

DISCUSSION PAPER SERIES

IZA DP No. 15477

**Labor Market Power, Self-Employment,
and Development**

Francesco Amodio
Pamela Medina
Monica Morlacco

AUGUST 2022

DISCUSSION PAPER SERIES

IZA DP No. 15477

Labor Market Power, Self-Employment, and Development

Francesco Amodio

McGill University, CIREQ and IZA

Pamela Medina

UTSC and Rotman School of Management

Monica Morlacco

University of Southern California

AUGUST 2022

Any opinions expressed in this paper are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but IZA takes no institutional policy positions. The IZA research network is committed to the IZA Guiding Principles of Research Integrity.

The IZA Institute of Labor Economics is an independent economic research institute that conducts research in labor economics and offers evidence-based policy advice on labor market issues. Supported by the Deutsche Post Foundation, IZA runs the world's largest network of economists, whose research aims to provide answers to the global labor market challenges of our time. Our key objective is to build bridges between academic research, policymakers and society.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

Labor Market Power, Self-Employment, and Development*

This paper shows that self-employment opportunities shape the market power of employers in low-income countries, with implications for industrial development. Using data from Peru, we document substantial employer concentration and high self-employment rates across manufacturing local labor markets. Where employer concentration is higher, wages are lower, and self-employment is more prevalent but less remunerative. To interpret these facts, we build a general equilibrium model where labor market power in each market arises from (i) strategic interactions among employers and (ii) sorting of heterogeneous workers across wage work and self-employment. We structurally estimate the model and quantify the relevance of these mechanisms for rent-sharing between workers and firms and for the effect of policies promoting manufacturing wage employment. We show that changes in concentration magnify the pass-through of productivity and profitability shocks to wages, but worker sorting across wage and self-employment mitigates these effects. We find that policies that increase firm productivity are more effective in expanding wage employment and increasing workers' earnings than other interventions that improve workers' skills or decrease firm entry cost.

JEL Classification: J2, J3, J42, L10, O14, O54

Keywords: labor market power, monopsony, self-employment, sorting, development

Corresponding author:

Francesco Amodio
McGill University
Leacock Building
855 Sherbrooke St. West
Montreal QC H3A 2T7
Canada

E-mail: francesco.amodio@mcgill.ca

* We thank Jacob Schmidt and Paola Marino Moreno for excellent research assistance. We are thankful to Vittorio Bassi, Arin Dube, Rafael Dix-Carneiro, Jonas Hjort, Isabela Manelici, Christian Moser, Suresh Naidu, Tommaso Porzio, Todd Schoellman, Gabriel Ulyssea, Jose Vasquez, Daniel Xu, and all participants to seminars at Boston University, JLAEE-University of Chicago, McGill University, UC Riverside, University of Bologna, University of Milan, Universidad de Montevideo, University of Nottingham, University of Southern California, Western University, University of Toronto, as well as participants in the Barcelona GSE Summer Forum Workshop "Market Power and the Labor Market," the STEG Theme 1 Workshop, the CEPR-ENS-Lyon Firms' Behavior in Hostile Environments Workshop, the FIMAD Seminar Series, LACEA, the MEA meetings, the NYU-Columbia Trade day, PacDev, the RES Annual Conference, Seminario MEP for their comments. Errors remain our own. We gratefully acknowledge financial support from the CEPR/DFID PEDL initiative – ERG 6500.

1 Introduction

The creation of salaried jobs has been one of the hallmarks of sustained economic growth (Diao, Ellis, McMillan, and Rodrik, 2020; McMillan and Zeufack, 2022). This has motivated a large set of industrial and labor market policies promoting wage employment sector growth in poor countries, with generally modest results (World Bank, 2012; Bandiera, Elsayed, Smurra, and Zipfel, 2022).¹ Understanding the factors determining the success of these policies remains a pressing open question. This paper shows that the market power of employers and its interaction with self-employment are critical for the effectiveness of industrial policy in low-income countries.

The combined presence of barriers to firm entry and high self-employment rates motivates our focus on labor market power. On the one hand, high entry costs, lack of a skilled workforce, and poor infrastructure limit the size and productivity of the wage employment sector (Djankov, La Porta, Lopez-de Silanes, and Shleifer, 2002; Hsieh and Klenow, 2010; Rud and Trapeznikova, 2021), increasing concentration among employers. Non-atomistic firms compete strategically for workers, reducing job opportunities and wages. On the other hand, self-employment is common, mostly informal (Gollin, 2008; Poschke, 2019), and features high rates of transition to and from wage employment (Donovan, Lu, and Schoellman, 2021). Workers choose to be self-employed when posted wages are too low (Blattman and Dercon, 2018; Breza, Kaur, and Shamdasani, 2021), limiting the scope of labor market power. Strategic competition among employers and workers' behavior jointly determine the extent of labor market power in the economy as well as the effectiveness and distributional impact of industrial policy.

We study these issues in the context of Peru where, similarly to several other low and middle-income countries, limited wage employment opportunities coexist with high rates of informal self-employment. Our empirical analysis combines balance sheet data on manufacturing firms with worker-level survey data on employment and earnings from 2004 to 2011. We start by presenting several novel stylized facts about the Peruvian manufacturing labor market. First, most local labor markets exhibit high levels of concentration among medium and large manufacturing employers, a local labor market being defined as an industry-commuting zone combination. On average, and across local labor markets, the probability that two randomly chosen wage workers work for the same firm is as high as 63%.

Second, both self-employment rates and labor market flows to and from self-employment are high, even within manufacturing. The share of manufacturing workers that are self-employed is about 40%. Around 15% (18%) of all wage (self-employed) workers in a given year were self-employed (wage) workers in the previous year. Transitions to wage employment are systematically more common among low-earning self-employed. In contrast, transitions to self-

¹See also OECD (2006); Rodrik (2008); Card, Kluve, and Weber (2010); McKenzie (2017).

employment are more prevalent among high-earning wage workers. Despite substantial mobility, more than 80% of manufacturing workers are repeatedly observed in the same 2-digit industry.

Third, concentration relates systematically to features of both wage work and self-employment. Across local labor markets, concentration among medium and large employers is strongly *positively* correlated with self-employment rates. It is also strongly *negatively* correlated with the average earnings and education levels of both wage workers and self-employed workers. The elasticity of wages to employer concentration – measured as wage-bill HHI – is -0.10 while the one of self-employment earnings to concentration is even larger, equal to -0.16.

Motivated by these facts, we build a general equilibrium model of the Peruvian economy to study the implications of labor market power and the role of self-employment. In the model, (i) strategic interactions among employers and (ii) worker sorting between wage work and self-employment jointly determine the extent of labor market power. Non-atomistic firms understand the effect of their hiring decisions on aggregate earnings and employment levels and strategically choose labor demand. Worker sorting gives rise to an upward-sloping supply of wage workers, which non-atomistic firms internalize.

On the firm side, multiple forces affect entry into the wage employment sector: firm productivity, fixed entry costs, and workers' ability or skill level. On the worker side, the sorting decision depends on their comparative advantage and posted wages (Roy, 1951): the higher the wage posted by employers, the more workers opt for wage employment as opposed to self-employment. In equilibrium wage employment is below competitive levels, driving a wedge between the marginal product of labor and unit wages. We take this wedge, or wage markdown, as the measure of labor market power in a given local labor market.

Our model yields several novel insights about the relationship between employer concentration, labor market power, and self-employment. First, we show theoretically that the relationship between employer concentration and labor market power is ambiguous. Firm entry directly lowers labor market power by reducing the scope of strategic interactions among employers. At the same time, however, it can raise labor market power indirectly because of worker sorting. Higher wages draw workers away from the self-employment sector into the wage employment sector, and the aggregate elasticity of wage work can decrease as a result. This finding can potentially reconcile the mixed empirical evidence on the relationship between employer concentration and labor market power (Bassier, Dube, and Naidu, 2022; Berger, Herkenhoff, and Mongey, 2022; Yeh, Macaluso, and Hershbein, 2022).

The second insight is that the impact of labor market power on the economy depends on the shape of the workers' joint ability or skill distribution, specifically on the degree of correlation between workers' abilities for wage and self-employment, and their relative dispersion.

When abilities are highly positively correlated and more dispersed in self-employment than in wage work, average worker productivity decreases in both sectors as labor market power increases (Heckman and Sedlacek, 1985; Heckman and Honoré, 1990; Young, 2014; Alvarez-Cuadrado, Amodio, and Poschke, 2020). Intuitively, this happens because the more skilled wage workers are better off as self-employed, implying negative selection in wage work and positive selection in self-employment. When labor market power increases, wages decrease and the most productive wage workers transition to self-employment, decreasing the productivity of wage work. The average productivity of self-employed workers also decreases because the newly self-employed are less skilled than those already active. This is not the case when the correlation between abilities is low or negative: both sectors feature positive selection, and labor productivity is inversely related to their employment share. When labor market power increases, the wage employment sector shrinks, but its productivity increases.

We estimate the model using our firm- and worker-level data. We consider a parameterization of the model that allows for heterogeneity across local labor markets in the fundamental forces affecting firm entry: entry costs, employers' productivity, and workers' absolute advantage in the wage sector. 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.

Identification of the model's parameters relies on a combination of direct and indirect inference methods. We first use observed sectoral employment shares, average earnings, and their variance to infer part of the workers' ability distribution parameters, notably the dispersion in workers' ability in the two sectors and their correlation. We find that abilities are strongly positively correlated across the wage and self-employment sectors and more dispersed in the latter. The data are thus consistent with a scenario where employers' labor market power negatively affects workers' productivity and earnings in both sectors, in line with our reduced-form evidence. We rely on a Simulated Method of Moments (SMM) strategy to estimate the remaining model's parameters. We discipline our model by matching moments derived by our stylized facts on cross-sectional dispersion of concentration and self-employment.

We use the estimated model to conduct two sets of counterfactual exercises. First, we study the pass-through of productivity and profitability shocks onto average wages. We show that the pass-through (or rent-sharing) elasticity of wages critically depends on labor market power, and in particular on the interlinkages between strategic interactions among employers and workers' sorting. We find that changes in concentration magnify the pass-through of productivity and profitability shocks to wages, yet worker sorting mitigates these effects. That is, worker sorting dampens the effect of labor market power on rent-sharing through its effects on the aggregate supply elasticity of wage work.

The second counterfactual exercise studies the aggregate and distributional consequences of

industrial and labor market policies in our economy with labor market power. We consider and evaluate three common types of policies promoting industrialization and manufacturing wage employment. These policies include: (i) reducing fixed entry costs for employers, for instance by streamlining regulations for business registration (e.g., one-stop shops) or reducing financial constraints of firms (Branstetter, Lima, Taylor, and Venâncio, 2014; Buera, Kaboski, and Shin, 2011); (ii) increasing productivity and connectivity of firms, e.g., by ameliorating infrastructures or through market integration policies (Allen, 2014; Startz, 2016; Brooks, Kaboski, Kondo, Li, and Qian, 2021); (iii) improving workers' skills through vocational and on-the-job training programs (Alfonsi, Bandiera, Bassi, Burgess, Rasul, Sulaiman, and Vitali, 2020). We find that these policies have very different aggregate and distributional effects as changes in firm entry and worker sorting simultaneously affect the size of wage markdowns and the extent of labor market power. In particular, policies that increase firm productivity are more effective in expanding wage employment and increasing workers' earnings than other interventions that improve workers' skills or decrease firm entry cost. Altogether, these results highlight the need to consider labor market power and its interaction with self-employment in the design of industrial development policy.

Related Literature This paper contributes to several strands of the literature. The first one studies the determinants of self-employment in low-income countries and its implications for labor market outcomes and flows. Gollin (2008) first observed that self-employment rates decline with income per capita and productivity across countries. Relatedly, Poschke (2019) shows that labor market search and matching frictions are a strong determinant of differences in unemployment and self-employment rates across countries.² Donovan, Lu, and Schoellman (2021) focus on transitions between employment states, finding that labor market flows are two to three times higher in the poorest countries than in the richest. Rud and Trapeznikova (2021) incorporate a search framework into a two-sector model of development to assess the relative importance of different obstacles to job creation and productivity. Herreño and Ocampo (2021) incorporate unemployment risk in a standard macro model of occupational choice and show that, in equilibrium, poor-unemployed individuals become self-employed for self-insurance purposes. These subsistence entrepreneurs coexist with high productivity individuals who positively select into self-employment, consistent with our conceptualization of manufacturing self-employment in theory and our empirical findings on transitions from wage work to self-employment. We contribute to this literature by providing a new framework that highlights the role of worker sorting and imperfect competition among employers in shaping equilibrium wage employment and self-employment rates and earnings.

Second, our paper belongs to the literature on monopsony power in the labor market. Recent evidence shows that employers in the U.S. may enjoy a certain degree of labor market

²Albrecht, Navarro, and Vroman (2009) and Narita (2020) also extend search and matching models to include features relevant for developing countries.

power, which the labor economics literature measures with the (inverse) elasticity of labor supply faced by the individual firm (Azar, Berry, and Marinescu, 2019; Berger, Herkenhoff, and Mongey, 2022).³ Some studies use employer concentration as proxy for labor market power showing that it correlates negatively with wages (Azar, Marinescu, and Steinbaum, 2022; Benmelech, Bergman, and Kim, 2022).⁴ Yet, using matched employer-employee data from Oregon, Bassier, Dube, and Naidu (2022) find no evidence that labor supply elasticities are decreasing with concentration. Similarly, using data on US manufacturers, Yeh, Macaluso, and Hershbein (2022) find stark differences between the evolution of aggregate labor markdown and the one of employer concentration over the last two decades. We provide a new micro-foundation for the positive slope of firm-level labor supply curves, due to worker sorting between wage work and self-employment. We show that, with sorting, employer concentration can have a non-monotonic relationship with labor market power, reconciling the different findings in the literature.

The literature on labor market power in low-income countries is much more limited. Muraidharan, Niehaus, and Sukhtankar (2017) show through experimental evidence that the labor market effects of public employment programs in rural India are consistent with the existence of monopsony power in private-sector employment. Using a combination of theory and data, Brooks, Kaboski, Li, and Qian (2021) find evidence of labor market power in the Chinese and Indian context, held by either individual firms or collusive groups of firms. Brooks, Kaboski, Kondo, Li, and Qian (2021) show that the expansion of the national highway system in India reduced labor markdowns significantly. Amodio and De Roux (2021) use plant and customs data from Colombia to generate plant-specific shocks to marginal revenue product, finding that workers produce around 40% more than their wage level.⁵ In Costa Rica, Alfaro-Ureña, Manelici, and Vasquez (2021) show that the expansions of multinational companies has a small impact on wages of workers in domestic firms, indicating a low level of labor market power. We contribute to this literature by using Peru as a case study to document several new stylized facts on employer concentration, its within-country variation, and its relationship with labor market outcomes. Most importantly, we provide a new general equilibrium framework to study the relationship between labor market power, self-employment, earnings, and productivity of both wage workers and self-employed workers. We show how to use this framework to gauge

³The literature has studied several mechanisms that can give rise to an upward-sloping labor supply curve such as search frictions, firm-specific amenities, and limited geographic mobility of workers (Manning, 2003; Card, Cardoso, Heining, and Kline, 2018). Most of the available evidence of finite labor supply elasticity in the U.S. comes from specific settings such as the labor market of schoolteachers (Ransom and Sims, 2010), military hospitals (Staiger, Spetz, and Phibbs, 2010), nurses (Matsudaira, 2014), university faculty (Goolsbee and Syverson, 2019), online labor markets (Dube, Jacobs, Naidu, and Suri, 2020), and the construction industry (Kroft, Luo, Mogstad, and Setzler, 2020). One exception is Yeh, Macaluso, and Hershbein (2022) who use rich administrative data for U.S. manufacturers and estimate an average plant-level markdown of 1.53.

⁴See also Rinz et al. (2018) and Schubert, Stansbury, and Taska (2021).

⁵Tortarolo and Zárate (2020) use the same data from Colombia to disentangle the extent of imperfect competition in product and labor markets. See also Pham (2019), MacKenzie (2019), and Felix (2021) on the interactions between trade and labor market distortions in China, India, and Brazil respectively.

the aggregate and distributional impact of different policies that promote entry and increase competition among employers in the labor market.

Finally, our work speaks to the large literature on informality in low-income countries, recently reviewed by [Ulyssea \(2020\)](#). Both [Dix-Carneiro and Kovak \(2019\)](#) and [Ponczek and Ulyssea \(2021\)](#) argue that informality acts as an “unemployment buffer” that reduces trade-induced adjustment costs in the labor market. Yet, [Dix-Carneiro, Goldberg, Meghir, and Ulyssea \(2021\)](#) show that this unemployment buffer role of informality does not translate into a “welfare buffer” meaning that in the event of a negative economic shock welfare declines by less when the informality rates are more modest.⁶ Our analysis adopts the notion of informal self-employment as a potential outside option for workers. We show that the existence of an informal self-employment sector mitigates the effects of labor market power on earnings and employment.

The remainder of the paper is organized as follows. Section 2 introduces the data sources and definitions. Section 3 presents the stylized 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 results of counterfactual policy analyses. Section 7 concludes.

2 Data and Definitions

The empirical analysis draws from two main datasets carrying information on firms and workers, respectively.

Firms 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 data includes standard balance-sheet information, such as revenues and labor and material expenditures, as well as information on the location of each plant. 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 in the wage employment sector. 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 Nuevos Soles (PEN) – equal to around 700K USD in 2010. Our final dataset counts 8,138 firm-year observations. As explained below, we validate the information in our EEA sample by contrasting it with the one retrievable from the Peruvian 2007 Economic

⁶However, [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.

Census on all establishments in the manufacturing sector. We derive summary statistics at the local labor market level and find them remarkably close to those from the 2007 EEA.

Workers 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, and is representative at the national and regional levels. The data provide information on demographics, education, and other individual characteristics of all household members. Respondents aged 14 or older fill out a specific module that includes questions on employment status, pay, and occupation. To be consistent with the firm-level data, we focus on the period 2004 to 2011 and restrict the sample to working-age individuals aged 25 to 65 who have largely 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 5 consecutive years. We use the 2007-2011 panel to document workers' transitions across employment states and sectors.

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 any monetary compensation. The information contained in ENAHO also allows identifying informal workers. A worker is labeled as informal if she (i) is a paid worker but reports not having health insurance,⁷ or (ii) is self-employed, but she is not registered or follows the procedures demanded by the national tax authority and has five or less employees.

Local Labor Markets A local labor market is defined as an industry within a geographical area. Industries are defined as a 2-digit CIIU Rev. 4 code. For our analysis, we primarily focus on the 23 industries in the manufacturing sector.⁸ Province boundaries inform the relevant geographical areas. Provinces are the administrative subdivisions of departments, which are the primary geopolitical sub-national units in Peru. 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 about 10 million. We thus follow [Piselli \(2013\)](#) and define local labor markets within the Lima province. The Survey of Transport, Labor, and Technology Use assembled by the Peruvian Studies Institute (IEP) allows for the identification of five distinct zones in which people do most of their activities: Lima Center, Lima North, Lima South, Lima East, and Lima West. These five areas, combined with the rest of the country's provinces, constitute the 199 geographical units that, together with the 23 manufacturing industries, define Peruvian manufacturing local labor markets.

⁷Employers in Peru are required by law to provide health insurance to their employees.

⁸Examples 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.

