Skill Supply, Firm Size, and Economic Development

Charles Gottlieb [*] Markus Poschke [†] Michael Tueting [‡]

Preliminary.
This version: October 30, 2025

Abstract

The organization of production varies widely across countries, with firms being substantially smaller in low-income countries. At the same time, educational attainment is lower in low-income countries. How are these two patterns related? In this paper, we examine the relationship between skill endowment and firm size across different stages of economic development. Using harmonized labor force data from 54 countries, we measure the skill intensity of employment by firm size and document four key facts. First, we show that the share of employment in large firms is about twice as high in high-income countries as in low-income countries. Second, across countries, large firms employ more skilled workers than small firms. Third, small firms in high-income countries are nearly as skill-intensive as large firms, but small firms in low-income countries employ far fewer skilled workers than their larger counterparts. Fourth, the skill gap between small and large firms is narrow when the skill premium is low, but it widens substantially when the premium is high. These findings suggest that small firms can easily substitute high-skill for low-skill workers when skilled workers are scarce and expensive, whereas large firms are less flexible. As a result, lower availability of high-skilled workers restricts the prevalence of large firms in low-income economies. We then use a span-of-control model with worker skill heterogeneity and two technologies (large- and small-scale) to analyze the impact of skill endowments on firm size distribution and economic development. Calibrated to U.S. data and varying only skill endowments to match those of low-, middle-, and high-income countries, the model replicates observed employment patterns by firm size and the skill intensity of firms across different stages of development. We interpret this as evidence that differences in skill endowments are a central driver of firm size distribution. Our findings imply that skill accumulation promotes development directly, by increasing productivity, and indirectly, by enabling the expansion of larger, more productive firms.

The organization of production differs widely across countries. In the United States, more than 80% of employment is in firms with ten or more employees. This number is only 40% in low-income countries. It is often argued that the scarcity of large firms in low-income countries makes wage employment less attractive relative to self-employment, contributing to lower overall productivity.

In this paper, we argue that operating large firms requires a workforce with a certain level of education, and we examine how the relatively lower educational attainment in low-income countries influences employment patterns and productivity.

Managing larger organizations requires basic administrative functions such as record-keeping and written communication. These tasks require a certain level of skill from a broad base of workers, not just from managers or specialized staff. This makes worker skills, even at modest levels of education, relevant for the effective operation of large firms. Although the connection between education and the operation of large firms may seem intuitive, empirical analysis has been constrained by the lack of internationally comparable data linking worker skills to employer size. As a result, our understanding of how limited human capital in low-income countries shapes their firm size distributions and productivity remains incomplete. Previous research has had to rely on two separate sources of information: one

[*] Aix-Marseille School of Economics, Université Aix-Marseille; charles.gottlieb@univ-amu.fr

[†] Department of Economics, McGill University; markus.poschke@mcgill.ca

[‡] Department of Economics, University of St.Gallen; michael.tueting@unisg.ch

This version: 14 August 2025. Niklas Halfar provided outstanding research assistance. The authors gratefully acknowledge funding from STEG-CEPR LRG Grant 943, which supported this research.

measuring national "skill endowments" through surveys and administrative records (e.g., UNESCO's Harmonized Schooling Series, OECD's PIAAC Assessments, the World Bank's STEP Surveys and LSMS/DHS Household Modules), and another reporting employment by firm size via business registers and enterprise surveys (e.g., the World Bank Enterprise Surveys, UNIDO INDSTAT, OECD Structural Business Statistics and Orbis). Hence, existing data allow measuring countries' skill endowments and provide information on employment by firm size, but do not allow measuring the skill mix of employment in firms of different sizes.

To overcome this knowledge gap, we build a new dataset on the skill composition of employment in small and large firms, harmonizing information from nationally representative labor force and household surveys from 54 countries at all stages of development. These surveys primarily collect individual-level data on income, employment status, and education, but also include questions about the size of respondents' employers. We leverage this feature to harmonize firm size information alongside individual skill data. Our main specification defines "skilled" workers as those with more than 9 years of schooling, typically corresponding to education beyond lower secondary school. We choose this threshold to ensure comparability across countries, especially since many low-income countries have very limited numbers of college-educated workers. In a robustness check, we also define "skilled" as having completed more than high school (i.e., upper secondary education in international terms). This alternative skill definition does not meaningfully alter our results or main conclusions.

Using this new dataset, we establish four facts. First, we show that the share of employment in large firms is about twice as high in high-income countries as in low-income countries. This finding is consistent with previous research, which shows that firms tend to be larger in wealthier countries [Bento and Restuccia, 2017, 2021, Gollin, 2008, Poschke, 2018]. Our remaining findings leverage the joint observation of worker skills and firm size. Second, across countries, employees of large firms are more skilled than those of small firms. Third, whereas small firms in rich countries are almost as skill-intensive as large firms, small firms in low-income countries employ many fewer skilled workers than their larger counterparts. Fourth, whereas small firms are almost as skill-intensive as large firms when the wage premium of being skilled is low, they are much less skill-intensive when the skill premium is high.

These patterns are consistent with a world in which skills are scarce in less wealthy economies, and large firms adopt more skill-intensive technologies; therefore, they cannot readily substitute low-skill labor for high-skill labor. As a result, when skilled workers are relatively more costly, large firms cut back on their use of such workers less than small firms do. Small firms in low-income countries significantly lag behind large firms in the share of skilled workers, with a gap of over thirty percentage points (and around twenty points in middle-income countries). In contrast, in high-income countries, such as the United States, this gap narrows to just five percentage points.

To better understand the data and the role of skill supply in firm growth and the broader organization of production, we develop a new heterogeneous firm macro model. The model is in the tradition of Hopenhayn [1992]. It features two sectors that differ not only in optimal scale, as in Buera et al. [2011], but also in factor intensity and potentially also their elasticity of substitution between low- and high-skilled labor, as in the representative firm models of Acemoglu and Guerrieri [2008] and Alvarez-Cuadrado et al. [2017]. Firms in the model produce with both low- and high-skill workers and choose between a large-scale technology and a small-scale technology. The optimal choice of technology depends on a firm's productivity and the prices of its inputs. In this setting, a lower skill endowment in the labor market has two effects. First, it raises the price of skills and causes all firms to use fewer skilled workers. A second effect extends beyond this: a greater skill premium leads to fewer firms utilizing large-scale technology.

To quantify the strength of these effects, we follow a standard approach in the macroeconomic literature on cross-country productivity differences. We calibrate the model using U.S. data and then examine the impact of varying skill endowments on firms' demand for skills, the firm size distribution, and productivity.

Our first finding is that as greater skill scarcity raises the skill premium, small firms strongly substitute towards low-skill workers. Large firms reduce their employment of skilled workers much less. As a result, skill intensity varies little with firm size in rich countries, but strongly in low-income countries. Although the model has been calibrated only to U.S. data and we only vary the aggregate skill endowment, the skill intensity of large versus small firms predicted by the model aligns closely with the facts observed in cross-country data. Second, a higher cost of skill makes running the large-scale technology less attractive. As a result, there are fewer large firms in low-income countries. The share of employment in large firms drops from around 80% in rich to around 10% in low-income countries. Again, this contrast is very close to the data patterns and is only driven by the change in the skill endowment.

Changes in technology have implications for output and productivity. We ask: How much would U.S. output decline if firms chose the smaller-scale technologies that operate in low-income countries? We find that, without any change in aggregate productivity or skill endowments, U.S. output would decline by 6% simply due to the use of smaller-scale technologies. This number corresponds to 20% of the total difference in net output between low-income economies and the U.S. in our model.

Before concluding, we relate our work to the literature on misallocation. Many scholars have attributed differences in firm sizes across countries to size- or productivity-dependent distortions [Guner et al., 2008, Restuccia and Rogerson, 2008]. Similarly, it is common to interpret dispersion in labor productivity across firms in a country as evidence of distortions. Yet, in our model, such differences arise endogenously, since firms operating different technologies optimally choose different levels of labor productivity. We show that when skills are scarce and expensive, the optimal size of large-scale firms is reduced more, raising their relative labor productivity. This greater gap is entirely due to optimal choices. We also show that introducing productivity-dependent distortions into our model has a similar effect to reducing skill endowments.

The remainder of the paper is structured as follows. In section 1, we relate our contribution to the literature. We describe the data in section 2 and discuss firm size and skill measurement. In section 3, we document cross-country patterns of employment by skill and firm size, as well as wage premia. Section 4 introduces a new model of heterogeneous firms. In section 5, we lay out our calibration strategy. In section 6, we report results from our counterfactual analysis. Section 7 relates our findings to the literature on misallocation. Section 8 concludes.

1 Literature

Our work is motivated by a literature showing that the production structure in low-income countries differs from that in high-income countries. Low-income economies are characterized by high levels of self-employment [Gollin, 2008] and smaller firms [Bento and Restuccia, 2017, 2021, Poschke, 2018]. Policy makers consider a lack of "good jobs" in large firms a development challenge, as evidenced by the World Bank's Development Report on Jobs [World Bank Development Committee, 2025].

The common data sources used by these authors are inherently limited: firm register data omit individual employee characteristics, and firm-level surveys, although occasionally capturing workforce skills, typically exclude informal and/or small firms. For instance, the World Bank Enterprise Surveys only cover formal (registered) companies with five or more employees.[1] We argue that this is a selected sample of firms, which can be misleading for low-income countries, where small and informal firms are prevalent and often constitute the majority of employment. In contrast, all labor force and household surveys we use are nationally representative and include information on individuals' characteristics, educational attainment, employment type, and employer firm size independently of whether the firm is registered or not. By definition, we also only ever see operational firms, which is another advantage of our data and circumvents the issue of inflating the share of small firms by including non-operational firms, as is common in firm-level surveys.

An alternative branch of literature aiming to understand the sources of differences in production structure has either attributed these differences to a set of frictions or distortions or has seen them as an optimal reaction to a different environment [Davis et al., 2023]. A large literature explored the effects of specific distortions on the efficiency of resource allocation and aggregate productivity, in particular, entry costs [Moscoso Boedo and Mukoyama, 2012, Poschke, 2010], labor market regulation [Hopenhayn and Rogerson, 1993, Poschke, 2009, Ulyssea, 2010], financial frictions [Buera et al., 2011,

[1] See https://www.enterprisesurveys.org/en/methodology

Midrigan and Xu, 2014], or delegation frictions [Akcigit et al., 2021, Grobovšek, 2020, Guner et al., 2018]. A parallel literature has diagnosed the existence of generic wedges or distortions that reduce aggregate productivity, particularly for large firms [Bartelsman et al., 2013, Hsieh and Klenow, 2009, Restuccia and Rogerson, 2008]. At the same time, others have argued that small firm sizes may be an optimal response to a different environment, for example, in terms of the level of capital [Gollin, 2008] or technology [Poschke, 2018]. To our knowledge, none of this work has addressed the effect of skill differences on the production structure across countries.

