsamples. We obtain these estimates by implementing more flexible second and first-stage specifications where we interact 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. In Column 2, we estimate labor market power separately for markets featuring different levels of labor market concentration as measured by wage-bill HHI. For firms operating in the least concentrated labor markets ($HHI \leq 0.15$), we estimate an inverse labor supply elasticity that is both statistically and economically insignificant. As concentration increases, labor market power increases. In markets with a moderate level of concentration ($0.15 < HHI \leq 0.25$), workers take home almost 80 cents for every marginal dollar they generate. In highly concentrated markets ($HHI > 0.25$), the wage take-home share is 63%.

The relationship between labor market power and concentration is mediated by the presence of self-employment opportunities. We divide the markets into two groups based on whether the self-employment rate is lower or higher than the average across markets. The results are

[24]See Wong (2023) for a detailed explanation of the firm first-order condition in this case, which describes the relationship between the wage markdown, firm rents, worker bargaining power, and firm-level labor supply elasticity.

[25]Collective bargaining agreement coverage is similarly low, ranging between 1.8% and 2.6%.

[26]Online Appendix Table A.5 reports all first-stage regression results.

Table 2: Estimates of Labor Market Power

	(1)	(2)	Self-Employment Rate Low (3)	 High (4)
All Markets	0.427*** (0.052)			
$HHI^{wn} \in (0, 0.15]$		0.004 (0.138)		
$HHI^{wn} \in (0.15, 0.25]$		0.270** (0.097)		
$HHI^{wn} \in (0, 0.25]$			-0.113 (0.091)	-0.066 (0.133)
$HHI^{wn} \in (0.25, 1]$		0.603*** (0.148)	0.750*** (0.113)	0.104 (0.069)
SW F-statistics	181.60	223.51 145.35 3177.18	232.84 669.67	124.58 628.07
Observations	6,191	6,191	3,907	2,204

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 work 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.

in Columns 3 and 4. In less concentrated markets, labor market power is insignificant, regardless of the self-employment rate. In highly concentrated markets, the extent of labor market power varies according to the prevalence of self-employment. We estimate the highest level of wage markdown for firms operating in highly concentrated markets with low self-employment rates. The estimated firm-level inverse labor supply elasticity is 0.75 and statistically significant at the 1% level, indicating that workers in these markets take home only 57 cents for every marginal dollar they generate. In highly concentrated markets where self-employment is prevalent, labor market power is positive, but lower in magnitude and insignificant at standard levels. It is, however, statistically different at the 5% level from its homolog in less concentrated markets. Among highly concentrated markets, the difference between those with high vs. low self-employment rates is statistically significant at the 1% level, and so is the difference-in-differences between markets that are highly concentrated vs. not with high vs. low self-employment rates.

How does self-employment limit firms' wage-setting power? The evidence from worker transitions between sectors shows that workers actively switch between wage work and self-employment based on their earnings. When concentration is high, wage jobs are scarce and unappealing, and more workers opt for self-employment. Falling wages might make it easier to push workers out of wage employment, leading to an increase in the supply elasticity of wage work and a decrease in employers' labor market power. The following section presents a general equilibrium model that formalizes these thoughts.

4 Model

We propose a model of the Peruvian economy where employer concentration, self-employment rates, and labor market power are jointly determined in general equilibrium. The model yields insights into demand and supply determinants of labor market power, specifically the role of concentration and self-employment. We use the estimated model to reconcile the empirical evidence and conduct counterfactual policy experiments, with a particular emphasis on industrial development policies.

4.1 Environment

We model Peru as a continuum of segmented local labor markets of measure K, each indexed by k. In each market, there is a finite and endogenous number of heterogeneous firms $M_k \in \mathbb{Z}_+$ and a fixed measure $L_k \in \mathbb{R}_+$ of heterogeneous workers with identical homothetic preferences. These workers self-select into wage work (sector F) or self-employment (sector S). They consume a final good C consisting of a bundle of market-level goods $\{C_k\}_{k\in(0,1)}$. They own shares in local firms, so that their income consists of labor earnings and firm profits.

4.1.1 Demand

The representative agent in local labor market k has Cobb-Douglas preferences over the consumption of market-level output $\{C_k\}_{k\in(0,1)}$ with expenditure shares equal to $\{\theta_k\}_{k\in(0,1)}$, with $\int_0^1 \theta_k dk = 1$. Each good C_k comes in two varieties, $C_{F,k}$ and $C_{S,k}$, produced by firms and self-employed workers, respectively. Each firm i in sector F produces a differentiated variety of the good $C_{F,k}$, which we denote by $c_{iF,k}$. Consumers' preferences can be written as follows:

$$C_k = \left[\beta C_{F,k}^{\frac{\rho-1}{\rho}} + C_{S,k}^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}}, \tag{4}$$

$$C_{F,k} = \left(\sum_{i=1}^{M_k} c_{iF,k}^{\frac{\eta-1}{\eta}}\right)^{\frac{\eta}{\eta-1}}. \tag{5}$$

Consumers substitute across $C_{F,k}$ and $C_{S,k}$ with a constant elasticity $\rho > 1$; they substitute across firm-level varieties $\{c_{iF,k}\}_{i\in M_k}$ with a constant elasticity $\eta > 1$. We maintain the assumption that $\eta > \rho$, which means that the consumers substitute more *within* than *across* sector goods. Lastly, $\beta > 0$ denotes a preference shifter for good $C_{F,k}$.

We choose the final good C as the numeraire, and denote as $Y = C$ the aggregate expenditure in the economy. The price indices associated with equations (4) and (5) are $P_k = \left[\beta^{\rho} P_{F,k}^{1-\rho} + P_{S,k}^{1-\rho}\right]^{\frac{1}{1-\rho}}$ and $P_{F,k} = \left(\sum_{i=1}^{M_k} p_{iF,k}^{1-\eta}\right)^{\frac{1}{1-\eta}}$, respectively. With the above demand structure, the representative consumer expenditure on the market k goods can be written as:

$$P_{F,k}C_{F,k} = \alpha_{F,k}\theta_k Y \quad \text{and} \quad P_{S,k}C_{S,k} = (1-\alpha_{F,k})\theta_k Y, \quad \text{with } \alpha_{F,k} = \beta\left(\frac{P_{F,k}}{\beta P_k}\right)^{1-\rho} \tag{6}$$

$$r_{iF,k} \equiv p_{iF,k}c_{iF,k} = s_{iF,k}\alpha_{F,k}\theta_k Y \qquad \text{with } s_{iF,k} = \left(\frac{p_{iF,k}}{P_{F,k}}\right)^{1-\eta}, \tag{7}$$

where $\alpha_{F,k} \equiv \frac{P_{F,k}C_{F,k}}{P_k C_k}$ in equation (6) is the expenditure share of variety F of good k over total expenditure on market k, and $s_{iF,k} \equiv \frac{r_{iF,k}}{\sum_{i\in M_k} r_{iF,k}}$ in equation (7) is the expenditure share on firm $i's$ variety in sector F over total expenditure on sector F in market k.

4.1.2 Labor Supply

Workers in each local labor market k are heterogeneous in their sector-specific abilities, determined by their endowment of efficiency units of labor $\mathbf{a} \equiv (a_F, a_S) \in \mathbb{R}+^2$. These are i.i.d. draws from a joint distribution $G_k(a_F, a_S)$ specific to each market.

We denote by $W_{s,k}$ the wage per efficiency unit in sector s of market k, with $s = \{F, S\}$. In turn, $I_{s,k}^h = W_{s,k}a_s^h$ denotes the sector S-earnings of worker h with ability draw (a_F^h, a_S^h). We assume that all firms in sector F pay the same wage $W_{F,k}$ per efficiency unit. Workers take the sector wages as given and self-select into wage work or self-employment so to maximize their earnings. This means that worker h will self-select into wage work if:

$$\hat{W}_k \geq A(h)^{-1}, \tag{8}$$

where $\hat{W}_k \equiv \frac{W_{F,k}}{W_{S,k}}$ is the relative efficiency unit wage in sector F, and $A(h) \equiv \frac{a_F^h}{a_S^h}$ is the worker's comparative advantage in the wage sector. This means that workers with higher $A(h)$ have a lower reservation wage for choosing sector F over sector S.

The sorting rule in equation (8) implies that labor supply in the wage employment sector is an increasing function of the relative unit wage $\hat{W}_k$:

$$N_{F,k} \equiv N_F(\hat{W}_k) = L_k \int_0^{\infty}\int_0^{a_F\hat{W}_k} a_F\, g_k(a_F, a_S)da_F da_S, \quad \text{with } N'_{F,k} > 0. \tag{9}$$

Intuitively, the higher $\hat{W}_k$, the more workers sort into the wage employment sector.[27] We denote

[27] Similarly, the labor supply function in sector S is a decreasing function of the relative wage $\hat{W}_k$.

the labor supply elasticity associated with equation (9) as:

$$\epsilon_F(\hat{W}_k) \equiv \frac{\partial \ln N_{F,k}}{\partial \ln \hat{W}_k} > 0. \tag{10}$$

In principle, the labor supply elasticity can either increase or decrease with the relative unit wage $\hat{W}_k$. However, Appendix B shows that under some regularity conditions on $G_k(a_F, a_S)$, the elasticity in equation (10) is a decreasing function of the relative wage $\hat{W}_k$. This means that wage workers become more sensitive to a unit wage increase or decrease as the relative wage in the sector decreases. We will discuss below how this property of the labor supply function is critical to understanding labor market power in the economy.

4.1.3 Technology and Market Structure

We assume that goods in sector S and F are produced using a technology that is linear in efficiency units of labor. Total output in sector S is given by

$$Y_{S,k} = N_{S,k}, \tag{11}$$

where $N_{S,k}$ denotes total efficiency units of labor in sector S. There is perfect competition in sector S, such that goods are sold at marginal cost, i.e., $P_{S,k} = W_{S,k}$.

We let $n_{iF,k}$ and $y_{iF,k}$ denote the labor demand and output of firm i in market k, respectively. Firm i's production function for its final-good variety is given by:

$$y_{iF,k} = T_k z_{iF,k} n_{iF,k}, \tag{12}$$

where T_k is a market-specific productivity shifter specific to sector F and $z_{iF,k}$ is a firm-specific productivity term, both exogenous.

Labor market clearing at the local labor market level requires that aggregate labor demand equals aggregate labor supply, hence, $N_{F,k} = \sum_{i=1}^{M_k} n_{iF,k}$, where $N_{F,k}$ is as in equation (9). Combining the previous expression with equation (12) yields the following equilibrium relation for aggregate output in sector F :

$$Y_{F,k} = T_k Z_{F,k} N_{F,k}, \tag{13}$$

where $Z_{F,k} \equiv \left(\sum_{i=1}^{M_k} s^y_{iF,k}(z_{iF,k})^{-1}\right)^{-1}$ is a productivity index of market k, defined as the harmonic average of firm-level productivities with weights equal to the output share of firm i $s^y_{iF,k} \equiv \frac{y_{iF,k}}{Y_{F,k}} = s_{iF,k}^{\frac{\eta}{\eta-1}} \in (0,1)$, where $s_{iF,k}$ is defined in equation (7).

4.1.4 Market Structure and Firm-Level Equilibrium

The firms engage in Nash-Cournot competition in both the product and labor markets. In the product market, firms produce different varieties, which results in oligopolistic competition in the final goods market. On the other hand, in the labor market, firms are identical and pay the same unit wage $W_{F,k}$. As a result, workers view them as perfect substitutes, creating an oligopsonistic labor market structure.

The problem of each firm is to choose labor to maximize total variable profits, namely:

$$\max_{n_{iF,k}} r_{iF,k} - W_{F,k} n_{iF,k}$$

subject to equations (7), (12), and (9). Firms take as given the aggregate prices P, P_k, and P_S, but they internalize the effect of their output and labor demand affect the aggregate price $P_{F,k}$ and aggregate wage $W_{F,k}$.

Solving for the firm-level equilibrium yields the following first-order condition:

$$\begin{aligned} MRPL_{iF,k} &= W_{F,k}\psi_{iF,k}, & (14)\\ \implies p_{iF,k} &= \mu_{iF,k}\frac{W_{F,k}\psi_{iF,k}}{T_k z_{iF,k}}, \qquad \forall i \in M_k & (15)\end{aligned}$$

where the second line follows from the definition of $MRPL_{iF,k}$. The term $\mu_{iF,k}$ is the firm-level markup over marginal cost, defined as $\mu_{iF,k} = \frac{\varepsilon_{iF,k}}{\varepsilon_{iF,k}-1}$ where $\varepsilon_{iF,k} = \left[\frac{1}{\eta}(1 - s_{iF,k}) + \frac{1}{\rho}s_{iF,k}\right]^{-1}$ is the demand elasticity, which takes the standard Cournot formulation.

Equation (14) writes the unit efficiency wage in sector F as markdown $\psi_{iF,k} \geq 1$ below the marginal revenue product of workers at firm i. Because firms are heterogeneous in their productivity $z_{iF,k}$, they have different $MRPL_{iF,k}$, which yields heterogeneous markdown even if they pay the same wage. In turn, more productive firms have higher $MRPL_{iF,k}$ and thus higher markdowns in equilibrium, leading to equilibrium misallocation of employment (Berger, Herkenhoff and Mongey, 2022).