Measuring Concentration We measure concentration among medium and large formal employers in each local labor market in several ways. The baseline concentration measures are the Herfindahl indexes for payroll and employment (Berger, Herkenhoff, and Mongey, 2022). Let w_{ikt} and n_{ikt} denote wage bill and employment, respectively, of firm i in local labor market k in year t . Wage-bill and employment HHI are defined as

$$HHI_{kt}^{wn} = \sum_{i \in k} (s_{ikt}^{wn})^2 \quad \text{and} \quad HHI_{kt}^n = \sum_{i \in k} (s_{ikt}^n)^2 \quad (1)$$

where

$$s_{ikt}^{wn} = \frac{w_{ikt}n_{ikt}}{\sum_{i \in k} w_{ikt}n_{ikt}} \quad \text{and} \quad s_{ikt}^n = \frac{n_{ikt}}{\sum_{i \in k} n_{ikt}}. \quad (2)$$

Values of HHI_{kt}^{wn} (HHI_{kt}^n) closer to one indicate that a few firms account for a large share of wage bill (employment) in the local labor market. In the case of employment, the HHI has a simple interpretation: it measures the probability that two employees chosen at random in the same local labor market work for the same firm. Our third and final measure of employer concentration is the number of firms operating in a local labor market.

3 Stylized Facts

This section presents several new stylized facts on Peruvian manufacturing local labor markets. We show that wage employment is often concentrated among a few medium and large formal employers, and that high employer concentration correlates systematically with the labor market outcomes of both wage workers and self-employed workers. These facts are consistent with employers having some degree of labor market power and workers actively sorting between wage and self-employment.

3.1 Employer Concentration

Fact 1. *Peruvian manufacturing local labor markets exhibit high levels of concentration among medium and large formal employers.*

Table 1 reports summary statistics from firm-level data across Peruvian local labor markets, averaged over the period 2004-2011. For wage-bill and employment HHI, it shows both unweighted and weighted means, with weights equal to the market's share of nation-wide total payroll and employment.

Peruvian local labor markets exhibit high levels of employer concentration. The average local labor market counts about six medium to large firms. Within a local labor market, the probability of two randomly chosen employees working for the same firm is as high as 63%. Weighted

employer HHI measures are markedly lower than the unweighted ones, in the range of 30 to 35%, proving that highly concentrated markets account for a relatively small share of the total national payroll and employment. Yet, around 8% of formal manufacturing employment is in local labor markets with only one medium or large firm.⁹ These numbers are also larger than the corresponding ones for the United States. Berger et al. (2022) report 0.5 and 0.2 respectively as unweighted and payroll weighted average wage-bill HHI across local labor markets.

Location is more salient than industry in explaining the overall variation in employer concentration across local labor markets. Regressing the unweighted wage-bill HHI over location fixed effects, we find that location explains about 43% of the variation in employer concentration across markets. This fraction increases to 57% when we include 2-digit industry fixed effects as additional regressors.¹⁰ Figure 1 shows the geographical distribution of wage-bill HHI (left panel), and the number of firms (right panel) within the country averaged over the period 2004-2011. For each province, it reports the average concentration across local industries. As expected, the most densely populated areas of Lima and Callao are clearly distinguishable as the least concentrated in the country.

3.2 Self-employment and its Characteristics

Fact 2. *Self-employment is highly prevalent and mostly informal. Transitions between self-employment and wage employment are common, occur mostly within industry, and correlate systematically with earnings.*

Using worker-level data, the first two columns of Table 2 report national employment shares across different employment categories for all industries and the manufacturing sector only, averaged over the period 2004-2011. Across all industries, wage workers account for 45% of the workforce while the self-employed account for 46%.¹¹ Wage employment rates are relatively higher within the manufacturing sector, which sees the average share of wage workers at 55% and the average share of self-employed workers at 40%.

The nature of self-employment in low and middle-income countries differ substantially from that of high-income countries. First, essentially all of it is informal. Rows 3 to 5 of Table 2 show that, both across all industries and within manufacturing, over 90% of self-employment is informal. This is confirmed by Online Appendix Figures A.4 and A.5 that report the distributions of the share of self-employed workers and informal self-employed workers across

⁹Online Appendix Figure A.1 plots the distributions of the three concentration measures. For wage-bill and employment HHI, the figures report both distributions for the weighted and unweighted measures. Figure A.2 shows that these measures are strongly correlated among themselves. Figure A.3 plots these distributions as derived from the 2007 Peruvian Economic Census, showing similar patterns.

¹⁰When considering employment HHI and number of firms as concentration measures, the amount of variation explained by location only is 45% and 30%, up to 60% and 50%, respectively, when adding industry fixed effects.

¹¹The remaining 9% of the workforce are auxiliary family workers.

local labor markets. The distributions are virtually unchanged when focusing on informal self-employed workers only. While there is also informality in the wage employment sector (about half of wage workers are informal), it is less pervasive and declining over time.

Informality matters in that it makes self-employment flexible and easily accessible. As explained above, a self-employed worker is informal if she is not registered with the national tax authority. Wage workers are informal if they report not having mandated employer-sponsored health insurance. Avoiding tax and/or labor regulation decreases self-employment entry and starting costs. The very high levels of informality among the self-employed are consistent with it being an easily accessible outside option for workers when wage employment opportunities are scarce or wages are low.¹²

Also for this reason, self-employment is more prevalent in certain industries than in others. In labor-intensive industries, where the initial investment in physical capital is low, credit constraints bite less and the scope for informality is higher. These industries account for the largest share of Peruvian manufacturing GDP and are indeed those where self-employment rates are higher. For instance, the share of self-employed workers is higher than 50% in clothing and furniture manufacturing as well as other unclassified manufacturing which includes jewelry and trinkets, musical instruments, etc. In contrast, self-employment only accounts for around 10% of employment in more capital-intensive industries such as pharmaceuticals, metals, and oil and petroleum manufacturing.

These self-employed informal workers earn considerably less than their counterparts on wage employment. The third and fourth columns of Table 2 report average daily earnings for each worker category. Across all industries, the average wage worker earns about 50% more than the average self-employed worker. These gaps are even higher (~150%) when comparing formal wage workers with informal self-employed workers. Similar figures are observed within the manufacturing sector.

Self-employment is also a very dynamic sector. Exploiting the panel version of the worker-level data, we find that transition rates between self-employment and wage employment are high and occur mostly within an industry. Table 3 reports yearly transitions between unemployment, wage employment and self-employment. The top panel reports statistics in the full sample of workers, while the bottom panel focuses on the manufacturing workers. About 15% of all wage workers in a given year were self-employed in the previous year. Similarly, 18% of all self-employed workers in a given year were wage workers in the previous year. Transition

¹²Overall, informal workers account for 73% of the workforce. This figure is close to the 80% reported by INEI in 2007, whose calculations indicate that the informal sector accounts for about a fifth of aggregate GDP. Nearly 90% of agricultural GDP is informal, while for the manufacturing and service sectors, the number goes down to 14% and 32% respectively (INEI, 2014). 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% in the full sample and virtually zero in the sample of manufacturing workers, similar to reports from the International Labor Organization on Peru and other low and middle income countries.

rates are slightly lower in the manufacturing sector, with 8% (9%) of all wage (self-employed) workers in a given year being self-employed (wage) workers in the previous year. The vast majority of these transitions between employment states occurs within the same industry. Indeed, Online Appendix Figure A.6 shows that more than 80% of manufacturing workers are repeatedly observed in the same 2-digit industry.

Finally, transitions between self-employment and wage employment are systematically correlated with workers' earnings. The left panel of Figure 2 plots transition probabilities across rankings of workers' earnings within a local labor market and year. Transitions to wage work are systematically more frequent among low-earning self-employed workers than high-earning ones. The right panel of Figure 2 shows that transitions from wage work to self-employment exhibit the opposite pattern, with high-earning wage workers transitioning to self-employment at a higher rate than low-earning ones. We interpret this evidence as suggestive that the worker at the margin between sectors is more skilled than the average wage worker, but less skilled than the average self-employed worker.

3.3 Employer Concentration and Labor Market Outcomes

Fact 3. *Across local labor markets, employer concentration is strongly positively correlated with self-employment rates, and strongly negatively correlated with average earnings and education of both wage and self-employed workers.*

We last explore the relationship between concentration and the labor market outcomes of wage workers and self-employed workers. To begin, the left graph in Figure 3 plots the average (log) worker earnings against employer concentration across local labor markets.¹³ On average, higher values of wage-bill HHI are associated with lower earnings. Notably, the middle and right graphs show that the same pattern holds when looking separately at the earnings of wage workers and self-employed workers, respectively.

We investigate these patterns more systematically in a regression framework. We implement the following regression specification

$$y_{i(j,g)t} = \beta \ln HHI_{(j,g)t}^{wn} + \mathbf{X}'_{i(j,g)t} \theta + \gamma_j + \delta_t + u_{i(j,g)} \quad (3)$$

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 $\ln HHI_{(j,g)t}^{wn}$ 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 and δ_t stand for industry and year fixed effects respectively. The two sets of fixed effects net out time-invariant industry-level averages and aggregate yearly

¹³Online Appendix Figure A.7 plots the same figure using employment HHI and number of firms as alternative concentration measures.

trends. Thus, our coefficient of interest β is identified off variation in employer concentration across geographical areas, consistent with the previous findings and discussion of the salience of geography in explaining variation in employer concentration across local labor markets.

Table 4 reports the coefficient estimates that we obtain using OLS.¹⁴ We start by considering the log of earnings as the dependent variable, pooling together wage workers and self-employed workers. The specification in column 1 includes a dummy for female, age, age squared, and schooling as worker-level controls. The employer concentration coefficient is negative and significant at the 1% level, consistent with the left graph of Figure 3. In columns 2 and 3, we estimate the same regression specification separately for wage workers and for self-employed workers. Consistent with the center and right graphs of Figure 3, the estimated β is negative in both sub-samples, significant at the 1% level.

We then investigate the relationship between employer concentration and self-employment rates. In column 4, we estimate the regression specification in equation (3), where the dependent variable is a dummy equal to one if the worker is self-employed. The estimated coefficients show that employer concentration is strongly positively correlated with self-employment rates, with an elasticity of 0.05 significant at the 1% level.

In columns 5 and 6, the dependent variable is the worker's educational level. We find that where employer concentration is higher, the average education level of both worker categories is lower. The coefficient of interest is negative in both cases, although significant only at the 10% level when looking at wage workers.

We interpret these facts as follows. If firms compete strategically for workers, concentration among employers reduces job opportunities and wages. As a result, more workers decide to be self-employed. The previous findings on transitions across sectors suggest that workers at the margin between sectors are more skilled than the average wage worker, but less skilled than the average self-employed worker. When employer concentration increases, the most skilled wage workers transition to self-employment, where they are less skilled than those already active. This is consistent with Table 4 showing that earnings and average education level decrease in both sectors when employer concentration increases and the self-employment sector expands.

Electrification, Concentration, and Labor Market Outcomes One concern with the OLS evidence in Table 4 is that employer concentration is endogenous to supply and demand forces, such that the presence of a systematic relationship between the former and labor market outcomes does not in and of itself provide evidence on the role of labor market power. Finding an instrumental variable for changes in employer concentration is also challenging, since most shocks to the economic environment that affect firm entry likely affect other labor market out-

¹⁴Online Appendix Tables A.1 and A.2 report the coefficient estimates we obtain when using employment HHI and number of firms as alternative concentration measures, showing similar results.

comes through channels other than changes in employer concentration, thus invalidating the exclusion restriction.

To make progress, we consider the reduced-form effects of a nationwide electrification program and use the findings of Table 4 to frame the interpretation of the results. This evidence altogether informs the model that we develop in the next Section and its estimation.

In 1993, the Peruvian Ministry of Energy and Mining launched the Rural Electrification Program (*Programa de Electrificación Rural*, PER) with the objective of promoting social and economic development in rural areas (Dasso and Fernandez 2015). From 1994 to 2012, 628 electrification projects were concluded throughout rural Peru with a total cost of USD 657.5 million (Dasso, Fernandez, and Nopo 2015). Using available information on the program roll-out across the country up to 2010, we can gauge its impact on employer concentration and labor market outcomes, and check whether the effects are consistent with the reasoning outlined above.

For each local labor market and year, we count the cumulative number of PER projects completed up to that year, and evaluate its relationship with a number of market and individual-level outcomes. The first column of Table 5 shows that electrification increases the total sales of medium and large firms in the local labor market, with the correspondent coefficient estimate being significant at the 1% level. This is consistent with access to electricity increasing productivity (Abeberese, Ackah, and Asuming 2019). Column 2 shows that employer concentration decreases systematically with electrification, in line with increased productivity that induces firm entry. Column 3 to 8 show the impact of electrification on the same individual-level labor market outcomes analyzed in Table 4. Electrification increases both wages and earnings from self-employment, with the effect being larger for the latter. Yet, despite this differential increase in returns from self-employment, its prevalence decreases with electrification, and the average education level of self-employed workers increase.

We interpret the results in Table 5 as showing that electrification promotes firm entry, reduces concentration, and increases job opportunities and wages. The least skilled self-employed workers become wage workers, and the average skill level and productivity of self-employed workers increases as a result. Selection shapes the wage and earnings response to electrification. In the next Section, we formalize these intuitions and show that these results can be rationalized in a model where strategic interactions among employers co-exist with worker sorting between wage and self-employment.

4 Model

Motivated by the stylized facts, we develop a model where employer concentration as well as wage employment and self-employment rates and earnings are jointly determined in general

equilibrium. We use the model to reconcile the empirical evidence, provide new theoretical insights, and perform counterfactual policy analyses.

4.1 Environment

The economy consists of a continuum of local labor markets indexed by $k \in (0, 1)$. In each market there is a finite number of firms $M_k \in \mathbb{Z}_+$ that need labor to operate, and a number $L_k \in \mathbb{R}_+$ of workers with identical homothetic preferences.¹⁵ Individuals cannot move across local labor markets. They work and 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 income and firm profits.

Demand Each market-level good C_k comes in two varieties, $C_{F,k}$ and $C_{S,k}$, produced by employing firms and self-employed workers, respectively. As such, $j = \{F, S\}$ also identifies the wage employment and self-employment sector within local labor market k . The representative agent has nested CES preferences over final goods, which can be written as

$$C = \exp \left\{ \int_0^1 \theta_k \log C_k dk \right\}, \quad \text{where} \quad C_k = \left[\beta C_{F,k}^{1-\frac{1}{\rho}} + C_{S,k}^{1-\frac{1}{\rho}} \right]^{\frac{\rho}{\rho-1}}, \quad (4)$$

where $\rho > 1$ denotes the elasticity of substitution across varieties within a market. The parameter $\theta_k \in (0, 1)$ is the individuals' expenditure share on market k 's good, which is constant due to the Cobb-Douglas assumption on the outer nest and is such that $\int_0^1 \theta_k dk = 1$. Finally, $\beta > 0$ is a demand shifter for the variety F of good k . We choose the final good C as the numeraire, and denote as $Y = C$ the aggregate expenditure in the economy.

Using the properties of the CES demand aggregator, consumer expenditure on variety $j = \{F, S\}$ of good k is equal to

$$P_{j,k} C_{j,k} = s_{j,k} \theta_k Y, \quad \text{with} \quad s_{j,i} \equiv \gamma \left(\frac{P_{j,k}}{\gamma P_k} \right)^{1-\rho}, \quad (5)$$

where $\gamma \equiv (\mathbb{I}_{j=F}(\beta - 1) + 1)$ and $\mathbb{I}_{j=F}$ is an indicator function equal to one if $j = F$. The term $s_{j,k}$ is the (endogenous) expenditure share of variety j of good k , which depends on the variety price $P_{j,k}$, and the market-level price index $P_k = [\beta^\rho P_{F,k}^{1-\rho} + P_{S,k}^{1-\rho}]^{\frac{1}{1-\rho}}$.

Labor Supply Decisions Besides a consumption choice, workers face an occupational choice between working for a wage in sector F or being self-employed in sector S . Workers are heterogeneous in their ability or skills in each sector, defined by their endowment of efficiency

¹⁵We normalize the average market size to $\mathbb{E}(L) = 1$, as is required by a model with a continuum of labor markets.

units of labor. Abilities $\mathbf{a} \equiv (a_F, a_S) \in \mathbb{R}_+^2$ are i.i.d. draws from a market-specific joint distribution $G_k(a_F, a_S)$.

The ability vector defines the worker's absolute and comparative advantage of working in each sector. Let $W_{j,k}$ denote the earnings per efficiency unit in sector j of market k . All workers take these unit earnings as given and sort according to their comparative advantage. Specifically, a worker will sort into wage employment if and only if

$$a_F \hat{W}_k \geq a_S, \quad (6)$$

where $\hat{W}_k \equiv \frac{W_{F,k}}{W_{S,k}}$ is the earnings per efficiency unit in sector F relative to sector S . Worker sorting determines the labor supply in each sector, which is equal to

$$N^j(\hat{W}_k) = L_k \int_0^\infty \int_0^{a_j(\hat{W}_k \mathbb{1}_{j=F} + \hat{W}_k^{-1} \mathbb{1}_{j=S})} a_j g_k(a_j, a_{-j}) da_j da_{-j}, \quad \text{with } j = \{F, S\} \quad (7)$$

These labor supply equations implies that the supply of wage work is an increasing function of the relative unit wage \hat{W}_k ; vice-versa, the supply of self-employed workers is a decreasing function of \hat{W}_k . Intuitively, the higher the relative wage \hat{W}_k the more workers sort into wage employment. We denote the supply elasticities of labor as $\epsilon_j(\hat{W}_k) \equiv \frac{\partial \ln N_{j,k}}{\partial \ln \hat{W}_k}$ for $j = \{F, S\}$. It follows that $\epsilon_F(\hat{W}_k) > 0$, while $\epsilon_S(\hat{W}_k) < 0$.¹⁶

Technology In each market k , M_k operating firms produce variety F of the final good while self-employed workers produce variety S . The production technology in both sectors is linear in efficiency units of labor. The total supply of variety S of good k , which we denote as $Y_{S,k}$, is given by

$$Y_{S,k} = A_S N_{S,k}, \quad (8)$$

where A_S is a constant productivity term affecting production of self-employed workers everywhere in the economy.¹⁷

We assume that manufacturing firms are homogeneous within a local labor market.¹⁸ They

¹⁶The wage work supply elasticity, $\epsilon_F(\hat{W}_k)$, is defined as

$$\epsilon_F(\hat{W}_k) \equiv \frac{\partial \ln N_{F,k}}{\partial \ln \hat{W}_k} = \frac{\hat{W}_k \int_0^\infty a_F^2 g_k(a_F, a_F \hat{W}_k) da_F}{\int_0^\infty \int_0^{a_F \hat{W}_k} a_F g_k(a_F, a_S) da_F da_S} > 0.$$

A similar aggregate labor supply elasticity can be derived for the S sector.

¹⁷Note that average productivity in the self-employment sector will still vary across local labor markets due to endogenous selection of heterogeneous workers into this sector. Worker sorting affects production through the total supply of efficiency units of labor $N_{S,k}$.

¹⁸In Section 4.4 below we develop an extension of the model where we relax this assumption and allow each of the M_k firms to be heterogeneous in their productivity and produce a different variety of the $C_{F,k}$ good.

produce variety F of good k with common market-level productivity z_k , which we can also think of as a market-specific cost shifter. We let $n_{F,ik}$ and $y_{F,ik}$ denote the labor demand and output of firm i in market k , respectively. Because firms are homogeneous within a market, we can write aggregate labor demand as $N_{F,k} = \sum_{i=1}^{M_k} n_{F,ik}$. Total supply of variety F of the market good will then be equal to

$$Y_{F,k} = \sum_{i=1}^{M_k} y_{F,ik} = \sum_{i=1}^{M_k} z_k n_{F,ik} = z_k N_{F,k}. \quad (9)$$

Labor market clearing requires that, given the equilibrium relative wage \hat{W}_k , labor demand in the two sectors ($N_{S,k}$ and $N_{F,k}$) equals labor supply, i.e. $N^S(\hat{W}_k)$ and $N^F(\hat{W}_k)$ from equation (7).