Our work also closely relates to a recent literature that has revisited the importance of human capital for cross-country income differences [Bils et al., 2024, Caselli and Ciccone, 2013, Hendricks and Schoellman, 2023, Jones, 2014]. This literature has mostly taken an aggregate perspective, and not taken the analysis to the firm level. To the best of our knowledge, only Hjort et al. [2023] analyze the effect of skill costs on firm sizes and aggregate productivity, focusing specifically on middle managers, and exploiting evidence from a single global firm.

In related work, Engbom et al. [2024] and Gottlieb et al. [2024] study how skill supply shapes the occupational composition of employment and aggregate productivity across countries. Our work differs from these studies in both the facts we document and the theoretical and quantitative analysis.

2 Data and measurement

This section outlines the data sources we use for our empirical analysis, as well as the choices we make to measure skills, firm size, and wages.

2.1 Primary data sources

We build a harmonized dataset that provides information on employment by firm size and workers' characteristics for a large set of countries. The harmonized dataset draws on nationally representative household and labor force surveys. All the surveys we use provide information on (i) individual characteristics (age and sex), (ii) education level, and (iii) firm size of the employer. It encompasses other existing harmonized cross-country datasets, such as the European Union's Statistics on Income and Living Conditions (SILC) and the Integrated Public Use Microdata Series (IPUMS International). We expand on this by identifying numerous additional surveys that provide the required information, which we then source and harmonize.

Overall, our dataset consists of 450 country-year surveys across 54 countries and spans the income per capita distribution, ranging from USD PPP 871 (Rwanda 2000) to 62313 (Norway 2012).[2] Appendix Table A-1 lists the full set of countries, years, and survey names that our dataset entails.

2.2 Sample selection

We restrict our analysis to the working-age population (age 15-65). For our main results, we only consider wage workers. In a robustness exercise, we include self-employment in small firm employment to account for a significant share of "own-account" workers (self-employed individuals without employees) in low-income countries.

Although our dataset contains multiple survey-year observations for many countries, we primarily use the cross-section around 2015. This is done to compare countries at different stages of development at the same time and to avoid potentially confounding time-varying factors, such as economic cycles or global crises (e.g., the COVID-19 pandemic).

For a subset of countries, we have information on wages (31 countries). To increase the statistical power in documenting our main findings on employment by firm size and skill, we utilize all surveys around 2015, including those without wage information (54 countries). However, we demonstrate in Appendix Table B-1 that the main facts we document remain equally valid for the more restricted sample of countries with wage information.

Appendix Table A-2 shows which years are used in the cross-section and which surveys include wage information.

[2] Data on GDP per capita comes from Feenstra et al. [2015].

2.3 Measurement

Our two main variables of interest are firm size and education.

Worker skill. Our dataset contains information about a worker's demographics and educational background. We use data on the completed degree and years of education to determine whether a worker is skilled or not. We define individuals with nine or fewer years of formal education as "unskilled," and those with more than nine years as "skilled". In most countries, this coincides with completing lower secondary education as defined by the International Standard Classification of Education (ISCED category 2), which typically corresponds to the transition point in the education system from a generalist education to subject-oriented instruction.[3]

Establishment size. All surveys we draw from ask wage workers the following question: "*How many employees work in your place of work (establishment/work site)?*". The answers provided to this question are generally in bins. We harmonize answers to this question into two consistent categories: small and large. Small firms are defined as those with fewer than ten employees, and large firms are defined as those with at least ten employees. This is the most common method by which labor force and household surveys collect information on employer firm size. If a survey provides more bins, we assign individuals to either of these two categories, provided the bins are consistent with the above thresholds.[4] This threshold may be perceived as low to some; however, as we show, around 60% of total wage employment is in firms with fewer than 10 employees in low-income countries and around 20% of total wage employment in high-income countries. These employment shares also imply that the *number of small firms* exceeds the *number of large firms* by far. In the U.S., for example, approximately 20% of employment is in small firms, while 80% of all firms are small.

Job characteristics. Our dataset also provides information on the job type and sector of employment of an individual's main job. Thus, we can distinguish wage employment from self-employment, and further between unpaid work, own-account work, and employers. We also observe whether individuals work in agriculture, manufacturing, or services, which we use in robustness checks.

3 Cross-country evidence

In this section, we use our dataset to document how the organization of production varies across countries. We exploit the joint distribution of employment by firm size and skill to document four facts. Figure 1 and Table 1 show our main empirical results, which we discuss below.

3.1 Employment by firm size

Fact 1: Large firms account for a much greater share of employment in high-income countries - roughly twice as much - than in low-income countries. Figure 1a shows the share of low-skill and high-skill workers employed in small and large firms across four income groups. The income groups correspond to low-income [$0, $3,000], lower-middle-income [$3,000, $10,000], upper-middle-income [$10,000, $30,000], and high-income [$30,000, ∞] categories. The figure shows that in low-income countries, 43% of wage workers are employed in firms with more than nine employees.

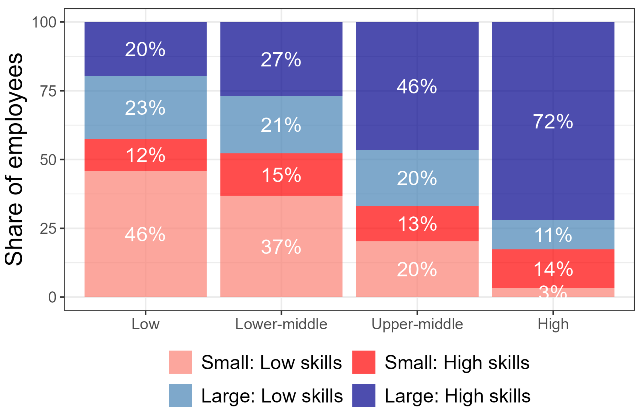
[3] An alternative would be to consider only college-educated individuals as skilled. This is a common convention in the analysis of labor markets in rich countries. Our choice of threshold responds to the fact that in low-income countries, only a very small share of the population is college educated. Moreover, our initial argument emphasizes the importance of skill in participating productively in a large organization. Plausibly, the most important skills for this dimension are sufficiently advanced literacy and numeracy, for which our threshold is a natural choice. In a robustness check, we also document the facts for a higher threshold of 12 years of schooling.

[4] We can also consistently capture employment in medium-sized firms of 10-49 employees for a subset of countries. These data reveal that the main variation with country income per capita is in the share of employment in small (<10) and large ($\geq$ 50) firms. The share of employment in medium-sized firms varies little with GDP per capita, and the skill intensity of medium-sized firms is similar to that of large firms. This implies that the facts we show in the next Section are not sensitive to the precise choice of threshold in the range of 10 to 50.

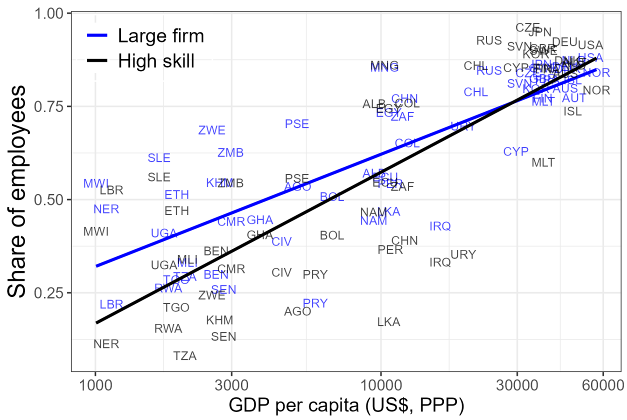
Figure 1: Size, skill, premia

Panel A: Employment by firm size & skill.

(a) Composition of employment.

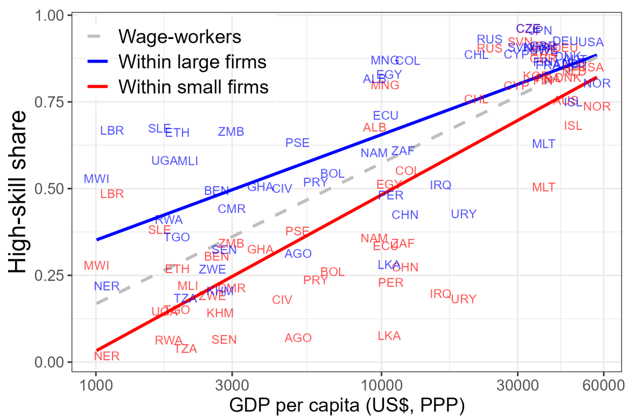


(b) Firm size & skill.

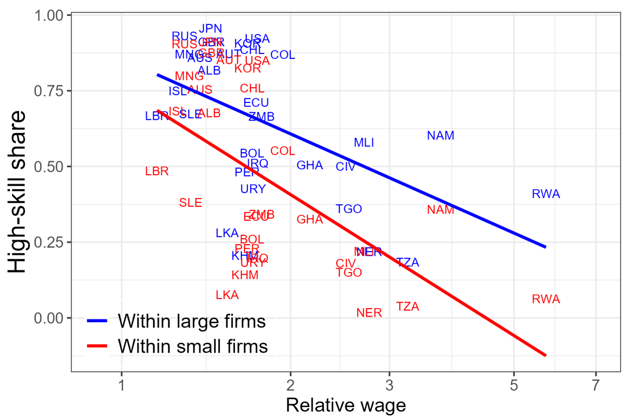


Panel B: Skill intensity by firm size & wage premia.

(c) Skill intensity by firm size (L_h^i/L^i).



(d) Skill intensity against relative wages.



Notes. Panel (a) shows the share of low-skill and high-skill workers employed in small and large firms across four country-income groups. Shares are averaged over the country-year observations in each income group. The income groups correspond to low-income [\$0, \$3,000], lower-middle-income [\$3,000, \$10,000], upper-middle-income [\$10,000, \$30,000], and high-income [\$30,000, ∞] categories. Due to rounding, the reported shares may not sum to 100%. Panel (b) plots the share of total wage employment in large firms (blue) and the share of high-skilled wage workers in total wage employment (black) against GDP per capita. Workers are classified as high-skilled if they have more than nine years of schooling. Panel (c) plots high-skill employment: i) as a share of all wage employment (grey dotted line), ii) as a share of employment in large firms (blue line), and iii) as a share of employment in small firms (red line) against GDP per capita. Panel (d) shows skill intensity (the fraction of high-skill employment in total employment within each firm size) against the relative wage of high- to low-skill workers. Skill intensity is shown for large firms (blue) and small firms (red). GDP per capita is in U.S. dollars at purchasing power parity as provided by Feenstra et al. [2015]. The lines represent the fitted values of a linear regression, and country-level observations are denoted by their corresponding ISO codes. The country-year observations and the underlying surveys are reported in Table A-2.

With rising incomes, this share is increasing to 48% in lower-middle, 67% in upper-middle, and 83% in high-income countries.[5]

This pattern aligns with the findings of Poschke [2018], Bento and Restuccia [2017, 2021], who have demonstrated that the average size of firms is larger in rich countries.