In equilibrium, the markdown $\psi_{iF,k}$ is given by:

$$\psi_{iF,k} = 1 + \frac{s^N_{iF,k}}{\epsilon_{F,k}(\hat{W_k})}, \tag{16}$$

where $s^N_{iF,k} \equiv \frac{n_{iF,K}}{N_{F,k}}$ is firm i's employment share in sector F, and $\epsilon_{F,k}(\hat{W_k})$ is the labor supply elasticity defined in equation (10). We take this markdown is our measure of the labor market power of employers. When $\psi_{iF,k} \to 1$, the wage equals the marginal revenue product of labor, as in a competitive equilibrium. On the contrary, when $\psi_{iF,k} > 1$, employer i pays their workers less than their $MRPL_{iF,k}$, i.e., they have labor market power. We will return to the markdown properties in Section 4.3.

4.2 Closing the Model

We now close the model in general equilibrium. The numerical implementation of the equilibrium solution is described in detail in Appendix D.

Market Equilibrium We first assume that the set of employers $\{M_k\}_{k\in(0,1)}$, exogenous productivity vectors $\{T_k, \{z_{iF,k}\}_{i\in M_k}\}_{k\in(0,1)}$, and aggregate variables (Y), are known. Given the vector of relative wages $\hat{\mathbf{W}} \equiv \{\hat{W}_k\}_{k\in(0,1)}$, equations (7), (15), and (16) define a fixed point problem that can be solved for market shares in the output and labor markets, markups, and markdowns: $\mathbf{\Lambda} \equiv \{\{s_{iF,k}, s^N_{iF,k}, \mu_{iF,k}, \psi_{iF,k}\}_{i\in M_k}\}_{k\in(0,1)}$. In turn, given the matrix $\mathbf{\Lambda}$, the relative wage vector $\hat{\mathbf{W}}$ solves labor and output market clearing conditions, from equations (6), (11), and (13). The resulting fixed point $(\hat{\mathbf{W}}, \mathbf{\Lambda})$ is the vector of market equilibria.

Entry To enter a given market, firms have to pay a fixed cost f^e_k in units of the final good. We consider a sequential entry game where firms with higher productivity move first. The equilibrium number of entrants can be solved with the following iterative procedure. Given a guess for the number of entrants $\{M_k\}_k$, we find the market equilibrium $(\hat{\mathbf{W}}, \mathbf{\Lambda})$ using the procedure outlined above and compute the profits of the marginal entrant. An equilibrium of the entry game is achieved when, for a subset of firms $i \in M_k$, equilibrium profits given by

$$\pi_{iF,k}(M_k;Y) \equiv s_{iF,k}\alpha_{F,k}\theta_k Y\left(1 - \frac{1}{\mu_{iF,k}(M_k)\psi_{iF,k}(M_k)}\right) - f^e_k \tag{17}$$

are non-negative, while for any additional entrant $j \notin M_k$, profits upon entry would be negative.[28] With sequential entry, this entry game has a unique cutoff equilibrium so only firms with productivity above some cutoff enter the market.

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

$$C = \int_k [E(W_{S,k}, W_{F,k}) + \Pi_{F,k} + PM_k f^e_k]\, dk, \tag{18}$$

where the three terms on the right-hand side correspond to (i) labor income in market k as given by $E(W_{S,k}, W_{F,k}) \equiv W_{S,k}N_S(\hat{W}_k) + W_{F,k}N_F(\hat{W}_k)$, (ii) aggregate profits of firms $\Pi_{F,k} = \sum_{i\in M_k} \pi_{iF,k}(M_k;Y)$ which are distributed to consumers, and (iii) entry of firms into the production stage. Product market clearing requires that the total demand of the final good equals the total value of production, i.e.,

$$Y = PC, \tag{19}$$

[28] Ignoring the integer problem that arises when the number of employers is finite, equation (17) is equivalent to a zero profit condition determining the equilibrium number of employers in each market.

with $P = 1$ by normalization. Conditional on the market equilibrium $\mathbf{K} = \left\{\mathbf{M}, \hat{\mathbf{W}}, \mathbf{\Lambda}\right\}$, the general equilibrium $\mathbf{X}$ solves equations (18)-(19). Conditional on the general equilibrium $\mathbf{X}$, the solutions to the market equilibrium and entry game described above yield the manufacturing equilibrium $\mathbf{K}$. The resulting fixed point $(\mathbf{X}; \mathbf{K})$ is the equilibrium in the economy.

4.3 Properties of the Model

4.3.1 Labor Market Power, Concentration, and Self-Employment

The reduced-form evidence in Table 2 of Section 3 demonstrates that labor market power increases with employer concentration, but self-employment moderates this correlation. Consistent with this evidence, our theory suggests that, by representing a readily available outside option for workers, self-employment acts as a check on the wage-setting power of employers.

Let $\bar{\psi}_{F,k} \equiv \sum_{i \in M_k} s_i^N \psi_{iF,k}$ denote the weighted average of firm-level markdowns in market k, with employment shares as weights. From equation (16), we can write:

$$\bar{\psi}_{F,k} = 1 + \frac{HHI_{F,k}^{wb}}{\epsilon_F(\hat{W}_k)}, \tag{20}$$

where $HHI_{F,k}^{wb}$ is the wage-bill HHI in sector F of market k, which in our model coincides with the employment-based HHI.

Equation (20) shows that the relationship between wage markdowns and concentration is mediated by the aggregate supply elasticity of wage work, $\epsilon_F(\hat{W}_k)$. In Appendix C.2, we show that, when worker abilities are drawn from a joint log-normal distribution, one could express this elasticity as an increasing function of the self-employment share in market k, i.e.,

$$\epsilon_F(\underbrace{\hat{W}_k}_{-}) = \tilde{\epsilon}_F(\underbrace{\textit{self rate}_k}_{+}), \quad \text{with} \quad \textit{self rate}_k \equiv \frac{L_{S,k}}{L_{F,k} + L_{S,k}} \tag{21}$$

where $L_{s,k} \in (0, L_k)$ denotes total employment in sector $s = \{F, S\}$.

Equations (20) and (21) summarize the role of concentration and self-employment in determining labor market power. On the one hand, concentration has a direct positive effect on the average markdown through standard oligopsony forces. On the other hand, when employer concentration is high and the relative wage is low, more workers opt for self-employment. Under the regularity conditions on the workers ability distribution discussed in Appendix B, the changing comparative advantage of the marginal worker increases the sensitivity of wage workers to unit wage changes. The associated increase in the elasticity $\epsilon_F(\hat{W}_k)$ reduces the aggregate markdown.[29]

[29] While the relative impact of these opposing forces is an empirical question, this variable elasticity channel can reconcile the conflicting evidence in the literature on the relationship between employer concentration and labor market power. See for instance Bassier, Dube and Naidu (2022); Berger, Herkenhoff and Mongey (2022); Yeh, Macaluso and Hershbein (2022).

4.3.2 Sectoral Shocks and Changes in Average Earnings

We now examine how labor market outcomes adjust in response to changes in sectoral wages per efficiency unit driven by either demand or supply factors. Our objective is twofold: first, to clarify the roles of comparative and absolute advantage in determining the distributional consequences of sectoral shocks; and second, to provide a framework for interpreting the results of our counterfactual experiments in Section 6.

Average Earnings from Wage Work Let the log change in the average earnings in the wage employment sector be denoted by:

$$d\ln \bar{E}_{F,k} = d\ln W_{F,k} + d\ln \bar{A}_{F,k},$$

where $\bar{A}_{F,k} \equiv \frac{N_{F,k}}{L_{F,k}}$ is average worker ability in sector F. From equation (14), we can write:

$$d\ln W_{F,k} = d\ln \overline{MRPL}_{F,k} - d\ln \bar{\psi}_{F,k},$$

where $\overline{MRPL}_{F,k}$ and $\bar{\psi}_{F,k}$ denote the payroll-weighted average MRPL and markdown of firms in market k, respectively, where the latter is defined in equation (20).

In turn, we can decompose the log change in the average market MRPL as:

$$d\ln \overline{MRPL}_{F,k} = d\ln Z_{F,k} + d\ln P_{F,k} - d\ln \Xi_{F,k}.$$

The first term captures changes in the productivity index defined in equation (13). The second and last term capture changes in the sectoral price index and in an average markup index, respectively, where the latter is defined as $\Xi_{F,k} \equiv \left(\sum_i s_{iF,k} \cdot (\mu_{iF,k})^{-1}\right)^{-1}$.

Putting pieces together, we can decompose the change in average sector F earnings as:

$$d\ln \bar{E}_{F,k} = -\underbrace{d\ln \bar{\psi}_{F,k}}_{\text{markdown}} + \underbrace{d\ln Z_{F,k}}_{\text{misallocation}} + \underbrace{d\ln P_{F,k} - d\ln \Xi_{F,k}.}_{\text{general equilibrium}} + \underbrace{d\ln \bar{A}_{F,k}}_{\text{selection}} \tag{22}$$

Equation (22) demonstrates how a shock to the economic environment can affect the average earnings in the wage employment sector, $\bar{E}_{F,k}$. There are two broad channels through which this can happen. The first is a *direct* effect on the wage per efficiency unit $W_{F,k}$, captured by the first three terms in the right-hand side. The second is the indirect effect on the average ability of wage workers. We refer to it as the *selection* channel since it occurs through worker self-selection across sectors. The latter channel depends critically on the parameters of the workers' ability distribution, as we explain below.

The direct effect on the unit efficiency wage can be further decomposed into several channels. The first one is a *markdown* channel, which captures how the average markdown in the labor market respond to the shock. The second one is a *misallocation* channel, which materializes through market share reallocation in the labor market. Due to equilibrium misallocation in the labor market, the productivity index $Z_{F,k}$ is below efficient levels. As the shock reallocates market shares towards (or away from) the most productive firms, aggregate productivity and wages increase (decrease). Lastly, the unit wage are also affected by changes in the aggregate price index and average markups. We define this channel the *general equilibrium* channel.[30]

Average Earnings from Self-Employment Similar forces also determine how earnings from self-employment are affected by shocks. In the self-employment sector, the markdown is always equal to one and self-employed workers have the same productivity. Hence, equation (22) can be simplified as:

$$d \ln \bar{E}_{S,k} = \underbrace{d \ln P_{S,k}}_{\text{general equilibrium}} + \underbrace{d \ln A_{S,k}}_{\text{selection}} .$$

Hence, only the selection and general equilibrium channel affect the earnings response of the self-employment sector.

The Selection Channel The selection channel determines how the average ability in the two sectors respond to demand or supply shocks, namely, $d \ln A_{F,k}$ and $d \ln A_{S,k}$. These effects depend on the parameters of the workers' ability skill distribution, particularly on the correlation of workers' abilities across wage work and self-employment and their relative dispersion (Adão, 2015; Alvarez-Cuadrado, Amodio and Poschke, 2020). If the two abilities are strongly positively correlated and more dispersed in self-employment than in wage work, mean ability will decrease in both sectors when the relative wage $\hat{W}_k$ falls and the wage employment sector shrinks.[31] This is because absolute and comparative advantage are negatively correlated in wage work but positively correlated in self-employment. The more skilled wage workers are better off as self-employed, implying negative selection in wage work and positive selection in self-employment. Vice versa, if abilities are negatively correlated or if the correlation is positive but low, the mean ability of wage workers will increase, and the one of self-employed workers will decrease when $\hat{W}_k$ is lower. In Appendix C.1, we formalize these insights for abilities drawn from a joint log-normal distribution.

5 Model Estimation

This section details how we estimate the theoretical model using the Peruvian data. First, we parameterize the model to make it tractable for empirical analysis. Second, we leverage the data to pin down the parameters of the joint ability distribution using direct inference. Third, we use the structure of the model to estimate all the remaining parameters.

[30] While the markup term may also reflect misallocation in the output market, our analysis focuses on labor market channels, and so we bulk them in the residual "general equilibrium" term.

[31] See also Heckman and Sedlacek (1985); Heckman and Honoré (1990); Young (2014).

5.1 Parameterization and Calibration Strategy

We consider a parameterization of the model that allows for heterogeneity across local labor markets in the fundamental forces affecting firms' and workers' decisions: market-level productivity T_k, fixed costs f_k^e, and the joint ability distribution G_k. Within each market, firm behavior is also contingent upon its productivity $z_{iF,k}$. We view each local labor market as a (multi-dimensional) observation from the structural data-generating process described by the model, with common parameters that need to be estimated.

To accurately capture the empirical properties of the firm sales distribution within each local labor market, we assume that $z_{iF,k}$ are drawn from a Pareto distribution with shape parameter Z_θ, while market productivities T_k are parameterized as draws from a log-normal distribution with parameters μ_T and σ_T:

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

The fixed entry cost f_k^e is specified as follows. We first let the market-level fixed cost f_k^e be a draw from a Weibull distribution with scale parameter F_λ and shape parameter F_k:

$$f^e \sim \text{Weibull}(F_\lambda, F_k).$$

Letting $f_k^{e,(n)}$ denote the fixed cost incurred by the n^{th} entrant in market k, we then impose:

$$f_k^{e,(n)} = \begin{cases} 0 & \text{if } n = 1 \\ f_k^e & \text{if } n > 1 \end{cases},$$

where $f_k^e > 0$ is the Weibull draw of market k. This formulation guarantees that at least one entrant exists in each local labor market, a necessary condition to hold fixed the number of "active" local labor markets.