Market Structure To keep the equilibrium tractable, we assume that both operating firms and self-employed workers are price-takers in the final good market, selling the final good at marginal cost and taking the prices $P_{F,k}$ and $P_{S,k}$ as given. They also take the economy-wide aggregates (Y, P) as given.

The problem of the firms is to choose labor demand to maximize total variable profits. We assume that employers fully internalize the effect of their labor demand on aggregate wages as in the standard Cournot formulation.

Given our assumptions on market structure, each employer i solves the following problem

$$\max_{n_{F,ik}} P_{F,k} z_k n_{F,ik} - W_{F,k} n_{F,ik} \quad (10)$$

where $W_{F,k} = \hat{W}(N_{F,k})W_{S,k}$ is the inverse function of equation (7), with $N_{F,k} = n_{F,ik} + \sum_{j \neq i} n_{F,jk}$. When solving the problem in equation (10), each employer takes as given the unit earnings in the self-employment sector $W_{S,k}$, and the total labor demand of its competitors $\sum_{j \neq i} n_{F,jk}$.

Because the problem in equation (10) is the same for all firms, we omit the subscript i in what follows. Solving the profit maximization problem yields the following first-order condition:

$$W_{F,k} = \frac{MRPL_{F,k}}{\psi_{F,k}}, \quad (11)$$

which says that employers set labor demand such that the unit wage $W_{F,k}$ is a markdown $\psi_{F,k}$ below the effective marginal revenue product of labor, $MRPL_{F,k} \equiv P_{F,k} z_k$.

This markdown measures the labor market power of employers and is given by

$$\psi_{F,k} \equiv \left(1 + \frac{1}{M_k \epsilon_F(\hat{W}_k)} \right), \quad (12)$$

where $\epsilon_F(\hat{W}_k)$ is the supply elasticity of wage work, as discussed above. The markdown is greater than one and it is a function of both employer concentration in the local labor market and the (endogenous) wage work supply elasticity. We will come back to the properties of the markdown in Section 4.3.

Market Equilibrium Given the set of employers and exogenous productivities $\{M_k, z_k\}_{k \in (0,1)}$, and aggregate variables (Y), we can characterize the equilibrium in each local labor market, which we will refer to as *market equilibrium*, by relative expenditure $\hat{s}_k \equiv \frac{s_{F,k}}{s_{S,k}}$ and relative wages $\hat{W}_k \equiv \frac{W_{F,k}}{W_{S,k}}$. The market equilibrium is unique and has the property that relative wages \hat{W}_k increase with the number of operating firms M_k and productivity z_k .¹⁹ We turn next to solving for the endogenous set of employers or entrants in each market.

Entry To enter a given market, firms have to pay a fixed cost f_k^e in units of the final good. Because firms are homogeneous, free entry in this economy implies that the firms will enter until it is profitable for the average firm to do so. Average profits can be written as

$$\pi_{F,k} = \pi(M_k; Y) = M_k^{-1} s_{F,k} \theta_k Y \left(1 - \frac{1}{\psi_k(M_k, \hat{W}_k)} \right) - f_k^e. \quad (13)$$

Because we are interested in contexts where the number of firms is finite, the entry condition can be restated as follows:

$$M_k : \quad \pi(M_k; Y) \geq 0 \quad \text{and} \quad \pi(M_k + 1; Y) < 0. \quad (14)$$

Ignoring the integer problem that arises when the number of employers is finite, equation (14) is equivalent to a zero profit condition determining the equilibrium number of employers in each market.

4.2 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 household expenditures and product markets clear. In particular,

$$Y = \int_k [E(W_{S,k}, W_{F,k}) + \Pi_{F,k} + P M_k f_k^e] dk, \quad (15)$$

¹⁹See Appendix B.1 for derivations.

where the three terms in 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 which are distributed to consumers, and (iii) entry of firms into the production stage. Product market clearing requires that total demand of the final good equals total value of production, i.e.

$$Y = PC, \tag{16}$$

where C is defined in (4). Our normalization assumption implies $P = 1$.

Conditional on the market equilibrium vector $\mathbf{K} = \left\{ \hat{s}_k, \hat{W}_k, M_k \right\}_{k \in (0,1)}$, the general equilibrium vector \mathbf{X} solves equations (15)-(16). Conditional on the general equilibrium vector \mathbf{X} , the solutions to the firms' problem and entry game in equations (11)-(14) yield the manufacturing equilibrium vector \mathbf{K} . The resulting fixed point $(\mathbf{X}; \mathbf{K})$ is the equilibrium in the economy.²⁰

4.3 Properties of the Model

Labor Market Power and Employer Concentration Equation (12) shows that, in our model with sorting of heterogeneous workers, the wage markdown and thus labor market power can have an ambiguous relationship with employer concentration. On the one hand, lowering concentration (higher M_k) decreases the wage markdown because each employer accounts for a lower employer share. On the other hand, an increase in M_k also affects the markdown indirectly by affecting the equilibrium relative wage and the labor supply elasticity. With higher entry the relative wage increases, but the supply elasticity of wage work may decrease due to worker sorting, mitigating the negative direct effect of firm entry on labor market power. While this mechanism is ultimately an empirical question, this insight can reconcile the mixed empirical evidence gathered in the literature on the relationship between employer concentration and labor market power.²¹

Labor Market Power and Earnings The effect of labor market power on labor productivity in each local labor market – as measured by output per worker – depends critically on the parameters of the workers' ability or skill distribution, and in particular on the correlation of workers' abilities across wage work and self-employment, and their relative dispersion. If the two abilities are strongly positively correlated and more dispersed in self-employment than in wage work, mean ability will increase in both sectors when the relative wage \hat{W}_k increases and the wage employment sector expands (Heckman and Sedlacek, 1985; Heckman and Honoré, 1990; Young, 2014; Alvarez-Cuadrado, Amodio, and Poschke, 2020). This is because the more skilled wage workers are better off as self-employed, implying negative selection in wage work and positive selection in self-employment. Labor market power decreases \hat{W}_k and limits the

²⁰See Appendix ?? for a detailed description of the solution algorithm.

²¹See for instance Bassier, Dube, and Naidu (2022); Berger, Herkenhoff, and Mongey (2022); Yeh, Macaluso, and Hershbein (2022).

size of the wage employment sector, pushing into self-employment individuals that are relatively more skilled than the inframarginal wage worker and less skilled than the inframarginal self-employed worker. Vice versa, if abilities are negatively correlated or if the correlation is positive but sufficiently low, the mean ability of wage workers will decrease and the one of self-employed workers will increase with \hat{W}_k . In Appendix B.3, we formalize these results in the case of abilities being drawn from a joint log-normal distribution.

Measuring Labor Market Power Labor market power in our model can be measured by the wage markdown below marginal revenue product of labor in equations (11)-(12). Rearranging the FOC in equation (11), it is straightforward to show that the markdown is also equal to

$$\psi_{F,k} = \frac{MRPL_{F,k}}{W_{F,k}} = \frac{P_{F,k}Y_{F,k}}{W_{F,k}N_{F,k}} \equiv (s_k^N)^{-1}, \quad (17)$$

namely, to the inverse of the share of labor expenditure of firms in market k over their total revenues, which we denote as $s_{F,k}^N$. The right-hand-side of equation (17) can be easily quantified given our firm-level data.²²

Online Appendix Table A.3 reports the mean and standard deviation of wage markdowns across two-digit industries, measured using the formula in equation (17). Figure 4 shows simple correlations across local labor markets between the average wage markdown and employer concentration as measured by wage-bill HHI (left panel), and between the average wage markdown and the share of self-employed workers (right panel). In line with the model's predictions, the evidence shows that the relationship between employer concentration and wage markdown is non-monotonic. It also shows a clear positive relationship between wage markdowns and the self-employment rates. We are going to leverage these correlations as untargeted moments for the model's estimation.

4.4 Model Extensions

Firm Heterogeneity We can allow each of the M_k firms to produce a different variety $c_{iF,k}$ of the $C_{F,k}$ good such that

$$C_{F,k} = \left(\sum_{i=1}^{M_k} c_{iF,k}^{1-\frac{1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} \quad (18)$$

²²The simplicity of the expression stems from our simplifying assumptions of (i) perfect competition in the output market, which implies that firms' markups $\nu_{F,k}$ are everywhere equal to one and (ii) linear technology in labor, which implies a unity output elasticity with respect to labor, i.e. $\theta_{F,k}^N \equiv \frac{\partial \ln Y_{F,k}}{\partial \ln N_{F,k}} = 1$. Absent these restrictions on market structure in the output market and technology, the formula would generalize to $\psi_{F,k} = \frac{\theta_{F,k}^N}{s_{F,k}^N \nu_{F,k}}$, which is reminiscent of the expression used in extensive empirical work estimating market power in input markets (e.g., Morlacco 2020; Yeh, Macaluso, and Hershbein 2022).

and

$$P_{F,k} = \left(\sum_{i=1}^{M_k} p_{iF,k}^{1-\sigma} \right)^{\frac{1}{1-\sigma}} \quad (19)$$

where $\sigma > 1$ denotes the elasticity of substitution across the varieties produced by the M_k firms and $p_{iF,k}$ is the firm-level output price.

Let $z_{i,k}$ denote the productivity of firm i in market k , each one adopting a linear production technology, i.e. $y_{iF,k} = z_{i,k}n_{iF,k}$. The aggregate supply of variety F is now equal to

$$Y_{F,k} = \tilde{A}_{F,k} N_{F,k} \quad (20)$$

where $\tilde{A}_{F,k}$ is a market-specific productivity index which weights firm-level productivities with their output shares $s_{i,k}$, and $N_{F,k}$ is the aggregate labor demand in the wage employment sector.

Solving the firm profit maximization problem yields the same first-order condition as in equation 11, but with

$$MRPL_{F,k} \equiv p_{iF,k} z_{i,k} \left(\frac{\epsilon_{iF,k} - 1}{\epsilon_{iF,k}} \right) \quad (21)$$

where $\epsilon_{iF,k} = \left[\frac{1}{\sigma}(1 - s_{i,k}) + \frac{1}{\rho}s_{i,k} \right]^{-1}$ is the price elasticity of demand.

The wage markdown is given by

$$\psi_{F,k} \equiv \left(1 + \frac{s_{i,k}^N}{\epsilon_F(\hat{W}_k)} \right), \quad (22)$$

where $s_{i,k}^N$ is employment share of firm i in local labor market k . Notice that with firm homogeneity we have $s_{iF,k}^N = 1/M_k$ and the markdown expression is equivalent to the one in equation 12.

We can solve for the equilibrium number of M_k operating firms in this model with an iterative procedure which considers a sequential entry game where firms with higher productivity move first. The entry game has a unique cutoff equilibrium whereby only firms with productivity above some cutoff enter the market and make non-negative profits while for any additional entrant profits upon entry would be negative.

5 Model Estimation

This section details the procedure we implement to estimate the theoretical model using the Peruvian data. First, we describe how we parameterize the model to make it tractable for empirical analysis. Second, we illustrate how we leverage the data to pin down the parameters of the joint ability distribution. Third, we discuss how we use the structure of the model to

jointly identify and estimate all other remaining parameters.

5.1 Parameterization

The scope and variation of labor market power across local labor markets depends on market-level measures of productivity z_k , fixed costs f_k^e , and the shape of the joint ability distribution G_k .

To match the sales and payroll distribution of firms across local labor markets, we assume that market productivities z_k are drawn from a log-normal distribution with parameters μ_Z and σ_Z , i.e.

$$z \sim \log \mathcal{N}(\mu_z, \sigma_z). \quad (23)$$

We also assume that fixed entry costs in market k are a linear function of the number of entrants, namely

$$f_k^e = \alpha_0 + \alpha_1 M_k, \quad (24)$$

where $\alpha_0, \alpha_1 > 0$ are constant and common across local labor markets. We find that a similar parameterization of the fixed costs across markets allows us to capture the distribution of entrants with a high degree of accuracy.

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

$$\log \mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}), \quad \text{where } \boldsymbol{\mu}_k = \begin{pmatrix} \mu_{F,k} \\ \mu_{S,k} \end{pmatrix}, \quad \text{and } \boldsymbol{\Sigma}_k = \begin{pmatrix} \sigma_{F,k}^2 & \rho_k \sigma_{F,k} \sigma_{S,k} \\ \rho_k \sigma_{F,k} \sigma_{S,k} & \sigma_{S,k}^2 \end{pmatrix}, \quad (25)$$

with $\boldsymbol{\mu}_k$ determining mean absolute advantage in wage work and self-employment, ρ_k capturing the correlation between the two, and $\boldsymbol{\Sigma}_k$ altogether governing the extent of comparative advantage.

We restrict both the variance-covariance matrix $\boldsymbol{\Sigma}_k$ as well as the relative mean absolute advantage in the two employment sectors $\Delta\mu_k \equiv \mu_{F,k} - \mu_{S,k}$ to be the same across local labor markets, i.e. $\boldsymbol{\Sigma}_k = \boldsymbol{\Sigma}$ and $\Delta\mu_k = \Delta\mu, \forall k$. As we explain below, this assumption, together with the assumption of log normality, enables us to use direct inference methods for identification of the parameters while greatly simplifying the estimation procedure.

Yet, we let mean absolute advantages $\mu_{F,k}$ and $\mu_{S,k}$ to be specific to each local labor market k , capturing heterogeneity in mean ability or skills across locations. Specifically, we assume that $\mu_{F,k}$ is itself drawn from a log-normal distribution with common parameters μ_{μ_F} and σ_{μ_F} as given by

$$\mu_F \sim \log \mathcal{N}(\mu_{\mu_F}, \sigma_{\mu_F}), \quad (26)$$

and $\mu_{S,k}$ can be recovered as $\mu_{S,k} = \mu_{F,k} - \Delta\mu$.

With these parametric assumptions, we estimate the model in three steps. In the first step, we calibrate market-level Cobb-Douglas expenditure shares θ_k from the data as equal to the income share of each local labor market. We also calibrate the share of population living in each market L_k from our household surveys. We report the histogram of the resulting θ_k and L_k in Online Appendix Figure A.8. 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, explained in Section 5.2, we use the observed employment shares, mean earnings and their variances to recover the (common) parameters of the variance matrix of the workers' ability distributions $\{\sigma_F, \sigma_S, \varrho\}$, as well as the relative mean absolute advantage in the two employment sectors $\Delta\mu$.

In the third step, described in Section 5.3, we implement a simulated method of moments (SMM) procedure to estimate the remaining parameters $(\mu_z, \sigma_z, \alpha_0, \alpha_1, \mu_{\mu_F}, \sigma_{\mu_F}, \beta, \rho, A_S)$.

5.2 Inferring Ability Distribution Moments

We estimate the common parameter vector governing the ability distribution across local labor markets by leveraging the properties of the joint log-normal distribution in equation (25). Following Heckman and Sedlacek (1985), let $\sigma^* = (\sigma_F^2 + \sigma_S^2 - 2\varrho\sigma_F\sigma_S)^{-\frac{1}{2}}$. The observed probability that a given worker in local labor market k operates in sector F is equal to

$$\Phi\left(\frac{\ln \hat{W}_k + \Delta\mu}{\sigma^*}\right) \equiv \Phi(c_{F,k}) \quad (27)$$

where $\hat{W}_k \equiv W_{F,k}/W_{S,k}$ is the relative unit wage and $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. The probability of operating in sector S is instead equal to $\Phi(c_{S,k}) = \Phi(-c_{F,k})$.

The observed mean of log earnings in each sector in market k can be written as

$$E(\ln a_{j,k}W_{j,k} | a_{j,k}W_{j,k} \geq a_{-j,k}W_{-j,k}) = \ln W_{j,k} + \mu_j + \left(\frac{\sigma_j^2 - \varrho\sigma_j\sigma_{-j}}{\sigma^*}\right) \lambda(c_{j,k}), \quad (28)$$

for $j = \{F, S\}$, with $\lambda(c)$ being the inverse Mills ratio.²³

The observed variance of log earnings is equal to

$$\text{Var}(\ln a_{j,k}W_{j,k} | a_{j,k}W_{j,k} \geq a_{-j,k}W_{-j,k}) = \sigma_j^2 + \left(\frac{\sigma_j^2 - \varrho\sigma_j\sigma_{-j}}{\sigma^*}\right)^2 [\lambda(c_{j,k})c_{j,k} - \lambda^2(c_{j,k})]. \quad (29)$$

²³Note that $-j$ refers to the sector other than j . For example, when $j = F$ then $-j = S$.

We can recover $c_{F,k}, c_{S,k}$ from (27) by inverting the observed employment shares, from which we can also get $\lambda(c_{F,k}), \lambda(c_{S,k})$. With K markets, the system of variances in (29) consists of $2 \times K$ equations in the vector of unknowns $\Theta = (\sigma_F, \sigma_S, \rho)$. We consider a Minimum Distance Estimation (MDE) procedure to recover Θ with non-negativity constraints on the parameters.²⁴

Given an estimate for the vector Θ , we proceed with the estimation of the relative mean absolute advantage parameter $\Delta\mu$. From equation (28) it follows that the difference across sectors in the mean of log earnings is equal to:

$$\begin{aligned} E(\ln a_{F,k} W_{F,k} | a_{F,k} W_k \geq a_{S,k}) - E(\ln a_{S,k} W_{S,k} | a_{F,k} W_k < a_{S,k}) = \\ \ln \hat{W}_k + \Delta\mu + \left(\frac{\sigma_F^2 - \rho \sigma_F \sigma_S}{\sigma^*} \right) \lambda(c_{F,k}) - \left(\frac{\sigma_S^2 - \rho \sigma_F \sigma_S}{\sigma^*} \right) \lambda(c_{S,k}), \end{aligned} \quad (31)$$

where the term in the left-hand side is observed, and the last two terms in the right-hand side are known given the estimated $\Theta = (\sigma_F, \sigma_S, \rho)$.

Estimating $\Delta\mu$ from (31) using OLS is challenging because the relative unit earnings $\ln \hat{W}_k$ is unobserved and related to $\Delta\mu$ via the joint ability distribution. However, equation (27) shows that, if $\Delta\mu$ and Σ are common across local labor markets, the observed share of self-employed (or wage) workers in each market is a monotonic function of \hat{W}_k only. We can therefore adopt a control function approach and approximate \hat{W}_k using a flexible polynomial in the share of self-employed workers in k . Alternatively, we can exploit the result in (12) showing that the markdown in k is a function of the number of firms M_k and the wage work supply elasticity, which is also a function of \hat{W}_k only.²⁵ It follows that \hat{W}_k can be also approximated using a flexible polynomial in the observed number of firms M_k and the markdown ψ_k , whose estimates we discussed in Section 4.3.