Figure 1b shows how employment in large firms and the share of high-skill workers in total wage employment vary with GDP per capita. In blue, we plot the share of total wage employment in large

[5]Due to rounding, the reported shares in panel a of Figure 1 may not sum to 100%.

firms, and in black, the share of high-skill wage workers in total wage employment. The figure confirms that low-income countries have fewer large firms and fewer high-skill workers.

Next, we ask how these patterns relate to the skill intensity of employment in large and small firms.

3.2 Skill intensity of employment by firm size

Fact 2: Across countries, large firms employ more skilled workers than small firms. Figure 1c shows the share of high-skill workers in total wage employment (grey dotted line), in large firms (blue line), and small firms (red line) against GDP per capita. The figure illustrates that the skill mix in employment varies systematically between small and large firms. Large firms always employ a larger share of high-skill workers than small firms. Large firms, hence, consistently employ workers who are more skilled than the average wage worker in the economy.

Fact 3: Small firms in high-income countries are nearly as skill-intensive as large firms, but small firms in low-income countries employ far fewer skilled workers than their larger counterparts. Panel c of Figure 1 also shows that the gap in skill intensity between small and large firms is much larger in low-income countries than in high-income countries. To ease comparability, Table 1 reports the average share of high-skill workers in large and small firms across income groups. In low-income countries, 45% of the total employment of large firms is high-skilled, whereas this is only 21% for small firms. For high-income countries, the corresponding numbers are 87% and 82%, respectively. The difference in skill intensity between large and small firms is thus 24 percentage points in low-income countries, but only five percentage points in high-income countries. In relative terms, the skill intensity of small firms in low-income countries is only 47% of that of large firms, whereas it is 94% in high-income countries. The gap in skill intensity between large and small firms closes smoothly as we move up the GDP per capita distribution.

Table 1: Skill intensity by firm size (L_h/L) & relative wages (w_h/w_l).

	Country Income Group			
Firm category	Low income	Lower-middle income	Upper-middle income	High income
Large firm	0.45	0.55	0.67	0.87
Small firm	0.21	0.31	0.46	0.82
N (Cross-section)	16	8	13	17
Relative wage (w_h/w_l)	2.54	2.3	1.63	1.5
N (Wage sample)	10	5	9	7

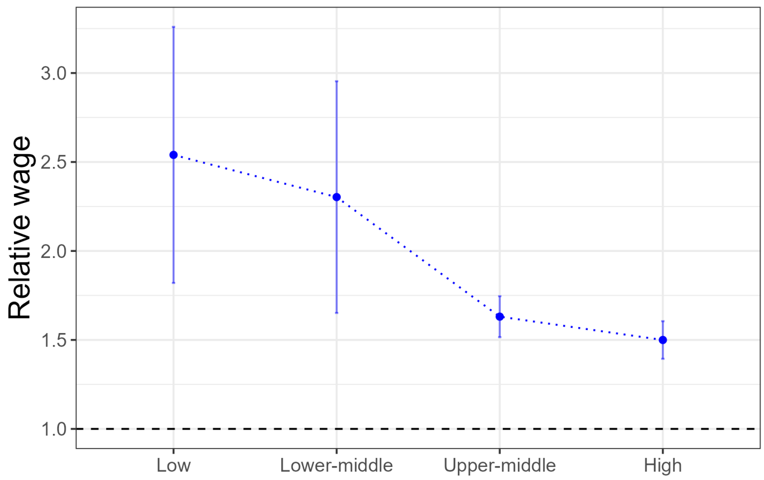
Notes. This table reports the average share of high-skill workers in large and small firms across income groups. The income groups correspond to low-income [\$0, \$3,000], lower-middle-income [\$3,000, \$10,000], upper-middle-income [\$10,000, \$30,000], and high-income [\$30,000, ∞] categories. The skill intensity is defined as the share of high-skill workers in large (L_h^b/L^b) and small firms (L_h^s/L^s). The relative wage is defined as the ratio of the average wage of high-skill workers to that of low-skill workers (w_h/w_l).

3.3 Skill premium and the skill intensity of small and large firms

Fact 4: The skill gap between small and large firms is narrow when the skill premium is low, but widens substantially when the premium is high. Figure 1d shows the skill intensity of small and large firms against the relative wage of high- to low-skill workers. The figure shows that the skill intensity of small firms is much lower (in relative terms) than that of large firms when the skill premium is high, but not when it is low. Hence, small firms can substitute low-skill for high-skill workers much more easily than large firms when the skill premium is high.

Figure 2 shows the average relative wage of high- to low-skill workers across income groups. Indeed, the skill premium is higher in low-income countries than in high-income countries. In low-income countries, the average skill premium is 2.54 (i.e., high-skill workers earn 2.54 times as much as low-skill workers), while it is only 1.5 in high-income countries. The average skill premium is thus 70% higher in low-income countries than in high-income countries.

Figure 2: Relative wage (w_h/w_l) by income group



Notes. This figure shows the average relative wage of high- to low-skill workers across income groups. The income groups correspond to low-income [$0, $3,000], lower-middle-income [$3,000, $10,000], upper-middle-income [$10,000, $30,000], and high-income [$30,000, ∞] categories. The relative wage is defined as the ratio of the average wage of high-skill workers to that of low-skill workers (w_h/w_l). Error bars represent the 90% confidence intervals of the group means.

3.4 Robustness

Summarizing, large firms are more skill-intensive everywhere. However, the extent to which their skill intensity varies with country income and skill endowments is smaller than for small firms, suggesting that large firms are less flexible in adjusting to the scarcity of skilled workers in low-income countries. We now discuss several robustness checks.

Controlling for sectoral composition. To ensure that our results are not merely picking up differences in firm size and skill intensity across sectors, we control for sectoral composition in Appendix Figure C-1. Echoing our main results, we find that the share of employment in large firms is increasing with income per capita within agriculture, manufacturing, and services. Similarly, we find that the share of high-skill workers in sectoral employment is growing with income per capita. While agriculture is less skill-intensive than manufacturing and services overall, large firms within each sector are consistently more skill-intensive than small firms (see Appendix Figure C-2). In all sectors, the gap in skill intensity between large and small firms is also much larger in low-income countries than in high-income countries.

Including the self-employed. Self-employment accounts for a large share of total employment in low-income countries. The majority of these self-employed individuals are "own-account" workers, that is, self-employed without employees, who can effectively be considered as operating single-person firms. We report our results for the case where we assign self-employment to our small firm category in Appendix Figure C-3. The main results persist, and the shift out of small firms to large firms along the GDP per capita spectrum becomes even more striking.

Alternative skill definition. In Figure 1, we define skilled workers as those with more than nine years of schooling. As a robustness check, Appendix Figure C-4 presents the results using an alternative definition: skilled workers are those with more than 12 years of schooling. Under this definition, low-skilled workers include those who have completed high school, while high-skilled workers are those with education beyond high school. Our main results persist.

Aggregate vs. per-country patterns. In Figure 1c, the fitted lines are the results of regressing the share of high-skilled employment in total employment within each firm type on log GDP per capita. The blue line, lying above the red line, hence indicates that, on average, large firms are more skill-intensive. Due to the varying slopes, this gap is particularly pronounced in low-income settings. Because the lines pick up average treatment effects, it is possible that outliers could bias their slopes. To control for this, we plot in Appendix Figure C-5 the gap in skill intensity between large and small firms for each country separately. We find that the patterns persist with large firms employing more high-skilled workers than small firms, and that gap is particularly large for low-income countries.

Mincer regression. In Figure 2, we document that the raw relative wages of high- vs. low-skilled workers are higher in low-income countries. This pattern persists when, instead of calculating the raw average relative wages of high- vs. low-skilled workers, we run a Mincer wage regression on the individual level, including controls [Mincer, 1974]. We estimate, for each country separately, the skill premium of being high-skilled:

$$\log w_{ict} = \beta \mathbf{1}\{\text{high-skill}\}_{ict} + X_{ict}\gamma + \varepsilon_{ict},$$

where w_{ict} is the hourly wage of individual i in country-survey c at time t. $\mathbf{1}\{\text{high-skill}\}_{ict}$ is a dummy variable that takes the value one if the individual i is high-skilled, and zero otherwise. X_{ict} are individual characteristics (sex, age, age-squared, marital status) of individual i in country c at time t, respectively. ε_{ict} is the error term. β is our coefficient of interest and represents the skill premium.

The results are shown in Appendix Figure C-6.

4 Model

We develop a simple model of firm size in the spirit of Hopenhayn [1992] to examine the role of skill supply in shaping the firm size distribution and aggregate productivity. In this section, we highlight the key components of the model. The full derivation can be found in Appendix D.

4.1 Households

The representative household derives utility from the consumption of a final good. The representative household consists of workers who can be low-skilled (l) or high-skilled (h). All workers supply labor inelastically.

4.2 Technology

Firms have access to two technologies, a small (s) and a big (b) technology.

In the empirical analysis, large (small) referred to firms with at least (less than) ten employees; here, big and small refer to technology choices. These will correlate with firm size, but the size threshold need not be at ten. When we take our model to the data, we compute model statistics both by size, as in the empirical analysis, and for each technology.

Firms differ in their productivity z. Each firm produces a final good using a CES production function that combines skilled L_h and unskilled labor L_l. The skill intensity is defined by the parameter μ^i, while ρ^i denotes the elasticity of substitution. Technology also differs across firms in terms of returns to scale, captured by the parameter γ^i. The parameters μ^i, ρ^i, and γ^i are indexed by i to allow for differences between small and large technologies. We also include a skill-bias parameter A in technology. Note that we abstract from capital; however, differences in the returns to scale for labor can be interpreted as differences in the stock of physical capital. The output of a firm with productivity z that uses technology i is then given by

$$y^i(z) = z\left[\mu^i {L_l^i}^{\frac{\rho^i-1}{\rho^i}} + (1-\mu^i)(AL_h^i)^{\frac{\rho^i-1}{\rho^i}}\right]^{\frac{\rho^i}{\rho^i-1}\gamma^i}.$$

Finally, we also allow for an output tax $\tau(z)$, which may vary with a firm's productivity z to capture firm size distortions as customary in the literature.