Lastly, we assume that the endowment of efficiency units of labor in the two sectors $\mathbf{a} = (a_F, a_S)$ is a draw from the following joint log-normal distribution:

$$\log \mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu_k}, \boldsymbol{\Sigma}), \quad \text{where } \boldsymbol{\mu_k} = \begin{pmatrix} \mu_{F,k} \\ \mu_{S,k} \end{pmatrix}, \quad \text{and } \boldsymbol{\Sigma}_k = \begin{pmatrix} \sigma_{F,k}^2 & \varrho_k \sigma_{F,k}\sigma_{S,k} \\ \varrho_k \sigma_{F,k}\sigma_{S,k} & \sigma_{S,k}^2 \end{pmatrix}, \tag{23}$$

where $\boldsymbol{\mu}_k$ determines the vector of mean absolute advantage in wage work and self-employment, ϱ_k captures the correlation between the two, and $\boldsymbol{\Sigma}_k$ is the variance-covariance matrix governing comparative advantage.

We impose the following restrictions on equation (23) to simplify the estimation procedure. First, we restrict the variance-covariance matrix $\boldsymbol{\Sigma_k}$ to be constant across local labor markets, i.e. $\boldsymbol{\Sigma_k} = \boldsymbol{\Sigma}$. As explained in Section 5.2, this assumption allows us to recover the matrix Σ from the available data using direct inference methods.

In contrast, we explain in Appendix C.3 that the absolute advantage moments cannot be easily recovered from cross-sectional data on worker earnings across the two sectors. Hence, we make progress as follows. We first normalize the mean comparative advantage in the two employment sectors to zero: $\hat{\mu}_k \equiv \mu_{F,k} - \mu_{S,k} = 0 \; \forall k$. To allow for heterogeneity in worker ability across local labor markets, we then assume that mean absolute advantage in market k, $\mu_k = \mu_{F,k} = \mu_{S,k}$ is itself a draw from a log-normal distribution with common parameters μ_μ and σ_μ as given by

$$\mu \sim \log \mathcal{N}(\mu_\mu, \sigma_\mu).$$

With these parametric assumptions, we estimate the model in three steps. In the first step, we calibrate market-level Cobb-Douglas expenditure shares $\{\theta_k\}_k$ from the data as equal to the income share of each local labor market. We also calibrate the percentage of population in each market $\{L_k\}_k$ to match the relative size of the workforce in market k. We report the histogram of the resulting $\{\theta_k, L_k\}$ in Online Appendix Figure A.5. The histograms show an unequal distribution of expenditure and population across local labor markets: the 90-10 ratio for expenditure shares lies above 700; the one for population is around 40.

In the second step, which we describe in Section 5.2, we use the observed employment shares and their variances to recover the (common) parameters of the variance-covariance matrix of the workers' ability distributions $\boldsymbol{\Theta} \equiv \{\sigma_F, \sigma_S, \varrho\}$. In the third step, described in Section 5.3, we implement a method of simulated moments (MSM) procedure to estimate the remaining parameters $(Z_\theta, \mu_T, \sigma_T, F_\lambda, F_k, \mu_\mu, \sigma_\mu, \eta, \rho, \beta)$.

5.2 Inferring Ability Distribution Moments

We estimate the variance-covariance matrix $\boldsymbol{\Sigma}$ by leveraging the properties of the joint log-normal distribution in equation (23). We provide a detailed description of the procedure in Appendix C.3, which we summarize here.

Let $\sigma* = (\sigma_F^2 + \sigma_S^2 - 2\varrho\sigma_F\sigma_S)^{-\frac{1}{2}}$. The (observed) probability that a given worker operates in sector F is equal to

$$\Pr\left(h \in \textit{wage sector}\right) = \Phi(c_F), \tag{24}$$

where $c_{F,k} \equiv \left(\frac{\ln \hat{W}_k + \hat{\mu}_k}{\sigma*}\right)$ is a market-level term and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. The probability of operating in sector S is $\Phi(c_{S,k}) = \Phi(-c_{F,k})$.

When the abilities $\mathbf{a} = (a_F, a_S)$ follow the distribution in equation (23) and $\boldsymbol{\Sigma_k} = \boldsymbol{\Sigma}$, the observed variance of log earnings in the two sectors can be written as:

$$\begin{aligned} Var\left(\ln a_F W_{F,k} | a_F \hat{W}_k \geqslant a_S\right) &= \sigma_F^2 + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[\lambda(c_{F,k})c_{F,k} - \lambda^2(c_{F,k})\right] \\ Var\left(\ln a_S W_{S,k} | a_F \hat{W}_k \leq a_S\right) &= \sigma_S^2 + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[\lambda(c_{S,k})c_{S,k} - \lambda^2(c_{S,k})\right] \end{aligned}, \tag{25}$$

where $\lambda(x) \equiv \frac{\phi(x)}{\Phi(x)} \geq 0$ is the inverse Mills ratio.

The left-hand side of equation (25) is observed, given cross sectional data on workers' earnings across the two sectors. Similarly, one can easily recover the terms $c_{F,k}$ and $c_{S,k}$ from simple inversion of the observed employment shares in the two sectors, from which we can also get $\lambda(c_{F,k}), \lambda(c_{S,k})$. This reduces the system in equation (25) to a system of $2 \times K$ equations in a vector of 3 unknowns, namely, $\mathbf{\Theta} = (\sigma_F, \sigma_S, \varrho)$. We recover the vector $\mathbf{\Theta}$ from a Minimum Distance Estimation (MDE) procedure with non-negativity constraints on the parameters.

Table 3 reports the estimates of the variance-covariance matrix $\mathbf{\Sigma}$. The two abilities strongly correlate, with $\hat{\varrho} = 0.93$, 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 between 0.02 and 0.07. As explained in Section 4.3, this means that we find a (small) negative correlation between the workers' comparative and absolute advantage in wage employment, and positive correlation of advantages in self-employment. As a result, mean ability increases in both sectors when the wage employment sector expands. These estimates are consistent with the evidence presented in Figure 1 showing that transitions into and from wage employment are systematically more common among low-earning self-employed, while transitions to self-employment are more prevalent among high-earning wage workers. These parameter values also indicate that at least part of the negative correlation between concentration and earnings (from self-employment in particular) documented in the right panel of Figure 2 and Online Appendix Table A.2 is driven by worker sorting, as average skills decrease in both sectors when employer concentration is high and the wage employment sector is small.

5.3 Moments and Identification of Remaining Parameters

We now describe the Method of Simulated Moments (MSM) procedure we employ to estimate the remaining parameter vector $\Phi = (Z_\theta, \mu_T, \sigma_T, F_\lambda, F_k, \mu_\mu, \sigma_\mu, \eta, \rho, \beta)$. The full MSM procedure is described in Appendix D.

We target 20 empirical moments corresponding to local labor market outcomes, informed by our empirical facts in Section 3. We focus on moments that are informative about the cross-sectional features of concentration, self-employment, and their relationships. While variation in any parameter tends to affect all moments simultaneously, certain parameters are more likely to affect specific moments. We now provide a discussion of the key factors ensuring identification.

First, we target cross-sectional moments related to concentration and firm performance across local labor markets. We target the mean and standard deviation of the (log of the) number of firms across local labor markets, the employment-based Herfindahl-Hirschman Index (HHI), both weighted and unweighted, and the concentration ratios CR1 and CR3. We also target the share of local labor markets with only one employer and the associated percentage of total wage employment. These moments are crucial for identifying productivity parameters and fixed cost distribution. Intuitively, the fixed cost parameters (F_λ and F_k) determine the

Table 3: Summary of Model Parameters

	Parameter	Description	Value
A. Externally Fixed			
1.	$\hat{\mu}_k$	Mean comparative advantage	0
2.	η	Substitution elasticity *within* wage sector	6
3.	β	Sector F demand shifter	1.5
B. Externally Estimated			
4.	σ_F	St. dev. of log ability as a wage worker	0.959
5.	σ_S	St. dev. of log ability as a self-employed	1.025
6.	ϱ (Roy)	Correlation of log abilities	0.935
C. Estimated via MSM			
7.	μ_T	Mean of market-level productivity	0.195
8.	σ_T	St. Dev. Of market-level productivity	0.553
9.	F (scale) (*10^6)	Fixed entry cost (Location)	2.05
10.	F (shape)	Fixed entry cost (Shape)	0.342
11.	θ	Firm-level productivity dispersion parameter	4.566
12.	μ_μ	Mean of market-level abs. adv. of workers' abilities	-0.2
13.	σ_μ	St. Dev. of market-level abs. adv. workers' abilities	0.353
14.	ρ	Substitution elasticity *across* sectors	3.468

Notes. This table reports the parameter values for the quantitative model. See Section 5 for details on parameter estimation.

average number of firms in each market, market concentration, and the share of markets with only one firm. At the same time, the standard deviation of concentration measures and the employment share of markets with only one firm are mostly informative about μ_T and σ_T. To capture the parameters of the productivity distribution both across labor markets (μ_T and σ_T) and across firms within a market (Z_θ), we also target the interquartile range (IQR) of log sales across markets.

Next, we target the wage employment share and the relative (log) worker earnings between wage and self-employed workers. Given the variance matrix (Σ), the number of entrants M_k, and market-level productivity T_k, these moments depend on the mean absolute advantage across markets, hence on (μ_μ, σ_μ), and the across-sector elasticity ρ. Finally, we target the correlations between the employment-based HHI and several labor market outcomes, including (log) earnings in both sectors, (log) education of wage and self-employed workers, the share of wage workers within a labor market, and (log) sales. The sensitivity of earnings and abilities of wage and self-employed workers to the number of firms depends on the across-sector elasticity (ρ) and the absolute advantage of workers, hence (μ_μ, σ_μ).

We look for the set of parameters Φ that solves $\hat{\Phi} = \arg\min_\Phi \hat{\mathbf{f}}(\mathbf{\Phi})'\mathbb{W}\hat{\mathbf{f}}(\mathbf{\Phi})$, where $\hat{f}_i(\Phi) \equiv [m_i(\Phi) - \hat{m}_i]$, $m_i(\Phi)$ is the value of moment i in our model given the parameter vector Φ, and

$\hat{m}_i$ is the properly normalized corresponding moment computed in the data. Finally, $\mathbb{W}$ is the weighting matrix, which we chose to be diagonal and inversely proportional to $\hat{\mathbf{m}}$. We give more weight to the means than the standard deviations of the variables and reduce the weights in the regression coefficients to account for estimation error. We also choose not to target the regression standard errors. The estimated parameters are reported in Table 3.

Normalizations To improve estimation precision, we adopt the following strategies. In our estimation, we find that the elasticity of substitution η and productivity parameter Z_θ are weakly separately identified. The moments tend to be sensitive to the ratio $\kappa = Z_\theta/(\eta - 1)$, which approximately corresponds to the Pareto tail of the sales distribution across firms (Gaubert and Itskhoki, 2021). To address this, we fix η to 6 and estimate Z_θ in the MSM routine.

Moreover, we fix the demand shifter for sector F good β to β=1.5 since it is not well identified by any of the moments that we have available. This reduces the parameter vector to $\Phi = (Z_\theta, \mu_T, \sigma_T, F_\lambda, F_k, \mu_\mu, \sigma_\mu, \rho)$, thereby improving the precision of estimation.

5.4 Model Fit

Table 4 compares the model-based values of the 20 moments targeted in estimation with their empirical counterparts. The ability of the model to closely replicate the distribution of the number of firms and self-employment is essential for the quantitative analysis of labor market power. Despite its parsimony, the model provides a good fit for the data overall.

The model effectively captures the different measures of concentration across local labor markets, such as the mean (log) number of medium-to-large firms in the average labor market, the unweighted and weighted mean employment-based HHI across local labor markets, and the high concentration ratios. While the model reasonably matches the high share of monopsonistic labor markets, which is 39% in the data but 48% in the model, the match is not perfect. However, the observed 8% payroll share of these markets is exactly replicated by our model.

The model also provides a satisfactory fit of the share of wage workers across markets. The model predicts that about 60% of workers are wage workers, while the data shows that around 70% of workers are in wage employment.[32] The earnings of wage workers are about 0.23 log points higher than that of self-employed workers, both in the model and the data.

Lastly, the model accurately replicates the cross-sectional correlation of earnings, ability, wage employment rates, and total sales with the payroll HHI. All coefficients are estimated to be economically and statistically significant, even though the standard errors were not targeted in estimation. The magnitudes are broadly similar between the model and the data, with the exception of those related to the correlation of mean wage sector earnings and total firm sales with HHI, which the model overshoots and undershoots, respectively.

[32]Table 3 shows that the wage-employment share in manufacturing is 60%, which is what our model predicts. The 70% number used in estimation comes from a different dataset merging information about firms and workers to ensure consistency with the model, while the 60% in Table 3 comes from the full sample of worker data.

