

# Business Size, Development, and Inequality in Latin America

## A Tale of One Tail

*Marcela Eslava*  
*Marcela Meléndez*  
*Laura Tenjo*  
*Nicolás Urdaneta*



**WORLD BANK GROUP**

Latin America and the Caribbean Region  
Office of the Chief Economist  
October 2023



**Reproducible Research Repository**

A verified reproducibility package for this paper is available at <http://reproducibility.worldbank.org>, click [here](#) for direct access.

## Abstract

Using official employment surveys for 45 advanced economies and Latin American countries, this paper shows that the positive cross-country correlation between business size and GDP per capita is tighter than previously found using firm-level datasets and finds a close negative business size-Gini relationship. The paper also finds a closer connection between individual income and business size for

workers in less developed countries compared with those in advanced economies. Because employment data address the bias against the smallest productive units that characterize firm-level datasets, our approach uniquely assesses and highlights the dominance of the left tail of the business size distribution in less developed countries.

---

This paper is a product of the Office of the Chief Economist, Latin America and the Caribbean Region. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at [mmelendez@worldbank.org](mailto:mmelendez@worldbank.org), [meslava@uniandes.edu.co](mailto:meslava@uniandes.edu.co), [n.urdaneta@duke.edu](mailto:n.urdaneta@duke.edu), [lauratenjo@worldbank.org](mailto:lauratenjo@worldbank.org). A verified reproducibility package for this paper is available at <http://reproducibility.worldbank.org>, click [here](#) for direct access.

*The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.*

# Business Size, Development, and Inequality in Latin America: A Tale of One Tail

Marcela Eslava\*    Marcela Meléndez†    Laura Tenjo ‡    Nicolás Urdaneta §

*Keywords: business size, developing economies, firm-level datasets, self-employment, micro-enterprises, income inequality, Latin America.*

*JEL Codes: O54, E24, J21, J46*

\*Universidad de Los Andes, Bogotá, meslava@uniandes.edu.co

†The World Bank, mmelendez@worldbank.org

‡The World Bank, lauratenjo@worldbank.org

§Duke University and Universidad de Los Andes, n.urdaneta@duke.edu

We are extremely thankful to Gabriela Bojanini and Andrés Soto for their excellent data management and processing work. We thank audiences at the Latin American and Caribbean Economic Review (LACIR) meetings, the 2022 LACEA Annual Meeting, the LACEA Network on Inequality, Universidad de Los Andes, Universidad de Chile, CAF's research seminar, and the World Bank for useful comments.

# 1 Introduction

Understanding the facts on the size of productive units in developing economies is critical for uncovering barriers to development. The business size distribution reflects the incentives and difficulties businesses face to innovate and grow and is thus crucial to identifying distortions hampering aggregate productivity for a given productivity distribution across businesses (Restuccia and Rogerson (2008); Hsieh and Klenow (2009); Hopenhayn (2016, 1992) and a long literature following their footsteps). Thus, an important dimension of the literature addressing the development problem has been the documentation of business size distribution in less developed countries (LDCs) in contrast to advanced economies (AEs). Thanks to these efforts, we know that LDCs have firms that are smaller on average, even within sectors, and that uncovering this relationship requires using sources of data that appropriately cover the bottom tail of the firm size distribution (Bento and Restuccia, 2021, 2017, 2021; Poschke, 2018). Business size is also crucial in the debate on within-country inequality, which has lately focused on how large firms' (labor and product) market power affects workers' incomes, vis-a-vis business owners.<sup>1</sup> Although the emphasis has been on AEs, some recent efforts extend the analysis to LDCs (Brooks et al. (2021), Amodio and de Roux (2021), Amodio et al. (2022)), where inequality has been historically very high and remains so, despite a slightly decreasing trend after the year 2000.

The study of the business size distribution has relied heavily on firm-level data. While firm data has been crucial in developing these branches of the literature and has brought a great deal of useful knowledge, it has limitations that are particularly severe for LDCs, because even the most comprehensive firm datasets exclude large segments of the productive sector in these countries. We aim to address this bias. Using official employment/household survey data, we characterize the (employment-weighted) size distribution of the universe of productive units from which workers obtain their income in 45 countries, including advanced economies and LDCs, particularly (but not only) those in Latin America. In doing so, we also quantify the magnitude of the mentioned bias in firm-level datasets. Finally, taking advantage of the nature of our data, where the unit of observation is the individual, and the information is representative of the entire private economy, we characterize the relationship between the business size distribution and individuals' earnings for the overall working population. By working population, we mean the complete set of individuals in the (private) economy who declare obtaining income from productive activities, either as employees, self-employed workers or firm owners.

<sup>1</sup> The literature has documented the concurrence of increasing wealth concentration and income inequality, falling labor shares in national income, and increasing market power in advanced economies over the last decades. Piketty (2014); Piketty et al. (2017); Saez and Zucman (2020a,b); Autor et al. (2020); Karabarbounis and Neiman (2013); De Loecker and Eeckhout (2018); De Loecker et al. (2020); Yeh et al. (2022).

Our comprehensive characterization of the weighted firm size distribution shows that not only is the average business size smaller in LDCs, but that the most outstanding contrast with AEs is an extreme skewness towards tiny productive units. While 70% of the workforce in AEs is engaged in firms of 10 or more employees, in LATAM, that fraction is only 32%. Roughly half of the (68%) remaining workers in the region are self-employed without employees, while the other half work at (or own) micro-businesses with fewer than 10 employees. Firm-level databases, even economic censuses, neglect a significant fraction of these two groups either as a deliberate design decision or due to the difficulties in identifying these businesses and recording their information. For instance, the Mexican Economic Census, the region's most comprehensive dataset on businesses, arguably covering all private businesses outside agriculture, covers only 60% of workers in the private non-agricultural sector. Most of the omission has to do with workers in self-employment and in businesses with fewer than five workers (Busso et al., 2012). Much of the activity left out in LDCs is informal because own-account workers and micro-firms in these countries most frequently do not make social security contributions or register their businesses (three-quarters of this segment in our LATAM data).<sup>2</sup> Still, the omission occurs even in datasets that also intend to cover informal firms.

The data also show a tight correlation between individuals' earnings and business size in LATAM. As an extreme manifestation of this close relationship, the share of individuals in LATAM's bottom income decile working at firms of 10 or more employees is only 3%, while over 60% of individuals in that income decile are own-account workers without employees. Firm size and individual earnings are also correlated in AEs, which is consistent with models of firms that are heterogeneous in productivity and pay wages correlated with the marginal productivity of their workers, but the gradient is much weaker. The allocation of workers across size categories, rather than productivity (income) differences across them, explains 60% of the starker earnings gap of individuals in decile one versus the median in LATAM compared to the US.

These stylized facts have crucial implications. Our findings regarding individuals' earnings imply that the bias in firm-level datasets against the self-employed without workers and micro-enterprises, tends to exclude low-income workers (including people in poverty). This omission implies a critical limitation for those interested in understanding inequality in LDCs and the causes of their income gap relative to advanced economies. Research not covering activity in the far left tail of the firm size distribution leaves out precisely the workers in the region for whom the income and productivity gaps relative to citizens of richer countries and the median worker in their own country are starkest. This calls for research integrating information on the universe of businesses in the economy, including self-employment initiatives without employees,

<sup>2</sup> In fact, the ILO operational definition of the informal sector corresponds precisely to all productive units under the 10-employee threshold.

when explaining the region’s high inequality or low productivity. It also requires that policymakers and academics recognize that the emphasis on superstar firms currently dominating the inequality debate in the developed world directs the spotlight away from the segment of the economy where the incomes of a vast majority of Latin Americans originate.

In terms of the business size distribution, our work is closest to Poschke’s (2018), which departs from the use of firm-level datasets by using information at the entrepreneur level from the Global Entrepreneurship Monitor and Amadeus. Our use of the official employment/household surveys, the most comprehensive data sources on workforces at the national level, permits a full coverage of businesses of all sizes and motivations –including survival entrepreneurship lacking employment generation potential–. It also enables us to tie business size to the earnings of individuals associated with different-sized businesses. The prevalence of self-employment without employees as a defining characteristic of less developed economies, and its relationship with survival –rather than transformational– entrepreneurship, has been documented, among others, by Schultz (1990); Pietrobelli et al. (2004); Gindling and Newhouse (2014). Our analysis merges that fact into a more comprehensive view of the business landscape in LDCs versus AEs, and of how its characteristics are reflected in individual incomes.

The paper proceeds as follows. Section 2 presents the data. Section 3 shows our results, first characterizing the overall business size distribution (Section 3.1.) and then the relationship between personal earnings and the size of businesses where those earnings originate (Section 3.2.). The implications of our results are discussed in Section 4. Throughout the paper, we maintain a comparative perspective, contrasting LATAM with European economies, the US, and some Asian economies, including both those more and less advanced. The sample of comparators is dictated by data availability.

## 2 Data

The use of comprehensive data on the productive activities where individuals obtain their income, either as self-employed or as salaried workers or firm owners, is crucial to our research question. For this reason, we rely on microdata from the surveys used in different countries to generate official labor market statistics, which we refer to as ”employment surveys” even though in some countries they are labeled as household surveys. Given their nature and purpose, they are designed to be representative of the entire labor force in each country. Beyond collecting information on the work status and income of each adult individual in the household, the surveys frequently ask for the size, in number of workers, of the business where the person works or, in the case of employers, the number of people they employ.