4.3 Skill demand

A firm demands skilled and unskilled labor to maximize its profit

$$\pi^i(z) = \max_{L_l^i, L_h^i}(1-\tau(z))z\left[\mu^i \left(L_l^i\right)^{\frac{\rho^i-1}{\rho^i}} + (1-\mu^i)(AL_h^i)^{\frac{\rho^i-1}{\rho^i}}\right]^{\frac{\rho^i}{\rho^i-1}\gamma^i} - w_l L_l^i - w_h L_h^i.$$

The optimal ratio of skilled to unskilled workers is common for all firms of a given size type, regardless of their productivity z, and is given by

$$\left(\frac{L_l}{L_h}\right)^i = \left(A^{\frac{\rho^i-1}{\rho^i}}\frac{1-\mu^i}{\mu^i}\frac{w_l}{w_h}\right)^{-\rho^i} \equiv \Omega^i.$$

The skill intensity of employment depends on relative wages, technology parameters, and the skill bias A. We use this expression to rewrite the total employment of a firm with productivity level z such that

$$L^i(z) \;=\; L_h^i(z) \;+\; L_l^i(z) \;=\; \left(1+\Omega^i\right) L_h^i(z).$$

We introduce Θ^i, the effective output multiplier from the CES bundle, and $\tilde{\Omega}^i$, the total wage bill in terms of total high-skilled labor cost.

$$\Theta^i \;=\; \left[\mu^i\,(\Omega^i)^{\frac{\rho^i-1}{\rho^i}} \;+\; (1-\mu^i)\,A^{\frac{\rho^i-1}{\rho^i}}\right]^{\frac{\rho^i}{\rho^i-1}\gamma^i}, \qquad \text{where} \quad \tilde{\Omega}^i \;=\; 1 \;+\; \frac{w_l}{w_h}\,\Omega^i.$$

With these terms, the equilibrium demand for high-skill labor at productivity z can be written as

$$L_h^i(z) \;=\; \left(\frac{z^{1-\nu}\,\Theta^i\,\gamma^i}{\tilde{\Omega}^i\,w_h}\right)^{\frac{1}{1-\gamma^i}}.$$

Since $\gamma^i < 1$ and $\nu < 1$, high-skill demand increases with the productivity level of the firm, and decreases with the high-skilled wage.

4.4 Technology choice

Firms choose the technology they use to produce goods. To do so, they compare the profits from operating both technologies.
The profit of a firm amounts to revenue net of taxes minus the total wage bill and amounts to

$$\pi^i(z) \;=\; (1-\tau(z))\,y^i(z) \;-\; w_h\,L_h^i(z) \;-\; w_l\,L_l^i(z).$$

We use the optimal skill intensity of labor demand to rewrite the profit function as

$$\pi^i(z) \;=\; (1-\tau(z))\,\Pi^i\,w_h^{-\frac{\gamma^i}{1-\gamma^i}}\;z^{\frac{1}{1-\gamma^i}},$$

where

$$\Pi^i \equiv \left(\Theta^i\right)^{\frac{1}{1-\gamma^i}}\left(\tilde{\Omega}^i\right)^{-\frac{\gamma^i}{1-\gamma^i}}\left[\left(\gamma^i\right)^{\frac{\gamma^i}{1-\gamma^i}} - \left(\gamma^i\right)^{\frac{1}{1-\gamma^i}}\right],$$

which depends on technology parameters $(\mu^i, \rho^i, \gamma^i, A)$, the skill intensity of employment $\tilde{\Omega}^i$, and the relative total cost of skilled employment Θ^i.
There exists a level of productivity z^* at which the firm is indifferent between the small and large technology. The cut-off level z^* is defined as follows

$$\pi^s(z^*) \ = \ \pi^b(z^*) \qquad \Longrightarrow \qquad z^* = w_h \left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{(1-\gamma^b)(1-\gamma^s)}{\gamma^b-\gamma^s}}$$

The threshold level of productivity to adopt the big technology is increasing in the wages of skilled workers. When the returns to scale of the big technology are above those of the small technology, the threshold is increasing in Π^s and decreasing in Π^b.

To summarize, firms choose their technology based on their productivity level: if $z \geq z^*$, they adopt the large-technology type ($i = b$); if $z < z^*$, they adopt the small-technology type ($i = s$).

4.5 Entry and exit

There is a large mass of potential entrants. Potential entrants have to pay a fixed entry cost c^e to draw a productivity level z from a distribution $G(z)$. For a given productivity level, firms then choose whether they produce and operate a technology or exit. They exit if their productivity level does not allow them to sustain the fixed operation cost c^f.

We assume that both the entry and fixed costs are in units of the aggregate wage bill, such that they amount to $c^e\Lambda w_h$ and $c^f\Lambda w_h$ where $\Lambda = w_l/w_h L_l + L_h,$. We make this choice for reasons of tractability, as in Klenow and Li [2025].

In the presence of fixed operation costs, not all firms are profitable. That is, a firm that draws a productivity z only operates if its profits are higher than its operating costs. For the small technology, there exists a threshold $\hat{z}$ such that the firm is indifferent between exit and operating the small technology:[6]

$$(1-\tau(\hat{z}))\,\Pi^s \hat{z}^{\frac{1}{1-\gamma^s}} w_h^{-\frac{\gamma^s}{1-\gamma^s}} = c^f \Lambda w_h.$$

This indifference condition defines a productivity threshold

$$\hat{z} = \left(\frac{c^f\Lambda}{(1-\tau(\hat{z}))\,\Pi^s}\right)^{1-\gamma^s} w_h.$$

A firm with a productivity level $z \leq \hat{z}$ will not operate. Firms with a productivity level above $\hat{z}$ will operate and make their technology choice. Note that although the "natural case" is $\hat{z} < z^*$, so that both large and small firms operate, it is also possible that $\hat{z} > z^*$, in which case only large firms operate. In this scenario, z^* is not relevant, and the conditions determining $\hat{z}$ involve γ^b and Π^b rather than γ^s and Π^s. In the "natural case," entrants who draw $z < \hat{z}_j$ do not enter, those with $z > z_j^*$ operate using the large technology, and those with $z \in [\hat{z}_j, z_j^*]$ operate using the small technology. Firms that draw $z = z_j^*$ are indifferent between operating as small or large firms, and we assume they choose the large technology. Similarly, firms with $z = \hat{z}_j$ are indifferent between entering and not entering, and we assume they enter.

4.6 Equilibrium allocation

We now derive the equilibrium of the model. To do so, we make parametric assumptions about the distribution of firm productivity and the size-dependent distortion. We assume that firm productivity z follows a Pareto distribution with scale z_m and shape parameter α, and denote its cumulative density function by $G(z) = 1 - (z_m/z)^\alpha$.

[6]In general, there are two thresholds, one for each technology. We only discuss the exit threshold for the small technology since this is the empirically relevant one. An exhaustive characterization of both thresholds is provided in section D.

We also assume that tax distortions take the form $(1-\tau(s)) = z^{-\nu}$, where ν governs how distortions vary with productivity. A higher value of ν implies that more productive firms face larger distortions, while a lower (or negative) ν implies that distortions fall—or rise more slowly—with productivity. In this sense, ν captures the progressivity or regressivity of size-dependent distortions. Positive values of ν introduce misallocation by penalizing high-productivity (and typically larger) firms, thereby reducing aggregate productivity.

Distribution of firms. We now solve for the share of firms that enter and the share of firms that operate each technology. Given the exit threshold, only firms with a productivity level $z \geq \hat{z}$ operate; among those, all firms with a productivity level above z^* operate the big technology. In the natural case where $z^* > \hat{z}$, the fraction of firms that adopt the large technology amounts to

$$m^b \;=\; \frac{1-G(z^*)}{1-G(\hat{z})} \;=\; \left(\tfrac{\hat{z}}{z^*}\right)^{\alpha}.$$

Since $\hat{z}$ and z^* scale with w_h, the ratio $\hat{z}/z^*$ depends only on technology parameters and the distortion $\tau(z)$.

Potential entrants pay an entry cost before drawing their productivity level and expect to earn profits upon entry. Free entry requires that, in expectation, the profits from entry are at least equal to the entry cost, such that

$$c^e\,\Lambda\,w_h \;=\; \int_{\hat{z}}^{z^*} \pi^s(z)\,dG(z) \;+\; \int_{z^*}^{\infty} \pi^b(z)\,dG(z).$$

The free entry condition pins down w_h (and thereby also $\hat{z}, z^*, m^b$) as a function of parameters $\{\mu^i, \rho^i, \gamma^i, \alpha, c^f, c^e, z_m\}$, the wage bill via Λ, and the size-dependent distortion τ.

Aggregate labor demand. The labor demand for low and high-skilled workers is the sum of labor demand for firms that operate both technologies. The aggregate demand for high-skill workers is

$$L_h \;=\; \frac{M}{1-G(\hat{z})}\Big[\Big(\tfrac{\Theta^s\,\gamma^s}{\tilde{\Omega}^s\,w_h}\Big)^{\frac{1}{1-\gamma^s}}\,\bar{z}^s \;+\; \Big(\tfrac{\Theta^b\,\gamma^b}{\tilde{\Omega}^b\,w_h}\Big)^{\frac{1}{1-\gamma^b}}\,\bar{z}^b\Big],$$

Similarly, aggregate low-skill demand is

$$L_l \;=\; \frac{M}{1-G(\hat{z})}\Big[\Omega^s\,\Big(\tfrac{\Theta^s\,\gamma^s}{\tilde{\Omega}^s\,w_h}\Big)^{\frac{1}{1-\gamma^s}}\,\bar{z}^s \;+\; \Omega^b\,\Big(\tfrac{\Theta^b\,\gamma^b}{\tilde{\Omega}^b\,w_h}\Big)^{\frac{1}{1-\gamma^b}}\,\bar{z}^b\Big],$$

where

$$\bar{z}^s \;=\; \int_{\hat{z}}^{z^*} z^{\frac{1-\nu}{1-\gamma^s}}\,dG(z), \qquad \bar{z}^b \;=\; \int_{z^*}^{\infty} z^{\frac{1-\nu}{1-\gamma^b}}\,dG(z),$$

and M^e denotes the mass of entrants, and $M = (1-G(\hat{z}))\,M^e$ the mass of firms that actually operate.

For a full characterization of the equilibrium, see Appendix section D-1.

5 Calibration

We calibrate our model to the U.S. economy. We need to calibrate 6 production parameters $(\mu^i, \rho^i, \gamma^i)$, entry and fixed cost (c^e, c^f), the productivity distribution (α and z_m), and distortions (ν). We directly extract aggregate skill endowments (L_h, L_l) from the data.

Set parameters. We set the elasticity of substitution and returns to scale in large firms, ρ^b and γ^b, to typical estimated values in the literature, 4 [Bils et al., 2024] and 0.85 [Atkeson and Kehoe, 2007, Katz and Murphy, 1992]. We normalize z_m to 1, and assume that there are no distortions in the U.S. economy ($\nu = 0$).

Calibrated parameters. We calibrate the following six parameters: $\mu^s, \mu^b, \gamma^s, c^e, c^f$, and α to match the skill premium, the share of employment in large firms, the skill intensity of large firms, the share of large firms, and the tail index of the firm size distribution in the United States.

These model moments are closely related to and informative about the parameters we are calibrating. Concretely, the skill intensity of large firms, given the skill premium, directly implies a value for μ^b. The share of employment in large firms then implies the skill intensity of small firms, and thus μ^s. Given γ^b, the tail index of the firm size distribution directly implies the value of α. Given the function for optimal employment, α equals the empirical tail index divided by $1-\gamma^b$. The fixed cost c^f closely affects the share of large firms, as a lower fixed cost implies a lower entry threshold and more small firms.