Table 4: Targeted Moments and Model Fit

	Moment	Model	Data
1	Log Number of Firms (Mean)	1.35	1.22
2	Log Number of Firms (St. Dev.)	1.61	1.17
3	HHI_k^e (Mean, Unweighted)	0.58	0.59
4	HHI_k^e (St. Dev.)	0.42	0.35
5	HHI_k^e (Mean, Weighted)	0.24	0.31
6	CR1 (Mean)	0.63	0.69
7	CR1 (St. Dev.)	0.37	0.29
8	CR3 (Mean)	0.79	0.88
9	CR3 (St. Dev.)	0.28	0.18
10	% Markets w/ 1 firm	0.48	0.39
11	Wagebill Share of Markets w/ 1 firm	0.08	0.08
12	(Log) Sales IQR	3.52	2.92
13	% Workers in Wage Employment (Mean)	0.6	0.71
14	% Workers in Wage Employment (St. Dev.)	0.21	0.32
15	(Log) (Earnings$_F$/Earnings$_S$) (Mean)	0.23	0.23
16	(Log) (Earnings$_F$/Earnings$_S$) (St. Dev.)	1.13	1.04
	Coefficients and Std. Errors from Regression of:		
17	Mean wage sector earnings on (log) HHI_k^{wb}	-0.67	-0.14
		0.1	0.02
18	Mean self-employment earnings on (log) HHI_k^{wb}	-0.11	-0.18
		0.11	0.06
19	% Workers in Wage Employment on (log) HHI_k^{wb}	-0.1	-0.04
		0.01	0.01
20	(Log) Total Sales on (Log) HHI_k^n	-0.52	-1.38
		0.1	0.07

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 5 for more details on the moments' construction.

Untargeted Moments We now examine the model's fit for moments not targeted during estimation. In particular, we evaluate the model's ability to replicate the reduced-form patterns of labor market power across local labor markets. Like Table 2, Table 5 displays the heterogeneous patterns of labor market power across different subsets of local labor markets. For each group, we report the mean of $\bar{\psi}_k - 1$ across all markets, where $\bar{\psi}_k$ is the average markdown in local labor market k defined in equation (20).

A direct comparison of the magnitude of the coefficients is not feasible as the results from Table 2 correspond only to local average treatment effect (LATE) estimates, and they only incorporate the response of a particular firm to its own MRPL shock, holding its competitors' employment fixed; i.e., they do not take into account cross-firm general equilibrium effects. However, these estimates, and the patterns they imply, are still informative as they represent a

Table 5: Model Estimates of Labor Market Power

			Self-Empl. Rate	
			Low	High
	(1)	(2)	(3)	(4)
All Markets	0.36			
$HHI^{wn} \in (0, 0.15]$		0.12		
$HHI^{wn} \in (0.15, 0.25]$		0.27		
$HHI^{wn} \in (0, 0.25]$			0.18	0.06
$HHI^{wn} \in (0.25, 1]$		0.48	0.72	0.35

Notes. This table shows the mean markdown $\bar{\psi}_k - 1$ across all markets (Column 1) and in different subsets of markets (Columns 2 to 4) in the estimated model to compare with the reduced-form markdown estimates in Table 2. As in the latter, low and high self-employment rates are defined as below and above the average self-employment rate across local labor markets, respectively.

lower bound of the model estimates (Berger, Herkenhoff and Mongey (2022)).

The parameters in Table 5 corroborate this hypothesis as they are all larger than those found in Table 2, albeit close and in the same order of magnitude. Moreover, the model performs well in reproducing the observed patterns. The average markdown across local labor markets is 0.36 in the model and 0.43 in the data. The model matches the correlation between average markdown and concentration for low and medium-concentrated markets. However, it somewhat understates the average markdown in highly-concentrated markets, which is 0.48 in the model and 0.60 in the data. Furthermore, the model reflects the mitigating role of self-employment. As in the data, labor market power is highest in markets with high employer concentration and low self-employment rates, the average markdown being 0.72 in the model and 0.75 in the data. Labor market power is much lower in markets with a higher than average self-employment share, consistent with the evidence in Table 2.

6 Counterfactual Policy Analysis

Armed with the estimated model, we provide two sets of counterfactual experiments. First, we assess quantitatively the role of labor market power in explaining Peruvian labor market outcomes. Second, we consider three policies promoting industrialization and increasing wage employment by targeting firms or workers. We quantify their aggregate and distributional impact, and the role of labor market power in shaping their effectiveness.

6.1 Impact of Labor Market Power

To investigate the consequences of labor market power on labor market outcomes, we introduce an indicator, $\iota \equiv \frac{dN_{F,k}}{dn_{iF,k}} = \{0, 1\}$, in our general equilibrium model. This indicator captures strategic interactions among firms in the labor market. When $\iota = 1$, firms fully internalize the impact of their labor demand on aggregate labor demand and wages. Conversely, when $\iota = 0$, firms do not consider such consequences and act as wage-takers in the labor market. Our baseline model assumes $\iota = 1$ as firms compete for workers *à la* Cournot.

Under more general assumptions on the market structure, the wage markdown of firm i in market k can be expressed as:

$$\psi_{iF,k} = 1 + \iota \frac{s_{iF,n}}{\epsilon(\hat{W}_k)}, \tag{26}$$

which is a generalization of equation (16).[33] This expression highlights that when employers act as wage-takers, the markdown always equals one. We can therefore gauge the role of labor market power in equilibrium by comparing the baseline economy with one that sets $\iota = 0$.

Figure 3 presents the first set of results. Panel (a) compares the distribution of the share of wage employment across markets in the baseline and in the counterfactual economy with no labor market power. It shows that labor market power significantly restricts wage employment in Peru. Specifically, if labor markets were perfectly competitive, the average share of wage employment would rise by more than ten percentage points, from 60 to 71%, and the number of markets where wage employment accounts for over 80% of jobs would double.

Panel (b) focuses on employer concentration and reveals that, despite increased market entry, high concentration persists and even increases in an economy without labor market power. This result is due to market share reallocation. In the baseline economy, the most productive firms have higher markdowns and thus are the most distorted ones.[34] However, in the absence of labor market power, this source of misallocation disappears, resulting in an increase in market share for the most productive firms and, ultimately, higher concentration.

Panels (c) and (d) show that labor market power depresses average earnings from both wage work and self-employment. Without labor market power, average earnings would increase by 31% in the wage employment sector, and by 27% in the self-employment sector.

Next, we leverage the insights from Section 4.3.2 to decompose the change in average earnings in both sectors into their different components. Equations (27) and (28) summarize the results, where the operator Δ_ι for a given outcome variable X is defined as the average log

[33] To arrive at this equation, notice that the wage markdown of firm i in market k can be represented as:

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

In the baseline economy $\iota = 1$ such that the markdown converges to equation (16).

[34] For a more detailed analysis of a similar misallocation effect, see Berger, Herkenhoff and Mongey (2022). For empirical evidence of this chanel see also Eslava, Haltiwanger and Urdaneta (2023).

Figure 3: Effects of Labor Market Power

Notes. The three 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.

difference in outcome X across local labor markets between the counterfactual and baseline economy, i.e., $\Delta_\iota \ln X \equiv K^{-1} \sum_{k=1}^{K} (\ln X_k|_{\iota=1} - \ln X_k|_{\iota=0})$.

$$\underset{0.3126}{\Delta_\iota \ln \bar{E}_{F,k}} = \Big(- \underset{0.2862}{\overset{\text{markdown}}{\Delta_\iota \ln \bar{\psi}_{F,k}}}\Big) + \underset{0.0653}{\overset{\text{misallocation}}{\Delta_\iota \ln Z_{F,k}}} + \Big(\underset{-0.0386}{\overset{\text{general equilibrium}}{\Delta_\iota \ln GE_{F,k}}}\Big) + \underset{-0.0003}{\overset{\text{selection}}{\Delta_\iota \ln A_{F,k}}} \tag{27}$$

$$\underset{0.2725}{\Delta_\iota \ln \bar{E}_{S,k}} = \Big(\underset{0.1794}{\overset{\text{general equilibrium}}{\Delta_\iota \ln P_{S,k}}}\Big) + \Big(\underset{0.0931}{\overset{\text{selection}}{\Delta_\iota \ln A_{S,k}}}\Big) \tag{28}$$

In the wage sector, the change in average worker earnings can be attributed entirely to the change in earnings per efficiency units, captured by the first three terms in equation (27). The absence of markdowns in the economy without labor market power explains most of the unit (and average) earnings differential, increasing wages by about 28.62%. The reduction in misallocation further increases average earnings by about 6.5% in the economy without labor market power, while general equilibrium effects slightly offset these effects, reducing average wages by 3.9%. The latter effect is due to an increase in the average markup in the economy without labor market power, which results from the reallocation of market shares towards more productive (and high-markup) firms. The selection channel plays virtually no role, consistent with there being only mild negative selection in wage employment.

Equation (28) shows that average earnings also increase in the self-employment sector. About two-thirds of this effect is explained by an increase in unit earnings per efficiency unit, fully attributable to general equilibrium effects on output prices. Unlike the wage sector, the selection channel plays here a crucial role in determining the change in earnings in self-employment. The average ability in self-employment is 9.3% higher in the counterfactual economy, as more workers opt for wage work. The positive worker selection in self-employment explains the increase in average ability as the self-employment share decreases.

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 and the stagnation of firms' size and reduces worker earnings in the wage and self-employment sector through markdowns, selection, misallocation, and general equilibrium effects. In addition, our findings highlight the critical role of worker self-selection as a crucial margin through which labor market power decreases worker earnings in the self-employment sector.

These results have important implications for scholars and policymakers interested in understanding the factors that shape labor market outcomes and designing effective interventions to promote inclusive economic growth and development. In the next section, we explore this idea further by simulating policy interventions in our economy and examining the impact of labor market power.

6.2 Industrial Policy

Despite the long-standing focus on increasing wage employment to promote inclusive growth, policy interventions in this direction have often had limited impact (Bandiera et al., 2022). This section seeks to investigate whether the interplay between labor market power and self-employment can help explain the limited impact of these policies.

In the model, the decisions of heterogeneous firms and workers across local labor markets depend on the following market-level fundamentals: (i) firm productivity shifters T_k, (ii) fixed firm entry cost f_k^e, and (iii) the workers' joint ability distribution G_k. Governments around the developing world have implemented numerous policies to create more job opportunities by targeting these underlying factors. We examine three of these policies and their estimated effects, then simulate and evaluate their impact within our model to evaluate the contribution of labor market power, both qualitatively and quantitatively.

6.2.1 Firm Productivity

To increase firms' productivity, policies have often focused on increasing market integration (McCaig and Pavcnik, 2018), primarily through improving infrastructure (Fiorini, Sanfilippo and Sundaram, 2021). The objective is to increase market access, which increases productivity by reducing information frictions and shipping costs on both the input and output sides.

To evaluate this kind of intervention, we use a road infrastructure project in Peru as a case study. Between 2003 and 2010, Peru invested an average of approximately 895 million USD per year in constructing over 5,000 kilometers of new roads, which expanded the leading road network by over 10%. Using a reduced-form analysis, Volpe Martincus, Carballo and Cusolito (2017) find that firm exports increased by 3.7% on average due to the new road construction. Building on this evidence, we calibrate a shock to market-level productivity (T_k) in our model to achieve an average 3.7% increase in firm sales across local labor markets to assess the potential impact of a comparable infrastructure development program in our model.

The experiment shows a productivity shock of 2.7% in T_k is needed to achieve the targeted sales increase. Despite the uniform shock, its effect varies significantly across local labor markets. Panel (a) of Figure 4 reports the estimated impact of the policy across local labor markets, which we group in six bins according to the policy-induced change in labor market power. The number on top of each blue bar reports the total payroll share of markets in that bin. The Figure shows a substantial increase in wage employment shares and average wages in some markets but only marginal changes in most markets. Changes in markdown appear to account for the observed patterns, as labor market power declines significantly in markets where the policy is most effective but only moderately changes (increases) everywhere else.

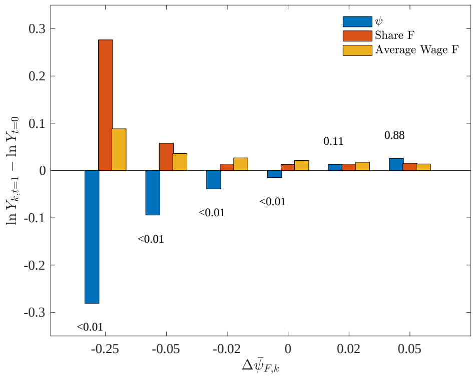
The heterogeneity in markdown changes results from the different impacts of the policy on two underlying factors: labor market concentration and labor supply elasticity. While concentration mostly decreases with increased productivity, the average labor supply elasticity also decreases due to more worker selection into wage work. In most markets, the latter suppresses any policy effect on labor market power, and markdowns increase slightly due to the policy. Our findings suggest that, on average, a 2.7% increase in productivity in local labor markets leads to a 0.8% increase in the average wage markdown, a 0.8% increase in the wage employment share, and a 1.5% increase in the average wage.[35]

We now ask how much the same infrastructure project would impact labor markets if those were perfectly competitive. We compare the policy effects predicted by the baseline model with those estimated in the counterfactual economy where strategic interactions among firms are eliminated, as in Section 6.1. Table 6 shows the results. For each policy exercise, it shows the average effects across markets in the baseline economy ($\Delta\bar{Y}_{\iota=1}$), in the counterfactual economy with no labor market power ($\Delta\bar{Y}_{\iota=0}$), and the percent difference between the two, which is, therefore, a difference-in-differences (% DID) estimate. Columns 1 to 3 focus on the pro-

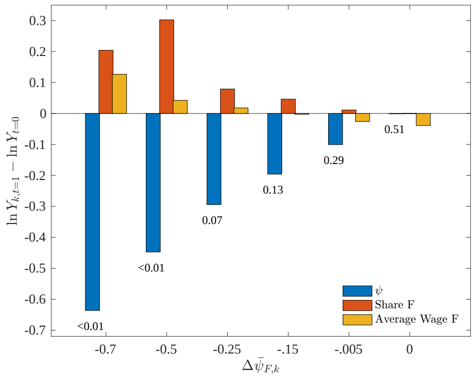
[35] The numbers are similar when markets are weighted by their nationwide payroll share.