These data are crucial for questions on how attributes of the business sector impact development or inequality. Their unparalleled strength is their comprehensiveness in covering all

work activities that generate income for the owners of production means in each economy. Its other unique advantage is the possibility of tying the individual income of the entire working population to the characteristics of the workplaces where such income is generated.

These strengths come with limitations. First, the unit of observation is not the business but the individual, and employers are not identified. Therefore, we can only characterize the employment-weighted business size distribution, not its unweighted version. Second, we are constrained to a reduced number of size bins, already limited in the original data sources and further reduced in our analysis for comparability across countries.<sup>3</sup> Third, because the data are at the level of individual workers, we have no information on capital stock or other business characteristics. Our only measurable business characteristics are size, sector, and labor productivity, measured from the earnings of employees and owners, which yield proxy measures of productivity distorted by the presence of labor market power and labor regulations.

We summarize here the basic characteristics of the data sources used for the different countries in our sample. Details are provided in the Appendix. For Latin America, we use the 2019 national household surveys, with the exception of Chile (2017) and Mexico (2018). The microdata for most countries are publicly available, and we use harmonized versions to ensure comparability across countries. We have data for 11 LATAM countries: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Mexico, Paraguay, Peru, and Uruguay.<sup>4</sup> We also have data for three LDCs in Asia (India, Pakistan, and Nepal) for which we use information from their 2019 labor force surveys (2018 in the case of Nepal).<sup>5</sup>

Our advanced economies include all countries in the European Union, the U.S., the Republic of Korea, Japan, and Australia. We use the 2019 microdata from the EU-SILC (Statistics on Income and Living Conditions) database for the European Union (EUROSTAT, 2019). We split E.U. countries between higher and lower income countries in the region.<sup>6</sup> For the remaining countries; our datasets are the U.S.' Current Population Survey Annual Social and Economic Supplement (CPS-ASEC; U.S. Census Bureau, 2018; Ruggles et al., 2023), Korea's Labour and Income Panel Study (KLIPS; Korea Labor Institute, 2019), Australia's Household Income and Labour Dynamics dataset (HILDA, Department of Social Services, 2019), and Japan's Labour Force Survey (Official

<sup>3</sup> In all countries, workers are asked whether they work on their own or in a business with more or fewer than 10 employees, but higher levels of disaggregation are captured using thresholds that are not always comparable across countries. We define comparable groups.

<sup>4</sup> The specific sources are: Argentina (INDEC, 2019), Bolivia (Bolivia, 2017), Brazil (IBGE, 2019), Chile (Ministerio de Desarrollo Social y Familia, 2017), Colombia (DANE, 2019), Costa Rica (INEC, 2019), the Dominican Republic (de la República Dominicana], 2019), Mexico (INEGI, 2018), Paraguay (INE Paraguay, 2019), Peru (INEI, 2019), and Uruguay (INE Uruguay, 2019)

<sup>5</sup> India's Periodic Labour Force Survey (National Statistical Office, 2018-2019), Pakistan's Labour Force Survey (Pakistan Bureau of Statistics, 2018-2019), and Nepal's Labour Force Survey (Central Bureau of Statistics, 2017-2018).

<sup>6</sup> Countries in each group are listed in the Appendix.

Statistics of Japan, 2019). For these countries, we rely on microdata from 2019, except for the U.S. (2018). For Japan we only have aggregates by business-size bins and thus can only produce the individual business size distribution (i.e., not the joint earnings-size distribution).

Our earnings variable includes all sources of income corresponding to individuals' occupations, all pre-tax. In all the data sources we use, workers classify themselves as self-employed, employees, or employers. For all employees and employers (business owners) we assign business sizes in number of workers as reported by the individual.<sup>7</sup> For consistency across countries, we restrict the sample to individuals who report deriving income from an economic activity, are at least 20 years old, and report positive earnings. We only report calculations including individuals in the private sector, but the stylized facts also hold when the public sector is included. Additional country-to-country details are provided in the Appendix (Table A.1.).

## 3 Results

### 3.1 Latin America's full business size distribution in perspective

Figure 1 depicts the employment-weighted business size distribution for Latin America and the other regions in our sample. Each bar shows the average across countries in the region for the corresponding figure, weighted by each country's workforce. We begin with Panel A, where businesses are classified according to whether they are self-employment initiatives or employer businesses above or below the 10-employee threshold. This is the most comprehensive panel in terms of country coverage because information on size bins defined by the 10-employee threshold is collected in most countries, including all those in our sample. The contrast between LATAM economies and all AEs in the sample is stark. While firms with at least 10 employees absorb, on average, 70% of the workforce in Europe, rich Asian economies, and the United States, they only explain 32% of employment in LATAM. The bulk of the region's employment —68%— is in micro-businesses or self-employment without employees. India, Pakistan, and Nepal display business size distributions similar to those in LATAM when only the urban workforce is considered.<sup>8</sup>

Of course, there is heterogeneity in business size distribution across countries within regions and across rich regions in the figure. Distributions by country are reported in the Appendix (Table

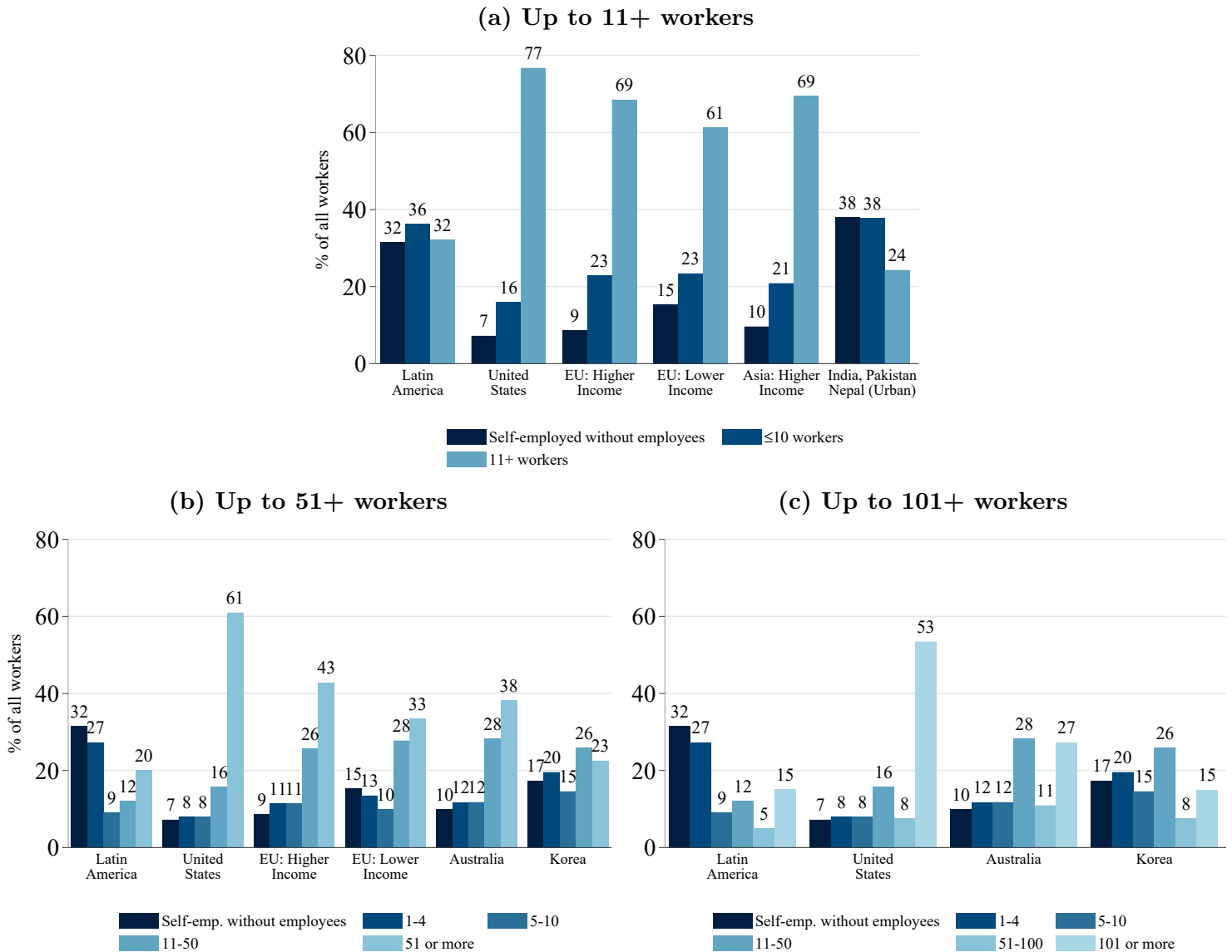
<sup>7</sup> In some countries, self-employed workers may answer the question of business size and may report a number of workers larger than one. This may happen for instance, for contractual workers or for co-owners of family businesses without employees. We do not use business size reports for self-employed workers.

<sup>8</sup> These three countries have over 60% of their overall population in the rural area, while Latin American countries have rates of urbanization similar to those in AEs (over 80%), and thus rural populations below 20%. We analyze only urban employment in these three countries but note that when including rural employment, the size distribution is entirely dominated by self-employment and micro-enterprises with fewer than 10 employees.