According to our data, the skill premium in the United States is 1.75, and the share of employment in large firms is 81%. The share of workers in large firms with high skills is 92%. In small firms, the percentage is 88%, and the share of high-skilled workers in the labor force is 91%. Using the Statistics of U.S. Businesses (SUSB), the share of firms with at least 10 employees is 21.4%. Using data from the U.S Census Business Dynamics Statistics (BDS) implies a tail index of employment of 1.13. We match all these moments exactly.

Finally, the elasticity of substitution in small firms, ρ^s, is not identified from data for a single country. We set it to 4 and explore sensitivity to this choice. It turns out that for reasonable values of ρ^s, results do not change much.

Table 2 reports the parameter values.

Table 2: Calibration to the U.S. economy

Parameter	Value	Target/Source
Pre-set:		
ρ_b	4	Bils, Kaymak and Wu (2024)
ρ_s	4	$= \rho_b$
γ_b	0.85	Atkeson and Kehoe (1995)
Calibrated internally:		
μ^b	0.771	Skill intensity large firms, U.S
μ^s	0.610	Share of employment in large firms, U.S.
γ_s	0.356	Skill premium, U.S
α	7.57	Tail index of employment of 1.13, U.S. (BDS)
c^f	0.729	Share large firms, U.S

6 Counterfactual: Skill endowments, the firm size distribution, and aggregate output

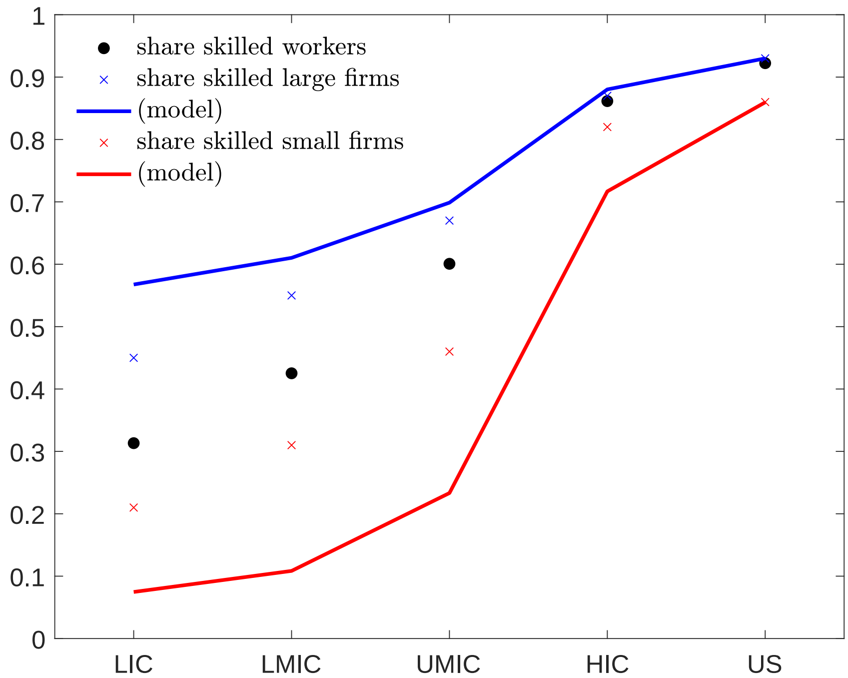
In section 3, we have shown very large differences in the skill composition of economies at different stages of development. In this section, we explore how the model economy reacts to changes in the aggregate skill endowment. We use the model, calibrated to U.S. data, and only vary the skill endowments to reflect those of different income groups to study how skill intensity by firm size, technology choice, and employment in large firms react and how well the model captures observed patterns in the data. For the aggregate skill endowments of our country income groups, we use the data reported in Appendix Table B-2. The average high-income economy (HIC) has a share of high-skilled workers of around 86%, which is slightly lower than the U.S. (91%). Upper-middle income economies (UIC) have a share of high-skilled workers of around 60%, and lower-middle income (LMIC) and low income economies (LIC) have shares of 42% and 31%, respectively.

Skill composition by firm size. We start by investigating how a change in the aggregate skill endowments affects skill intensity by firm size. The results are shown in Figure 3. The black dots show the share of skilled workers in the total workforce for each income group. The blue and red crosses show the share of skilled workers within the employment of large and small firms, respectively. We observe these values from our data. The solid lines show the skill intensity that small and large firms optimally choose according to our model, given the aggregate skill endowments.

The figure shows that the model broadly captures the fact that both large and small firms use less high-skilled labor in low-income countries. It also replicates the diverging trends in skill intensity across firm sizes. Large firms reduce their use of high-skilled labor somewhat less in the model than in the data, while small firms reduce it slightly more. It is important to note that we calibrate all parameters to the U.S. economy and vary only the skill endowments in this counterfactual exercise.[7]

The underlying mechanism that generates these results in the model is that a greater scarcity of high-skilled workers implies a larger skill premium (as observed in the data). As a result, all firms hire fewer skilled workers. Due to their greater ability to adjust, this difference is significantly larger for small firms. In low-income countries, where skilled workers make up only 31% of the labor force, small firms in the model hire almost none. At large firms, in contrast, slightly more than half of the workers are still highly skilled.

Figure 3: Skill intensity by firm size: Model and data



Notes. This figure shows the skill intensity by firm size as observed in the data and implied by the model when using the parameters calibrated to the U.S. economy (as discussed in section 5), and only varying the aggregate skill endowment across income groups. The x-axis represents our four income groups: low-income (LIC), lower-middle income (LMIC), upper-middle income (UIC), and high income (HIC), as well as the U.S. The y-axis is in shares. Black dots represent the average share of high-skilled workers for the different income groups as observed in the data (see panel a of Appendix Table B-2). The blue and red crosses represent the observed average skill intensity of small and large firms (see Table B-3) as observed in the data. The solid lines show the skill intensity for small and large firms that the model implies, when replacing the U.S. skill endowments with the respective income groups.

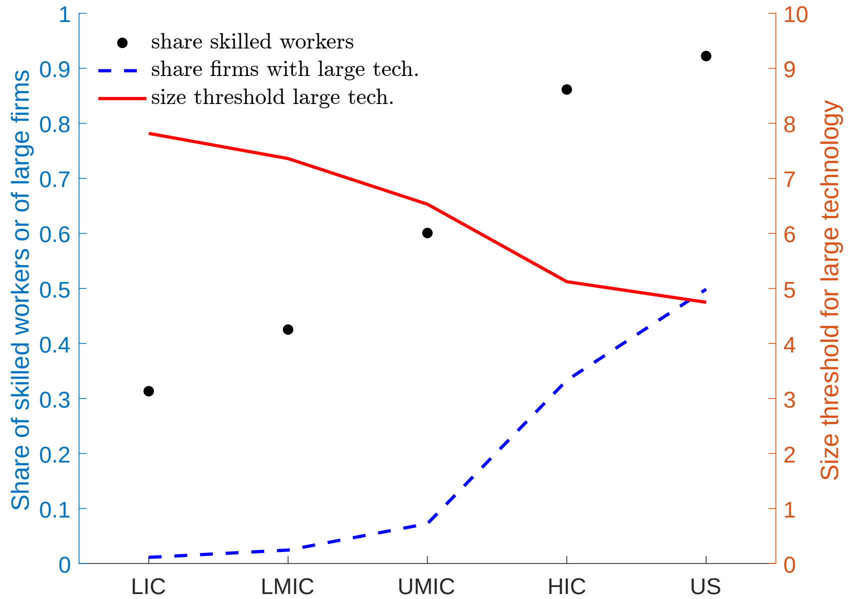
Technology choice. We now document how skill scarcity affects technology choice. While in the empirical analysis, large (small) referred to firms with at least (less than) ten employees, here, big and small refer to technology choices. These will correlate with firm size, but the size threshold need not be at ten. In the model, a higher skill premium makes running the big firm technology more costly, as it is more intensive in high-skilled labor. As a result, the productivity threshold for the big technology, z^*, rises with the skill premium and is not simply a fixed value.

[7] In that sense, a perfect fit would be surprising.

Figure 4 shows, in blue (left axis), the share of firms that choose to operate with the large-scale technology in the model, given the aggregate skill endowments. In red (right axis), it shows the size threshold at which firms switch to the large-scale technology. These moments are not directly observable in the data, so we cannot report a corresponding empirical counterpart.

The model implies that, in the U.S., the large technology is optimal for firms with more than 4.7 employees; this size threshold is around 8 in a low-income economy. As a result, very few firms adopt large-scale technology in low-income countries where the skill premium is high.

Figure 4: Technology choice



Notes. This figure shows the technology choice of firms implied by the model when using the parameters calibrated to the U.S. economy (as discussed in section 5), and only varying the aggregate skill endowment across income groups. The x-axis represents our four income groups: low-income (LIC), lower-middle income (LMIC), upper-middle income (UIC), and high income (HIC), as well as the U.S. The left y-axis (blue) represents shares. Black dots represent the average share of high-skilled workers for the different income groups (see panel a of Appendix Table B-2). The dashed blue line represents the share of firms that optimally chose to operate a large-scale technology when replacing the U.S. skill endowments with the respective income groups. The right y-axis shows the size threshold (implied by the threshold productivity z^*) after which a firm chooses to operate the large-scale technology.

We now recover the share of employment in firms with at least 10 employees (our empirical measure of "large" firms) and the model's implied share of large firms. From the threshold in Figure 4, it becomes clear that all large firms use the large-scale technology; however, because the threshold is below ten, some small firms will also use the large-scale technology.

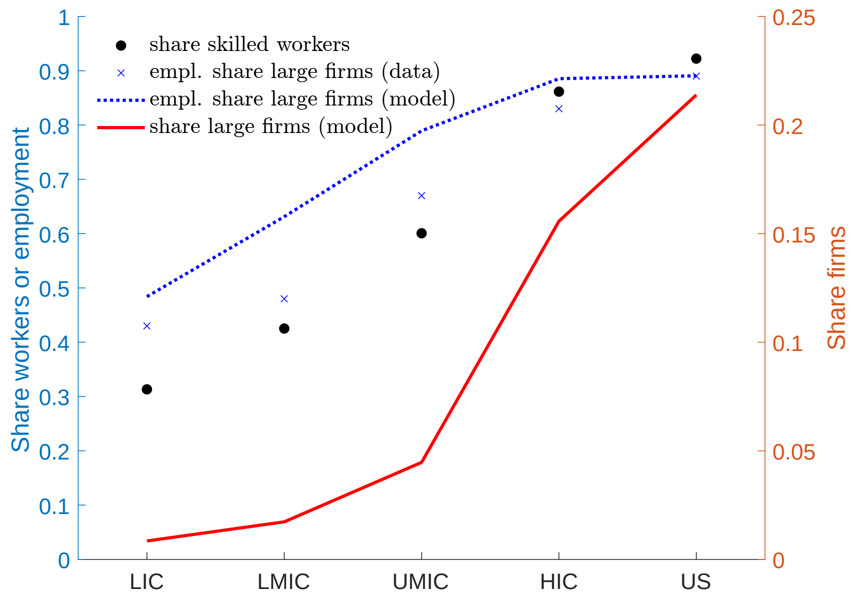
The results are shown in Figure 5. The blue crosses indicate the employment share in large firms, as measured from the data (see Table B-2). The dashed blue line shows the employment share in large firms implied by the model. The model captures the trend that employment in large firms is about twice as large in high-income countries as in low-income countries, although it slightly overstates employment in large firms.