Figure 4: Effects of Policy Shocks Across Markets

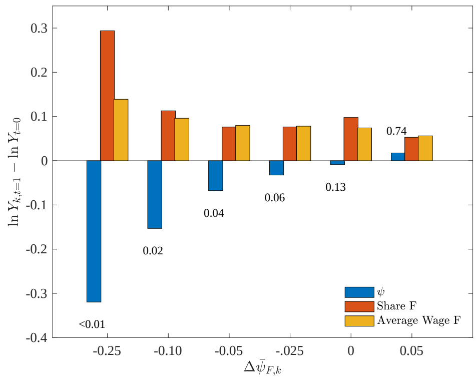
(a) Firm Productivity



(b) Firm Entry



(c) Worker Skills



Notes. The three panels illustrate the estimated change in wage markdown, wage employment share, and average wage across local labor markets resulting from the three policy experiments. It does so for separate bins determined by the size of the markdown change. The numbers next to each set of bars indicate the payroll share of local labor markets in that bin.

ductivity shock, showing that the presence of labor market power affects the pass-through of productivity shock to wages. In the baseline economy with labor market power, the average markdown increase illustrated in Panel (a) of Figure 4 reduces the wage take-home share despite increasing worker productivity. This results in a lower pass-through of productivity shock to wages (56%) and a minor increase in wages (1.5%) compared to the scenario without labor market power, where the pass-through is higher (81%) and the wage increase is almost 50% higher.

Table 6 also shows that labor market power affects earnings from self-employment, with a smaller increase (1.2%) in the baseline economy compared to the scenario without labor market power (1.9%). This result is due to two underlying mechanisms: the improvement in the selection and average ability when the self-employment sector shrinks and the demand shock in both sectors through general equilibrium forces. In the aggregate, the effect of the policy on aggregate labor income is 64% higher in the presence of perfectly competitive labor markets relative to the baseline economy with labor market power.

Table 6: Average Effect of Policy Shocks: Difference-In-Differences

	Firm Productivity ΔT_k			Firm Entry Δf_k^e			Worker Skills $\Delta\hat{\mu}_k$		
	$\Delta\bar{Y}_{\iota=1}$ (1)	$\Delta\bar{Y}_{\iota=0}$ (2)	% DID (3)	$\Delta\bar{Y}_{\iota=1}$ (4)	$\Delta\bar{Y}_{\iota=0}$ (5)	% DID (6)	$\Delta\bar{Y}_{\iota=1}$ (7)	$\Delta\bar{Y}_{\iota=0}$ (8)	% DID (9)
Wage Employment Share	0.008	0.009	0.125	0.012	0.005	-0.583	0.035	0.029	-0.171
Log Avg. Wage $\bar{a}_F W_F$	0.015	0.022	0.467	-0.026	-0.019	-0.269	0.065	0.079	0.215
Log Avg. Self-Empl. earnings $\bar{a}_S W_S$	0.012	0.019	0.583	-0.03	-0.021	-0.3	0.051	0.068	0.333
Log Labor Income	0.011	0.018	0.636	-0.031	-0.021	-0.323	0.047	0.066	0.404

Notes. This table reports, for the main outcomes and the three policy experiments, the average effect of the policy change across markets in the baseline economy ($\Delta\bar{Y}_{\iota=1}$), in the counterfactual economy with no labor market power ($\Delta\bar{Y}_{\iota=0}$), and the percent difference between the two (% DID).

6.2.2 Firm Entry Cost

Policies to reduce entry costs typically involve government programs that simplify entry regulation. A case in point is the Mexican Rapid Business Opening System (SARE), which aimed to simplify local business registration procedures across different Mexican municipalities at various times starting in May 2002. Both Kaplan, Piedra and Seira (2011) and Bruhn (2011) evaluate the impact of the reform on several economic outcomes at the municipality level.[36] Bruhn (2011) finds that the reform increased the number of registered businesses by 5% and the fraction of wage earners in eligible industries by 2.2%. Similarly to what we have done for productivity, we can target a similar wage employment increase, back up the underlying decrease in f_k^e, and evaluate the effects of such policy change within the estimated model.

We find that reducing fixed costs by 41.5% across local labor markets is required to achieve an average increase in wage employment of 2.2%. Despite the notable decline in entry costs, Panel (b) in Figure 4 shows that the labor market impact of this policy would be minor in most markets, and that labor market power is once again responsible for the muted effects. Wage employment and average wages increase the most in those markets where labor market power decreases substantially, but these markets account for a small fraction of nationwide payroll.

In more than half of the local labor markets, the policy does not change wage markdowns nor wage employment shares and lowers labor income. Two mechanisms are responsible for this result. First, the increase in wage work supply elasticity offsets the negative effect of decreasing concentration on markdowns. Second, the entry of inefficient firms leads to a decline in the workers' marginal revenue product and wages. Remarkably, Bruhn (2011) also finds that the income of incumbent businesses decreases by 3%, which is consistent with the adverse effects on wages and income due to increased entry that we estimate within the model.

Columns 4 to 6 of Table 6 show the impact of labor market power on policies aimed at

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

promoting business entry. The increase in wage employment is 58% higher in the baseline economy than in the counterfactual one, but the effect is small (1.2%) and does not come close to filling the initial 11 percentage point gap. Notably, while the earnings of both wage workers and self-employed always decrease when entry costs decline, the reduction is more significant in the economy with labor market power, with the decrease in labor earnings being about 32% lower in the absence of labor market power.

Decomposing the changes in average wages reveals that the observed results are due to differential changes in the workers' marginal revenue product of labor. The reduction in the average MRPL due to the policy is twice as large in our baseline economy (4%) than in the economy with no labor market power (2%). This result is attributed to labor misallocation in the baseline economy, where the most productive firms are also the most distorted ones. Their payroll share is lower than optimal, resulting in a more substantial reduction in average productivity following the entry of low-productivity firms. The reduction in productivity more than offsets the reduction in markdowns due to increased employer competition, leading to a decrease in average wages. Even in self-employment, where average ability increases due to favorable worker selection, earnings decrease due to general equilibrium demand effects.

6.2.3 Worker Skills

The policies discussed so far have focused on firms and labor demand. We now consider a distinct set of policies to enhance the supply of wage work by supporting workers, such as skill training and apprenticeship programs. These programs operate under the assumption that a lack of specific technical skills is responsible for the lack of employment and that these skills can be acquired through short-term training (McKenzie, 2017).

Several programs of this kind have been implemented in Latin America, including the Peruvian Job Youth Training Program, also known as *Projoven*. The program was created in 1996 and ran until 2010, aiming to provide young people with limited resources with short-term training and labor market experience related to the needs of the productive sector, thus oriented towards meeting employers' demand.[37] An experimental evaluation of Projoven by Díaz and Rosas-Shady (2016) found that around two years after the program's completion, randomly selected participants had a 3.6 percentage point higher chance of finding wage employment compared to the control group, although the effect was not statistically significant.[38]

To simulate a similar training program in our model, we consider a shock to the workers' mean comparative advantage in wage work $\hat{\mu}_k = \mu_{F,k} - \mu_{S,k}$. In order to increase wage employment by 3.6%, mean comparative advantage $\hat{\mu}_k$ needs to rise from 0 to 0.12. Similarly to firm productivity and entry policies, Panel (c) of Figure 4 shows that the effect of the training program would be highly heterogeneous across markets and would mostly depend on the

[37] The program cost per beneficiary (including operating costs and a stipend) was around 420 USD.

[38] However, they did find significant positive effects on the probability of having formal employment, such as jobs with health insurance and pensions.

changes in labor market power. The average markdowns can either increase or decrease. On the one hand, the higher availability of skilled wage work encourages firm entry, which leads to a rise in the share of wage employment and a decrease in market concentration and wage-setting power. On the other hand, as workers' comparative advantage in wage work increases and more workers sort into wage work, the supply elasticity of the latter decreases, which increases labor market power. As with the previous two policy exercises, in the average local labor market, these two effects compensate each other, and the average markdown is almost unchanged.

Columns 7 to 9 of Table 6 show how the effect of policies targeting worker skills would change in the absence of labor market power. The average ability of wage workers increases by 12% in both scenarios by construction. However, wages only increase by 6.5 and 8% in the baseline and counterfactual economy, respectively. This is because the policy reduces unit wages, with a greater reduction occurring when firms have labor market power. Interestingly, and consistent with Panel (c) of Figure 4, the policy's effect on unit wages is not due to increased labor market power. Instead, unit wages decline because the entry of inefficient firms affects the workers' marginal revenue product of labor. The latter effect is more pronounced in the presence of labor market power due to a misallocation effect, whereby low-productivity firms hold a huge share of payroll.

As more workers sort into wage work, selection improves in self-employment and average ability and wages increase. However, the latter increase differentially more in the absence of labor market power due to more substantial general equilibrium effects on sectoral demand and unit wages. Taken together, labor income increases by 40% more in the economy without labor market power. Labor market power significantly lowers the impact of policies aimed at improving worker skills, primarily through misallocation and general equilibrium effects.

6.2.4 Summary

This section shows that our framework provides a useful lens for understanding the effects of industrialization policies aimed at increasing wage employment and wages, which depend on their impact on labor market power. These policies are most effective when they reduce wage markdowns substantially, but this is rare. In most cases, worker sorting and its effect on aggregate labor supply elasticity fully offset the policy's negative effects on labor market power via concentration, reducing policy impact on labor market outcomes. Generally, these policies would be more effective in the absence of labor market power because of the selection, misallocation, and general equilibrium effects described above.

7 Conclusions

Understanding the persistently high self-employment rates in many emerging economies, despite substantial policy efforts to reduce them, is a pressing question for development economists.

This paper argues that understanding labor market power and its interplay with self-employment is essential to addressing this issue. To support our argument, we present new evidence on labor market power in Peru, a new general equilibrium model, and counterfactual policy experiments. Our model replicates key features of labor markets in low- and middle-income countries. Using the model, we demonstrate that labor market power depends on both traditional demand-side determinants, namely, employer concentration and oligopsony power, and less traditional supply-side determinants, and in particular the variable labor supply elasticity generated by the sorting of heterogeneous workers across wage work and self-employment.

We demonstrate that self-employment has a dual role in the presence of labor market power. While it can safeguard workers against the dominance of large firms, by offering a livelihood when wage employment opportunities are scarce, it can also hamper the effectiveness of industrial policies designed to make wage employment more attractive. This is due to the variable supply elasticity of wage work. This elasticity increases when concentration is high and wages are low, yet decreases when the wage employment sector expands.

These findings shed light on the role of self-employment in development. On the one hand, they challenge the traditional view of the self-employment sector as a source of inelastic labor supply to the industrial sector, as in Lewis (1954) and Rauch (1991). On the other hand, they suggest that labor market power may prevent the "capitalist" sector from absorbing an efficient share of workers, leading to an overreliance on self-employment and potentially hindering the development process. Therefore, the design of development policies must acknowledge the complex interplay between labor market power, self-employment, and industrial development.

The findings in this paper have implications beyond the context of low-income countries. In high-income countries, the advent of the digital economy, though still recent (Farrell and Greig, 2016), has fundamentally changed the nature of firms and work, leading to increased availability of self-employment and flexible work arrangements and a decline in traditional employment relationships. At the same time, there is growing evidence that firms hold a substantial degree of labor market power over wages. Understanding the interplay between the rise of the digital economy and labor market power of more traditional employers is a promising direction for future work.

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. 2015. "Worker heterogeneity, wage inequality, and international trade: Theory and evidence from brazil." 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.

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

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

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

Azar, José, Steven T Berry, and Ioana Marinescu. 2022. "Estimating Labor Market Power." National Bureau of Economic Research Working Paper 30365.

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.

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 Eco-*

nomic Review, 111(10): 3184–3224.

Brooks, Wyatt J., Joseph P. Kaboski, Illenin O. Kondo, Yao Amber Li, and Wei Qian. 2021*a*. "Infrastructure Investment and Labor Monopsony Power." *IMF Economic Review*, 69(3): 470–504.

Brooks, Wyatt J., Joseph P. Kaboski, Yao Amber Li, and Wei Qian. 2021*b*. "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.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin. 2011. "Finance and development: A tale of two sectors." *American economic review*, 101(5): 1964–2002.

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

Dasso, Rosamaría, Fernando Fernandez, and Hugo R. Nopo. 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.

Diao, Xinshen, Mia Ellis, Margaret S. McMillan, and Dani Rodrik. 2021. "Africa's Manufacturing Puzzle: Evidence from Tanzanian and Ethiopian Firms." National Bureau of Economic Research, Inc NBER Working Papers 28344.

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 K Goldberg, Costas Meghir, and Gabriel Ulyssea. 2021. "Trade and Informality in the Presence of Labor Market Frictions and Regulations." National Bureau of Economic Research Working Paper 28391.

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, Will Jianyu Lu, and Todd Schoellman. 2023. "Labor Market Dynamics and Development." *Quarterly Journal of Economics*. Accepted.

Eslava, Marcela, John Haltiwanger, and Nicolas Urdaneta. 2023. "The Size and Life-Cycle Growth of Plants: The Role of Productivity, Demand, and Wedges." *The Review of Economic Studies*.

Farrell, Diana, and Fiona Greig. 2016. "Paychecks, paydays, and the online platform economy." Vol. 109, 1–40, JSTOR.