A.2.). The U.S. and Japan show the largest fraction of workers in firms above 10 employees with close to 80%. The figure is close to 70% in our higher-income countries in the E.U. and Asia, and 61% in E.U. countries with a per capita GDP below the regional median. Polar cases within their regions, for a low fraction of workers absorbed by firms with 10+ employees, are Korea (49%) in high-income Asia, Bolivia (13%) in LATAM, and Greece (37%) in lower-income E.U. In LATAM, this figure ranges from Bolivia's 13% to Chile's 50%. But, in all LATAM countries, this share is below the corresponding figure for any country in the more affluent regions of the world covered by our sample, with the only exception of Greece. Also, Korea comes close to Chile. Countries in South Asia are similar to LATAM in the large shares of self-employment but have a thicker segment of wage employment in the 1 to 10 workers range, and a thinner segment at the top.

**Figure 1: Full business size distribution (employment weighted): Latin America (11 countries) versus other regions.**



**Note:** Each bar is a weighted average of countries, weighting by total workers. Firm size categories for Chile, Costa Rica, Uruguay, Europe, Asia, and the United States are categorized as 1-9 and 10 or more. Panel B: rather than cut off at 50, Argentina cuts at 40 and Costa Rica at 30. Panel C: For Brazil, Chile, and Uruguay, 51-100 versus 100+ split is imputed using the average distribution across those two categories in the rest of LATAM. **Sources:** Authors' calculations using employment or household surveys representative of the labor force. Details in the Appendix.

The other two panels of Figure 1 present the business size distribution for more disaggregate size bins. Care must be exercised when comparing panels since we lose progressively more countries as we use more disaggregate size bins, but this more disaggregate comparison sheds light on interesting differences across regions. First, in LATAM, the vast majority (27% out of 36%) of workers in businesses of 1-10 employees fall in the 1-4 workers category, while more prosperous

regions display roughly similar employment shares in the 1-4 versus 5-10 worker bins. Second, the share of workers in LATAM is furthest from those in AEs in the largest firm category in each panel. For instance, in the 50+ employees group, the figure is 20% for LATAM compared to 33% in lower-income E.U. countries, 43% in higher-income E.U., and 61% in the U.S. And, in the 100+ category, the figure is 15% in LATAM and 53% in the U.S. (with data lacking for Europe for this cutoff). Finally, these results suggest a missing middle in LATAM, at least in terms of the employment-weighted business size distribution. The distribution has a U-shape, even if asymmetric (i.e., a larger mass in the far left than in the far right tail, although data limitations keep us from separately capturing the truly superstar corporations).<sup>9</sup> The idea of a missing middle in LATAM’s business size distribution, put forward by Tybout (2000), has been disputed on the basis that this pattern is identifiable only for specific sectors and when specific size bins are used to characterize the data (Hsieh and Olken, 2014). Beyond an adherence or lack of to the “missing middle” terminology, it is apparent in Figure 1 that the business size distribution in LATAM is not simply shifted to the left relative to richer economies; the much lower employment absorption by large firms in the region is not replaced by employment in SMEs but rather by own-account work and micro-businesses.

The much higher weight of self-employment and micro-businesses in LATAM’s employment relative to wealthier economies remains true within economic sectors (detailed results reported in the Appendix). Both categories weigh particularly high in LATAM’s services, with self-employment carrying a higher weight in wholesale, retail, and hospitality services and 1-4 employee businesses capturing a large share of employment in high-skilled services. Manufacturing has larger businesses in all regions (Bento and Restuccia (2021)). Even in LATAM, 10+ employee businesses absorb the majority of manufacturing employment (55%), but this is still much lower than in more prosperous economies (above 85% in Europe and the U.S.).

The previous results align with the positive cross-country correlation between business size and GDP per capita estimated by Bento and Restuccia (2021) and Poschke (2018) but imply an even stronger correlation when self-employment activities without employees are accounted for.<sup>10</sup> In Figure 2, which pools together all countries in our sample, the share of self-employment without workers and workers in micro-businesses of up to 10 employees, falls sharply as income increases (top two left panels). The regression coefficient for the standardized share of workers who are self-employed on standardized per capita GDP is -0.799 (top left panel), while that for workers in businesses with under 10 workers is negative but smaller (-0.598, second left panel from top). Inversely, the workforce share in larger businesses of 10-50 employees or 50+ employees increases

<sup>9</sup> The U-shape in Figure 1 is underestimated because two of the LATAM countries in the sample use smaller cutoffs for the 10-50 categories: Costa Rica cuts at 30 and Argentina at 40, rather than 50 employees.

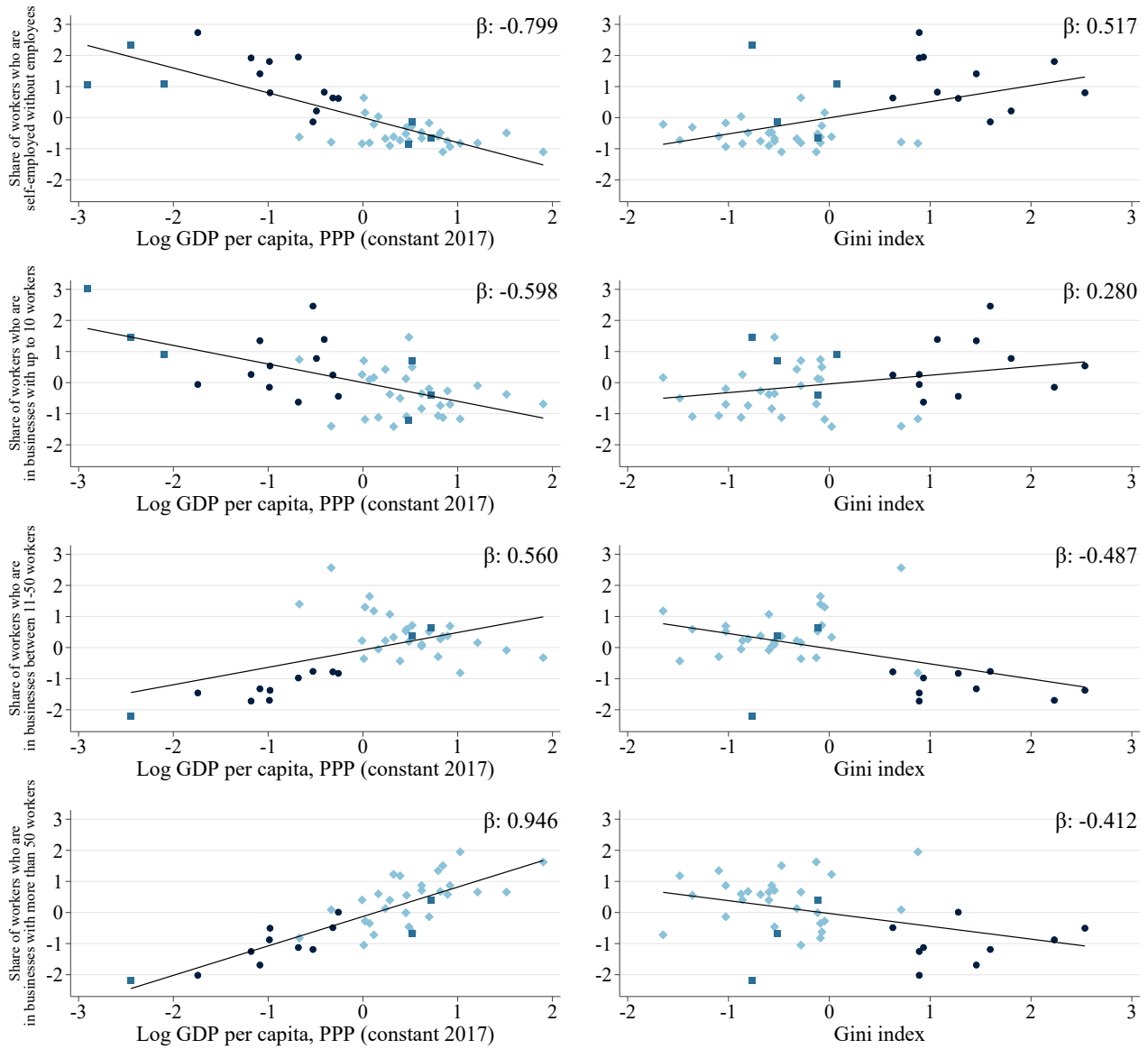
<sup>10</sup> Because our data are at the level of individuals and not businesses, we cannot reproduce the elasticity of average size to GDP. However, we can produce an elasticity for the share of employment in each size bin to GDP.

with income. The (standardized) coefficient of employment share on GDP per capita is large and positive for both size categories, as observed in the two bottom left panels of Figure 2. The relationship is stronger for workers in businesses with 50+ employees: a coefficient of 0.946 for the 50+ category versus 0.560 for the employment share in firms between 11 and 50 employees.

These correlations with GDP per capita, and the difference in Figure 1 across more and less advanced economies, make it tempting to conclude that the differences in size distributions reflect stages of development. However, economies that are rich today did not display distributions as skewed as those in our sample's LDCs when at similar GDP levels. For instance, in 1940, the U.S. had a per capita GDP similar (in purchasing power parity, PPP, terms) to that of Peru today. Yet, today 48% of Peru's workers are self-employed without employees, while the corresponding figure in the U.S. in 1940 was 21% (see Appendix table A3 for a more detailed historical comparison). Additionally, the distribution of salaried employment across business sizes in the U.S. shows little change from that of the 1970s, the first decade for which we have this detailed information.