The top panel of Table 6 reports the estimates of the ability distribution parameters. We find that the two abilities are strongly highly correlated, with $\hat{\rho} = 0.98$. Also, ability for self-employment is more dispersed than ability for wage work, i.e. $\hat{\sigma}_S > \hat{\sigma}_F$. Our preferred estimate of $\Delta\mu$ is equal to 1.02, implying that the mean absolute advantage of workers is higher on average in wage employment than in self-employment.²⁶

²⁴Let $h_{j,k}(\Theta)$ denote the RHS of (29). The constrained minimum distance estimator of Θ is given by:

$$\begin{aligned} \hat{\Theta} : \arg \min f_k(\Theta)' \mathbb{W} f_k(\Theta), \\ \text{subject to } \sigma_j > 0, \text{ and } \rho \in (-1, 1), \quad j = \{F, S\}, \end{aligned} \quad (30)$$

where $f_k(\Theta) \equiv \text{Var}(\ln a_{j,k} W_{j,k} | a_{j,k} W_{j,k} \geq a_{-j,k} W_{-j,k}) - h_{j,k}(\Theta)$ 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.

²⁵Under the assumption of common $\Delta\mu$ and Σ , the $\epsilon_F(\hat{W}_k)$ function does not vary across markets. See Appendix B.4 for a formal proof.

²⁶Online Appendix Table A.4 shows both the estimates obtained by implementing OLS over the system in (29) and the constrained MDE estimates. Notice that the implied value of ρ is below -1 when using OLS estimates, which motivates the use of a constrained estimator. The bottom panel presents the OLS estimates of the $\Delta\mu$

As explained in Section 4.3, these parameter values imply that mean ability increases in both sectors when the relative wage \hat{W}_k increases and the wage employment sector expands, because the worker at the margin between sectors is of more skilled than the inframarginal wage worker and less skilled than the inframarginal self-employed. This is consistent with the evidence presented in Section 3. On the one hand, Figure 2 shows that transitions to wage employment are systematically more common among low-earning self-employed, while transitions to self-employment are more prevalent among high-earning wage workers. On the other hand, Table 4 shows that the average education levels of both self-employed and wage workers are higher in less concentrated local labor markets. These parameter values also indicate that at least part of the negative correlation between concentration and earnings in both sectors 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 Simulated Method of Moments (SMM) procedure we employ to estimate the remaining parameter vector $\Phi = (\mu_z, \sigma_z, \alpha_0, \alpha_1, \mu_{\mu_F}, \sigma_{\mu_F}, \beta, \rho, A_S)$.

We target 12 empirical moments, which correspond to local labor markets outcomes. In line with our motivating evidence in Section 3, we target moments that are informative about cross-sectional features of concentration, self-employment, and their relationships. As it is well-known with this type of estimation, variation in any parameter tends to affect all moments simultaneously. Nonetheless, some parameters are susceptible to specific moments (see Andrews, Gentzkow, and Shapiro 2017). We provide here a discussion of the leading forces ensuring identification.

First, we target the cross-sectional mean and standard deviation of the log of the number of firms across local labor markets, as well as the share of local labor markets with only one employer and the associated percentage of total wage employment (Table 1). These moments are critical for identifying the parameters of productivity and fixed cost distribution. Intuitively, the fixed cost distribution parameters (α_0, α_1) determine the average number of firms in each market and the share of markets with only one firm. On the other hand, the standard deviation of the number of firms and the employment share of markets with only one firm are mostly informative of μ_z and σ_z .

We then target the relative self-employment employment shares, earnings, and abilities across markets (Table 2). The corresponding moments include the cross-sectional mean of the share of wage workers inside a local labor market, the relative earnings of wage workers, and their relative ability as measured by years of education. Given the common parameters of the work-

parameter, with standard errors obtained by block bootstrapping over the entire procedure. It shows results when we use the two different polynomials mentioned above to control for the unobserved \hat{W}_k .

ers' ability distribution $(\Sigma, \Delta\mu)$, the number of entrants M_k , and market-level productivity z , these moments can be written as functions of the model parameters β, ρ and A_s .²⁷

Finally, we target the correlations between the number of firms in the wage employment sector and several market outcomes, including the average earnings and abilities of wage and self-employed workers and the average share of wage workers within a labor market (Table 4). The sensitivity of earnings and abilities of wage and self-employed workers to the number of firms depend on the workers' absolute advantage in the two sectors is informative of the absolute advantage of workers, and thus of $\mu_{\mu_F}, \sigma_{\mu_F}$.

We solve for the set of parameters Φ that solves $\hat{\Phi} = \arg \min_{\Phi} \hat{\mathbf{f}}(\Phi)' \mathbb{W} \hat{\mathbf{f}}(\Phi)$, 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 $\hat{\mathbf{m}}$. The estimated parameters are reported in Table 6.

5.4 Model Fit

Table 7 reports the model-based values of the 12 moments used in estimation and compares them with their empirical counterparts. The table further reports the OLS standard errors for the regression coefficients, which the model can reproduce despite not being targeted in estimation. 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. Overall, despite its parsimony, the model provides a good fit to the data.

The model matches the distribution of the number of firms across local labor markets accurately, with 4 medium-to-large firms in the average labor market, a large share of monopsonistic markets ($\sim 30\%$), and sizeable cross-sectional variation. The total employment share of monopsonistic markets is 7% in both the model and the data.

The model also provides a reasonable, although not perfect, fit of the average incidence of self-employment across markets, which sees around 48% vs. 30% of workers as self-employed in the data and the model, respectively. The earnings of wage workers are about 70% as high as self-employed workers' earnings, both in the model and in the data. The relative average ability of wage workers is higher than that of self-employed workers. The relative ability is higher in the model than in the data; we can justify this discrepancy by noticing that in the data, we only observe a noisy proxy of workers' abilities, that is, years of education.

Last but not least, the model can replicate with a high degree of accuracy the cross-sectional correlation of earnings, ability, and self-employment rates with the number of firms. Except for average ability in the two sectors, all the coefficients are estimated economically and statis-

²⁷See equation (7) in the Appendix.

tically significant.

Untargeted Moments We conclude by exploring the fit of the model for other untargeted moments. We focus on the wage markdown and its correlation with number of firms and self-employment share. Figure 5 plots the model implied distribution of markdowns. Close to 35% of markets have wages as high as 98-100% of marginal revenue product of labor. These relatively competitive local labor markets account for the vast majority of wage employment in the country, in line with the descriptives in Table 1. Yet, at the other end of the spectrum, about a third of local labor markets have relatively high wage markdowns, with wages approximately 20% *below* the marginal revenue product of labor, although these markets only account for a small share of total payroll wages. Taken together, the average unweighted markdown is estimated around 1.07, which decreases to 1.01 once we account for the labor market's share of total payroll.

We compare these estimates of the wage markdown at the market level ψ_k with their data counterpart, obtained as described in Section 4.3. We do so by implementing the following regression specification

$$\ln \psi_k = \alpha + \beta \ln M_k + \gamma \ln \theta_k + u_k \quad (32)$$

with M_k being the number of firms and θ_k being the labor market payroll share. The β estimate and associated standard error are -0.02 and 0.005 in the data and -0.04 and 0.006 in the model respectively.²⁸ When replacing the number of firms in equation (32) with the share of self-employed workers, we find its coefficient estimate and associated standard error to be 0.92 and 0.073 in the data and 0.58 and 0.09 in the model respectively.

In the data, conditional on payroll share, wage markdowns and labor market power are lower in markets with more employers. Also, self-employment shares correlate positively with the wage markdown measures. These patterns hold within the estimated model, although wage markdowns were not an explicit target for estimation.

6 Counterfactual Policy Analysis

In this section, we provide two sets of counterfactual analyses that we deem interesting for policy making. First, we assess quantitatively the role of labor market power for rent sharing between firms and workers and the salience of employer concentration and worker sorting into wage or self-employment as defining mechanisms. Second, we consider three sets of policies that promote industrialization and expand the wage employment sector. We then quantify their aggregate and distributional impact on labor market power and labor market outcomes.

²⁸Because the markdown estimates in the data are obtained under the simplifying assumption of a linear technology, in the data-based regressions we also control for industry fixed effects that partially absorb differences in markdown estimates related to industry-specific technology.

6.1 Rent Sharing

Labor market power shapes the pass-through of firm productivity or profitability shocks to wages. In our model, from equation (11), it follows that the wage elasticity to the marginal revenue product of labor is equal to

$$\varepsilon_W = 1 - \varepsilon_\psi \quad (33)$$

where ε_ψ is the markdown elasticity, and ε_W is the wage elasticity, both relative to changes in an aggregate productivity or profitability shock. Therefore, the percentage change in wages following a productivity shock crucially depends on how that shock affects labor market power. In turn, as discussed in Section 4.3, the wage markdown depends on employer concentration and the extent to which firms interact strategically as well as on worker sorting and how it shapes the supply elasticity of wage work.

To quantify the salience of these mechanisms, we simulate within the estimated model the impact of an aggregate productivity shock Δz , and a nation-wide wage subsidy Δt_F .²⁹ Let $\bar{\psi}_0$ be the (average, unweighted) markdown in the baseline economy. We consider aggregate shocks to productivity and profitability of different sizes and compute the new model equilibrium and the corresponding average markdown $\bar{\psi}_1$. Furthermore, for each level of the aggregate shock, we also compute $\bar{\psi}_{01}$ as the value of the average markdown we would observe if labor supply in the wage and self-employment sector did not change compared to baseline. This allows us to decompose the overall impact on labor market power $\ln(\bar{\psi}_1/\bar{\psi}_0)$ into the two separate effects of (i) changes in the number of firms and employer concentration $\ln(\bar{\psi}_{01} - \bar{\psi}_0)$ and (ii) changes in worker sorting and wage work supply elasticity $\ln(\bar{\psi}_1 - \bar{\psi}_{01})$.

Productivity Shock The left panel of Figure 6 shows the average percentage change in markdowns following a productivity shock of different sizes. The solid line shows that, following a positive (negative) shock to productivity, the average markdown decreases (increases). Therefore, changes in labor market power *amplify* the transmission of productivity shocks to wages. For instance, when productivity increases by 50%, the average markdown decreases by 0.5%. This signifies a markdown elasticity of about -0.01, resulting in a wage elasticity (or pass-through) of 1.01. When productivity increases by 50%, wages increase by almost 51% due to the simultaneous decrease in markdown. Conversely, if productivity decreases by 50%, wages would decrease by about 51% due to the markdown increasing.

The dashed line shows the counterfactual change in markdown in the intermediate equilibrium where labor supply is kept fixed and only strategic interactions among firms matter. Comparing the solid and dashed lines, we can assess the role of changes in employer concentration and worker sorting in shaping the overall effect of productivity shocks on labor market power.

²⁹The aggregate shocks are defined such that market-level productivity in the counterfactual economy is equal to $z_{k,1} = z_{k,0}(1 + \Delta z)$, and the effective unit cost of labor is equal to $(1 + \Delta t_F)W_{F,k}$, with $\Delta t_F < 0$.

Following a positive (negative) productivity shock, the number of operating firms increases (decreases), which has a direct negative (positive) effect on labor market power. Worker sorting *mitigates* these effects as the supply elasticity of wage work changes in the opposite direction, as shown in the right panel of Figure 6. As the wage employment sector expands, the supply elasticity of wage work decreases. This is because, at the new equilibrium, the worker at the margin between wage work and self-employment belongs to a different portion of the joint ability distribution and is now less sensitive to wage changes. The opposite happens in response to a negative productivity shock.

Wage Subsidy We now turn in Figure 7 to consider shocks to firm profitability due to a nationwide decrease of labor costs as driven by the introduction of a wage subsidy. As in the previous case, we consider shocks of different sizes.

We document that wage subsidies have qualitatively similar effects to productivity shocks. Unlike productivity shocks, however, the impact of wage subsidies is non-linear and tends to decrease in magnitude with the size of the subsidy. A wage subsidy of 20% results in a decrease in average markdowns of 3%. This signifies a markdown elasticity of about -0.06, resulting in a wage elasticity (or pass-through) of 1.06. When marginal revenue product of labor increases by 20%, wages increase by 26% due to the simultaneous decrease in markdowns. Notably, the pass-through of wage subsidies to wages is higher for lower level of subsidies. Similarly to the case of productivity shocks, effects would be larger in the absence of worker sorting, which induces changes in the labor supply elasticity, and markdowns thereof, as shown in the right panel figure.

Our findings indicate that changes in labor market power increase the rent-sharing elasticity and thus amplify the transmission of productivity and profitability shocks to wages. At the same time, changes in the supply elasticity of wage work mitigate the labor market power response to these shocks thus making the rent-sharing elasticity lower than what it would be in the absence of worker sorting.

6.2 Industrial Policy

In the model, the scope and variation of labor market power across local labor markets is shaped by (i) fixed costs of entry, (ii) productivity or cost shifters at the local labor market level, and (iii) the workers' joint ability or skill distribution. These are all possible targets for policies that promote industrialization and expand the wage employment sector. With the estimated model, we can simulate and compare their impact and evaluate the role of labor market power in determining their overall impact on the economy.

Industrial development policies of this kind are routinely implemented in low-income countries. Examples of policies that reduce the fixed costs of entry include interventions aimed at

decreasing the monetary and non-monetary cost to set up a formal manufacturing firm, e.g., one-stop shops (Branstetter, Lima, Taylor, and Venâncio, 2014), or at reducing financial constraints of entrepreneurs (Buera, Kaboski, and Shin, 2011). Similarly, policies that increase the productivity of all firms in a market include efforts to improve market access or more stable access to electricity and other complementary inputs (Allen, 2014; Startz, 2016; Brooks, Kaboski, Kondo, Li, and Qian, 2021). Lastly, workers' ability is the target of workers' training programs aimed to increase workers' productivity and employability in the labor market, such as vocational training programs and on-the-job training (Alfonsi, Bandiera, Bassi, Burgess, Rasul, Sulaiman, and Vitali, 2020).

For each one of these three possible targets, we consider two different policy experiments. In the first one, we calibrate each intervention in order to achieve a reduction of 50% in the economy-wide average level of employer concentration. This allows us to study whether conditional on promoting firm entry to the same extent intervening with policy on entry cost, firm productivity, or workers' skills yields different aggregate and distributional outcomes.

The second thought experiment eliminates differences across local labor markets in either fixed costs, productivity, or workers' ability. To do so, we set each of them, in turn, equal to a least frictional benchmark corresponding to either the minimum (fixed cost) or maximum (productivity of ability) observed in the baseline estimated model.³⁰ This allows us to quantify the separate contribution of each feature in determining labor market power in the aggregate and across local labor markets.

Promoting Firm Entry Top panel of Table 8 and Figure 8 illustrate the results from our first counterfactual policy experiment. The Table shows the impact of the different policy interventions on aggregate outcomes, which include: average HHI across local labor markets, average markdown, the average share of wage and self-employed workers, average earnings of wage and self-employed workers, and aggregate income.

By construction, the three interventions promote firm entry and decrease employer concentration to the same extent, as shown by the first row of Table 8 and the top left graph in Figure 8. The change in average markdown is also very similar across the three scenarios. Yet, the top right graph in Figure 8 show that the change in labor market power is not uniform across local labor markets, and different depending on whether the policy targets fixed cost, productivity, or workers' ability. Markdown reductions are generally larger in those local labor markets where markdowns are larger to begin with. Promoting firm entry by reducing fixed cost or improving workers' skills yield a larger reduction in markdowns in least competitive local labor markets compared to increasing firm productivity.

³⁰In the case of fixed cost of entry, in the counterfactual economy we assume that fixed costs do not depend on the number of entrants in the market.

The bottom left graph in Figure 8 shows that the wage employment sector expands in all three cases as the share of wage employment increases and the share of self-employment decreases compared to baseline, the more so when firm productivity is targeted.

Table 8 also shows that the three policies have very different impacts in terms of total earnings and income (rows 3-5). For example, a policy that reduces fixed entry costs has a negligible effect on aggregate income and earnings of both groups of workers. On the contrary, policies that aim to improve firms' or workers' productivity substantially impact both aggregate income and the wages of both groups of workers.

Overall, while all the policies successfully promote entry and decrease concentration in the wage employment sector to the same extent, they have potentially very different effects on labor market power, the relative prevalence of wage employment vs. self-employment, earnings and aggregate income.

Eliminating Heterogeneity The bottom panel of Table 8 and Figure 9 show results from our second thought experiment, which eliminates heterogeneity across markets in either fixed costs, firm productivity, or workers' skills. The Table and Figure are constructed symmetrically to the previous case.

Table 8 shows that differences across markets in labor market power and labor market outcomes, more generally, are not related to heterogeneity in fixed costs across markets. This is consistent with the patterns in the top graphs of Figure 9. This has to do, in part, with the fact that heterogeneity in fixed costs in our model stems from heterogeneity in the number of firms; evidence shows that results would only marginally change if fixed costs were unrelated to the number of firms.

Policies that eliminate heterogeneity in firm productivity or workers' ability have sizable effects. They have overall similar impacts on labor market outcomes, decreasing concentration and labor market power to a larger extent in less competitive local labor markets. Policies that increase workers' skills have a small effect on the relative prevalence of wage employment vs. self-employment, but increase earnings for both groups of workers. Raising firm productivity has a larger positive (negative) effect on the share of wage employment (self-employment) and on earnings in both sectors. The bottom graphs of Figure 9 show that these effects are more pronounced in markets where the share of wage employment (self-employment) at baseline is smaller (larger).

7 Conclusions

The presence of a competitive labor market is crucial for the benefits of economic growth to be shared by workers as much as by firms and capital owners, lifting the poorest segment of

the population out of poverty. This paper studies how self-employment impacts the scope and implications of labor market power in low-income countries.

Using data on Peru, we document a systematic relationship between employer concentration, self-employment rates, and labor market outcomes across manufacturing local labor markets. We interpret the facts using a general equilibrium model of the Peruvian economy where strategic interactions among employers and sorting of heterogeneous workers between wage work and self-employment jointly determine the scope of labor market power.

We estimate the model and use it to study the aggregate and distributional effects of productivity and profitability shocks as well as different industrial policy instruments. We find that changes in concentration increase the rent-sharing elasticity and thus magnify the pass-through of productivity and profitability shocks to wages. Yet, worker sorting mitigates the labor market power response to these shocks. We also find that industrial policies targeting fixed costs, firm productivity, or workers' skills may have very different aggregate and distributional effects as changes in firm entry and worker sorting simultaneously affect the size of wage markdowns and the extent of labor market power. Counterfactual policy analyses reveal that policies that increase firm productivity are more effective in expanding wage employment and increasing workers' earnings than other interventions that improve workers' skills or decrease firm entry cost.

Our analysis and results highlight the role of self-employment opportunities in shaping labor market power, and the importance of taking both into consideration in the design of industrial development policies. More research is needed to understand the extent to which labor market power interacts with other frictions that may affect worker sorting patterns within manufacturing as well as transitions out of agriculture along the structural transformation path.