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

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

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.

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." National Bureau of Economic Research, Inc 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, and Peter J Klenow. 2010. "Development accounting." *American Economic Journal: Macroeconomics*, 2(1): 207–23.

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

Kahn, Matthew E, and Joseph Tracy. 2019. "Monopsony in spatial equilibrium." *NBER Working Paper, No. w26295*.

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.

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. 2019. "Trade and Market Power in Product and Labor Markets." Mimeo.

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.

McCaig, Brian, and Nina Pavcnik. 2018. "Export Markets and Labor Allocation in a Low-Income Country." *American Economic Review*, 108(7): 1899–1941.

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.

Meghir, Costas, Renata Narita, and Jean-Marc Robin. 2015. "Wages and Informality in Developing Countries." *American Economic Review*, 105(4): 1509–46.

Méndez-Chacón, Esteban, and Diana Van Patten. 2022. "Multinationals, Monopsony and Local Development: Evidence from the United Fruit Company." *Econometrica*. Forthcoming.

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2017. "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India." National Bureau of Economic Research, Inc NBER Working Papers 23838.

OECD. 2006. *OECD employment outlook 2006: Boosting jobs and incomes.* Organisation for Economic Co-operation and Development, Paris, France.

Pham, Hoang. 2019. "Trade Reform, Oligopsony, and Labor Market Distortions: Theory and Evidence." Mimeo.

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.

Wong, Horng Chern. 2023. "Understanding High-Wage Firms." Mimeo.

World Bank. 2012. *World development report 2013: Jobs.* The World Bank.

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

Young, Alwyn. 2014. "Structural transformation, the mismeasurement of productivity growth, and the cost disease of services." *American Economic Review*, 104(11): 3635–67.

Appendix for Online Publication

A Additional Tables and Figures

Table A.1: Employer Concentration Across Local Labor Markets

	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 (Weighted by LLM payroll share)	0.37	0.03	0.37	0.03
Wage-bill HHI (Weighted by LLM employment share)	0.34	0.03	0.34	0.03
Employment HHI	0.63	0.35	0.59	0.35
Employment HHI (Weighted by LLM payroll share)	0.33	0.03	0.33	0.03
Employment HHI (Weighted by LLM employment share)	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 reports summary statistics and employer concentration measures from firm-level data from EEA across Peruvian local labor markets, averaging across 2004-2011. It shows them separately for the entire sample and for the one merged with worker-level data from ENAHO. Local labor markets are defined by 2-digit industries within locations, the latter corresponding to Peruvian provinces or commuting zones.

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 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: Concentration, Self-employment and Earnings – Employment HHI

						Log of Earnings			
	Self-employed {0,1}				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 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 A.4: 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

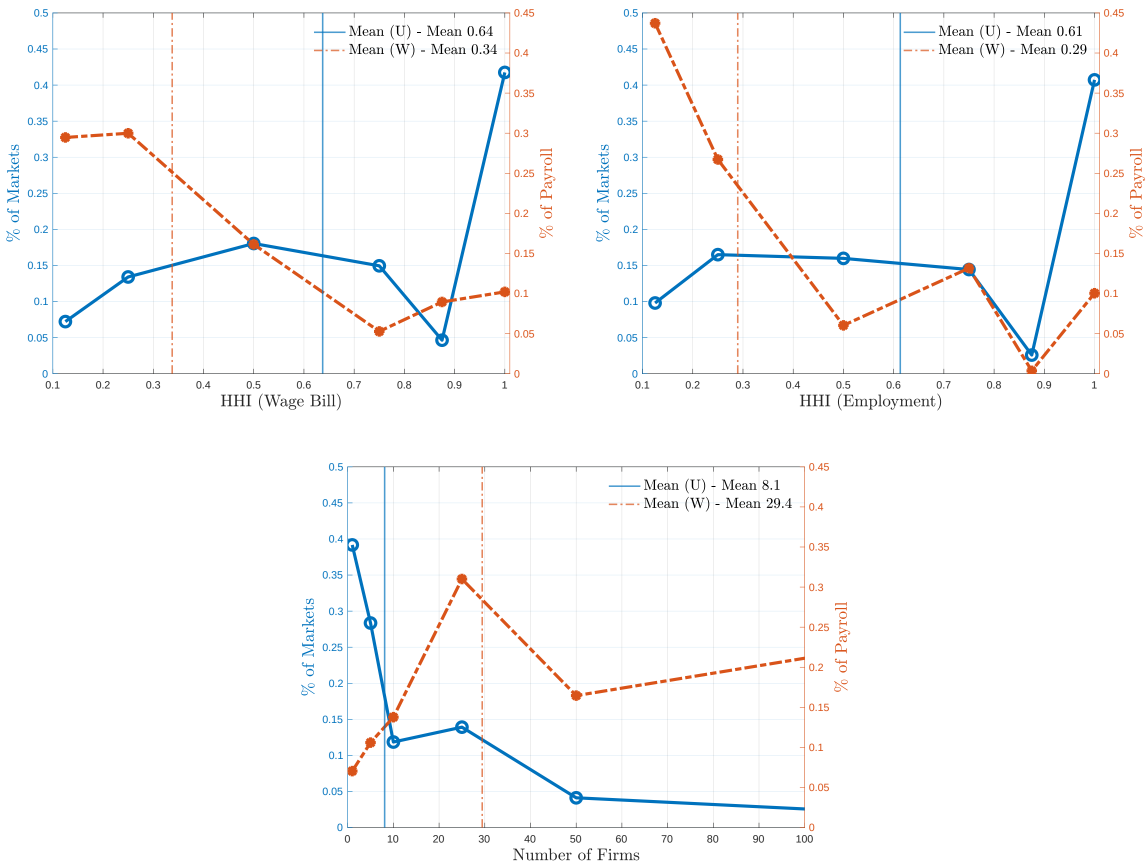
Notes. * p-value< 0.1; ** p-value<0.05; *** p-value<0.01. Unit of observation is a working-age individual surveyed in ENAHO. A local labor market k is defined by a 2-digit industry j within a province or commuting zone g. This table reports the coefficient estimates obtained when estimating the regression 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 A.5: Estimates of Labor Market Power – First Stage Regression Results

$HHI^{wn} \in$ Self-employment Rate	$(0,1]$ All (1)	$(0,0.15]$ All (2)	$(0.15,0.25]$ 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							
$PER_{gt} \times EC_{i(j,g)}$	0.005***							
	(0.000)							
$\times \mathbb{I}\{HHI^{wn} \in (0,0.15]\}$		-0.001	0.002***	0.004***				
		(0.002)	(0.001)	(0.001)				
$\times \mathbb{I}\{HHI^{wn} \in (0.15,0.25]\}$		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)	(0.001)				
$\times \mathbb{I}\{HHI^{wn} \in (0,0.25]\} \times \mathbb{I}\{\text{Low SE}\}$					-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***	-0.027***
					(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

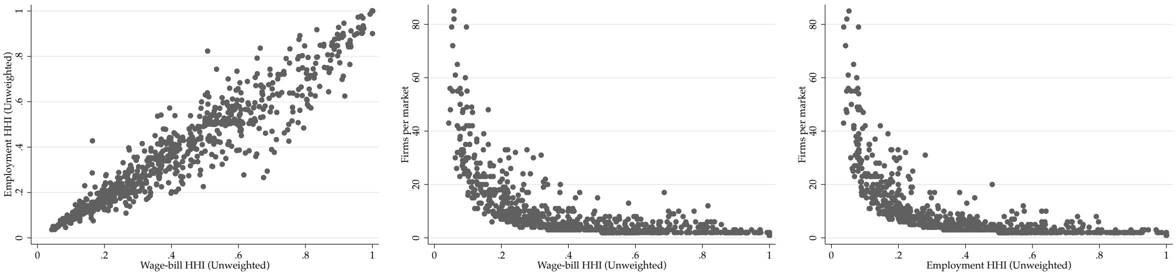
Notes. * p-value$<$ 0.1; ** p-value$<$0.05; *** p-value$<$0.01. Unit of observation is a medium to large firm in EEA. The table reports the first stage estimates corresponding to the second stage estimates reported in Table 2. The dependent variable is the log of firm-level employment $\ln l_{i(j,g)t}$. 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 $\times$ year fixed effects are included in all specifications. Standard errors are clustered at the level of location g, i.e. province or commuting zone.

Figure A.1: Employer Concentration Across Local Labor Markets



Notes. The figure plots the distribution of the three employer concentration measures – wage-bill HHI (top left panel), employment HHI (top right panel), and number of firms (bottom center panel) – across local labor markets. The blue lines represent the unweighted measures, whereas the orange lines plot the share of payroll accounted by those local labor markets. The blue solid line 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 A.2: Correlation Between Employer Concentration Measures

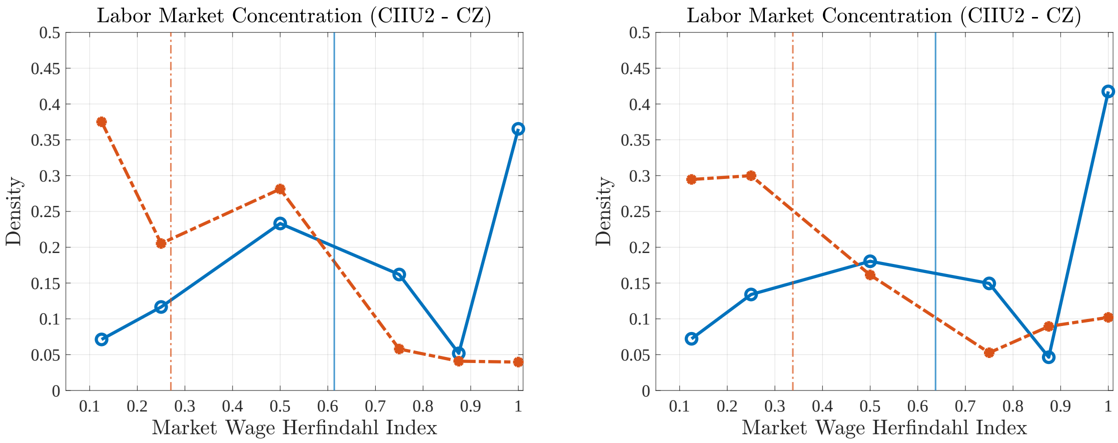


Notes. The figure plots the raw correlation of the three employer concentration measures – wage-bill HHI, employment HHI, and number of firms (bottom center panel) – one against the other across all local labor market-level observations. Wage-bill and employment HHI are strongly positively correlated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

Figure A.3: Employer Concentration Across Local Labor Markets - Census Data vs EEA Data

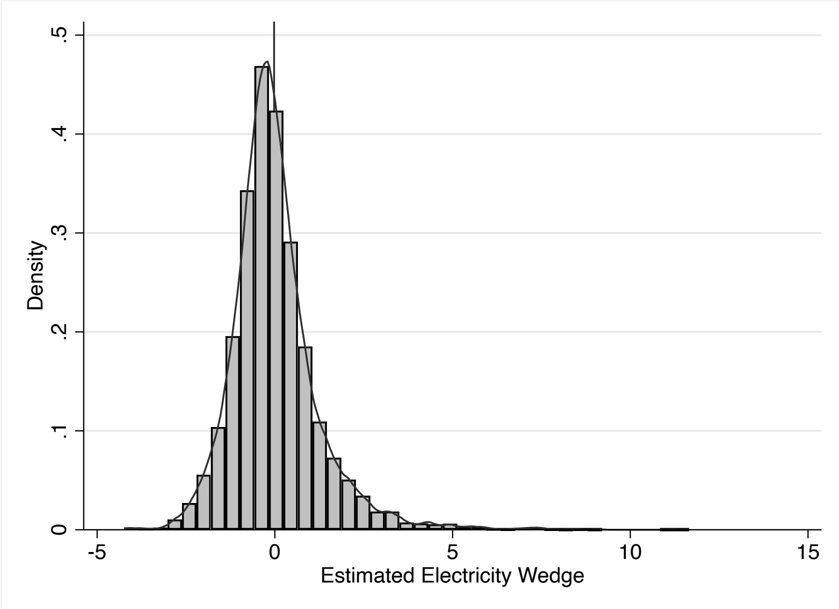
(a) 2007 Census Data

(b) 2007 EEA dataset



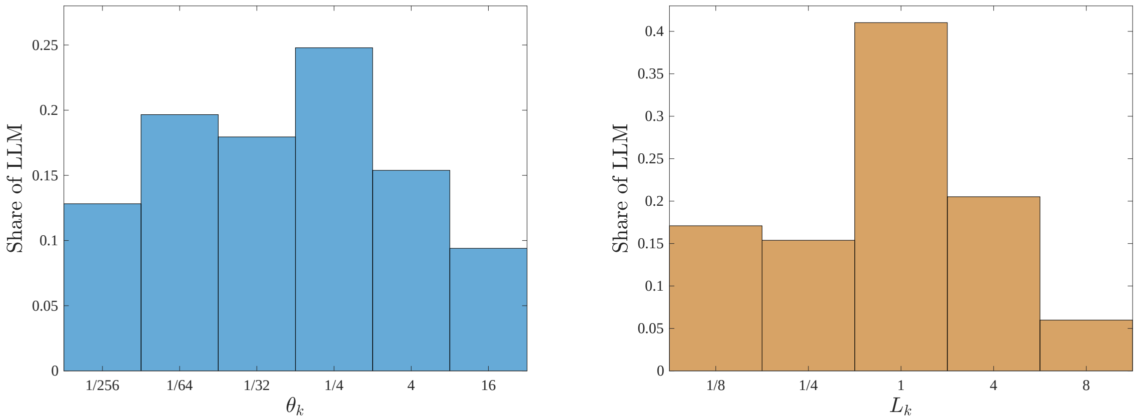
Notes. The figure plots the distribution of wage-bill HHI computed from the 2007 Peruvian Economic Census (left panel) and the same distribution computed from 2007 EEA dataset (right panel) across local labor markets in the manufacturing sector. The blue solid line in both panels corresponds to the unweighted 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 A.4: Distribution of the Estimated Electricity Wedge



Notes. The figure plots the distribution of the estimated electricity wedges $\ln\widehat{(1+\tau^e_{ij})}$ obtained following equation (1) as the estimated residuals from the regression of the (log of) inverse electricity expenditure share $\ln(p_{ij}y_{ij}/e_{ij})$ over the full set of 4-digit industry fixed effects and second-degree polynomials of output market shares in both the local labor market and economy-wide. The vertical bar indicates the value of the median across firms at baseline, i.e. as observed in the first year in which they appear in the data.