Correlograms on the right-hand side panel of Figure 2 present the complementary story of the inequality-business size relationship. The share of workers in smaller businesses has a strong positive correlation to the Gini coefficient. The regression coefficient is 0.517 for the standardized share of self-employed workers without employees on the Gini coefficient, and 0.280 when the dependent variable is the share of employment in micro-businesses of up to 10 employees (top two right panels). From the complementary size categories, countries with more workers in small, medium, and large firms, and less micro-businesses have more equal income distributions measured by Gini coefficients (bottom right panels). There is no contradiction between these observations and recent literature emphasizing the role played by large firms and market concentration in increasing inequality. That literature emphasizes the role of giant firms on inequality, while our findings point to a negative correlation between inequality and the prevalence of small, medium, and large firms. This as opposed to the relationship between inequality and self-employment and micro-businesses, in a context where these tend to exhibit very low relative productivity, within their countries and compared to AEs. In fact, the regression coefficient is slightly smaller for the 50+ category than for the 11-50 one (-0.412 and -0.487, respectively, bottom two panels), which already points to a nonlinearity at the top end of business sizes. Our results, however, do emphasize that recent findings highlighting the risks growing giant corporations may pose to equality, cannot be oversimplified into a generalized statement that the weight of larger firms in employment is broadly associated with higher inequality. This is the case especially in the context of LDCs, where even small firms of, say, 10-50 workers are relatively large for their context.

**Figure 2: Share of working population in different business size classes, versus GDP per capita and Gini coefficient**



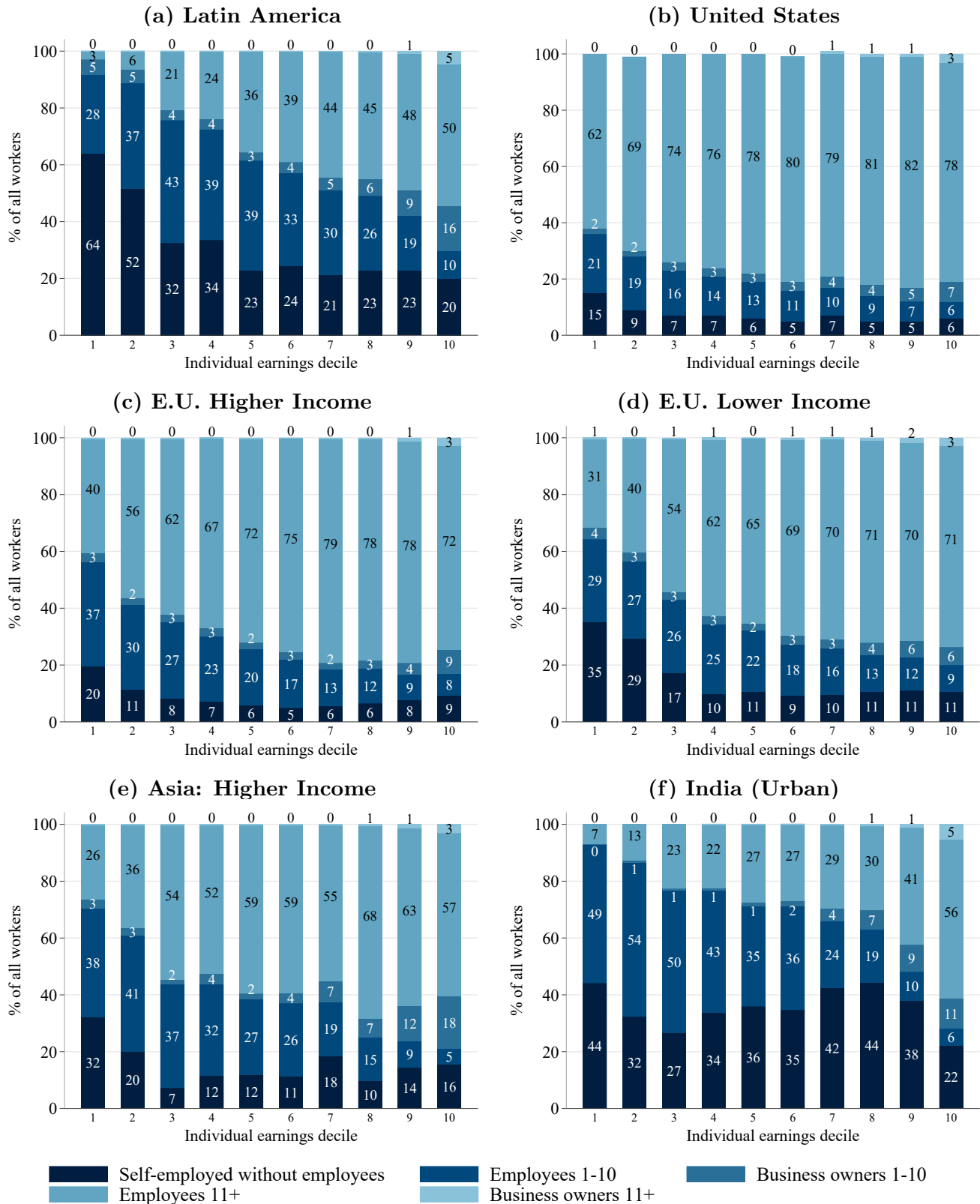
● Latin America    ◆ Europe and U.S.    ■ Asia and Australia    — Linear fit

**Note:** Both variables in each graph have been standardized. Each dot represents a country in the sample. Japan and Nepal are missing in the right-hand panels because their Gini indices are unavailable in the WDI. Argentina, Costa Rica, Japan, India, and Nepal are missing in the bottom two panels because of the lack of separate reports for the 50+ category. The coefficient below each panel corresponds to the linear regression coefficient of the y-axis variable on the x-axis variable. **Sources:** Data on Gini and GDP per capita are from World Development Indicators (World Bank, 2023). GDP per capita and Gini Indices are the last publicly available for each country between 2016 and 2019.

### 3.2 The joint distribution of individual earnings and business size

Figure 3 displays the joint distribution of business size, individual earnings, and functional position (employee-owner) for the different regions in our sample. LATAM displays a very steep gradient of individual earnings to business size class, in contrast to the U.S. and Europe. For instance, while for LATAM, the 10+ category (owners+employees) goes from capturing only 3% of the workforce in the bottom earnings decile to 55% in the top decile, a much weaker transition is observed in the U.S. (62% to 81%) and in above-mean-income E.U. economies (40% to 75%). Another interesting peculiarity of LATAM versus richer countries is the much higher prevalence of business ownership, even beyond self-employment. This is also a reflection of the fragmentation of the business sector in micro-sized units in LDCs. The higher fraction of individuals in the workforce who own businesses is true in all earnings deciles as well as for the different business size categories in the figure (1-10 employees, 10+ employees, and self-employment without employees). Interestingly, in India (the only LDC outside LATAM for which we have microdata on earnings), self-employment without employees carries a similar weight across all income deciles, rather than being strongly (negatively) correlated with earnings, as in LATAM.

Figure 3: Cross distributions of individual earnings and firm size (employment-weighted).



Note: This Figure shows the distribution of business sizes and employer/owner roles by deciles of personal earnings. The figure for each region shows the average across countries, weighted by the workforce. We do not include Pakistan and Nepal because comparable income measures cannot be produced.

Table 1 displays the results of regressing individual monthly earnings against employer size for our LATAM sample (self-employment without employees is the left-out category). It shows that the close connection between individual earnings and employer size remains after controlling for individual characteristics. Regression coefficients for size categories increase monotonically with size, with or without controls. The average earnings in the 50+ size bin roughly double those of the self-employed with no workers group, in regressions without additional controls. Evaluating these numbers at average individual characteristics, the gap is barely reduced when incorporating controls for gender, age, education, or economic sector. The additional regressor (other than employer size) with the highest explanatory power is years of schooling. It increases the regression R2 from 0.06 to 0.16 in regressions that already control for age and sex, and reduces the mean predicted earnings gap between the self-employed and the 50+ categories to a factor close to 1.8, as opposed to the 2.1 obtained in the unconditional regression.



**Table 1: Individual monthly earnings in PPP 2019 U.S. dollars as a function of business size and other characteristics. Latin America (11 countries).**

	1	2	3	4	5	6
<i>Business size</i>						
1 to 4	72.28 (0.97)	114.05 (0.97)	172.53 (0.91)	175.84 (1.05)	172.86 (0.99)	161.28 (1.02)
5 to 10	343.96 (1.35)	376.38 (1.35)	287.81 (1.26)	321.42 (1.34)	277.02 (1.27)	260.83 (1.30)
11 to 50	505.85 (1.14)	538.09 (1.15)	381.18 (1.08)	454.91 (1.15)	363.51 (1.09)	342.85 (1.16)
51 or more	690.66 (1.05)	717.52 (1.06)	500.58 (1.01)	601.98 (1.09)	470.45 (1.05)	452.09 (1.18)
Years of schooling			-118.93 (0.30)		-117.42 (0.31)	-112.83 (0.31)
Years of schooling <sup>2</sup>			11.92 (0.01)		11.65 (0.02)	11.21 (0.02)
Sex (woman=1)		-248.61 (0.70)	-349.21 (0.66)	-298.08 (0.77)	-345.80 (0.73)	-342.27 (0.76)
Age		43.85 (0.15)	47.79 (0.14)	43.77 (0.14)	46.70 (0.14)	46.16 (0.13)
Age <sup>2</sup>		-0.40 (0.00)	-0.40 (0.00)	-0.39 (0.00)	-0.38 (0.00)	-0.38 (0.00)
Constant	612.58 (0.63)	-361.61 (3.22)	-644.90 (3.36)	-590.70 (3.76)	-644.83 (3.81)	-116.32 (16.83)
R <sup>2</sup>	0.04	0.06	0.16	0.09	0.17	0.19
Obs with factor weights	168,830,240	168,830,240	168,830,240	168,830,240	168,830,240	168,830,240
Obs in dataset	758,388	758,388	758,388	758,388	758,388	758,388
Sector FE	NO	NO	NO	YES	YES	YES*

**Note:** The table presents coefficients from regressions of individuals' earnings on business size categories and other characteristics. Regressions are estimated separately for each country. Reported coefficients correspond to a weighted average of individual country coefficients, where the weight for each country is the total number of country observations. Standard errors (in parentheses) are the square root from the weighted average of the squares of the standard errors. Regressions exclude workers still enrolled in education. Columns 4 and 5 include one-digit sector fixed effects for all countries. Column 6 includes sector-level fixed effects at the most disaggregate level available for each country. **Sources:** Authors' calculations done using National Household Surveys for 2019 (2017 for Chile and 2018 for Mexico).