References

- Abeberese, A. B., C. G. Ackah, and P. O. Asuming (2019). Productivity Losses and Firm Responses to Electricity Shortages: Evidence from Ghana. *The World Bank Economic Review* 35(1), 1–18.
- Albrecht, J., L. Navarro, and S. Vroman (2009). The effects of labour market policies in an economy with an informal sector*. *The Economic Journal* 119(539), 1105–1129.
- Alfaro-Ureña, A., I. Manelici, and J. P. Vasquez (2021). The Effects of Multinationals on Workers: Evidence from Costa Rican Microdata. Mimeo.
- Alfonsi, L., O. Bandiera, V. Bassi, R. Burgess, I. Rasul, M. Sulaiman, and A. Vitali (2020). Tackling youth unemployment: Evidence from a labor market experiment in uganda. *Econometrica* 88(6), 2369–2414.
- Allen, T. (2014). Information frictions in trade. *Econometrica* 82(6), 2041–2083.
- Alvarez-Cuadrado, F., F. Amodio, and M. Poschke (2020). Selection and Absolute Advantage in Farming and Entrepreneurship. CEPR Discussion Papers 14269, C.E.P.R. Discussion Papers.
- Amodio, F. and N. De Roux (2021). Labor Market Power in Developing Countries: Evidence from Colombian Plants. CEPR Discussion Papers 16180, C.E.P.R. Discussion Papers.
- Andrews, I., M. Gentzkow, and J. M. Shapiro (2017). Measuring the sensitivity of parameter estimates to estimation moments. *The Quarterly Journal of Economics* 132(4), 1553–1592.
- Azar, J., S. Berry, and I. Marinescu (2019). Estimating labor market power. Mimeo.
- Azar, J., I. Marinescu, and M. Steinbaum (2022). Labor Market Concentration. *Journal of Human Resources* 57(S), 167–199.
- Bandiera, O., A. Elsayed, A. Smurra, and C. Zipfel (2022). Young adults and labor markets in africa. *Journal of Economic Perspectives* 36(1), 81–100.
- Bassier, I., A. Dube, and S. Naidu (2022). Monopsony in Movers: The Elasticity of Labor Supply to Firm Wage Policies. *Journal of Human Resources* 57(S), 50–86.
- Benmelech, E., N. K. Bergman, and H. Kim (2022). Strong Employers and Weak Employees: How Does Employer Concentration Affect Wages? *Journal of Human Resources* 57(S), 200–250.
- Berger, D., K. Herkenhoff, and S. Mongey (2022). Labor market power. *American Economic Review* 112(4), 1147–93.
- Blattman, C. and S. 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.
- Branstetter, L., F. Lima, L. J. Taylor, and A. Venâncio (2014). Do entry regulations deter entrepreneurship and job creation? evidence from recent reforms in portugal. *The Economic Journal* 124(577), 805–832.
- Breza, E., S. Kaur, and Y. Shamdasani (2021). Labor rationing. *American Economic Review* 111(10), 3184–3224.
- Brooks, W. J., J. P. Kaboski, I. O. Kondo, Y. A. Li, and W. Qian (2021). Infrastructure Investment and Labor Monopsony Power. *IMF Economic Review* 69(3), 470–504.
- Brooks, W. J., J. P. Kaboski, Y. A. Li, and W. Qian (2021). Exploitation of labor? classical monopsony power and labor’s share. *Journal of Development Economics* 150, 102627.
- Buera, F. J., J. P. Kaboski, and Y. Shin (2011). Finance and development: A tale of two sectors. *American economic review* 101(5), 1964–2002.
- Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics* 36(S1), S13–S70.
- Card, D., J. Kluve, and A. Weber (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal* 120(548), F452–F477.
- Dasso, R. and F. Fernandez (2015). The effects of electrification on employment in rural peru. *IZA Journal of Labor & Development* 4(6).
- Dasso, R., F. Fernandez, and H. R. Nopo (2015). Electrification and Educational Outcomes in Rural Peru. IZA Discussion Papers 8928, Institute of Labor Economics (IZA).
- Diao, X., M. Ellis, M. McMillan, and D. Rodrik (2020). Africa’s Manufacturing Puzzle: Evidence from Tanzanian and Ethiopian Firms. Mimeo.
- Dix-Carneiro, R., P. K. Goldberg, C. Meghir, and G. Ulyssea (2021). Trade and informality in the presence of labor market frictions and regulations. Working Paper 28391, National Bureau of Economic Research.
- Dix-Carneiro, R. and B. K. Kovak (2019). Margins of labor market adjustment to trade. *Journal of International Economics* 117(C), 125–142.
- Djankov, S., R. La Porta, F. Lopez-de Silanes, and A. Shleifer (2002). The regulation of entry. *The Quarterly Journal of Economics* 117(1), 1–37.
- Donovan, K., W. Lu, and T. Schoellman (2021). Labor market dynamics and development. Mimeo.

- Dube, A., J. Jacobs, S. Naidu, and S. Suri (2020). Monopsony in online labor markets. *American Economic Review: Insights* 2(1), 33–46.
- Felix, M. (2021). Trade, labor market concentration, and wages. Mimeo.
- Gollin, D. (2008). Nobody’s business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics* 55(2), 219 – 233.
- Goolsbee, A. and C. Syverson (2019). Monopsony power in higher education: A tale of two tracks. Working Paper 26070, National Bureau of Economic Research.
- Heckman, J. J. and B. E. Honoré (1990). The Empirical Content of the Roy Model. *Econometrica* 58(5), 1121–1149.
- Heckman, J. J. and G. 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.
- Herreño, J. and S. Ocampo (2021). Self-employment and development. Mimeo.
- Hsieh, C.-T. and P. J. Klenow (2010). Development accounting. *American Economic Journal: Macroeconomics* 2(1), 207–23.
- INEI (2014). *Producción y empleo informal en el Perú, cuenta satélite de la economía informal 2007-2012*. Instituto Nacional de Estadística e Informática, Lima.
- Kroft, K., Y. Luo, M. Mogstad, and B. Setzler (2020). Imperfect competition and rents in labor and product markets: The case of the construction industry. Working Paper 27325, National Bureau of Economic Research.
- MacKenzie, G. (2019). Trade and market power in product and labor markets. Mimeo.
- Manning, A. (2003). *Monopsony in Motion: Imperfect Competition in Labor Markets*. Princeton University Press.
- Matsudaira, J. D. (2014). Monopsony in the low-wage labor market? evidence from minimum nurse staffing regulations. *The Review of Economics and Statistics* 96(1), 92–102.
- McKenzie, D. (2017). How Effective Are Active Labor Market Policies in Developing Countries? A Critical Review of Recent Evidence. *The World Bank Research Observer* 32(2), 127–154.
- McMillan, M. and A. Zeufack (2022, February). Labor productivity growth and industrialization in africa. *Journal of Economic Perspectives* 36(1), 3–32.
- Meghir, C., R. Narita, and J.-M. Robin (2015). Wages and informality in developing countries. *American Economic Review* 105(4), 1509–46.

- Morlacco, M. (2020). Market power in input markets: Theory and evidence from french manufacturing. Mimeo.
- Muralidharan, K., P. Niehaus, and S. Sukhtankar (2017). General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India. NBER Working Papers 23838, National Bureau of Economic Research, Inc.
- Narita, R. (2020). Self-employment in developing countries: A search-equilibrium approach. *Review of Economic Dynamics* 35, 1 – 34.
- OECD (2006). *OECD employment outlook 2006: Boosting jobs and incomes*. Organisation for Economic Co-operation and Development, Paris, France.
- Pham, H. (2019). Trade reform, oligopsony, and labor market distortions: Theory and evidence. Mimeo.
- Piselli, R. (2013). Patronos de transporte en lima metropolitana: adónde, cuánto y por qué viajan los limenos,. *Revista Argumentos*, 25–28.
- Ponczek, V. and G. Ulysea (2021). Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil. *The Economic Journal*. ueab052.
- Poschke, M. (2019). Wage employment, unemployment and self-employment across countries. Mimeo.
- Ransom, M. R. and D. P. Sims (2010). Estimating the firm’s labor supply curve in a “new monopsony” framework: Schoolteachers in missouri. *Journal of Labor Economics* 28(2), 331–355.
- Rinz, K. et al. (2018). Labor market concentration, earnings inequality, and earnings mobility. *Center for Administrative Records Research and Applications Working Paper 10*.
- Rodrik, D. (2008). Normalizing industrial policy.
- Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford economic papers* 3(2), 135–146.
- Rud, J. P. and I. Trapeznikova (2021). Job creation and wages in least developed countries: Evidence from sub-saharan africa. *The Economic Journal* 131(635), 1331–1364.
- Schubert, G., A. Stansbury, and B. Taska (2021). Employer concentration and outside option. Mimeo.
- Staiger, D. O., J. Spetz, and C. S. Phibbs (2010). Is there monopsony in the labor market? evidence from a natural experiment. *Journal of Labor Economics* 28(2), 211–236.
- Startz, M. (2016). The value of face-to-face: Search and contracting problems in nigerian trade. Available at SSRN 3096685.

- Tortarolo, D. and R. Zárate (2020). Imperfect competition in product and labor markets. a quantitative analysis. Mimeo.
- Ulyssea, G. (2020). Informality: Causes and consequences for development. *Annual Review of Economics* 12(1), 525–546.
- World Bank (2012). *World development report 2013: Jobs*. The World Bank.
- Yeh, C., C. Macaluso, and B. Hershbein (2022). Monopsony in the US Labor Market. *American Economic Review* 112(7), 2099–2138.
- Young, A. (2014). Structural transformation, the mismeasurement of productivity growth, and the cost disease of services. *American Economic Review* 104(11), 3635–67.

Tables and Figures

Table 1: Employer Concentration Across Local Labor Markets

	Mean	St. Dev.
Number of Firms	6.39	10.37
Wage-bill HHI	65.31	33.49
Wage-bill HHI (Weighted by LLM share of total payroll)	37.02	2.99
Wage-bill HHI (Weighted by LLM share of total employment)	33.89	2.65
Employment HHI	62.80	34.71
Employment (Weighted by LLM share of total payroll)	32.97	2.74
Employment (Weighted by LLM share of total employment)	31.11	2.47
Percent of LLM with 1 firm	38.78	2.27
Payroll share of LLM with 1 firm	7.94	1.79
Employment share of LLM with 1 firm	7.80	1.23
Number of Local Labor Markets	280	
Number of Locations	61	
Industries	23	

Notes. This table reports summary statistics of employer concentration measures across local labor markets, averaging across 2004-2011. Local labor markets are defined by 2-digit industries within locations, the latter corresponding to Peruvian provinces or commuting zones.

Table 2: Wage Employment, Self-employment and Earnings

	Employment Shares		Average Daily Earnings	
	All	Manufacturing	All	Manufacturing
Employed	45.46	55.73	32.6	33.51
Self-Employed	46.34	40.46	21.71	24.02
Employed Formal	22.32	28.17	47.33	46.81
Self-Employed Informal	41.9	35.9	17.96	16.85
Informal	73.24	67.27	15.59	13.51
Unemployed	3.28	.15		
All			24.57	23.51

Notes. This table reports summary statistics of worker-level data from ENAHO, averaged over the period 2004-2011. Sample is restricted to working-age individuals aged 25 to 65. Average earnings are daily and nominal, reported in PEN (1 PEN \approx 0.35 USD in 2010).

Table 3: Workers' Mobility: Wage Employment and Self-Employment

<i>All Industries</i>				
		Time t		
		Unemployed	Employed	Self-Employed
Time $t - 1$	Unemployed	19.92	50.94	29.14
	Employed	1.93	79.88	18.18
	Self-Employed	0.93	15.26	83.82
<i>Manufacturing</i>				
		Time t		
		Unemployed	Employed	Self-Employed
Time $t - 1$	Unemployed	-	-	-
	Employed	-	91.04	8.58
	Self-Employed	-	8.08	91.92

Notes. This table reports statistics on transition rates of workers between self-employment and wage employment, using the panel dataset from ENAHO over the period 2007-2011. Sample is restricted to working-age individuals aged 25 to 65.

Table 4: Employer Concentration, Self-employment and Earnings

	Log of Earnings			Self-employed	Schooling	
	All Workers (1)	Employees (2)	Self-employed (3)	{0,1} (4)	Employees (5)	Self-employed (6)
Employer Concentration (Log of Wage-bill HHI)	-0.086*** (0.027)	-0.100*** (0.021)	-0.158*** (0.048)	0.049*** (0.014)	-0.128* (0.066)	-0.264*** (0.075)
Female	-0.530*** (0.042)	-0.381*** (0.036)	-1.228*** (0.075)	0.121*** (0.023)	0.001 (0.080)	-0.269** (0.119)
Age	0.074*** (0.009)	0.027*** (0.009)	0.115*** (0.017)	0.017*** (0.005)	0.054** (0.024)	0.117*** (0.036)
Age sq.	-0.001*** (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
Schooling	0.237*** (0.010)	0.161*** (0.008)	0.109*** (0.017)	-0.001 (0.004)		
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6792	4706	2054	7637	4706	2054
R^2	0.296	0.363	0.383	0.132	0.108	0.124

Notes. * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. This table reports the coefficient estimates obtained when estimating equation (3) using OLS. 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 . Standard errors are clustered at the local labor market level.

Table 5: Electrification, Concentration, and Labor Market Outcomes

	Log of Total Sales		Log of Wage-bill HHI		Log of Earnings		Self-employed {0,1}		Schooling	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Electrification	0.061*** (0.011)	-0.033*** (0.005)	0.009*** (0.003)	0.006*** (0.001)	0.021*** (0.007)	-0.007*** (0.002)	0.001 (0.001)	0.007*** (0.002)		
Female			-0.768*** (0.034)	-0.441*** (0.028)	-1.254*** (0.044)	0.126*** (0.017)	-0.005 (0.006)	-0.155*** (0.015)		
Age			0.087*** (0.007)	0.031*** (0.007)	0.096*** (0.009)	0.035*** (0.003)	-0.001 (0.002)	0.009*** (0.003)		
Age sq.			-0.001*** (0.000)	-0.000*** (0.000)	-0.001*** (0.000)	-0.000*** (0.000)	-0.000 (0.000)	-0.000*** (0.000)		
Schooling			0.224*** (0.007)	0.164*** (0.007)	0.161*** (0.009)	-0.026*** (0.003)				
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes		
Observations	1040	1040	18669	8922	9688	20802	8922	9688		
R ²	0.068	0.120	0.238	0.238	0.254	0.117	0.008	0.040		

Notes. * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. This table reports the coefficient estimates obtained when estimating equation (3) using OLS. Unit of observation in columns 1 and 2 is the local labor market k , as defined by a 2-digit industry j within a province or commuting zone g . Unit of observations in columns 3 to 8 is a working-age individual surveyed in ENAHO. Standard errors are clustered at the local labor market level.

Table 6: Summary of Model Parameters

Parameter	Description	Value
<i>Externally Estimated</i>		
σ_F	St. dev. of log ability as a wage worker	1.05
σ_S	St. dev. of log ability as self-employed	1.10
ϱ	Correlation of log abilities	0.98
$\Delta\mu \equiv \mu_{F,k} - \mu_{S,k}$	Relative absolute advantage	1.02
<i>Estimated via SMM</i>		
μ_z	Mean of market-level productivity	1.27
σ_z	St. dev. of market-level productivity	0.43
$\alpha_0 (\times 10^3)$	Fixed cost (Intercept)	0.72
$\alpha_1 (\times 10^3)$	Fixed cost (Slope)	0.08
μ_{μ_F}	Mean of absolute advantage distribution	-1.19
σ_{μ_F}	St. dev. of absolute advantage distribution	0.47
β	Demand shifter for variety F	1.22
ρ	Elasticity of Substitution between Varieties F and S	1.17
A_S	Productivity in sector S	1.2

Notes. This table reports the estimated parameters. See Section 5 for more details.

Table 7: Targeted Moments and Model Fit

Moment	Model	Data
1. Log of number of firms (mean)	1.46	1.46
2. Log of number of firm (st. dev.)	1.32	1.27
3. % of local labor markets with 1 firm	0.26	0.27
4. % workers in markets with 1 firm	0.07	0.07
5. Share of wage workers across markets (mean)	0.70	0.52
6. Relative earnings in wage sector (mean)	1.64	1.70
7. Relative ability of wage workers (mean)	1.71	1.16
<i>Regression Coefficients</i>		
8. Mean wage sector earnings on (log) no. of firms	0.03*** (0.004)	0.08*** (0.012)
9. Mean self-employment sector earnings on (log) no. of firms	0.03*** (0.004)	0.12*** (0.03)
10. Share of wage workers on (log) no. of firms	0.02*** (0.004)	0.03*** (0.007)
11 Mean ability of wage workers on (log) no. of firms	0.009 (0.04)	0.006** (0.003)
12. Mean ability of self-employed workers on (log) no. of firms	0.02 (0.04)	0.02*** (0.006)

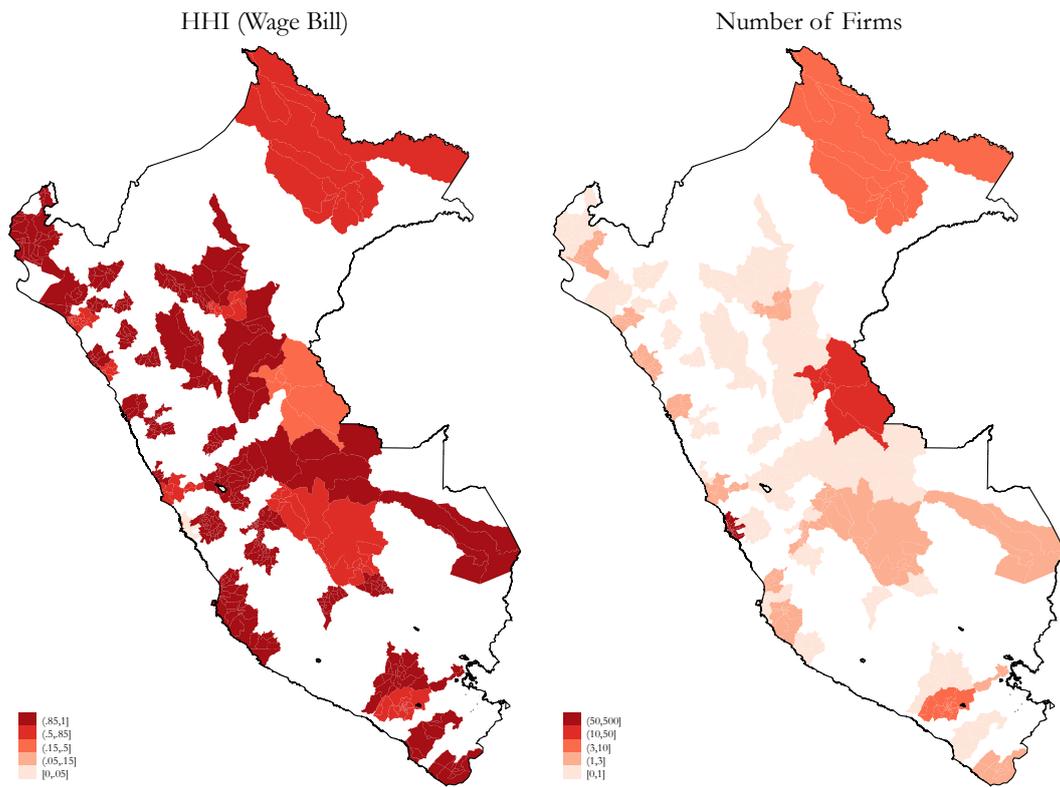
Notes. This table reports the moments used in the estimation, and compares them with the same moments calculated from the estimated model. The data moments are computed in the sample of local labor markets where at least one formal firm is active and both the share of self-employed workers and wage workers is strictly between 0 and 1. See Section 5 for more details on the moments' construction.