Figure A.5: Cobb-Douglas and Population Shares in the Data



Notes. $X_k = N_k \tilde{X}_k$ so that $\mathbb{E}X_k = 1$ for $X = \{\theta, L\}$. Summary statistics are as follows. For the $\tilde{\theta}_k$: mean is 0.85%. The largest Cobb-Douglas share is 11.7%, the 90th percentile is 2.2%, the median is 0.03% and the 10th percentile is 0.003%. For the $\tilde{L}_k$: mean is 0.85%. The largest population share is 6%, the 90th percentile is 2.5%, the median is 0.35% and the 10th percentile is 0.06%. The correlation coefficient from a regression of expenditure shares on population shares is 1.18, with standard error 0.14.

B Aggregate Supply Elasticity of Wage Work

In this section, we discuss conditions under which the aggregate supply elasticity of wage work decreases with the relative unit wage. For expositional clarity, we omit the local labor market subscript, and we write the labor supply elasticity in the wage employment sector as:

$$N(x) = \int_0^\infty \int_0^{a_F \cdot x} a_F \; g(a_F, a_S) da_F da_S,$$

where x denotes the relative wage per efficiency unit in the wage employment sector. The previous function corresponds to equation (9) in the main text.

The first and second derivatives of $N(x)$ can be found as:

$$N'(x) = \int_0^\infty a_F^2 g(a_F, a_F x) da_F > 0$$

and

$$N''(x) = \int_0^\infty a_F^3 g_{a_S}(a_F, a_S) da_F.$$

The aggregate elasticity of wage work supply is thus given by:

$$\epsilon(x) \equiv N'(x)\frac{x}{N(x)} = \frac{x \int_0^\infty a_F^2 g(a_F, a_F x) da_F}{\int_0^\infty \int_0^{a_F \cdot x} a_F \; g(a_F, a_S) da_F da_S} > 0. \tag{1}$$

To determine when $\epsilon(x)$ is decreasing in x, we need to examine the sign of its derivative with respect to x:

$$\epsilon'(x) = N''(x)\frac{x}{N(x)} + N'(x)\frac{N(x) - N'(x)x}{N^2(x)}.$$

We find:

$$\begin{aligned}
&\epsilon'(x) < 0 \\
\frac{N''(x)}{N'(x)} x \left(\frac{N'(x)}{N(x)}\right) &< \left(\frac{N'(x)}{N(x)}\right)^2 x - \frac{N'(x)}{N(x)} \\
\iff \frac{N''(x)}{N'(x)} x &< \frac{N'(x)}{N(x)} x - 1 \\
\iff \rho(x) &< \epsilon(x) - 1
\end{aligned}$$

where $\rho(x) \equiv \frac{N''(x)}{N'(x)} x$ is the super-elasticity of labor supply, measuring the percentage change in the aggregate supply elasticity following a one percent increase in the relative wage. The latter inequality says that, for the elasticity to be decreasing in x, the super-elasticity of labor supply must be sufficiently lower than the labor supply elasticity $\epsilon(x)$. We verify numerically that this condition is always satisfied when the distribution $g(a_F, a_S)$ is joint log-normal.

C Sorting and Joint Log Normality

C.1 Ability Distribution and Sectoral Earnings

In Section 4.3.2 we argued that the scope and sign of the selection channel depend on the parameters of the workers' ability distribution, and in particular on the schedules of absolute and comparative advantage. This section illustrates this point under a standard functional form restriction for the joint ability distribution $G_k(a_F, a_S)$, namely, joint log-normality. For simplicity, we focus the discussion on a single market k and drop the market-level subscript hereafter.

Let the abilities $\mathbf{a} = (a_F, a_S)$ be drawn from the following joint log-normal distribution:

$$\log \mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad \text{where } \boldsymbol{\mu} = \begin{pmatrix} \mu_F \\ \mu_S \end{pmatrix}, \quad \text{and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_F^2 & \varrho\sigma_F\sigma_S \\ \varrho\sigma_F\sigma_S & \sigma_S^2 \end{pmatrix} \tag{A.2}$$

where $\boldsymbol{\mu}$ governs the absolute advantage of workers $\boldsymbol{\Sigma}$ governs the extent of comparative advantage, and ϱ captures the correlation between a_F and a_S.

Following Heckman and Sedlacek (1985), let us define the term $\sigma^* = \sqrt{\sigma_F^2 + \sigma_S^2 - 2\varrho\sigma_F\sigma_S}$. The probability that a given worker operates in sector F is equal to

$$\Pr\left(h \in \textit{wage sector}\right) = \Phi(c_F), \quad c_F \equiv \frac{\ln \hat{W} + \hat{\mu}}{\sigma^*} \tag{A.3}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a normal standard variable, $\hat{W} \equiv \frac{W_F}{W_S}$ is the relative wage per efficiency unit of labor in the wage employment sector, and $\hat{\mu} \equiv \mu_F - \mu_S$ is mean comparative advantage. The probability of operating in sector S is instead equal to $\Phi(c_S) = 1 - \Phi(c_F) = \Phi(-c_F)$.

The mean average ability in each sector, which is defined as the log endowment of efficiency units of labor, is given by

$$\begin{aligned} A_F &\equiv E\left(\ln a_F | a_F\hat{W} \geqslant a_S\right) = \mu_F + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_F) \\ A_S &\equiv E\left(\ln a_S | a_F\hat{W} < a_S\right) = \mu_S + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_S) \end{aligned} \tag{A.4}$$

where $\lambda(x) \equiv \frac{\phi(x)}{\Phi(x)} \geq 0$ is the inverse Mills ratio, which is a convex monotone decreasing function of x.

Consider now a shock ϑ to the economic environment that lowers the relative wage per efficiency unit $\hat{W}$, and therefore shrinks the wage employment sector. Given the system in

(A.4), we can write the response of average ability in the two sectors as:

$$\begin{aligned}\frac{dA_F}{d\vartheta} &= \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right) \cdot \frac{d\lambda(c_F)}{dc_F} \cdot \frac{dc_F}{d\vartheta} \\ \frac{dA_S}{d\vartheta} &= \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right) \cdot \frac{d\lambda(c_S)}{dc_S} \cdot \frac{dc_S}{d\vartheta}.\end{aligned} \tag{A.5}$$

We know that $\frac{dc_F}{d\vartheta} < 0$ and $\frac{dc_S}{d\vartheta} > 0$ by construction. We also know that $\frac{d\lambda(c_F)}{dc_F} < 0$ and $\frac{d\lambda(c_S)}{dc_S} < 0$. This implies that the signs of $\frac{dA_F}{d\vartheta}$ and $\frac{dA_S}{d\vartheta}$ are uniquely determined by the the two terms $\left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)$ and $\left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)$.

If the two abilities are uncorrelated ($\varrho = 0$) or negatively correlated ($\varrho < 0$), the numerators of both terms are strictly positive, implying that $\frac{dA_F}{d\vartheta} > 0$ and $\frac{dA_S}{d\vartheta} < 0$. In this case, average ability will increase in the wage employment sector and decrease in self-employment.

If instead $\varrho > 0$ 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 $\left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right) > 0$, from which it follows $\frac{dA_F}{d\vartheta} < 0$ and $\frac{dA_S}{d\vartheta} < 0$. This means that if the correlation between abilities is positive and high enough, and if abilities are more dispersed in the self-employment sector compared to the wage employment sector, average ability will decrease in both the wage employment and the self-employment sector.

C.2 Self-Employment and Relative Wage

The assumption of jointl log-normality on the sectoral abilities implies a tractable characterization of the aggregate supply elasticity of wage work.

Let us define the self-employment share in a generic local labor market as

$$\textit{self rate} = \frac{L_S}{L_F + L_S}.$$

From equation (A.3), we can write:

$$\textit{self rate} = 1 - \Phi\left(\frac{\ln \hat{W} + \hat{\mu}}{\sigma^*}\right) \tag{A.6}$$

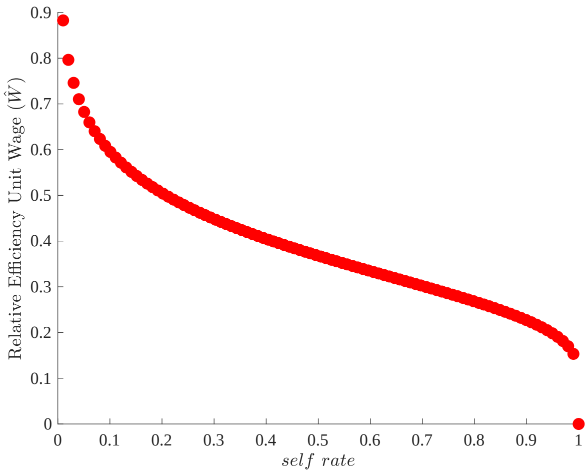
Equation (A.6) shows that there exists a one-to-one negative relationship between the equilibrium self-employment share, and the market-level relative wage per efficiency unit $\hat{W}$.

Because the cdf function $\Phi(\cdot)$ is monotone and increasing, it is also invertible. By simple algebra, we could further rearrange equation (A.6) to write:

$$\hat{W} = \exp\left(-\hat{\mu} + \sigma^*\Phi^{-1}\left(1 - \textit{self rate}\right)\right), \tag{A.7}$$

where $\Phi^{-1}(x)$ is the quantile function or the inverse normal distribution function, which is a

Figure A.6: Relative Wage Per Efficiency Unit and Self-Employment



monotone increasing function of x.

Figure A.6 uses equation (A.7) to plot the relative wage $\hat{W}$ as a function of the self-employment rate, showing a monotone negative relationship between the two. We thus write the latter in compact form as $\hat{W} = h(\underbrace{\textit{self rate}}_{-})$. Given that the aggregate supply elasticity of wage work is a function of $\hat{W}$, it follows that we can also express it as:

$$\begin{aligned}\epsilon_F\left(\hat{W}\right) &= \epsilon_F\underbrace{\left(h(\underbrace{\textit{self rate}}_{-})\right)}_{-}\\ &= \tilde{\epsilon}_F\left(\underbrace{\textit{self rate}}_{+}\right).\end{aligned}$$

C.3 Inferring Ability Distribution Moments

We now show that the assumption of joint log-normality also facilitates the estimation of the ability parameters, and in particular of the variance matrix Σ.

When the abilities $\mathbf{a} = (a_F, a_S)$ follow the distribution in equation (A.2), the observed variance of log earnings is equal to

$$\begin{aligned}Var\left(\ln a_F W_{F,k} | a_F \hat{W}_k \geqslant a_S\right) &= \sigma_F^2 + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[\lambda(c_{F,k})c_{F,k} - \lambda^2(c_{F,k})\right]\\ Var\left(\ln a_S W_{S,k} | a_F \hat{W}_k \leq a_S\right) &= \sigma_S^2 + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)^2 \left[\lambda(c_{S,k})c_{S,k} - \lambda^2(c_{S,k})\right]\end{aligned}, \tag{A.8}$$

where

$$c_{F,k} = \Phi^{-1}(\textit{wage share}) \quad \text{and} \quad c_{S,k} = -c_{F,k}, \tag{A.9}$$

and where $\lambda(c)$ is the inverse Mills ratio defined as above.

The left-hand side of equation (A.8) is observed in a cross section of workers' earnings across the two sectors. Similarly, equation (A.9) shows that we can easily recover the terms $c_{F,k}$ and $c_{S,k}$ from simple inversion of the observed employment shares in the two sectors, from which we can also get $\lambda(c_{F,k})$, $\lambda(c_{S,k})$.

Hence, under the assumption of common variance matrix across local labor markets, i.e., $\mathbf{\Sigma}_k = \Sigma$ for all k, the system in equation (A.8) is a system of $2 \times K$ equations in a vector of 3 unknowns, namely, $\mathbf{\Theta} = (\sigma_F, \sigma_S, \varrho)$.