To quantify the importance of workforce allocation across different business sizes in overall earnings inequality, we decompose inequality into its extensive (i.e., allocation) and intensive margins. We do so comparatively for LATAM and for one of our reference economies, the US. The extensive margin captures the role played by the differential allocation of workers across business size categories, while the intensive margin captures the role of earnings differentials between those categories, proxying for productivity differentials. In particular, we decompose the gap between the overall median of the earnings distribution and the average earnings of each earnings decile into within -and between- size category components. We index these categories by  $j$ , which indicates

size according to the following five-category classification:

$size = \{ \text{self-employed without employees, employee in a business with 1-10 employees, owner of a business with 1-10 employees, employee in a business with 10+ employees, owner of a business with 10+ employees} \}$

More formally, indexing individuals by  $i$ , denoting as  $y_i$  the ratio between the earnings distribution median and individual  $i$ 's earnings, and denoting the mean of  $y_i$  within a given size class as  $\bar{y}_{j,D1} = \frac{\sum_{i \in (j,D1)} y_i}{N_{j,D1}}$ , the average median-to-own earnings gap for a given decile of the earnings distribution (D1) can be written as:

$$\bar{y}_{D1} = \sum_{j \in size} \frac{N_{j,D1}}{N_{D1}} \bar{y}_{j,D1} = \underbrace{\frac{\sum_{j \in size} \bar{y}_{j,D1}}{5}}_{within} + \underbrace{\sum_{j \in size} (\bar{y}_{j,D1} - \frac{\sum_{j \in size} \bar{y}_{j,D1}}{5}) (\frac{N_{j,D1}}{N_{D1}} - \frac{\sum_{j \in size} \frac{N_{j,D1}}{N_{D1}}}{5})}_{between} \quad (1)$$

where  $N_{D1}$  is the number of individuals in D1, and  $N_{j,D1}$  is the number of individuals in category  $j$  within D1. For brevity, we focus on the distribution's extremes, analyzing deciles 1 and 10 exclusively, denoted respectively as D1 and D10. For D10, we define  $y_i$  as the ratio between the earnings of individual  $i$  and the median of the earnings distribution (i.e., we define it inversely so that the absolute gap is always measured vis-a-vis the median).

The within term captures the average deviation from median earnings across business size categories, while the between term captures the allocation of workers across categories with larger/smaller earnings deviations from the median. Equation 1 underlines high inequality, measured by the gap between average earnings in the extreme (top or bottom) deciles and median earnings, can be explained by a large gap versus median earnings in the average size category, by the concentration of workers in categories with the largest earnings gaps, or by a combination of these. We counterpose the LATAM and U.S. results and ask to what extent greater inequality in LATAM is explained by the cross-category allocation of workers, or by earnings gaps versus the median in each category.

The first key takeaway from this decomposition is that, in LATAM, the between term contributes positively to the gap between earnings and median earnings in decile 1, increasing inequality, but diminishes that gap in decile 10. This is seen in Table 2, where we report results for the different terms of decomposition 1. Earning gaps are reported only for categories that absorb at least 0.5% of employment. The top panel of the table corresponds to LATAM, weighing country-level decompositions by the corresponding number of workers, while results for the U.S. are reported in the middle panel. Columns 1-3 show results for the bottom earnings decile, while columns 4-6 do so for decile 10. In the row marked "Total", columns 3 and 6 report  $\bar{y}_{D1}$  (left-hand side of equation 1), while columns 2 and 5 report the within term,  $\frac{\sum_{j \in size} \bar{y}_{j,D1}}{5}$ .

The between component (residual between the two just mentioned) contributes positively to the gap between median earnings and D1 earnings in LATAM: the total gap is 10.97, of which

9.81 is attributable to the within component of equation 1 and the remaining 1.08 to the between one. That is, the allocation of workers across categories adds to the huge distance between the earnings of the most disadvantaged individuals and the median. The positive contribution of the between term arises because most D1 workers fall under the own-account category. This segment exhibits the largest earnings gap versus the median, as observed in the other rows of the top panel each reporting, for size category  $j$ ,  $\frac{N_{j,D1}}{N_{D1}}$  in column 1,  $\bar{y}_{j,D1}$  in column 2 and  $\frac{N_{j,D1}}{N_{D1}}\bar{y}_{j,D1}$  in column 3. Median earnings in the region are a huge 11.89 times the average earnings of own-account workers without employees, which is by far the category that absorbs the largest share of workers in the decile (64%).

Meanwhile, the between component *reduces* this inequality measure for D10 in LATAM, from 6.59 of the within component to 5.51 (column 5 versus 6). The large weight of self-employment without employees and employment at micro-enterprises plays the role of reducing inequality in this top decile precisely because these categories have less productivity, and thus display a lower gap between mean and median earnings. In the U.S., meanwhile, allocation contributes negatively to the gap versus median earnings, both at the top and bottom deciles. It is only in the specific case of LATAM’s lower-income workers that their allocation across business sizes exacerbates inequality for given average gaps versus median earnings in each size category  $j$ . The between term adds to inequality in LATAM’s D1 more than it reduces inequality in the region’s D10 or in either of the extremes of the US’ earnings distribution.

Another main message from this decomposition is that differences in the between terms explain a large part (about one-third) of the larger average earnings gap versus the median that LATAM’s D1 workers exhibit relative to those in the U.S. As reported in the “Total” row, our results imply a much larger earnings gap relative to the median in D1 in LATAM versus the U.S. (column 3). While, as mentioned, the earnings median for LATAM is 10.97 times the average; in the U.S., the factor is a much milder 6.21. This is a 4.76  $\bar{y}_{D1,LATAM} - \bar{y}_{D1,US}$  difference (bottom panel, column 2). This large positive gap is consistent with the greater inequality in LATAM and with poverty’s crucial role in the intensity of inequality in the region.<sup>11</sup> Moreover, around two-thirds of that LATAM versus U.S. difference is accounted for by a larger LATAM gap in earnings versus the median for the average segment  $j$ : 9.89 in LATAM versus 6.49 for the U.S. (column 2), accounting for 3.41 (or 71%) of the 4.76 difference between the two regions. The remaining 29% (close to one-third) is explained by the between component of equation 1.

<sup>11</sup> Close to 30% of the population in the region today falls below the poverty line.

**Table 2: Bottom and top decile earnings versus the median: within and between size categories**

Size class ( $j$ )	Decile 1			Decile 10		
	$\frac{N_{j,D1}}{N_{D1}}$	$\bar{y}_{j,D1}$	$\frac{N_{j,D1}}{N_{D1}} * \bar{y}_{j,D1}$	$\frac{N_{j,D10}}{N_{D10}}$	$\bar{y}_{j,D10}$	$\frac{N_{j,D10}}{N_{D10}} * \bar{y}_{j,D10}$
	(1)	(2)	(3)	(4)	(5)	(6)
<b>Latin America</b>						
Self-employed without employees	0.640	11.89	7.59	0.199	4.78	0.95
Worker in $\leq 10$ employee business	0.278	8.71	2.22	0.098	4.15	0.40
Worker in 11+ employee business	0.029	8.38	0.21	0.498	5.29	2.61
Owner of $\leq 10$ employee business	0.053	10.59	0.95	0.159	6.08	0.98
Owner of 11+ employee business	–	–	–	0.046	12.65	0.57
<b>Total</b>	<b>100%</b>	<b>9.89</b>	<b>10.97</b>	<b>100%</b>	<b>6.59</b>	<b>5.51</b>
<b>United States</b>						
Self-employed without employees	0.151	7.11	1.07	0.056	7.27	0.41
Worker in $\leq 10$ employee business	0.210	6.31	1.33	0.062	5.69	0.36
Worker in 11+ employee business	0.615	5.94	3.65	0.780	5.24	4.09
Owner of $\leq 10$ employee business	0.024	6.61	0.16	0.073	6.02	0.44
Owner of 11+ employee business	–	–	–	0.029	5.93	0.17
<b>Total</b>	<b>100%</b>	<b>6.49</b>	<b>6.21</b>	<b>100%</b>	<b>6.03</b>	<b>5.46</b>
<b>Latin America - United States</b>						
Self-employed without employees	0.489	4.78	6.51	0.142	-2.49	0.54
Worker in $\leq 10$ employee business	0.068	2.40	0.89	0.036	-1.54	0.05
Worker in 11+ employee business	-0.587	2.44	-3.44	-0.281	0.05	-1.48
Owner of $\leq 10$ employee business	0.030	3.98	0.79	0.086	0.06	0.54
Owner of 11+ employee business	–	–	–	0.017	6.72	0.40
<b>Total</b>	<b>–</b>	<b>3.40</b>	<b>4.76</b>	<b>–</b>	<b>0.56</b>	<b>0.05</b>

**Note:** This table presents results for the different components of the decomposition equation (1), for each business size class. The “Total” row in column (3) reports the overall own-to-median income gap in the respective decile, while in column (2) it reports its within component. The gap between the two is the “between” or “allocation” component. In the other rows, corresponding to individual size categories, column (3) is the product of columns (1) and (2) for a given country (not directly evident in the top panel because the figure shown is an average across countries). The bottom panel presents differences between the middle and top panels. Sources and list of LATAM countries included in the top panel are as in Figure 1 .