Table 8: Policy Counterfactuals

	Baseline	Fixed Entry Costs (f^e)	Market Access (z)	Abilities (μ_F)
<i>Promoting Firm Entry</i>				
HHI	1	0.5	0.5	0.5
Markdown ψ	1	0.97	0.98	0.97
Wage-Employment Share	1	1.01	1.1	1.01
Self-Employment Share	1	0.97	0.77	0.97
Earnings Wage Employed	1	1.04	4.27	5.36
Earnings Self-Employed	1	1	2.97	5.16
Total Income	1	1	3.73	5.17
<i>Eliminating Heterogeneity</i>				
HHI	1	0.96	0.64	0.76
Markdown ψ	1	1	0.98	0.98
Wage-Employment Share	1	1	1.06	1.01
Self-Employment Share	1	1	0.86	0.99
Earnings Wage Employed	1	1	2.47	2.15
Earnings Self-Employed	1	1	1.99	2.11
Total Income	1	1	2.27	2.11

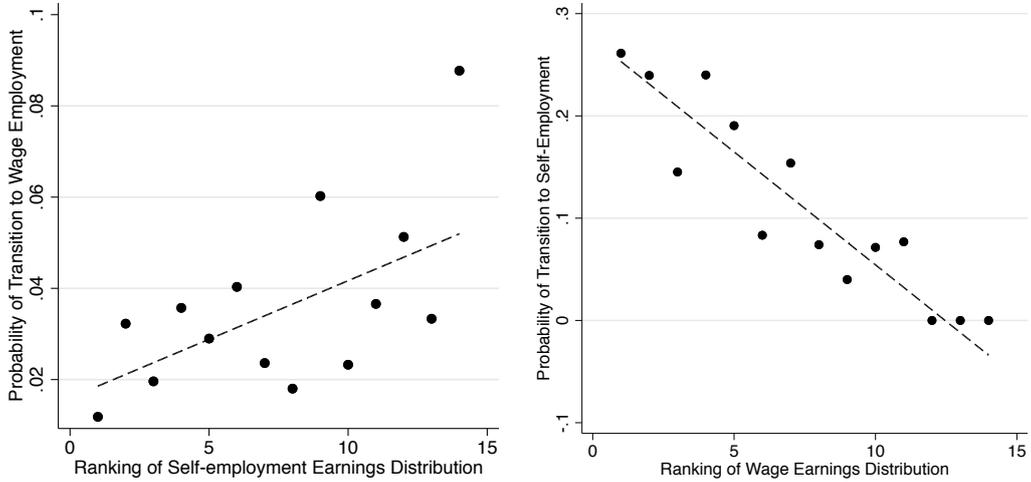
Notes. The table reports the results of the two counterfactual policy experiments described in Section 6.2. In the first one (top panel), we consider a reduction in fixed entry costs, an increase in firm productivity and an increase in workers' skills that reduce that reduce by 50% the economy-wide average level of employer concentration. In the second one (bottom panel) we eliminate differences across local labor markets in either fixed costs, productivity, or workers' ability. The table reports average values across local labor markets relative to baseline.

Figure 1: Employer Concentration Map



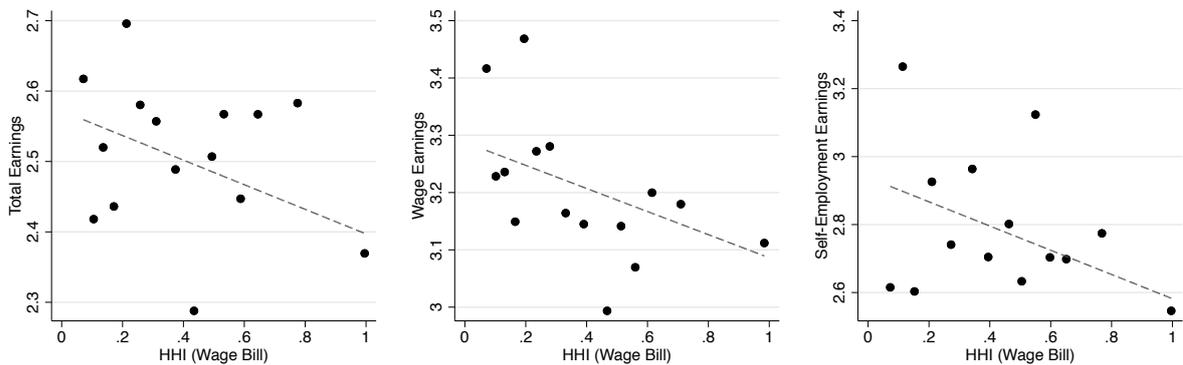
Notes. The figures shows maps of employer concentration across Peruvian provinces, averaged over the period 2004-2011. The left panel uses the wage-bill HHI as measure of concentration, whereas the right panel considers the number of firms. The figure shows the average concentration as calculated for each industry within province or commuting zone (local labor market) and then averaged within provinces.

Figure 2: Transitions Probabilities And Earnings



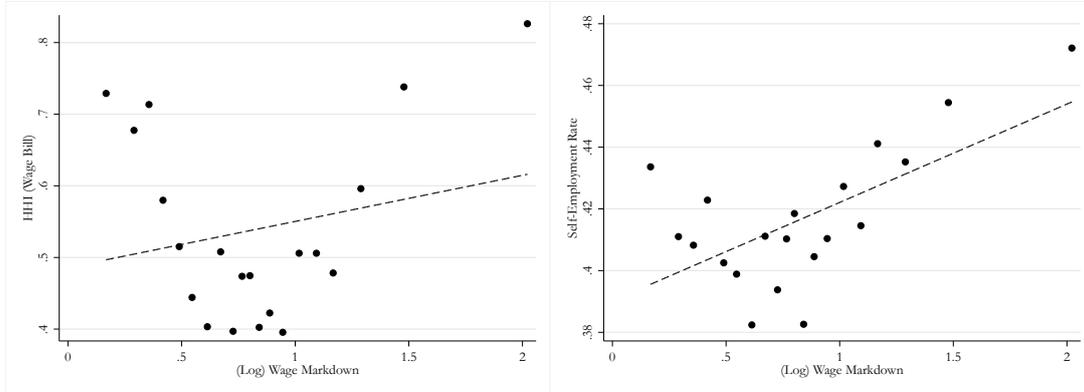
Notes. The figures illustrate the relationship between the likelihood of transitioning across sectors and earnings. The left graph plots average yearly transition probabilities from self-employment to wage employment within the same local labor market against (rankings of) the self-employment earnings distribution in the previous year. The right graph plots average yearly transition probabilities from wage employment to self-employment within the same local labor market against (rankings of) the wage employment earnings distribution in the previous year.

Figure 3: Employer Concentration And Earnings



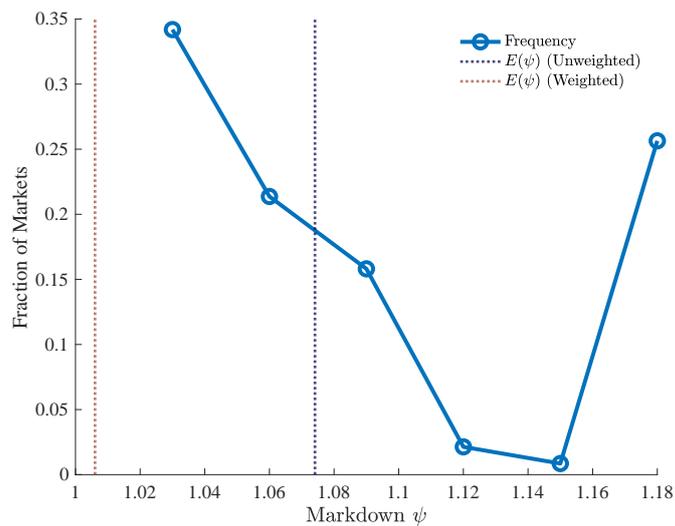
Notes. The figures illustrate the relationship between employer concentration and earnings for all workers (left graph), wage workers (centre graph), and self-employed workers (right graph) and using wage-bill HHI as measure of employer concentration. Each figure shows both binned scatterplots and linear fit. Figure A.7 in Appendix A shows the same graphs using employment HHI and number of firms per local labor market as alternative measures of employer concentration.

Figure 4: Wage Markdowns - Correlations



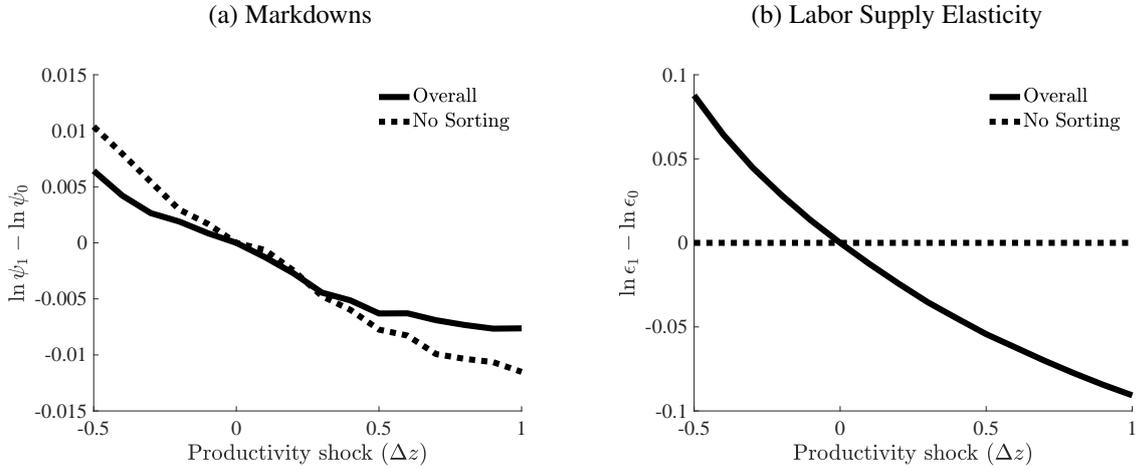
Notes. The figures illustrate the relationship across local labor markets between employer concentration as measured by wage-bill HHI and average wage markdown (left panel), and between the share of self-employment and average wage markdown (right panel). Markdowns are measured using the formula in equation (17).

Figure 5: Wage Markdowns - Distribution (Model)



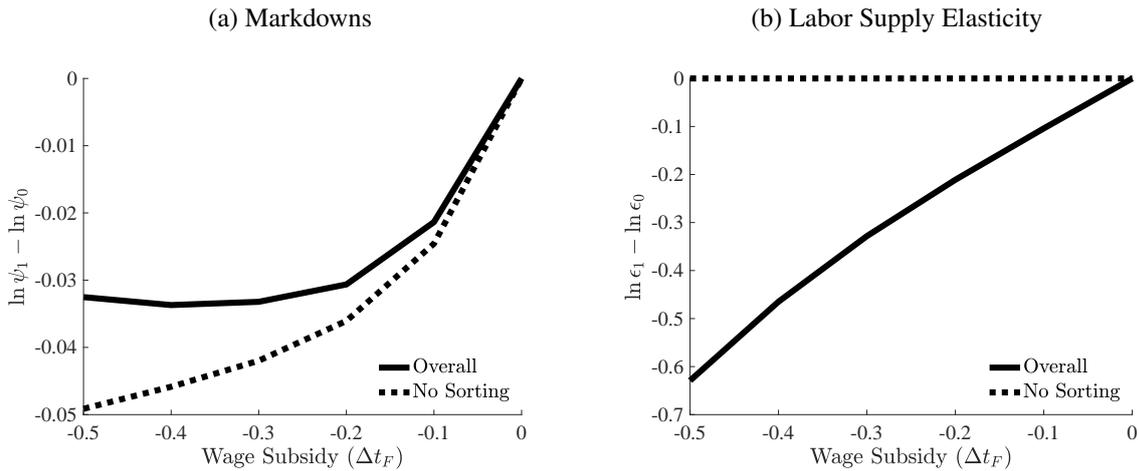
Notes. The figure illustrates the model-implied distribution of average markdown across local labor markets. The average unweighted wage markdown is estimated equal to 1.07, while the average weighted wage markdown is equal to 1.01.

Figure 6: Rent Sharing with Aggregate Productivity Shock



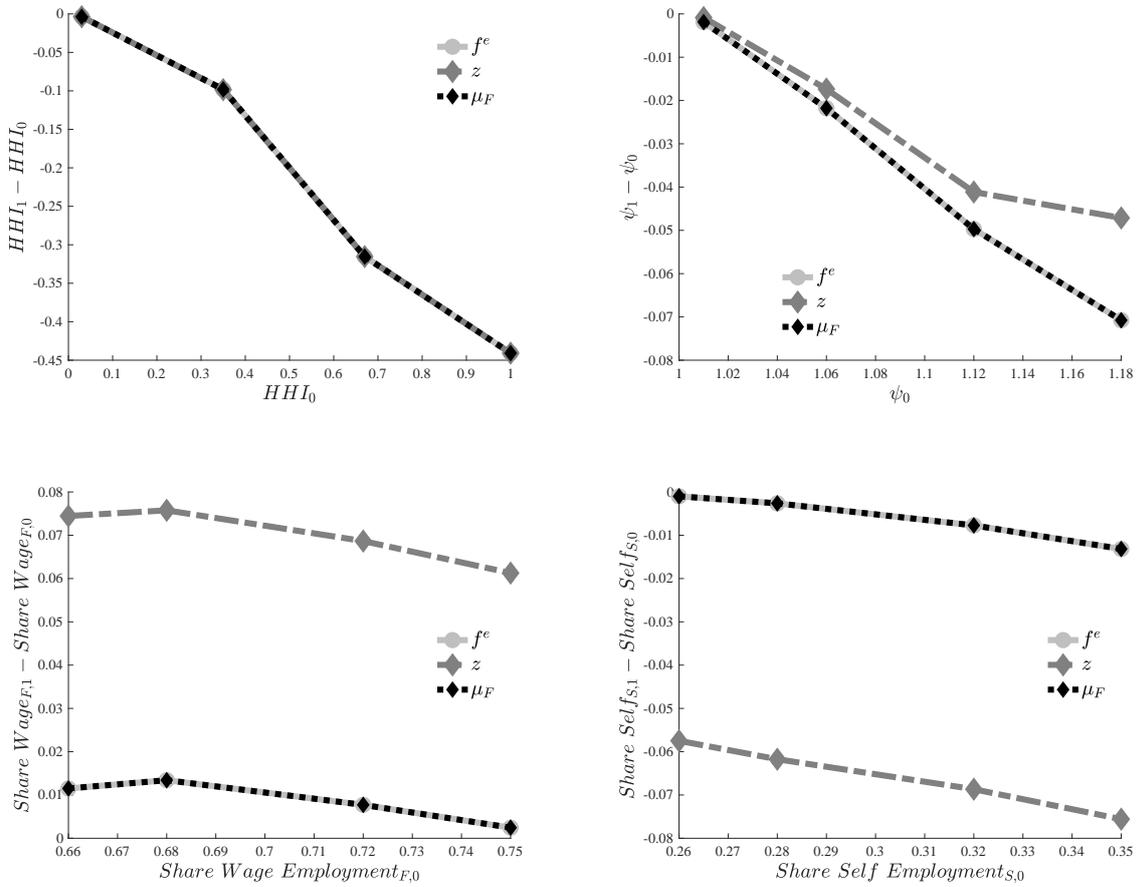
Notes. The figures show the average percentage change in average markdown and labor supply elasticity following a productivity shock of different sizes, as described in Section 6.1. The solid line in both figures represents the overall effect. The dashed line is the counterfactual change in markdown or supply elasticity in the "no-sorting" equilibrium, where labor supply is kept fixed and only strategic interactions among firms matter.

Figure 7: Rent Sharing with Wage Subsidy



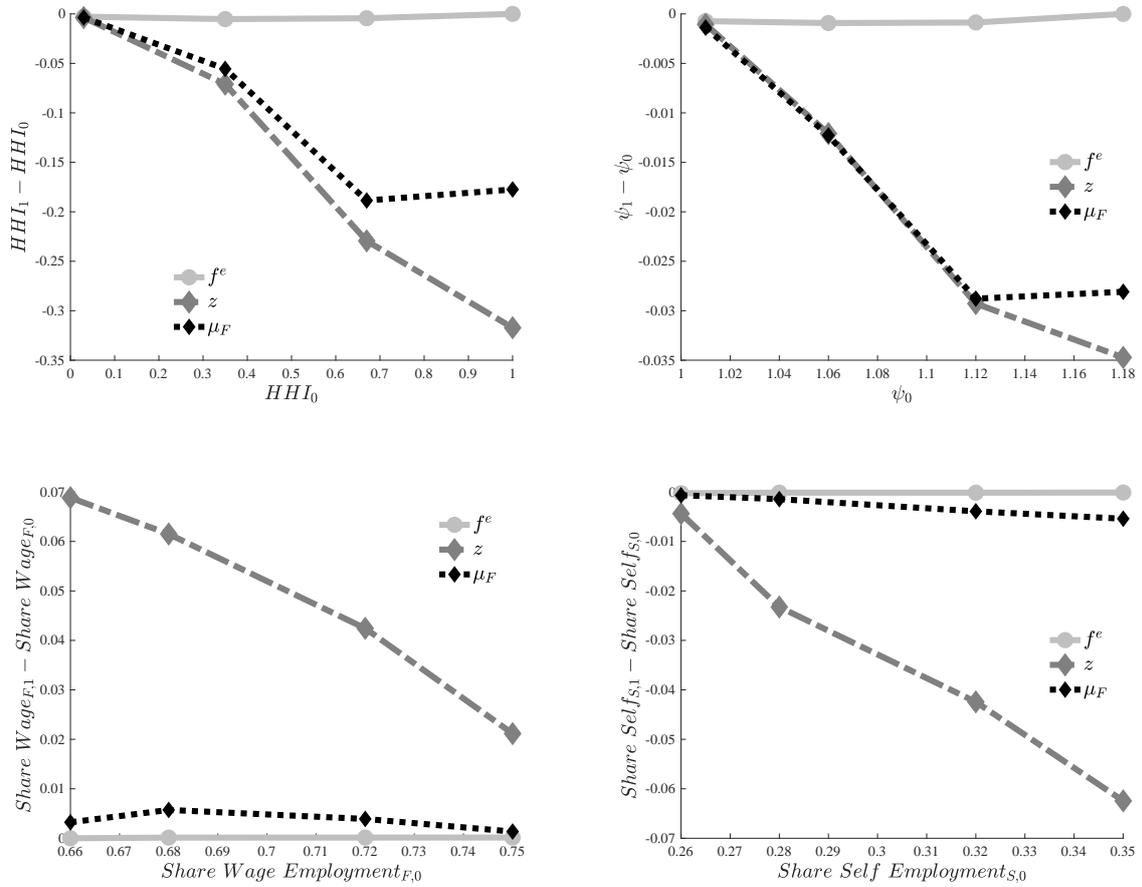
Notes. The figures show the average percentage change in average markdown and labor supply elasticity following the implementation of a wage subsidy of different sizes, as described in Section 6.1. The solid line in both figures represents the overall effect. The dashed line is the counterfactual change in markdown or supply elasticity in the "no-sorting" equilibrium, where labor supply is kept fixed and only strategic interactions among firms matter.

Figure 8: Policy Experiment: Promoting Firm Entry



Notes. The figures illustrate the results of the first counterfactual policy experiment described in Section 6.2, where we consider a reduction in fixed entry costs, an increase in firm productivity and an increase in workers' skills that reduce that reduce by 50% the economy-wide average level of employer concentration. The figures show baseline level and changes in employer concentration, wage markdowns, wage employment and self-employment shares in the three policy scenarios.

Figure 9: Policy Experiment: Eliminating Heterogeneity



Notes. The figures illustrate the results of the second counterfactual policy experiment described in Section 6.2, where we eliminate differences across local labor markets in either fixed costs, productivity, or workers' ability. The figures show baseline level and changes in employer concentration, wage markdowns, wage employment and self-employment shares in the three policy scenarios.