We recover the vector $\mathbf{\Theta}$ from a Minimum Distance Estimation (MDE) procedure with non-negativity constraints on the parameters. Let $h_{j,k}(\mathbf{\Theta})$ denote the RHS of equation (A.8). The constrained minimum distance estimator of $\mathbf{\Theta}$ is given by:

$$\begin{gathered}\hat{\mathbf{\Theta}} : \quad \arg\min f_k(\mathbf{\Theta})' \mathbb{W} f_k(\mathbf{\Theta}), \\ \text{subject to } \ \sigma_j > 0, \ \text{ and } \varrho \in (-1, 1),\end{gathered} \tag{A.10}$$

where $f_k(\mathbf{\Theta}) \equiv Var\left(\ln a_{s,k} W_{s,k} | a_{s,k} W_{s,k} \geqslant a_{-s,k} W_{-s,k}\right) - h_{s,k}(\mathbf{\Theta})$ for $s = \{F, S\}$ with $\mathbb{W} = \mathbb{I}$. To obtain standard errors, we iterate the minimum distance estimation procedure 1000 times using bootstrapped samples of local labor markets with replacement, thus keeping the sample size equal to the original one.

Identification of Mean Absolute Advantage Terms Given the system of abilities in equation (A.4), one could also write the mean of log earnings in each sector in market k as:

$$\begin{aligned} E\left(\ln a_F W_{F,k} | a_F \hat{W}_k \geqslant a_S\right) &= \ln W_{F,k} + \mu_{F,k} + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_{F,k}) \\ E\left(\ln a_S W_{S,k} | a_F \hat{W}_k \leq a_S\right) &= \ln W_{S,k} + \mu_{S,k} + \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_{S,k}) \end{aligned}, \tag{A.11}$$

which implies the following expression for the difference across sectors in the mean of log earnings:

$$\begin{gathered} E\left(\ln a_{F,k} W_{F,k} | a_F \hat{W}_k \geqslant a_S\right) - E\left(\ln a_{S,k} W_{S,k} | a_F \hat{W}_k \leq a_S\right) = \\ \ln \hat{W}_k + \hat{\mu}_k + \left(\frac{\sigma_F^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_{F,k}) - \left(\frac{\sigma_S^2 - \varrho\sigma_F\sigma_S}{\sigma^*}\right)\lambda(c_{S,k}), \end{gathered} \tag{A.12}$$

Equations (A.11) and (A.12) show that the absolute advantage terms μ_s for $s = \{F, S\}$, and the mean absolute advantage $\hat{\mu}_k$, are not identified given cross-sectional data on mean earnings. Even if the left-hand side in both equations is observed and the last (two) terms in the right-hand side are known given the estimated $\mathbf{\Theta} = (\sigma_F, \sigma_S, \varrho)$ and $\{c_{F,k}, c_{S,k}\}$, the relative unit wage $(\ln \hat{W}_k)$ is unobserved and related to $\hat{\mu}_k$ via the joint ability distribution, hampering identification.

D Estimation Appendix

1. For given parameter values of (μ_T, σ_T), (F_λ, F_k), and (μ_μ, σ_μ), we draw K local labor market productivities T_k from the log-normal distribution with parameters (μ_T, σ_T), K local labor market fixed entry costs f_k^e from the Weibull distribution with parameters (F_λ, F_k) and location parameter 0, and K mean absolute advantage parameters μ_k from the log-normal distribution with parameters (μ_μ, σ_μ). We keep the seed of all random draws constant throughout estimation.[39]
2. For given values of parameter Z_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, which follows a $Frechet(Z_\theta, T_k)$ distribution, 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. Eaton and Kortum (2010) show that $U_k^{(1)}$, $(U_k^{(2)} - U_k^{(1)})$, $(U_k^{(3)} - U_k^{(2)})$, etc., are i.i.d. exponential with cdf $G_U(u) = 1 - e^{-u}$. 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.
3. With the calibrated value of local labor market shares and populations $\{\theta_k, L_k\}_{k=1}^K$, the normalization $P = 1$, and given the estimates of the variance matrix of ability distribution Σ, the draws of $\{T_k, f_k^e, \mu_k, \{z_{iF,k}\}_{i=1}^{\bar{M}}\}_{k=1}^K$, and the remaining model parameters (ρ), we implement the following fixed point procedure:
   (a) Take an initial guess for aggregate income Y_0, which completes the general equilibrium vector $\mathbf{X} = (Y, P)$.
   (b) Given $\mathbf{X}$, we solve for the market equilibrium $\mathbf{K} = \{\mathbf{M}, \hat{\mathbf{W}}, \mathbf{\Lambda}\}$, as described in Section 4.2 in the main text and detailed in Section D.1 below. Because entry is a computationally complex problem, we consider a simplified entry game, while verifying that the approximation error is negligible.
   (c) Given $\mathbf{K}$, use the general equilibrium conditions (18) and (19) to solve for the new values of Y.
   (d) 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.
   (e) 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 $\Phi = (Z_\theta, \mu_T, \sigma_T, F_\lambda, F_k, \mu_\mu, \sigma_\mu, \rho)$.

[39]To avoid mechanical correlations between the draws from different distribution, we consider three different (random) seed values for the three distributions. The correlation matrix is:

$$\begin{array}{c} T \\ \mu \\ f^e \end{array} \begin{bmatrix} 1 & . & . \\ -.0037 & 1 & . \\ -.1132 & .0459 & 1 \end{bmatrix}$$

4. On a grid for parameters Φ with 10,000 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 $\tilde{\mathbf{m}}$. 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.
5. With the results from the first Halton grid, we recompute a second finer Halton grid of 10,000 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.
6. 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.1 Solving for the *Market* Equilibrium

We now describe in details step 3(b), which solves for the market equilibrium. We first derive the key equilibrium expression needed to solve for the relative wage in each market, and then detail the algorithm.

D.1.1 Main Equilibrium Equation

The price of each firm i in market k is:

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

with associated price index:

$$P_{F,k} \equiv \left(\sum (p_{iF,k})^{1-\sigma}\right)^{\frac{1}{1-\sigma}} = \frac{W_{F,k}}{T_k\Phi_{F,k}},$$

where $\Phi_{F,k} \equiv \left(\left[\sum_i \left(\frac{z_{iF,k}}{\mu_{iF,k}\psi_{iF,k}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}\right)$ is a market-level index of aggregate productivity, which reflects both true productivity and misallocation effects. The self-employment sector is competitive, such that aggregate prices reflect marginal cost: $P_S = W_S$.

Given equation (6), we write:

$$\frac{P_{F,k}Y_{F,k}}{P_{S,k}Y_{S,k}} = \beta^{\rho}\left(\frac{P_{F,k}}{P_{S,k}}\right)^{1-\rho}$$
$$\implies \frac{Y_{F,k}}{Y_{S,k}} = \beta^{\rho}\left(\frac{P_{F,k}}{P_{S,k}}\right)^{-\rho}.$$

Substituting for $Y_{F,k}$ and $Y_{S,k}$ using (11) and (13), using the equations above to substitute for prices, and rearranging terms, we obtain:

$$\hat{W_k}^{\rho}\hat{N}\left(\hat{W}_k\right) = \beta^{\rho}T_k^{\rho-1}\Phi_{F,k}^{\rho}\tilde{Z}_{F,k}^{-1}, \tag{A.13}$$

where $\hat{N}(\hat{W}_k) \equiv \frac{N_F(\hat{W}_k)}{N_S(\hat{W}_k)}$ is the relative labor supply (in efficiency units of labor) in sector F. Equation (A.13) gives a vector of K equilibrium conditions in K unknowns $\{\hat{W_k}\}$.

The profit of firm i are:

$$\pi_{iF,k} = s_{iF,k}\alpha_{F,k}\theta_k Y\left(1 - \frac{1}{\mu_{iF,k}\psi_{iF,k}}\right) - f^e,$$

where the sector market share $\alpha_{F,k}$, solves:

$$\left(\frac{\alpha_{F,k}}{1-\alpha_{F,k}}\right) = \beta^{\rho}\left(\hat{W}_k\right)^{1-\rho}T_k^{\rho-1}\Phi_{F,k}^{\rho-1}.$$

D.1.2 Solution Algorithm (for given M_k)

With equilibrium condition (A.13), we can now lay out an algebraic algorithm to solve for the market equilibrium $\{\hat{\mathbf{W}}, \mathbf{\Lambda}\}$, given general equilibrium variables $(Y, 1)$ and the number of entrants $\{M_k\}_k$. We will solve for the number of entrants next.

1. Given the GE vector $\mathbf{X} = (Y, 1)$, we first compute expenditures on each sector using the Cobb-Douglas formula:
$$Y_k = \theta_k Y. \tag{A.14}$$
2. Given Y_k and the set of entrants $\{M_k\}$, and market-level draws $\{T_k, f_k^e, \mu_k, \{z_{iF,k}\}_{i=1}^{M_k}\}_{k=1}^{K}$, the equilibrium in each LLM is a relative efficiency wage $\hat{\mathbf{W}} = \{\hat{W}_k\}$, and vector $\mathbf{\Lambda} = \{\{s_{iF,k}, s_{iF,k}^N, \mu_{iF,k}, \psi_{iF,k}\}_{i\in M_k}\}_{k\in(0,1)}$, such that:
   (a) Given $\hat{W}_k$, the vector $\mathbf{\Lambda}$ solves the following fixed point problem:
$$s_i = \frac{\left(\mu_i\frac{\psi_i}{z_i}\right)^{1-\sigma}}{\sum_{i=1}^{M_k}\left(\mu_i\frac{\psi_i}{z_i}\right)^{1-\sigma}},$$

where the markup is

$$\mu_{iF,S} = \frac{\varepsilon_{iF,k}}{\varepsilon_{iF,k}-1}, \quad \text{with} \quad \varepsilon_{iF,k}=\left[\frac{1}{\sigma}(1-s_{iF,k})+\frac{1}{\rho}s_{iF,k}\right]^{-1},$$

and the markdown is

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

(b) Given $\mathbf{\Lambda}$, the wage $\hat{W}_k$ solves (A.13).

3. We solve for $\{\hat{\mathbf{W}}, \mathbf{\Lambda}\}$ using an iterative fixed point procedure.

D.1.3 Entry (Solving for M_k)

Solving for exact equilibrium values of M_k is computationally costly, therefore, we adopt the following approximation procedure. We assume that upon entry, firms consider a simplified problem by expecting that they will charge the minimum markup and markdown. That is, we assume that $\mu = \frac{\sigma}{\sigma-1}$ and $\psi = 1$. The profits are:

$$\pi_{iF,k} = s_{iF,k}\alpha_{F,k}\frac{\theta_k Y}{\sigma} - f^e. \tag{A.15}$$

With constant markups, the firm's market share is only a function of productivity:

$$s_i = \frac{(z_i)^{\sigma-1}}{\sum_{i=1}^{M_k}(z_i)^{\sigma-1}} = s_i^N.$$

So to solve for profits above we have everything but the shares α. However, notice that under the assumption of constant markups and markdowns, the index $\Phi(\hat{W}_k)$ becomes: $\Phi_{F,k} \equiv \left(\left[\sum_i \left(\frac{z_{iF,k}}{\mu_{iF,k}\psi_{iF,k}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}\right) = \left(\frac{\sigma}{\sigma-1}\right)^{-1}\tilde{Z}_{F,k}$, where $\tilde{Z}_{F,k} \equiv \left(\left[\sum_i (z_{iF,k})^{\sigma-1}\right]^{\frac{1}{\sigma-1}}\right)$, such that equilibrium equation (A.13) becomes:

$$\hat{W}_k^{\rho}\hat{N}_k = \left(\frac{\sigma}{\sigma-1}\right)^{-\rho}\beta^{\rho}\left(T_k Z_{F,k}\right)^{\rho-1}, \tag{A.16}$$

which means that $\hat{W}_k$ can be solved as only a function of known variables, for given M_k. With $\hat{W}_k$, the sector market share $\alpha_{F,k}$, solves:

$$\left(\frac{\alpha_{F,k}}{1-\alpha_{F,k}}\right) = \frac{\sigma}{\sigma-1}\hat{W}_k\hat{N}_k, \tag{A.17}$$

which allows to compute the profits of all entrants given equation (A.15). Hence, the number of entrants can be solved as follows:

1. For each value of $M = 1, ..., \bar{M}$ and each $k = 1, .., K$, compute the relative wage, sector market share, and profits of marginal entrants from equations (A.15), (A.16), and (A.17).
2. An equilibrium of the entry game is achieved when, for a given number of entrants M, the profits of the marginal entrant are non-negative, while for any additional entrant $M + 1$, profits upon entry would be negative.

D.1.4 Solution Algorithm (General)

We can finally put pieces together and describe the algorithm that solves for the market equilibrium $\mathbf{K} = \{\mathbf{M}, \hat{\mathbf{W}}, \boldsymbol{\Lambda}\}$.

1. First, given the general equilibrium vector $(Y, 1)$, solve for the number of entrants in each market $\{M_k\}_{k=1}^{K}$ following the simplified entry game described in D.1.3.
2. Given the general equilibrium vector $(Y, 1)$ and the number of entrants $\{M_k\}_{k=1}^{K}$, solve for the vectors $\{\hat{\mathbf{W}}, \boldsymbol{\Lambda}\}$ following the procedure described in D.1.2.

D.1.5 Approximation Error

To assess the approximation error of adopting the simplified entry procedure instead of the full procedure, we proceed as follows. With the estimated model parameters that we obtain following the procedure in Section 5, we compute the equilibrium in the simplified entry world as well as in the full entry model. We check numerically that our procedure recovers a M_k which differs from the exact solution by at most one firm. This differences does not affect the main results in the paper, neither qualitatively nor quantitatively. For this reason, we view this approximation error as small, and adopt the simplified entry assumption throughout.