The large role played by this between component in LATAM’s larger inequality versus the U.S. reflects the fact that, within D1 workers, the region assigns a much larger fraction of the workforce to the categories with the lowest earnings (i.e., the largest earnings gaps versus the median), compared to the U.S. For D1 individuals, self-employment and ownership of businesses with below 10 employees are the categories that exhibit the largest (negative) gaps versus median

earnings. These gaps are much larger in LATAM than in the U.S.: 11.89 versus 7.11 times, for self-employed workers without employees, and 10.59 versus 6.61 for micro-business owners (column 2). It is precisely these large-gap categories that LATAM devotes many more workers to than the U.S.: self-employment without workers absorbs 64% in LATAM versus 15% in the U.S., while the figures for micro-business ownership are 5.3% in LATAM and 2.4% in the U.S. Meanwhile, salaried workers in businesses with above 10 employees are closest to the median (a 5.94 gap in the U.S. and 8.38 in LATAM), but the category only absorbs 3% of D1 workers in the region versus 62% in the U.S.

The story is quite different at the top of the earnings distribution (columns 4-6 of Table 2). There, the allocation of workers across size categories (the between component) **reduces** the gap between LATAM and the U.S. in the average earnings of decile 10 relative to the median, from 0.56 in the within component to 0.05 in total (bottom panel). This is because the region devotes a much larger fraction of its workforce to self-employment and salaried employment in micro-businesses. Within D10, these groups exhibit much smaller earnings gaps versus the median than in the U.S. In both regions, the largest earnings gaps versus the median in D10 are for owners of businesses with above 10 employees, followed by owners of businesses with less than 10 employees (column 5). LATAM displays larger gaps in both categories, although a much larger one for owners of 11+ employees businesses (12.65 LATAM versus 5.93 U.S.) than for micro-business owners (6.08 versus 6.02). It also displays a slightly larger gap for workers in businesses with above 10 employees (5.29 versus 5.24).

In summary, the employment distribution in LATAM, highly skewed towards own-account workers without employees and micro-businesses, adds to inequality in the region in absolute terms and to the fact that inequality is starker than in the U.S. Moreover, because this effect is explained by the large negative earnings gaps of the poorest workers, disproportionately absorbed by these business categories, the implicit misallocation of resources towards tiny low-productivity business units is particularly problematic.

## 4 Discussion

Our results have important implications for pressing policy debates in Latin America. In the quest to identify the most critical barriers to growth, our findings call for a focus on factors that can explain the vast prevalence of self-employment and micro-businesses and their impact on aggregate productivity and workers' incomes. This, as well as for empirical and quantitative research based on data comprehensive of all segments of businesses, simultaneously covering self-employment, microenterprises, small and medium-sized enterprises (SMEs), and large corporations.

Regarding the inequality debate, which is crucial for LATAM as one of the world's most unequal regions, our results call for an emphasis broader than the focus on the monopsony power

of superstar firms in the labor market that dominates the current global policy debate. Superstar firms are rarer in LATAM, and low individual incomes from self-employment and micro-entrepreneurship (where monopsony power in the labor market is not the relevant factor) play a more critical role in explaining income inequality in the region. Larger firms in LATAM indeed have market power, and there is evidence that they operate with higher markdowns in the labor market than their smaller peers (De Loecker et al., 2020; Amodio and de Roux, 2021; Amodio et al., 2022).<sup>12</sup> But our results show that workers in these firms fall into the middle and upper sections of the income distribution in LATAM, while the lowest incomes are those of individuals who do not have access to jobs at firms that can exercise market power. In the region, thus, a first-order issue in terms of the policy response to inequality and poverty is facilitating the emergence of a layer of small, medium, and large firms able to absorb workers who would otherwise earn low incomes as self-employed or work in low-productivity micro-businesses.

The stylized facts we provide must also be factored into the discussion on entrepreneurship policy in developing economies. While in rich countries, the self-employed tend to receive high earnings (e.g., Bhandari et al., 2022), most of those among the vast mass of entrepreneurs and own-account workers in LATAM belong at the bottom of an income distribution that is already significantly to the left of that in advanced economies. The policy debate and part of the economic literature on LATAM frequently overlook this differential characteristic of entrepreneurship in the region, leading to an emphasis on programs that subsidize (or otherwise foster) entrepreneurship irrespective of its income and employment potential.

## References

- AMODIO, F. AND N. DE ROUX (2021): “Labor Market Power in Developing Countries: Evidence from Colombian Plants,” Documentos CEDE 019267, Universidad de los Andes. Facultad de Economía CEDE.
- AMODIO, F., P. MEDINA, AND M. MORLACCO (2022): “Labor Market Power, Self-Employment, and Development,” IZA Discussion Papers 15477, Institute of Labor Economics (IZA).
- AUTOR, D., D. DORN, L. F. KATZ, C. PATTERSON, AND J. VAN REENEN (2020): “The Fall of the Labor Share and the Rise of Superstar Firms,” *The Quarterly Journal of Economics*, 135, 645–709.

<sup>12</sup> Market power of the larger firms and corporations in LATAM also translates into political power and the capacity to influence policymaking perpetuating and amplifying inequality (Schneider (2021), Fairfield (2015)).

- BENTO, P. AND D. RESTUCCIA (2017): “Misallocation, Establishment Size, and Productivity,” *American Economic Journal: Macroeconomics*, 9, 267–303.
- (2021): “On average establishment size across sectors and countries,” *Journal of Monetary Economics*, 117, 220–242.
- BHANDARI, A., T. J. M. TOBEY KASS, E. R. MCGRATTAN, AND E. SCHULZ (2022): “On the Nature of Entrepreneurship,” *SOI Working Papers, Internal Revenue Service*.
- BOLIVIA, I. (2017): “Encuesta de Hogares,” Instituto Nacional de Estadística, Bolivia.
- 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.
- BUSSO, M., M. V. FAZIO, AND S. LEVY (2012): “(In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico,” Research Department Publications 4789, Inter-American Development Bank, Research Department.
- CENTRAL BUREAU OF STATISTICS (2017-2018): “Nepal Labour Force Survey,” National Planning Commission, Central Bureau of Statistics, Government of Nepal.
- DANE (2019): “Gran Encuesta Integrada de Hogares, GEIH,” Departamento Administrativo Nacional de Estadística, Colombia.
- DE LA REPÚBLICA DOMINICANA], B. C. (2019): “Encuesta Nacional Continua de la Fuerza de Trabajo, ENCFT,” Banco Central, República Dominicana.
- DE LOECKER, J. AND J. EECKHOUT (2018): “Global Market Power,” Working Paper 24768, National Bureau of Economic Research.
- DE LOECKER, J., J. EECKHOUT, AND G. UNGER (2020): “The Rise of Market Power and the Macroeconomic Implications,” *The Quarterly Journal of Economics*, 135, 561–644.
- DEPARTMENT OF SOCIAL SERVICES (2019): “Household, Income and Labour Dynamics in Australia Survey, HILDA,” Department of Social Services, Government of Australia.
- EUROSTAT (2019): “European Union Statistics on Income and Living Conditions, EU-SILC,” EUROSTAT.
- FAIRFIELD, T. (2015): *Private Wealth and Public Revenue in Latin America: Business Power and Tax Politics*, Cambridge University Press.
- GINDLING, T. H. AND D. NEWHOUSE (2014): “Self-employment in the developing world,” *World development*, 56, 313–331.

- HOPENHAYN, H. (2016): “Firm Size and Development,” *Economía Journal*, 17, 27–49.
- HOPENHAYN, H. A. (1992): “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” *Econometrica*, 60, 1127–1150.
- HSIEH, C.-T. AND P. J. KLENOW (2009): “Misallocation and manufacturing TFP in China and India,” *The Quarterly journal of economics*, 124, 1403–1448.
- HSIEH, C.-T. AND B. A. OLKEN (2014): “The Missing “Missing Middle”,” *Journal of Economic Perspectives*, 28, 89–108.
- IBGE (2019): “Pesquisa Nacional por Amostra de Domicílios Contínua,” .
- INDEC (2019): “Encuesta Permanente de Hogares, EPH,” Instituto Nacional de Estadística y Censos, República Argentina.
- INE PARAGUAY (2019): “Encuesta Permanente de Hogares Continua, EPHC,” Instituto Nacional de Estadística, Paraguay.
- INE URUGUAY (2019): “Encuesta Continua de Hogares,” Instituto Nacional de Estadística, Uruguay.
- INEC (2019): “Encuesta Nacional de Hogares, ENAHO,” Instituto Nacional de Estadística y Censos, Costa Rica.
- INEGI (2018): “Encuesta Nacional de Ingresos y Gastos de los Hogares,” Instituto Nacional de Estadística y Geografía, México.
- INEI (2019): “Encuesta Nacional de Hogares, ENAHO,” Instituto Nacional de Estadística e Informática, Perú.
- KARABARBOUNIS, L. AND B. NEIMAN (2013): “The Global Decline of the Labor Share,” *The Quarterly Journal of Economics*, 129, 61–103.
- KOREA LABOR INSTITUTE (2019): “Korean Labor & Income Panel Study, KLIPS,” Korea Labor Institute.
- MINISTERIO DE DESARROLLO SOCIAL Y FAMILIA (2017): “Encuesta de Caracterización Socioeconómica Nacional, CASEN,” Ministerio de Desarrollo Social y Familia, Chile.
- NATIONAL STATISTICAL OFFICE (2018-2019): “Periodic Labour Force Survey,” Ministry of Statistics and Programme Implementation, National Statistical Office, Government of India.
- OFFICIAL STATISTICS OF JAPAN (2019): “Labour Force Survey,” Ministry of Internal Affairs and Communication.