The red line represents the model-implied share of firms with at least 10 employees, which increases along the development spectrum. Our data does not directly allow us to measure the firm size distribution (only employment in firms of different sizes) without additional distributional assumptions; hence, we do not report a data counterpart. Consistent with intuition, this share is much larger in high-income countries than in low-income countries.

Overall, although no cross-country data were used in the calibration, the model closely approximates differences in firm size patterns across countries, varying only the skill endowment.

Output. These changes in the size distribution also affect output. In adjusting their technology choices, firms react optimally to different skill prices and thus, differences in thresholds. Technology choices across economies in our model are efficient. Nevertheless, we can evaluate the impact of

Figure 5: Employment in large firms ($L \geq 10$)



Notes. This figure shows the employment in large firms as observed in the data and implied by the model when using the parameters calibrated to the U.S. economy (as discussed in section 5), and only varying the aggregate skill endowment across income groups. The x-axis represents our four income groups: low-income (LIC), lower-middle income (LMIC), upper-middle income (UIC), and high income (HIC), as well as the U.S. The left y-axis (in blue) shows the share of employment in large firms. The right y-axis (red) shows the share of large firms. Black dots represent the average share of high-skilled workers for the different income groups (see panel a of Appendix Table B-2). The blue crosses represent the observed share of employment in large firms (see Appendix table Table B-2). The solid blue line shows the models' implied share of employment in large firms. The red line shows the models' implied share of large firms in all firms.

differences in technology choices across countries by assessing their effect on output, keeping skill endowments fixed. For this, we impose the cutoffs z^* and $\hat{z}$ from the low-income group on the U.S. economy.

In doing so, we find that imposing the technology and entry choices of a low-income economy in the U.S. implies a reduction in output (net of fixed and entry costs) of 6%. That is, changes in technology account for 20% of the difference in net output between low-income economies and the U.S. in our model.[8]

The reaction of technology choices to the skill endowment also implies that increasing workforce skills has benefits beyond simply those of having more skilled workers for a given production structure. Skill accumulation enables the more effective use of large-scale technology.

7 Discussion: Endowments, firm sizes, and misallocation

In this section, we briefly relate our results to those of a rich literature on misallocation. This literature often interprets average product dispersion and differences in the firm size distribution as evidence of distortions.

7.1 Average and marginal products

In work on misallocation, it is common to use labor productivity or a firm's average product as a proxy for its marginal product, since the two are proportional to each other with Cobb-Douglas production functions. Then, dispersion in measured labor productivity across firms is often taken to indicate marginal product dispersion and interpreted as evidence of the presence of distortions. Such an approach usually suggests greater distortions to large firms. It is well known that size- or productivity-dependent distortions induce misallocation and significantly reduce an economy's output.

[8]To ensure labor market clearing, we allow the number of firms, M, and the skill premium to adjust. This implies a lower skill premium in low-income economies, since their technology choices imply a lower value of skill.

With CES production functions, average and marginal products are not uniformly proportional to each other. As a result, in the competitive equilibrium of the model analyzed here, labor productivity is not equated across firms using different technologies, even in the absence of distortions. Hence, in the present context, dispersion in labor productivity does not necessarily indicate the presence of distortions.

To see this, consider the optimal labor productivity of each type of firm. At optimal employment, given wages, the average output of a firm of technology i is

$$\left(\frac{y}{L}\right)^i = \frac{w_h}{\gamma^i} \frac{1 + \left(\frac{1-\mu^i}{\mu^i}\right)^{-\rho^i} \left(\frac{w_l}{w_h}\right)^{1-\rho^i}}{1 + \left(\frac{1-\mu^i}{\mu^i}\right)^{-\rho^i} \left(\frac{w_l}{w_h}\right)^{-\rho^i}}.$$

Optimal labor productivity varies with a firm's technology, as well as with factor prices in an economy.

First, the average products of large and small firms differ because of differences in returns to scale, γ. Greater returns to scale imply a slower decline in the marginal product with firm size, thus a greater optimal size for any given z and input prices, and consequently a lower optimal average product. This channel, captured via γ^i, affects the optimal average product of large vs small firms in the same way in all economies, regardless of input prices.

Second, if firms differ in factor intensity, optimal average products depend on relative input prices. An increase in the skill premium increases the unit cost of large relative to small firms, reducing their relative size, and thus increases their optimal labor productivity.

This implies that in economies where skills are more scarce, generating a greater skill premium, the average product of large firms relative to small firms is greater – simply because a high price of skill curtails the size of large firms, and without any size- or productivity-dependent distortions. Across our country groups, labor productivity in large relative to small firms, $(y/L)^b/(y/L)^s$, is more than 50% greater in a middle-income country compared to the U.S., and more than twice as large in a low-income country compared to the U.S. This increase occurs solely because skill scarcity raises the skill premium, which curtails the size of large firms and increases their average productivity relative to small firms, without any size- or productivity-dependent distortions.

7.2 Productivity- or size-dependent distortions

In the model, a lower skill endowment has a strong impact on the firm size distribution. How do these effects compare to those of the productivity- or size-dependent distortions that have frequently been analyzed in the literature?

To illustrate, consider increasing the parameter ν, which controls productivity-dependent distortions, from its benchmark value of zero to 0.25. This implies an elasticity of the net of tax on revenue with respect to z of 0.25.

Because this distortion primarily affects large firms, it makes running the large technology less attractive and raises the threshold z^* relative to the continuation threshold $\hat{z}$. As a result, the share of large firms in our model declines by 40%, and the share of firms with at least ten workers by more than half. The share of employment in large firms falls by about half.

In addition, because large firms intensively employ skilled workers, distortions also reduce the skill premium. Conditional on the technology choice, this implies a greater skill intensity among all firms – but, of course, this is balanced by the smaller share of large firms.

This brief analysis illustrates that differences in the firm size distribution per se do not necessarily indicate the presence of productivity-dependent distortions. These can instead stem from differences in skill endowments, which have a powerful effect on the firm size distribution. Nor does variation in average product dispersion across countries necessarily indicate the presence of distortions, given that it can be the direct consequence of differences in skill prices when firms differ in skill intensity. Plausibly, differences in the firm size distribution across countries reflect both differences in skill endowments and distortions.

8 Conclusion

This paper provides new facts on the relationship between skill endowments, firm size distribution, and economic development. First, we demonstrate that the share of employment in large firms in high-income countries is approximately twice that in low-income countries. Second, we demonstrate that employees of large firms are generally more skilled than those of small firms across countries. Third, we demonstrate that in low-income countries, employment in small firms is less skilled than in large firms, whereas in high-income countries, skilled workers are distributed similarly across all firm sizes. Fourth, the skill gap between small and large firms is narrow when the skill premium is low, but it widens substantially when the premium is high. This evidence suggests that higher levels of education are associated with larger firm sizes and that high-skilled workers in large firms generate higher incomes.

Building on these facts, we develop a quantitative heterogeneous firm model with skill heterogeneity and endogenous technology choice to understand how differences in skill endowments shape the firm size distribution, technology adoption, and aggregate productivity across countries. Calibrated to the U.S. economy, the model, when varying only the aggregate skill endowment to levels observed in low- and middle-income countries, closely replicates the empirical patterns we document: large firms employ a higher share of skilled workers everywhere, but the gap in skill intensity between large and small firms is much wider in low-income settings. The model also replicates the striking decline in the employment share of large firms as skill endowments fall, aligning with observed differences across the development spectrum.

Our findings highlight that the scarcity of skilled workers in low-income countries not only raises the skill premium but also constrains the ability of firms to scale up. Because large-scale technologies are more skill-intensive and less flexible in substituting away from skilled workers, high skill premia make the adoption of such technologies prohibitively expensive for all but the most productive firms in low-income economies. Consequently, differences in skill endowments indirectly limit the prevalence of large, high-productivity firms, contributing to persistent productivity gaps across countries.

Beyond providing a new explanation for differences in the firm size distribution, our analysis has important implications for understanding cross-country income differences and the role of skill accumulation in development. We are the first to show that skill accumulation promotes development not only directly, by increasing individual productivity, but also indirectly, by enabling the expansion of large, more productive firms. Our model quantifies this channel, showing that shifting the technology choices of firms in a high-skill economy like the U.S. to those observed in low-skill economies can reduce output by around 6%, even without changing aggregate productivity or skill endowments. This accounts for a substantial fraction of the observed output differences between rich and low-income countries.

Finally, our findings speak to the interpretation of observed differences in firm size distributions and productivity dispersion across countries. While these differences are often attributed to misallocation or size-dependent distortions, our results show that they can arise endogenously from differences in skill endowments and prices, even in the absence of distortions.

In summary, this paper advances our understanding of how human capital shapes the organization of production and aggregate productivity across countries. It highlights the importance of considering the interaction between skill endowments, firm size, and technology adoption in development and structural transformation. Future research could build on these insights by incorporating capital and financial frictions, sectoral differences, and the dynamics of skill accumulation to better understand the pathways through which human capital development drives economic growth.

References

Daron Acemoglu and Veronica Guerrieri. Capital deepening and non-balanced economic growth. *Journal of Political Economy*, 116(3):467–498, 2008.

Ufuk Akcigit, Harun Alp, and Michael Peters. Lack of selection and limits to delegation: Firm dynamics in developing countries. *American Economic Review*, 111(1):231–75, January 2021. doi: 10.1257/aer.20180555. URL https://www.aeaweb.org/articles?id=10.1257/aer.20180555.

Francisco Alvarez-Cuadrado, Ngo Van Long, and Markus Poschke. Capital-labor substitution, structural change and growth. *Theoretical Economics*, 12(3):1229–1266, 2017.

Andrew Atkeson and Patrick J Kehoe. Modeling the transition to a new economy: lessons from two technological revolutions. *American Economic Review*, 97(1):64–88, 2007.

Eric Bartelsman, John Haltiwanger, and Stefano Scarpetta. Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, 103(1):305–34, February 2013. doi: 10.1257/aer.103.1.305. URL https://www.aeaweb.org/articles?id=10.1257/aer.103.1.305.

Pedro Bento and Diego Restuccia. Misallocation, establishment size, and productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303, July 2017. doi: 10.1257/mac.20150281. URL https://www.aeaweb.org/articles?id=10.1257/mac.20150281.

Pedro Bento and Diego Restuccia. On average establishment size across sectors and countries. *Journal of Monetary Economics*, 117:220–242, 2021.