Online Appendix

A Additional Tables and Figures

Table A.1: Employer Concentration, Self-employment and Earnings – Employment HHI

	Log of Earnings			Self-employed	Schooling	
	All Workers (1)	Employees (2)	Self-employed (3)	{0,1} (4)	Employees (5)	Self-employed (6)
Employer Concentration (Log of Employment HHI)	-0.075*** (0.026)	-0.096*** (0.019)	-0.146*** (0.045)	0.050*** (0.013)	-0.115* (0.061)	-0.237*** (0.074)
Female	-0.529*** (0.042)	-0.381*** (0.036)	-1.229*** (0.075)	0.121*** (0.023)	0.002 (0.080)	-0.271** (0.119)
Age	0.073*** (0.009)	0.027*** (0.009)	0.115*** (0.017)	0.017*** (0.005)	0.054** (0.024)	0.116*** (0.035)
Age sq.	-0.001*** (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
Schooling	0.238*** (0.009)	0.162*** (0.008)	0.109*** (0.017)	-0.001 (0.004)		
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6792	4706	2054	7637	4706	2054
R^2	0.296	0.363	0.382	0.133	0.108	0.123

Notes. * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. This table reports the coefficient estimates obtained when estimating equation (3) using OLS. 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 . Standard errors are clustered at the local labor market level.

Table A.2: Employer Concentration, Self-employment and Earnings – Number of Firms

	Log of Earnings			Self-employed	Schooling	
	All Workers (1)	Employees (2)	Self-employed (3)	{0,1} (4)	Employees (5)	Self-employed (6)
Employer Concentration (Log of Number of Firms)	0.055*** (0.017)	0.068*** (0.014)	0.109*** (0.030)	-0.033*** (0.009)	0.072 (0.044)	0.170*** (0.050)
Female	-0.527*** (0.042)	-0.379*** (0.036)	-1.224*** (0.075)	0.120*** (0.023)	0.004 (0.081)	-0.263** (0.119)
Age	0.073*** (0.009)	0.028*** (0.009)	0.113*** (0.017)	0.017*** (0.005)	0.054** (0.024)	0.113*** (0.035)
Age sq.	-0.001*** (0.000)	-0.000** (0.000)	-0.001*** (0.000)	-0.000 (0.000)	-0.001*** (0.000)	-0.002*** (0.000)
Schooling	0.237*** (0.010)	0.162*** (0.008)	0.108*** (0.017)	-0.001 (0.004)		
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6792	4706	2054	7637	4706	2054
R^2	0.296	0.364	0.384	0.133	0.107	0.124

Notes. * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. This table reports the coefficient estimates obtained when estimating equation (3) using OLS. 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 . Standard errors are clustered at the local labor market level.

Table A.3: Wage Markdown Across Industries

Industry		Wage Markdown	
		Mean	St. Dev.
10	Food Products	3.08	2.13
11	Beverages	2.73	1.33
13	Textiles	2.45	1.64
14	Apparel	2.1	1.49
15	Leather Products	2.76	1.79
16	Wood	3.39	2.13
17	Pulp & Paper Products	2.66	.85
18	Printing & Recorded Media	2.71	1.26
19	Coke and Refined Petroleum	3.35	1.31
20	Chemical Products	3.59	2.31
21	Pharma Products	3.03	1.9
22	Rubber & Plastic Products	3.49	2.17
23	Non metallic Mineral Products	5.08	3.54
24	Basic Metals	3.11	1.68
25	Fabricated Metal Products	2.25	.71
26	Computer & Electronics	2.61	.8
27	Electrical Equipment	2.65	1.27
28	Machinery & Equipment	2.54	1.02
29	Motor Vehicles & (Semi-)Trailers	2.5	.92
30	Other Transport Equipment	2.61	2.51
31	Furniture	2.61	1.41
32	Other Manufacturing	2.52	1.39
33	Repair & Installation of Machinery & Equipment	2.05	.97

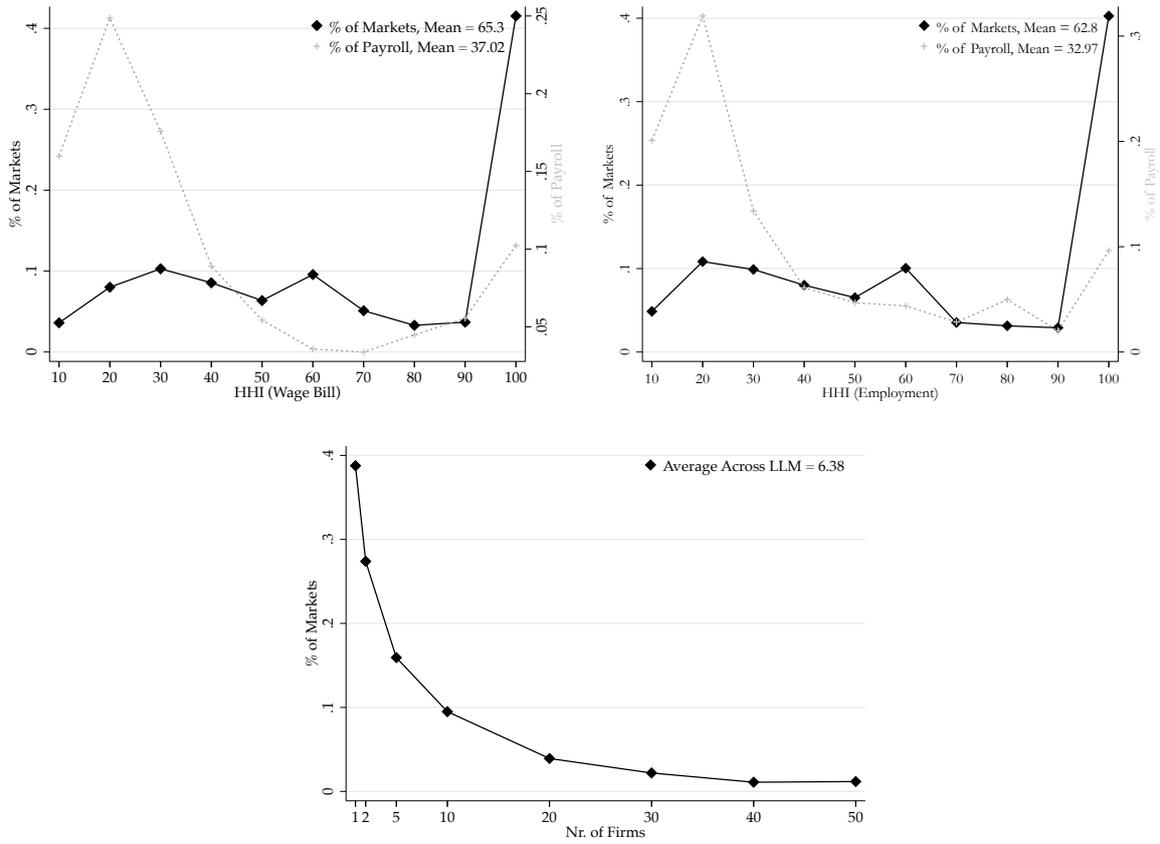
Notes. The table reports the mean and standard deviation of the wage markdown across manufacturing industries. Wage markdowns are computed at the local labor market level as value added divided by total wage bill, as in equation (17), and then aggregated at the industry level.

Table A.4: Ability Distribution Parameter Estimates

Parameter	OLS		MDE Estimates
	<i>(Unconstrained)</i>		<i>(Constrained)</i>
	(1)	(2)	(3)
σ_F	0.63*** (0.02)		1.05*** [0.016]
σ_S	1.32*** (0.03)		1.10*** [0.03]
ρ	-2.17		0.98*** [0.03]
$\Delta\mu$	1.02*** [0.10]	1.03*** [0.24]	

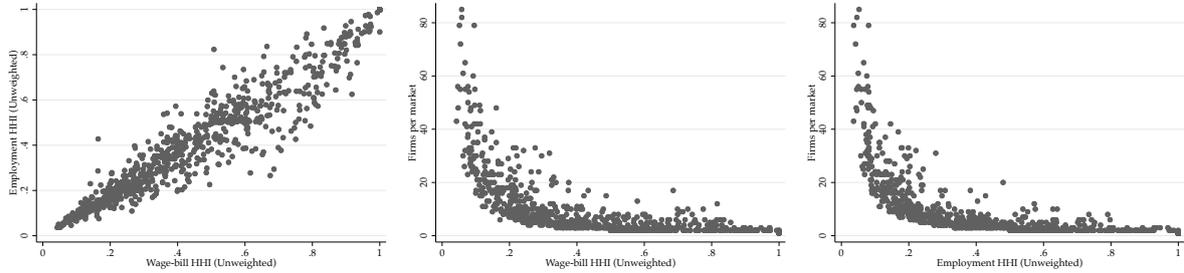
Notes. * p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01. The table reports the parameter estimates of the vector $\Theta = (\sigma_F, \sigma_S, \rho)$ and of $\Delta\mu$, that we obtain with the procedure detailed in Section 5.2. Square brackets are used when standard errors are obtained via a bootstrap procedure, while round brackets are used for analytical standard errors. Columns (1) and (2) reports estimates using standard OLS estimators, while column (3) reports the estimates obtained via the minimum-distance estimation. The value of ρ implied by standard OLS estimates is greater than 1 in absolute value, which motivates the use of a constrained estimator. The estimate of $\Delta\mu$ in column (1) is obtained by approximating \hat{W}_k with a polynomial control function in the market share of self-employed workers; the one in column (2) is the one obtained by approximating \hat{W}_k with a polynomial control function of the number of firms in the market and the wage markdown. Standard errors for $\Delta\mu$ are obtained by block bootstrapping over the entire procedure.

Figure A.1: Employer Concentration Across Local Labor Markets



Notes. The figure plots the distribution of the three employer concentration measures – wage-bill HHI, employment HHI and number of firms – across local labor markets. The dark lines represent the unweighted measures, whereas the light grey lines plot the concentration measure as weighted by the local labor market’s share of total payroll.

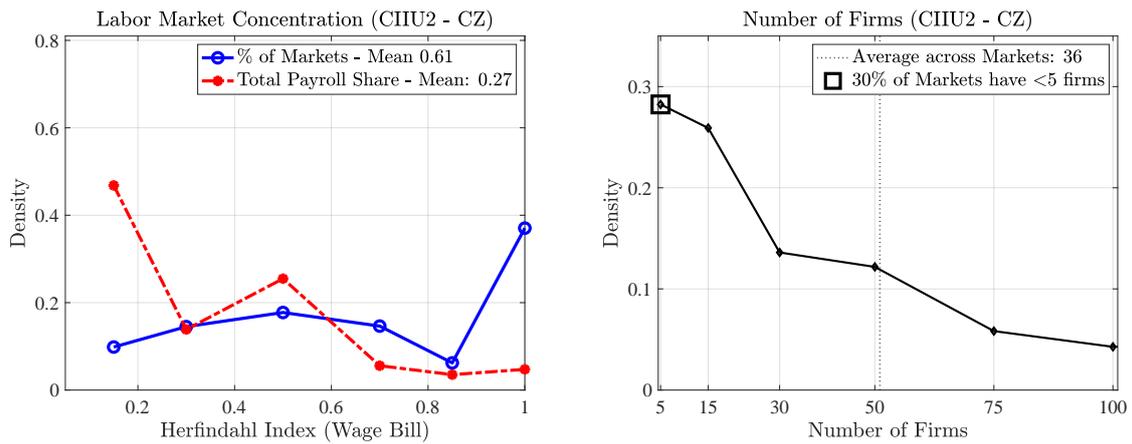
Figure A.2: Correlation Between Employer Concentration Measures



wa

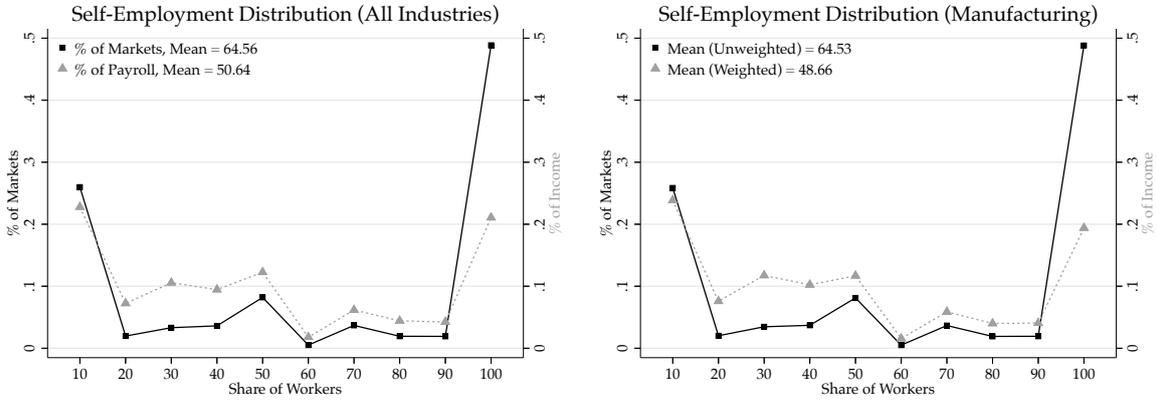
Notes. The figure plots the three employer concentration measures – wage-bill HHI, employment HHI and number of firms – one against the other across all local labor market-level observations. Wage-bill and employment HHI are strongly positively correlated and they are both strongly negatively correlated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

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



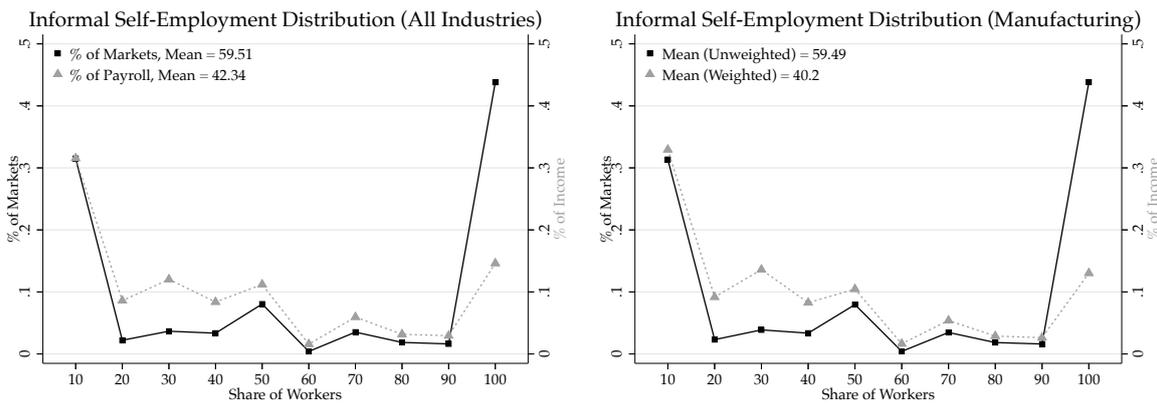
Notes. The figure plots the distribution of wage-bill HHI and number of firms across local labor markets according to the 2007 Peruvian Economic Census, focusing on all establishments in the manufacturing sector. The solid line in the left graph corresponds to the unweighted measure while the dash line corresponds to the weighted measure having as a weight the local labor market's share of nation-wide payroll.

Figure A.4: Self-Employment Rates Across Local Labor Markets



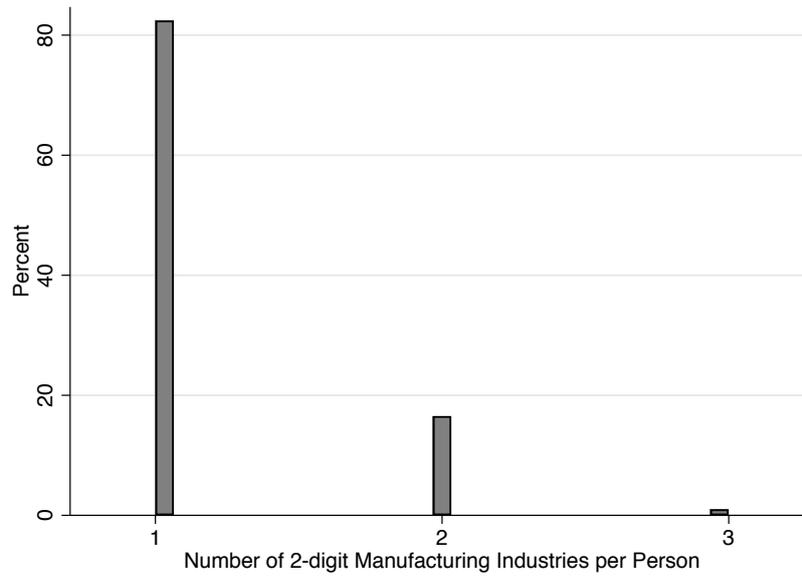
Notes. The figure plots the distribution of the three employer concentration measures – wage-bill HHI, employment HHI and number of firms – across local labor markets. The dark lines represent the unweighted measures, whereas the light grey lines plot the concentration measure as weighted by the local labor market’s share of total payroll.

Figure A.5: Informal Self-Employment Rates Across Local Labor Markets



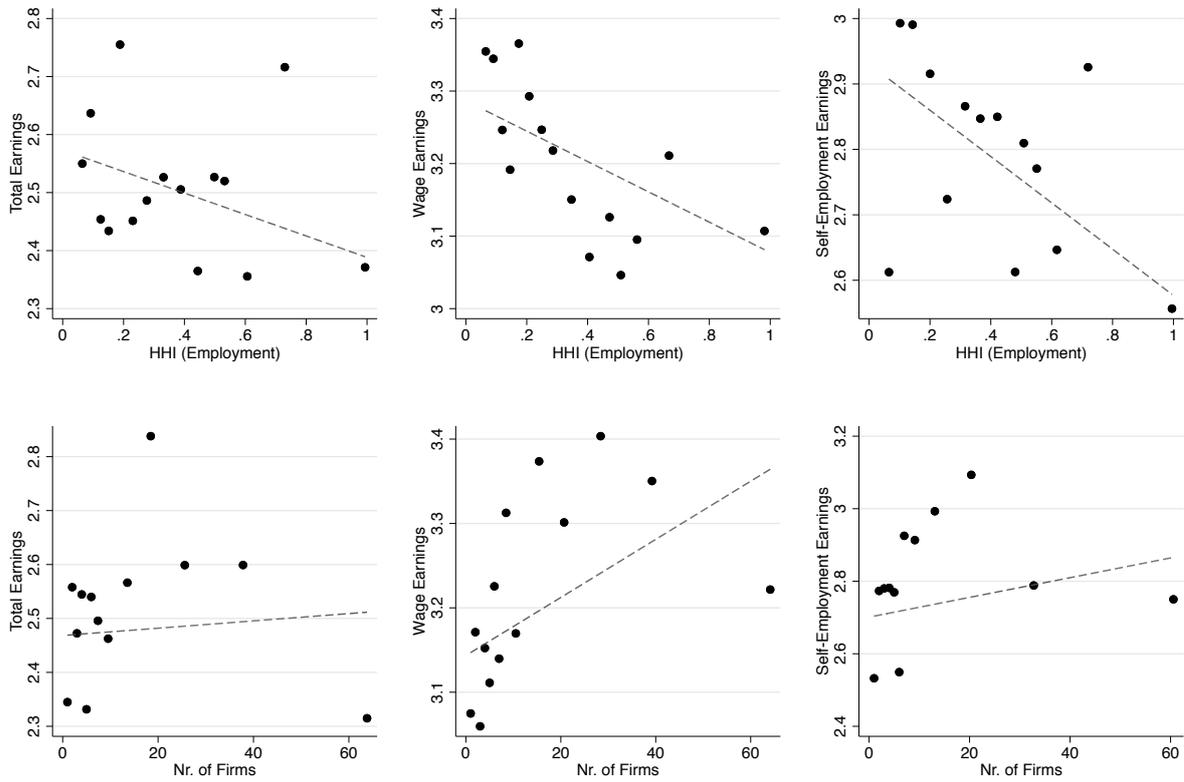
Notes. The figure plots the distribution of the three employer concentration measures – wage-bill HHI, employment HHI and number of firms – across local labor markets. The dark lines represent the unweighted measures, whereas the light grey lines plot the concentration measure as weighted by the local labor market’s share of total payroll.