- PAKISTAN BUREAU OF STATISTICS (2018-2019): “Labour Force Survey,” Pakistan Bureau of Statistics, Government of Pakistan.
- PIETROBELLI, C., R. RABELLOTTI, AND M. AQUILINA (2004): “An empirical study of the determinants of self-employment in developing countries,” *Journal of International Development*, 16, 803–820.
- PIKETTY, T. (2014): *Capital in the Twenty-First Century*, Harvard University Press.
- PIKETTY, T., E. SAEZ, AND G. ZUCMAN (2017): “Distributional National Accounts: Methods and Estimates for the United States,” *The Quarterly Journal of Economics*, 133, 553–609.
- POSCHKE, M. (2018): “The firm size distribution across countries and skill-biased change in entrepreneurial technology,” *American Economic Journal: Macroeconomics*, 10, 1–41.
- RESTUCCIA, D. AND R. ROGERSON (2008): “Policy distortions and aggregate productivity with heterogeneous establishments,” *Review of Economic Dynamics*, 11, 707–720.
- RUGGLES, S., S. FLOOD, M. SOBEK, D. BROCKMAN, G. COOPER, S. RICHARDS, AND M. SCHOUWEILER (2023): “IPUMS USA: Version 13.0 [dataset],” Minneapolis, MN: IPUMS, 2023.
- SAEZ, E. AND G. ZUCMAN (2020a): “The Rise of Income and Wealth Inequality in America: Evidence from Distributional Macroeconomic Accounts,” *Journal of Economic Perspectives*, 34, 3–26.
- (2020b): “Trends in US Income and Wealth Inequality: Revising After the Revisionists,” Working Paper 27921, National Bureau of Economic Research.
- SCHNEIDER, B. R. (2021): “Big business and the inequality trap in Latin America: Taxes, Collusion, and Undue Influence,” *UNDP LAC Working Paper, 5, 2021, Background paper for the Regional Human Development Report for Latin America and the Caribbean 2021*.
- SCHULTZ, T. P. (1990): “Women’s changing participation in the labor force: a world perspective,” *Economic Development and Cultural Change*, 38, 457–488.
- TYBOUT, J. R. (2000): “Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?” *Journal of Economic Literature*, 38, 11–44.
- U.S. CENSUS BUREAU (1943): “Employment and Personal Characteristics,” 1940 Census of Population: The Labor Force (Sample Statistics).
- (1953): “General Characteristics,” 1950 Census of Population: Volume 2. Characteristics of the Population.

——— (1962): “Employment Status, Weeks Worked, Occupation, and Industry for the Population of the United States: 1960,” 1960 Census of Population, Supplementary Report.

——— (1972a): “Industry of Employed Persons for the United States: 1970,” 1970 Census of Population, Supplementary Report.

——— (1972b): “Weeks Worked, Class of Worker, Last Occupation of the Experienced Unemployed, and Labor Mobility for the United States: 1970,” 1970 Census of Population, Supplementary Report.

——— (2018): “Consumer Population Survey, ASEC,” United States Census Bureau.

WORLD BANK (2023): “World Development Indicators, WDI,” World Bank.

YEH, C., C. MACALUSO, AND B. HERSHBEIN (2022): “Monopsony in the US Labor Market,” *American Economic Review*, 112, 2099–2138.

# Appendix

## A.1 Country-by-country data description and analysis

Table A.1: Size categories and surveys by country.

Country	Region	Size categories					Source
		1-4	5-10	11-50	51-100	101 or more	
Argentina	Latin America	1-4	5-10	11-40	41-100	101 or more	Encuesta Permanente de Hogares 2019
Bolivia	Latin America	1-4	5-10	11-50	51-100	101 or more	Encuesta de Hogares 2019
Brazil	Latin America	1-5	6-10	11-50		51 or more	Pesquisa Nacional por Amostra de Domicílios Contínua 2019
Chile	Latin America	1-5	6-9	10-49		50 or more	Encuesta de Caracterización Socioeconómica Nacional 2017
Colombia	Latin America	1-5	6-10	11-50	51-100	101 or more	Gran Encuesta Integrada de Hogares 2019
Costa Rica	Latin America	1-4	5-9	10-29	30-99	100 or more	Encuesta Nacional de Hogares 2019
Dominican Republic	Latin America	1-4	5-10	11-50	51-99	100 or more	Encuesta Nacional Continua de la Fuerza de Trabajo 2019
Mexico	Latin America	1-5	6-10	11-50	51-100	101 or more	Encuesta Nacional de Ingresos y Gastos de los Hogares 2018
Paraguay	Latin America	1-5	6-10	11-50	51-100	101 or more	Encuesta Permanente de Hogares Continua 2019
Peru	Latin America	1-4	5-10	11-50	51-100	101 or more	Encuesta Nacional de Hogares 2019
Uruguay	Latin America	1-4	5-9	10-49		50 or more	Encuesta Continua de Hogares 2019
Austria	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Belgium	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Cyprus	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Denmark	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Finland	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
France	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Germany	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Ireland	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Italy	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Luxembourg	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Malta	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Sweden	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Switzerland	EU: Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Bulgaria	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Croatia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Czech Republic	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Estonia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Greece	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Latvia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Lithuania	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Poland	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Portugal	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Romania	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Serbia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Slovakia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Slovenia	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
Spain	EU: Less Advanced	1-4	5-10	11-49		50 or more	EU SILC 2019
United States	United States	1-4	5-9	10-49	50-99	100 or more	IPUMS CPS-ASEC 2018 + BDS 2018 to open the 1-9 category
Australia	Asia: High Income	1-4	5-9	10-49	50-99	100 or more	HILDA Survey 2019
Japan	Asia: High Income	1-4	5-9		10-99	100 or more	Labor Force Survey 2019 - Public Statistics
Korea	Asia: High Income	1-4	5-9	10-49	50-99	100 or more	KLIPS Survey 2019
India	South Asia	1-5	6-9	10-19		20 or more	Periodic Labour Force Survey 2018-19
Nepal	South Asia	1-5	6-9	10-19		20 or more	Nepal Labour Force Survey 2017-18
Pakistan	South Asia	1-5	6-9	10-49	50-100	101 or more	Pakistan Labour Force Survey 2018-19

**Table A.2: Size distribution by country**

	Argentina	Bolivia	Brazil	Chile	Colombia	Costa Rica
Self-employed without employees	32	60	31	29	46	23
1-4	29	19	24	17	20	25
5-10	12	8	9	7	6	10
11-50	11	10	11	16	8	14
51-100 / 51+ (Brazil, Chile)	5	2	25	33	2	10
101 or more	11	1	-	-	17	18
	Dominican Republic	Mexico	Paraguay	Peru	Uruguay	Latin America
Self-employed without employees	48	17	40	48	29	32
1-4	13	41	32	20	20	27
5-10	8	10	9	10	10	9
11-50	14	16	11	8	16	12
51-100 / 51+ (Uruguay)	3	4	3	3	25	5
101 or more	13	11	5	11	-	15
	Austria	Belgium	Cyprus	Denmark	Germany	Finland
Self-employed without employees	8	11	8	6	3	17
1-4	11	8	24	9	6	11
5-10	14	9	18	12	11	15
11-50	26	20	24	29	26	27
51 or more	41	52	26	45	55	30
	France	Ireland	Italy	Luxembourg	Malta	Switzerland
Self-employed without employees	10	12	16	3	12	7
1-4	14	11	19	10	13	13
5-10	10	12	13	10	6	13
11-50	24	22	29	20	23	24
51 or more	43	42	23	56	45	42
	Sweden	EU: Higher Income	Bulgaria	Croatia	Czech Republic	Estonia
Self-employed without employees	12	9	8	7	15	6
1-4	9	11	9	18	7	12
5-10	12	11	5	12	9	12
11-50	25	26	45	25	28	32
51 or more	42	43	34	38	41	38
	Greece	Latvia	Lithuania	Poland	Portugal	Romania
Self-employed without employees	29	7	10	20	9	22
1-4	21	15	8	9	18	11
5-10	13	13	6	7	14	5
11-50	20	37	26	22	25	34
51 or more	17	27	50	41	34	28
	Serbia	Slovakia	Slovenia	Spain	EU: Lower Income	United States
Self-employed without employees	10	16	9	12	15	7
1-4	19	15	14	17	13	8
5-10	15	14	9	12	10	8
11-50	35	33	19	27	28	16
51 or more (EU) / 51-100 (US)	20	22	50	32	33	8
101 or more	-	-	-	-	-	53
	Australia	Korea	Japan	India	Pakistan	Nepal
Self-employed without employees	10	17	7	35	54	35
1-4	12	20	8	26	30	44
5-10	12	15	8	10	12	13
11-50 / 11-19 (India, Nepal)	28	26	13	7	4	4
51-100 / 20+ (India, Nepal)	11	8	15	21	0	4
101 or more	27	15	49	-	0	-