Mark Bils, Bariş Kaymak, and Kai-Jie Wu. Labor substitutability among schooling groups. *American Economic Journal: Macroeconomics*, 16(4):1–34, 2024.

Francisco J. Buera and Roberto N. Fattal-Jaef. *The dynamics of development: Innovation and reallocation*. 2018. doi: 10.1596/1813-9450-8505. URL https://elibrary.worldbank.org/doi/abs/10.1596/1813-9450-8505.

Francisco J. Buera, Joseph P. Kaboski, and Yongseok Shin. Finance and development: A tale of two sectors. *American Economic Review*, 101(5):1964–2002, August 2011. doi: 10.1257/aer.101.5.1964. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.5.1964.

Francesco Caselli and Antonio Ciccone. The contribution of schooling in development accounting: Results from a nonparametric upper bound. *Journal of Development Economics*, 104(C):199–211, 2013. URL https://ideas.repec.org/a/eee/deveco/v104y2013icp199-211.html.

Austin Davis, Eric Hsu, and Mitchell VanVuren. Self employment, micro-entrepreneurship, and development, 2023.

Niklas Engbom, Hannes Malmberg, Tommaso Porzio, Federico Rossi, and Todd Schoellman. Economic development according to Chandler. *mimeo*, 2024.

Robert C. Feenstra, Robert Inklaar, and Marcel P. Timmer. The next generation of the penn world table. *American Economic Review*, 105(10):3150–3182, 2015.

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

Charles Gottlieb, Jan Grobovšek, and Alex Monge-Naranjo. Occupational choices, human capital and cross-country income differences. *mimeo*, 2024.

Jan Grobovšek. Managerial delegation, law enforcement, and aggregate productivity. *The Review of Economic Studies*, 87(5):2256–2289, 05 2020. ISSN 0034-6527. doi: 10.1093/restud/rdaa025. URL https://doi.org/10.1093/restud/rdaa025.

Nezih Guner, Gustavo Ventura, and Yi Xu. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, 11(4):721–744, 2008.

Nezih Guner, Andrii Parkhomenko, and Gustavo Ventura. Managers and productivity differences. *Review of Economic Dynamics*, 29:256–282, 2018.

Lutz Hendricks and Todd Schoellman. Skilled labor productivity and cross-country income differences. *American Economic Journal: Macroeconomics*, 15(1):240–68, January 2023. doi: 10.1257/mac.20200256. URL https://www.aeaweb.org/articles?id=10.1257/mac.20200256.

Jonas Hjort, Hannes Malmberg, and Todd Schoellmann. The missing middle managers: Labor costs, firm structure, and development. *mimeo*, 2023.

Hugo Hopenhayn. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica*, 60(5):1127–1150, September 1992. URL https://ideas.repec.org/a/ecm/emetrp/v60y1992i5p1127-50.html.

Hugo Hopenhayn and Richard Rogerson. Job turnover and policy evaluation: A general equilibrium analysis. *Journal of Political Economy*, 101(5):915–938, 1993. ISSN 00223808, 1537534X. URL http://www.jstor.org/stable/2138602.

Chang-Tai Hsieh and Peter J. Klenow. Misallocation and manufacturing tfp in china and india. *The Quarterly Journal of Economics*, 124(4):1403–1448, 2009.

Benjamin F. Jones. The human capital stock: A generalized approach. *American Economic Review*, 104(11):3752–77, November 2014. doi: 10.1257/aer.104.11.3752. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.11.3752.

Lawrence F. Katz and Kevin M. Murphy. Changes in relative wages, 1963-1987: Supply and demand factors. *The Quarterly Journal of Economics*, 107(1):35–76, 1992.

Peter J. Klenow and Huiyu Li. Entry costs rise with growth. *Journal of Political Economy Macroeconomics*, 3(1):43–74, 2025.

Virgiliu Midrigan and Yi Xu. Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–58, 2014. URL https://EconPapers.repec.org/RePEc:aea:aecrev:v:104:y:2014:i:2:p:422-58.

Jacob Mincer. *Schooling, Experience, and Earnings.* National Bureau of Economic Research, New York, 1974.

Hernan Moscoso Boedo and Toshihiko Mukoyama. Evaluating the effects of entry regulations and firing costs on international income differences. *Journal of Economic Growth*, 17(2):143–170, 2012. URL https://EconPapers.repec.org/RePEc:kap:jecgro:v:17:y:2012:i:2:p:143-170.

Markus Poschke. Employment protection, firm selection, and growth. *Journal of Monetary Economics*, 56(8):1074–1085, 2009. URL https://EconPapers.repec.org/RePEc:eee:moneco:v:56:y:2009:i:8:p:1074-1085.

Markus Poschke. The regulation of entry and aggregate productivity. *The Economic Journal*, 120(549):1175–1200, 2010. doi: https://doi.org/10.1111/j.1468-0297.2010.02367.x. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1468-0297.2010.02367.x.

Markus Poschke. The firm size distribution across countries and skill-biased change in entrepreneurial technology. *American Economic Journal: Macroeconomics*, 10(3):1–41, July 2018. doi: 10.1257/mac.20140181. URL https://www.aeaweb.org/articles?id=10.1257/mac.20140181.

Diego Restuccia and Richard Rogerson. Policy distortions and aggregate productivity with heterogeneous establishments. *Review of Economic Dynamics*, 11(4):707–720, 2008.

Gabriel Ulyssea. Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1):87–99, 2010. ISSN 0304-3878. doi: https://doi.org/10.1016/j.jdeveco.2009.07.001. URL https://www.sciencedirect.com/science/article/pii/S0304387809000686.

World Bank Development Committee. Jobs: The Path to Prosperity. Technical Report DC2025-0002, World Bank Development Committee, April 24 2025. URL https://www.devcommittee.org/content/dam/sites/devcommittee/doc/documents/2025/Final_DC2025-0002.pdf.

A Data appendix

A-1 Sources

The primary data sources for the household and labor force surveys used to assemble our final dataset are listed in Table A-1. The table includes the country, the survey name, and the earliest and latest survey-year observations that satisfy our selection criteria.

Although our dataset contains multiple observations for many countries, we focus on surveys closest to the year 2015. This is driven by the motivation to compare different countries at different stages of development around the same time and to avoid potentially confounding time-varying factors, such as economic cycles or global crises (e.g., the COVID-19 pandemic). The survey-year combinations we use for our main results are shown in Table A-2.

Table A-1: Household and labor force surveys (sources)

Country	Survey Name	Min Year	Max Year
Albania	Labour Force Survey	2007	2013
Angola	Inquerito Integrado Sobre O Bem-Estar Da Populacao	2008	2008
Australia	Household, Income And Labour Dynamics In Australia	2001	2017
Austria	European Union Statistics On Income And Living Conditions	2004	2020
Benin	Enquête Modulaire Intégrée sur les Conditions de Vie des Ménages	2010	2015
Benin	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	2018
Bolivia	Encuesta Continua de Empleo	2015	2018
Bolivia	Encuesta de Hogares	2005	2020
Cambodia	Cambodia Labor Force And Child Labor Survey	2012	2019
Cambodia	Cambodia Labor Force Survey	2012	2019
Cambodia	Labor Force Survey	2012	2019
Cameroon	Fourth Cameroon Household Survey	2014	2014
Chile	Encuesta de Caracterización Socioeconómica Nacional	1990	2017
China	Family Panel Studies	2014	2016
Colombia	Gran Encuesta Integrada de Hogares	2007	2019
Cyprus	European Union Statistics On Income And Living Conditions	2005	2020
Czechia	European Union Statistics On Income And Living Conditions	2011	2020
Côte d'Ivoire	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	2018
Denmark	European Union Statistics On Income And Living Conditions	2004	2020
Ecuador	Encuesta de Condiciones de Vida	2005	2005
Ecuador	Encuesta Nacional de Empleo, Desempleo y Subempleo	2007	2018
Egypt	Labor Market Panel Survey	2006	2006
Egypt	Harmonized Labor Force Survey	2007	2017
Egypt	Labor Force Survey	2017	2017
Ethiopia	Ethiopia Socioeconomic Survey	2018	2018
Ethiopia	Socioeconomic Survey	2018	2018
Finland	European Union Statistics On Income And Living Conditions	2005	2006
France	Enquête Emploi Annuelle	2003	2019
France	Enquête Emploi en Continu	2003	2019
France	European Union Statistics On Income And Living Conditions	2004	2019
Germany	European Union Statistics On Income And Living Conditions	2005	2020
Germany	Socio-Economic Panel	2005	2019
Ghana	Ghana Living Standard Survey	1987	2008
Ghana	Living Standard Survey	1987	2008
Iceland	European Union Statistics On Income And Living Conditions	2004	2018
Iraq	Household Socio-Economic Survey	2007	2012
Japan	Employment Status Survey	1997	2017
Liberia	Household Income And Expenditure Survey	2014	2016
Malawi	Integrated Household Survey	2019	2019
Mali	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	2018
Malta	European Union Statistics On Income And Living Conditions	2008	2018
Mongolia	Labor Force Survey	2007	2021
Namibia	Labor Force Survey	2012	2018
Netherlands	European Union Statistics On Income And Living Conditions	2005	2020
Niger	National Survey On Household Living Conditions And Agriculture	2011	2011
Niger	Enquete Nationale sur L'emploi et Le Secteur Informel	2012	2012
Norway	European Union Statistics On Income And Living Conditions	2004	2020
Palestinian Territories	Harmonized Labor Force Survey	2009	2014
Paraguay	Integrated Public Use Microdata Series - International	2002	2002
Peru	Encuesta Nacional de Hogares	2007	2019
Russia	Russia Longitudinal Monitoring Survey	1994	2017
Rwanda	Enquête Intégrale sur les Conditions de Vie des Ménages	2000	2000
Rwanda	Labor Force Survey	2017	2020
Senegal	Enquête Nationale sur L'emploi Au Sénégal	2017	2019
Sierra Leone	Integrated Household Survey	2018	2018
Slovenia	European Union Statistics On Income And Living Conditions	2005	2020
South Africa	Labor Market Dynamics	2010	2019
South Africa	Quarterly Labor Force Survey	2008	2022
South Korea	Korean Labor And Income Panel Study	2003	2018
Sri Lanka	Labor Force Survey	2011	2022
Sweden	European Union Statistics On Income And Living Conditions	2004	2005
Tanzania	Living Standards Measurement Survey	2008	2019
Tanzania	National Panel Survey	2008	2019
Togo	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	2018
Uganda	Labor Force Survey	2012	2017
United Kingdom	British Household Panel Survey	1991	2008
United Kingdom	European Union Statistics On Income And Living Conditions	2008	2011
United States	Current Population Survey	2010	2017
Uruguay	Encuesta Continua de Hogares	2006	2017
Uruguay	Integrated Public Use Microdata Series - International	2006	2006
Zambia	Labour Force Survey	2017	2017
Zimbabwe	Labour Force And Child Labour Survey	2014	2019

Notes. This table lists the primary data sources for the household and labor force surveys used to assemble our final dataset. The table includes the country, the survey name, and the earliest and latest survey-year observation that satisfy our selection criteria. There may be multiple surveys for a single country. Our final dataset uses one survey for each country, which we report in Table A-2.