Figure A.6: Distribution of Number of Industries per Person



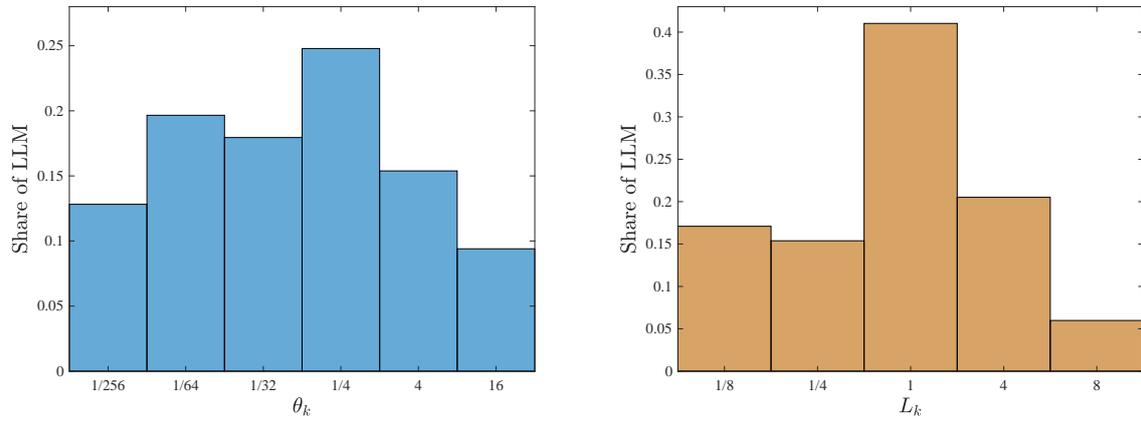
Notes. The figure shows the distribution of the number of 2-digit industries per person among respondents that are observed at least twice in the 2007-2011 ENAHO panel data. More than 80% of manufacturing workers are repeatedly observed in the same 2-digit industry.

Figure A.7: Employer Concentration And Earnings – Alternative Measures



Notes. The figures illustrate the relationship between employer concentration and earnings for all workers (left graph), wage workers (centre graph), and self-employed workers (right graph) and using two alternative measures of employer concentration: employment HHI and number of firms per local labor market. Each figure shows both binned scatterplots and linear fit.

Figure A.8: 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 Theory Appendix

B.1 Market Equilibrium: Uniqueness and Properties

From equation (5), we can write:

$$\frac{P_{F,k} C_{F,k}}{P_{S,k} C_{S,k}} = \left(\frac{s_{F,k}}{s_{S,k}} \right) \quad (1)$$

Market clearing at the good level implies that $C_{j,k} = Y_{j,k}$, such that:

$$\frac{P_{F,k} z_k N_{F,k}}{P_{S,k} A_S N_{S,k}} = \left(\frac{s_{F,k}}{s_{S,k}} \right) \quad (2)$$

$$\frac{z_k^{-1} W_{F,k} \psi_{F,k}}{A_S^{-1} W_{S,k}} \frac{z_k N_{F,k}}{A_S N_{S,k}} = \left(\frac{s_{F,k}}{s_{S,k}} \right) \quad (3)$$

$$\hat{W}_k \psi_{F,k}(\hat{W}_k, M_k) \hat{N}(\hat{W}_k) = \left(\frac{s_{F,k}}{s_{S,k}} \right) \quad (4)$$

where $\hat{N}(\hat{W}_k) \equiv \frac{N_{F,k}}{N_{S,k}}$. Equation (4) gives the first (market) equilibrium condition.

Dividing both sides of the first order condition of wage firms by $P_{S,k} = A_S^{-1} W_{S,k}$, we get:

$$\frac{P_{F,k}}{P_{S,k}} z_k = A_S \hat{W}_k \psi_{F,k}. \quad (5)$$

We substitute the term $\left(\frac{P_{F,k}}{P_{S,k}} \right)$ using equation (5) to obtain:

$$\left(\frac{s_{F,k}}{s_{S,k}} \right) = \beta^\rho z_k^{\rho-1} A_S^{1-\rho} \hat{W}_k^{1-\rho} \left(\psi(\hat{W}_k, M_k) \right)^{1-\rho}. \quad (6)$$

Equations (4) and (6) constitute the system of equilibrium equations for a market k . Equation (4) captures a demand condition, whereby expenditure on sector F is monotonically decreasing in the relative wage: the higher the relative wage, the higher the price of good F . On the contrary, equation (6) captures a supply condition featuring an increasing relationship between expenditures and wages: the higher the sector F wage, the higher the labor supply in the sector. As labor supply (and output) increases, prices adjust downwards to clear markets, increasing expenditures in the sector. Given the set of entrants M_k and productivities z_k , the market equilibrium $\{\hat{s}_k, \hat{W}_k\}_{k=1,\dots,K}$ is unique.

To see this, notice that the system yields the following equation:

$$\beta^\rho z_k^{\rho-1} A_S^{1-\rho} = \hat{W}_k^\rho \psi_k^\rho \hat{N}(\hat{W}_k), \quad (7)$$

which can be solved for a unique value of \hat{W}_k .

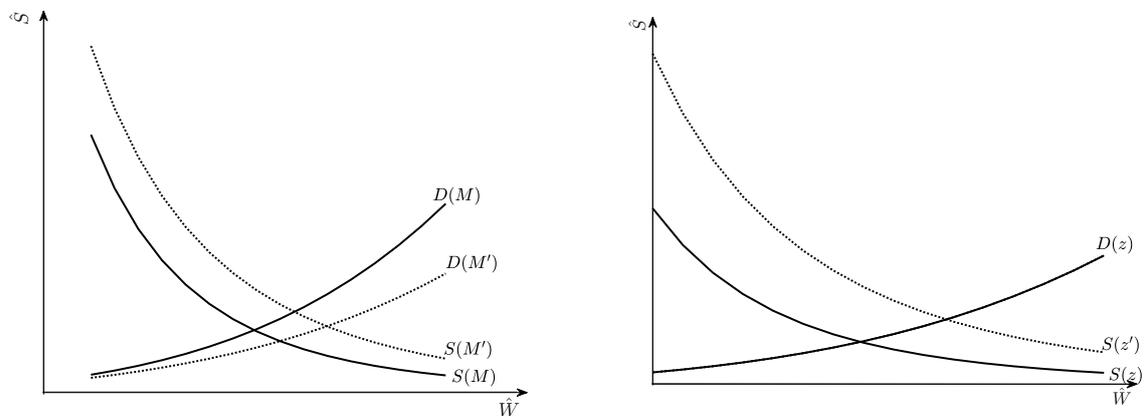
We now investigate how the equilibrium relative expenditure and wages depend on the number of entrants and market-level productivity. The left panel of Figure B.1 depicts the market equilibrium for two different values of M , with $M' > M$. As the number of firms increases, the wage markdown decreases, leading firms to produce more (and the supply curve to shift upwards) and lowering the relative price of sector F good (such that the demand curve shift downwards). The new equilibrium features a higher relative wage, needed to increase supply of labor in sector F . The effect on expenditure is instead uncertain, due to the contrasting effect of demand and supply forces.

An increase in productivity has a different effect on the equilibrium, as it only affects the supply curve directly. We plot comparative statics with respect to z_k in the right panel of Figure B.1. High productivity increases supply in the formal sector, which shifts the supply curve upwards similar to an increase in the number of entrants. To rebalance the equilibrium, both relative wages and expenditures must increase to clear the labor and good market, respectively.

Figure B.1: Market Equilibrium: Comparative Statics

(a) Number of Firms

(b) Productivity



B.2 Model Solution: Algorithm

General Equilibrium Given the productivity draws and the model parameters Θ , the model general equilibrium can be found by implementing the following fixed point procedure:

1. Take an initial guess for Y , which completes the GE vector $\mathbf{X} = (Y, P)$ where $P = 1$ by normalization.
2. Given \mathbf{X} , solve for the market equilibrium in each local labor market, characterizing $\mathbf{K} = \{M_k, s_{M,k}, \hat{W}_k\}$, as described below.
3. Given \mathbf{K} , use the general equilibrium conditions to solve for the new values of \mathbf{X} .
4. Update the initial values of Y , and loop over until convergence.

Market Equilibrium To solve for the market equilibrium in local labor market k given the GE vector $\mathbf{X} = (Y, P)$ and parameters, we proceed as follows:

1. Given the GE vector $\mathbf{X} = (Y, P)$, we first compute market-level expenditures using the Cobb-Douglas formula:

$$Y_k = \theta_k Y. \quad (8)$$

2. Given Y_k , the number of entrants M_k , and their marginal costs $\{z_k\}$, the market equilibrium (\hat{s}_k, \hat{W}_k) solves the system of equations (4) and (6).
3. Given (\hat{s}_k, \hat{W}_k) as a function of M_k , we can solve for the equilibrium number of entrants M_k using the following iterative procedure:

- (a) Guess that $M_0 = 1$.
- (b) Set $M_k = M_0$, and compute the market equilibrium as in point 2.
- (c) Given $(\hat{s}_k, \hat{W}_k, M_k)$ and unitary (per-firm) fixed costs f_k^e , profits in the sector as:

$$\Pi(M_k) = s_{M,k} \theta_k Y \left(1 - \frac{1}{\psi_k(M_k, \hat{W}_k)} \right) - M_k f_k^e. \quad (9)$$

- (d) If $\Pi(M_k) > 0$, then set $M_0 = M_k + 1$.
- (e) Repeat steps (b)-(d) until

$$\Pi(M_k) \geq 0 \quad \& \quad \Pi(M_k + 1) < 0, \quad (10)$$

setting the equilibrium number of firms equal to M_k .

4. Given the number of entrants M_k , recompute market equilibrium as in 2.

B.3 Labor Market Power and Sectoral Earnings

In Section 4.3 we argued that the effect of labor market power on labor productivity in each sector – as measured by output per worker – depends critically on the parameters of the workers’ ability distribution, and in particular on the correlation of workers’ abilities across sectors and the population variance.

To see this, consider the case in which abilities $\mathbf{a} = (a_F, a_S)$ are drawn from a joint log-normal distribution.

$$\log \mathbf{a} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \text{ where } \boldsymbol{\mu} = \begin{pmatrix} \mu_F \\ \mu_S \end{pmatrix}, \text{ and } \boldsymbol{\Sigma} = \begin{pmatrix} \sigma_F^2 & \rho\sigma_F\sigma_S \\ \rho\sigma_F\sigma_S & \sigma_S^2 \end{pmatrix} \quad (11)$$

where $\boldsymbol{\mu}$ governs the absolute advantage of workers in the two industry sectors, $\boldsymbol{\Sigma}$ governs the extent of comparative advantage, and ρ captures the correlation between a_F and a_S . Following Heckman and Sedlacek (1985), let then $\mathbf{u} = (u_F, u_S)$ be a mean zero normal vector and $\sigma^* = (\sigma_F^2 + \sigma_S^2 - 2\rho\sigma_F\sigma_S)^{\frac{1}{2}}$. The probability that a given worker in local labor market k operates in sector M is equal to

$$\Phi(c_{F,k}) = \Phi\left(\frac{\ln \hat{W}_k + \mu_F - \mu_S}{\sigma^*}\right) \quad (12)$$

where $\Phi(\cdot)$ is the cumulative distribution function of a normal standard variable and $\hat{W}_k \equiv \frac{W_{F,k}}{W_{S,k}}$ is the relative wage per efficiency unit of labor in sector M . The probability of operating in sector S is instead equal to $\Phi(c_{S,k}) = \Phi(-c_{F,k})$.

The mean log endowment of efficiency units of labor in each sector is given by

$$\begin{aligned} E\left(\ln a_{F,k} | a_{F,k} \hat{W}_k \geq a_{S,k}\right) &= \mu_F + \left(\frac{\sigma_F^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right) \lambda(c_{F,k}) \\ E\left(\ln a_{S,k} | a_{F,k} \hat{W}_k < a_{S,k}\right) &= \mu_S + \left(\frac{\sigma_S^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right) \lambda(c_{S,k}) \end{aligned} \quad (13)$$

where $\lambda(x) = \frac{\phi(x)}{\Phi(x)} \geq 0$ is a convex monotone decreasing function of x .

The above equations provide several useful insights. First, the mean endowment of efficiency units of labor or ability is higher than the population mean in both sectors only if $\rho\sigma_F\sigma_S < \min\{\sigma_F^2, \sigma_S^2\}$. This is the case if, for instance, the two abilities are uncorrelated as in the classical Roy model, or negatively correlated. If the correlation between abilities is positive and sufficiently high, i.e. $\rho\sigma_F\sigma_S > \min\{\sigma_F^2, \sigma_S^2\}$, mean ability is lower than the population mean in the sector where abilities are less dispersed, and higher than the population mean in the sector where abilities are more dispersed. Second, when the employment share of one sector increases the corresponding $\lambda(\cdot)$ decreases, and the difference between the sectoral mean and the population mean decreases. This means that the relationship between sectoral size and

mean ability depends on the sign and size of the correlation between abilities in the two sectors as well as their variance in the population.

In particular, it follows from the equations above that the effect of an increase in the number of entrants M_k on average ability in each sector is given by

$$\begin{aligned} \frac{\partial E \left(\ln a_{F,k} | a_{F,k} \hat{W}_k \geq a_{S,k} \right)}{\partial M_k} &= \left(\frac{\sigma_F^2 - \rho \sigma_F \sigma_S}{\sigma^*} \right) \frac{\partial \lambda(c_{F,k})}{\partial M_k} \\ \frac{\partial E \left(\ln a_{S,k} | a_{F,k} \hat{W}_k < a_{S,k} \right)}{\partial M_k} &= \left(\frac{\sigma_S^2 - \rho \sigma_F \sigma_S}{\sigma^*} \right) \frac{\partial \lambda(c_{S,k})}{\partial M_k} \end{aligned} \quad (14)$$

As concentration decreases, the wage employment share increases and the self-employment share decreases, so that $\partial \lambda(c_{F,k}) / \partial M_k < 0$ and $\partial \lambda(c_{S,k}) / \partial M_k > 0$. If $\rho \sigma_F \sigma_S < \min \{ \sigma_F^2, \sigma_S^2 \}$, mean ability will decrease in the wage employment sector and increase in the self-employment one. If $\sigma_F^2 < \rho \sigma_F \sigma_S < \sigma_S^2$, mean ability will increase in both sectors. If $\sigma_F^2 > \rho \sigma_F \sigma_S > \sigma_S^2$, mean ability will decrease in both sectors.

From the above it follows that employer concentration also affect the earning gap between self-employed and wage workers. From equation (11) in the paper it follows that

$$\frac{\partial \hat{W}_k}{\partial M_k} = \frac{1}{\psi_{F,k} W_{S,k}} \frac{\partial MRPL_{F,k}}{\partial M_k} - \frac{MRPL_{F,k}}{W_{S,k}^2} \frac{1}{\psi_{F,k}} \frac{\partial W_{S,k}}{\partial M_k} - \frac{MRPL_{F,k}}{W_{S,k}} \frac{1}{\psi_{F,k}^2} \frac{\partial \psi_{F,k}}{\partial M_k} \quad (15)$$

where the first and second term are negative as $\partial MRPL_{F,k} / \partial M_k = z_k \partial P_{F,k} / \partial M_k < 0$ and $\partial W_{S,k} / \partial M_k > 0$, but, as explained above, the sign of $\partial \psi_{F,k} / \partial M_k$ in the third term is ambiguous.

A high number of entrants (higher M_k) decreases the sectoral earning gap, but it does so to a lower extent if labor market power decreases with concentration, i.e. $\partial \psi_{F,k} / \partial M_k < 0$. If instead labor market power increases with concentration ($\partial \psi_{F,k} / \partial M_k > 0$) then a higher number of entrants decreases the sectoral earning gap even further. Intuitively, more competition among formal firms decreases the earning gap, but the effect is shaped by changes in their labor market power which could potentially go in the opposite direction.

Equation (15) describes the effect of a change in the number of entrants on the earning gap per efficiency units of labor. The effect on the log of average earning gap at the worker level is given by

$$\frac{\partial \ln \hat{W}_k}{\partial M_k} + \frac{\partial E \left(\ln a_{F,k} | a_{F,k} \hat{W}_k \geq a_{S,k} \right)}{\partial M_k} - \frac{\partial E \left(\ln a_{S,k} | a_{F,k} \hat{W}_k < a_{S,k} \right)}{\partial M_k} \quad (16)$$

which makes clear how the effect of employer concentration on the average earning gap is given

by changes in relative earnings per efficiency unit of labor as well as changes in the composition of wage workers and self-employed workers and their average abilities. As shown by equation (14) and discussed above, these effects depend on the sign and size of the correlation between abilities as well as their variance in the population.

B.4 Proof of Common $\epsilon_F(\hat{W}_k)$ Across Markets

Under the assumption of common $\Delta\mu$ and Σ , the mean log ability in sector F is given by

$$E\left(\ln a_F | a_F \hat{W}_k \geq a_S\right) = \mu_{F,k} + \left(\frac{\sigma_F^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right) \lambda(c_{F,k}) \quad (17)$$

and the variance is given by

$$Var\left(\ln a_F | a_F \hat{W}_k \geq a_S\right) = \sigma_F^2 + \left(\frac{\sigma_F^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right)^2 [\lambda(c_{F,k})c_{F,k} - \lambda^2(c_{F,k})] \quad (18)$$

with $c_{F,k} = \frac{\ln \hat{W}_k + \Delta\mu}{\sigma^*}$. We want to show that the elasticity $\epsilon_F(\hat{W}_k) \equiv \partial \ln N^F(\hat{W}_k) / \partial \ln \hat{W}_k$ is independent of $\mu_{F,k}$.

Consider the following approximation

$$E(\log x) \approx \log[E(x)] - \frac{Var(x)}{2E(x)^2} \quad (19)$$

Consider also a first order Taylor series approximation of $\log x$ around $E(x)$

$$\log x \approx \log[E(x)] + \frac{x - E(x)}{E(x)} \quad (20)$$

from which, taking the variance of both sides, we get

$$Var(\log x) \approx \frac{Var(x)}{E(x)^2} \quad (21)$$

Combining (19) and (21) we get

$$E(\log x) \approx \log[E(x)] - \frac{1}{2}Var(\log x) \quad (22)$$

It follows that

$$\log N^F(\hat{W}) \approx E\left(\ln a_F | a_F \hat{W} \geq a_S\right) + \frac{1}{2}Var\left(\ln a_F | a_F \hat{W} \geq a_S\right) \quad (23)$$

and after plugging in the expressions in (17) and (18) we get

$$\log N^F(\hat{W}) \approx \mu_F + \left(\frac{\sigma_F^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right) \lambda(c_F) + \frac{1}{2} \left\{ \sigma_F^2 + \left(\frac{\sigma_F^2 - \rho\sigma_F\sigma_S}{\sigma^*}\right)^2 [\lambda(c_F)c_F - \lambda^2(c_F)] \right\} \quad (24)$$

whose derivative with respect to $\ln \hat{W}_k$ is independent of $\mu_{F,k}$.