**Note:** Criteria as in Figure 1. **Sources:** National household or employment surveys 2019 (other year in some cases). Details in Table A.1

## A.2 Self-employment along the development path

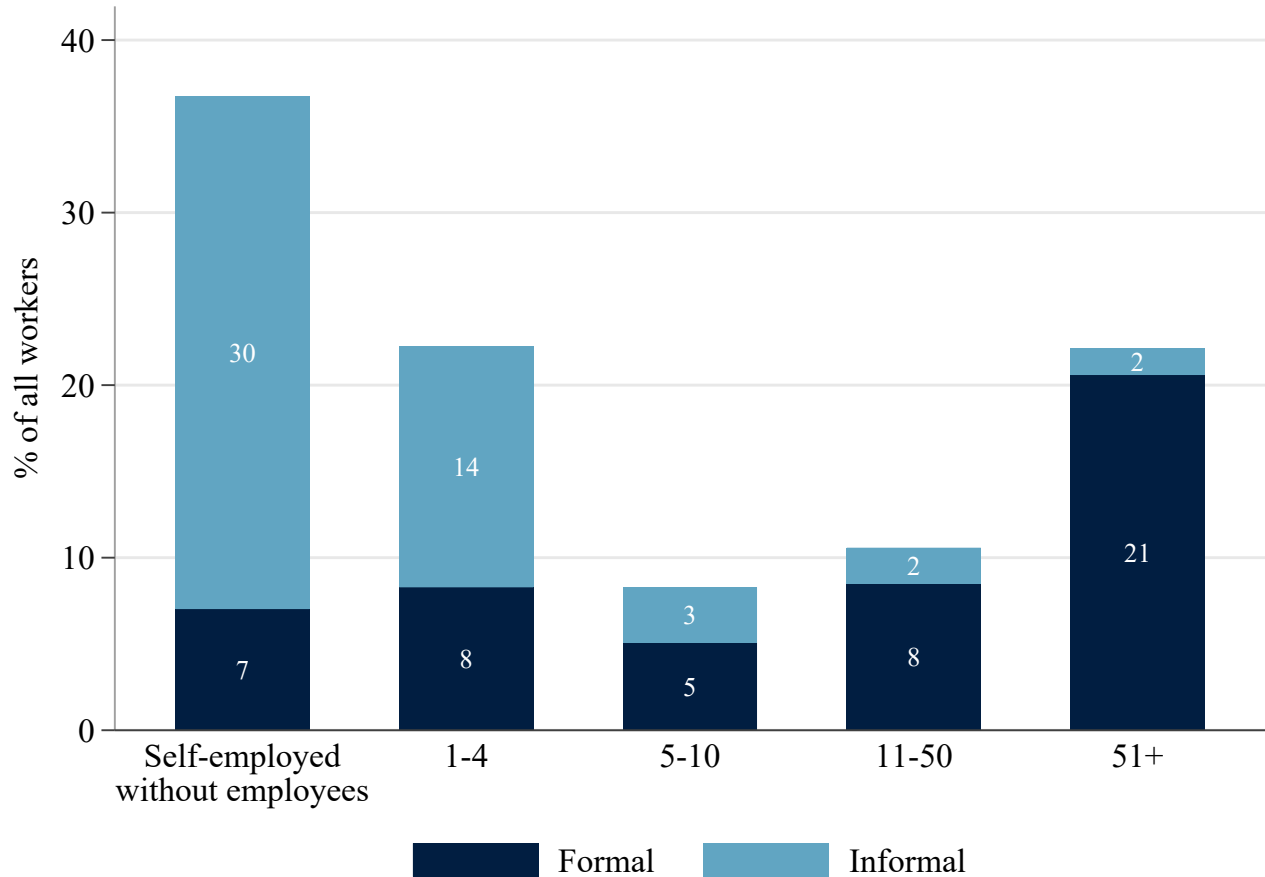
**Table A.3: Self-employment in Latin America in 2019 compared to self-employment in the United States between 1940 and 1970**

	US self-employment	US self-employment without workers in government	Per capita PPP GDP equivalent in 2018	GDP equivalent self-employment (2019)	Latin America self-employment (2019)
<b>1940</b>	21	-	Peru	48	32
<b>1950</b>	17	19	Costa Rica	23	32
<b>1960</b>	12	14	Argentina	32	32
<b>1970</b>	10	13	Chile	29	32

**Note:** GDP equivalent refers to the country in the Latin American sample with the PPP GDP per capita in 2018 closest to the GDP per capita in the US in 1940, 1950, 1960 and 1970. **Sources:** Self-employment: US Census of Population [U.S. Census Bureau (1943), U.S. Census Bureau (1953), U.S. Census Bureau (1962), U.S. Census Bureau (1972a), and U.S. Census Bureau (1972b)], Latin America: National household or employment surveys 2019 (other year in some cases - Details in Table A.1). GDP per capita: Our World in Data.

### A.3 Informality

Figure A.1: Firm size distribution and labor informality in Latin America (8 countries).

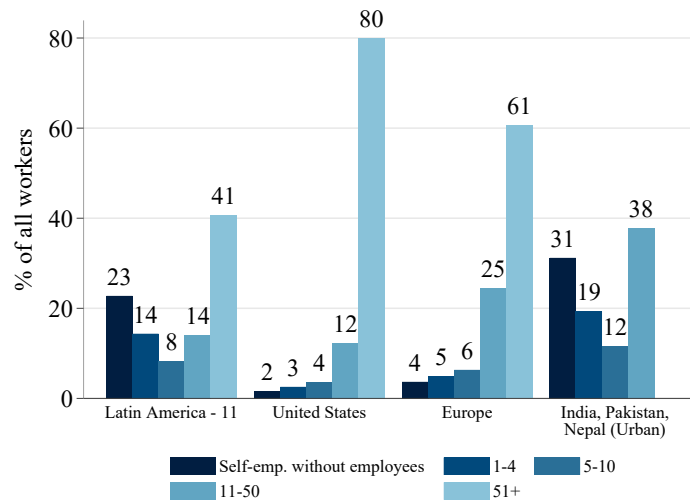


**Note:** Criteria as in Figure 1. **Sources:** National household or employment surveys 2019 (other year in some cases). Details in Table A.1. This figure excludes Argentina, Costa Rica and Mexico for Latin America due to comparability in informality. Informal workers defined as those not contributing to a pension system.

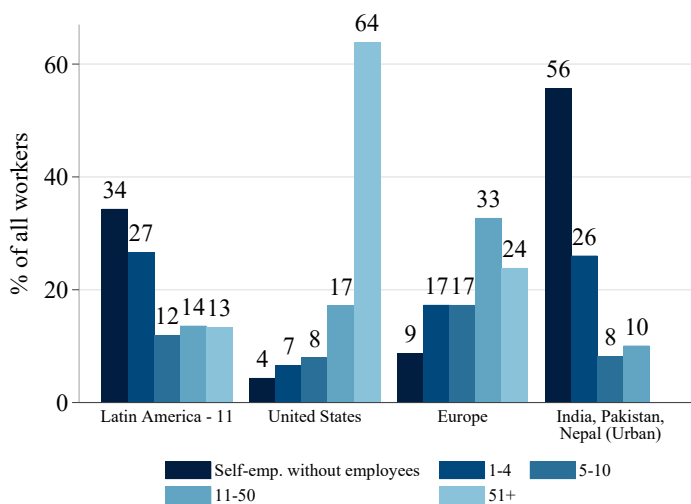
## A.4 Size distribution by economic sectors

Figure A.2: Full business size distribution by economic sectors (employment weighted): Latin America (11 countries) vs. other regions. Up to 51+ workers.

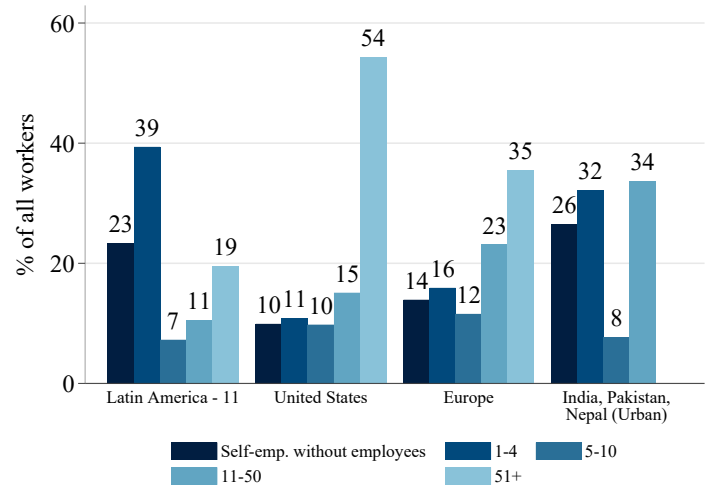
(a) Industry: Mining, Manufacturing, and Utilities



(b) Wholesale, Retail, Accommodation, and Food services



(c) High Skilled Services



**Note:** Each bar is a weighted average of countries, weighting by total workers in the sector of each country. For India, Pakistan, and Nepal the fourth bar groups 11-50 and 51+ categories as their surveys do not have this breakdown. High Skill Services include: Information and Communication, Financial and Insurance Activities, Real Estate Activities, Professional, Scientific and Technical Activities, Administrative and Support Service Activities, Arts, Entertainment and Recreation, Other Service Activities, Activities of Households, and Activities of Extraterritorial Organisations. Definitions of sectors are limited by region-specific definitions in the data. Other criteria as in Figure 1. Details in Table A.1.