Table A-2: Household and labor force surveys (sample selection)

Country	Survey Name	Year	Cross-section	Wage Sample
Albania	Labour Force Survey	2013	✓	✓
Angola	Inquerito Integrado Sobre O Bem-Estar Da Populacao	2008	✓	-
Australia	Household, Income And Labour Dynamics In Australia	2015	✓	✓
Austria	European Union Statistics On Income And Living Conditions	2015	✓	✓
Benin	Enquête Modulaire Intégrée sur les Conditions de Vie des Ménages	2015	✓	-
Bolivia	Encuesta de Hogares	2015	✓	✓
Cambodia	Cambodia Labor Force And Child Labor Survey	2012	✓	✓
Cameroon	Fourth Cameroon Household Survey	2014	✓	-
Chile	Encuesta de Caracterización Socioeconómica Nacional	2015	✓	✓
China	Family Panel Studies	2016	✓	-
Colombia	Gran Encuesta Integrada de Hogares	2015	✓	✓
Cyprus	European Union Statistics On Income And Living Conditions	2015	✓	-
Czechia	European Union Statistics On Income And Living Conditions	2015	✓	-
Côte d'Ivoire	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	✓	✓
Denmark	European Union Statistics On Income And Living Conditions	2015	✓	-
Ecuador	Encuesta Nacional de Empleo, Desempleo y Subempleo	2015	✓	✓
Egypt	Harmonized Labor Force Survey	2015	✓	-
Ethiopia	Socioeconomic Survey	2018	✓	-
Finland	European Union Statistics On Income And Living Conditions	2006	✓	-
France	Enquête Emploi en Continu	2015	✓	-
Germany	European Union Statistics On Income And Living Conditions	2015	✓	-
Ghana	Living Standard Survey	2008	✓	✓
Iceland	European Union Statistics On Income And Living Conditions	2015	✓	✓
Iraq	Household Socio-Economic Survey	2012	✓	✓
Japan	Employment Status Survey	2017	✓	✓
Liberia	Household Income And Expenditure Survey	2016	✓	✓
Malawi	Integrated Household Survey	2019	✓	-
Mali	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	✓	✓
Malta	European Union Statistics On Income And Living Conditions	2015	✓	-
Mongolia	Labor Force Survey	2015	✓	✓
Namibia	Labor Force Survey	2018	✓	✓
Netherlands	European Union Statistics On Income And Living Conditions	2015	✓	-
Niger	Enquete Nationale sur L'emploi et Le Secteur Informel	2012	✓	✓
Norway	European Union Statistics On Income And Living Conditions	2015	✓	-
Palestinian Territories	Harmonized Labor Force Survey	2014	✓	-
Paraguay	Integrated Public Use Microdata Series - International	2002	✓	-
Peru	Encuesta Nacional de Hogares	2015	✓	✓
Russia	Russia Longitudinal Monitoring Survey	2015	✓	✓
Rwanda	Labor Force Survey	2017	✓	✓
Senegal	Enquête Nationale sur L'emploi Au Sénégal	2017	✓	-
Sierra Leone	Integrated Household Survey	2018	✓	✓
Slovenia	European Union Statistics On Income And Living Conditions	2015	✓	-
South Africa	Labor Market Dynamics	2015	✓	-
South Korea	Korean Labor And Income Panel Study	2015	✓	✓
Sri Lanka	Labor Force Survey	2015	✓	✓
Sweden	European Union Statistics On Income And Living Conditions	2005	✓	-
Tanzania	Living Standards Measurement Survey	2014	✓	✓
Togo	Integrated Regional Survey On Employment And The Informal Sector In Member States Of UEMOA (ERI -ESI)	2018	✓	✓
Uganda	Labor Force Survey	2017	✓	✓
United Kingdom	European Union Statistics On Income And Living Conditions	2011	✓	✓
United States	Current Population Survey	2015	✓	✓
Uruguay	Encuesta Continua de Hogares	2015	✓	✓
Zambia	Labour Force Survey	2017	✓	✓
Zimbabwe	Labour Force And Child Labour Survey	2014	✓	-

Notes. This table shows the survey-year combinations we use for our main results. The table includes the country, the survey name, and the year of the survey. The survey-year combinations are chosen to be closest to the year 2015, allowing for comparisons between different countries at various stages of development around the same time and avoiding potentially confounding time-varying factors, such as economic cycles or global crises (e.g., the COVID-19 pandemic). The last column indicates whether the survey includes information on wages and is included in our wage sample.

B Additional tables and figures

Wage sample: To increase the statistical power in documenting facts 1 through 3, we rely on the full set of surveys that allow measuring skill intensity by firm size, including those surveys that lack wage information. Fact 4, however, which documents how the relative price of skills shapes differences in skill intensity between firms of different sizes, relies on a restricted sample of surveys that provide wages.

One concern in splitting our data into a wage sample and a non-wage sample is that the two samples may differ in terms of skill endowment, employment in large firms, and skill intensity by firm size and income group. To address this concern, we conduct t-tests to compare the two samples. Table B-1 shows the results of the t-tests for the share of high-skill workers, the share of employment in large firms, and the skill intensity by firm size and income group. The table shows that the two samples are not statistically different from each other in terms of i) skill endowment by income group, ii) employment in big firms by income group, and iii) skill intensity by firm size and income group.

Table B-1: T-test

Panel A: Share of high-skilled workers (L_h/L)

	Country Income Group			
Sample	Low income	Lower-middle income	Upper-middle income	High income
Cross-section	0.31	0.42	0.59	0.86
Wage sample	0.30	0.47	0.58	0.87
p-value	0.92	0.67	0.88	0.74

Panel B: Share of employment in large firms (L^b/L)

	Country Income Group			
Sample	Low income	Lower-middle income	Upper-middle income	High income
Cross-section	0.43	0.48	0.67	0.83
Wage sample	0.41	0.47	0.65	0.82
p-value	0.76	0.92	0.76	0.85

Panel C: Skill intensity by firm size (L_h^i/L^i)

		Country Income Group			
Sample	Firm category	Low income	Lower-middle income	Upper-middle income	High income
Cross-section	Large firm	0.45	0.55	0.67	0.87
Wage sample	Large firm	0.45	0.59	0.66	0.88
p-value	Large firm	0.99	0.62	0.94	0.68
Cross-section	Small firm	0.21	0.31	0.46	0.82
Wage sample	Small firm	0.20	0.36	0.45	0.82
p-value	Small firm	0.89	0.65	0.94	1.00

Notes. This table shows the results of the t-tests for the share of high-skill workers, the share of employment in large firms, and the skill intensity by firm size and income group for our cross-sectional and restricted wage sample.

Further results. Table B-2 shows the share of high- and low-skill workers in total wage employment by income group (panel (a)) and the share of employment in large and small firms by income group (panel (b)). The table shows the average results by income group corresponding to panel (b) of Figure 1.

Table B-3 shows the skill intensity by firm size and income group. The skill intensity is defined as the share of high-skill workers in total employment by firm size. This table complements the results shown in panel (c) of Figure 1.

Table B-2: Skill shares and employment by firm size and income group

Panel A:
Share of high- and low-skill workers in total wage employment by income group.

	Country Income Group			
Skill category	Low	Lower-middle	Upper-middle	High
High skill	0.31	0.42	0.59	0.86
Low skill	0.69	0.58	0.41	0.14
Number of countries	16	8	13	17

Panel B:
Share of employment in large and small firms by income group.

	Country Income Group			
Firm size category	Low	Lower-middle	Upper-middle	High
Large firm	0.43	0.48	0.67	0.83
Small firm	0.57	0.52	0.33	0.17
Number of countries	16	8	13	17

Notes. This table shows the share of high- and low-skill workers in total wage employment by income group (panel (a)) and the share of employment in large and small firms by income group (panel (b)). The table is showing the average results by income group corresponding to panel (b) of Figure 1. The income groups are low-income, lower-middle-income, upper-middle-income, and high-income countries and are based on a GDP per capita of [$0, $3,000], [$3,000, $10,000], [$10,000, $30,000], and [$30,000, ∞] respectively.

Table B-3: Skill intensity by firm size and income group

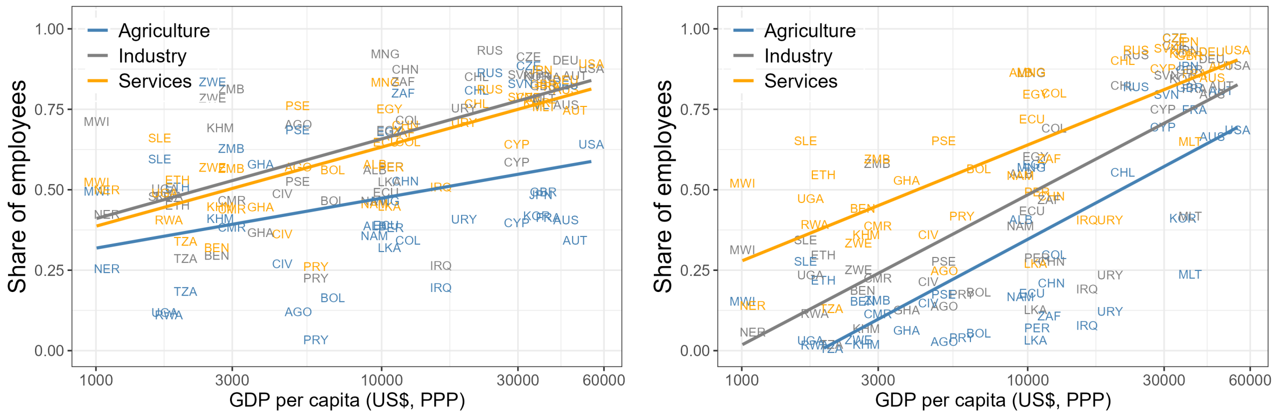
		Country Income Group			
Firm size category		Low	Lower-middle	Upper-middle	High
Large firm	High skill	0.45	0.55	0.67	0.87
	Low skill	0.55	0.45	0.33	0.13
Small firm	High skill	0.21	0.31	0.46	0.82
	Low skill	0.79	0.69	0.54	0.18
Number of countries		16	8	13	17

Notes. This table shows the skill intensity by firm size and income group. The skill intensity is defined as the share of high-skill workers in total sectoral wage employment by firm size. The income groups are low-income, lower-middle-income, upper-middle-income, and high-income countries and are based on a GDP per capita of [$0, $3,000], [$3,000, $10,000], [$10,000, $30,000], and [$30,000, ∞] respectively.

C Robustness

Sectoral composition. In this section, we check the robustness of our results to including the sectoral composition of employment. Figure C-1 shows the share of employment in large firms and the share of high-skill workers in total wage employment by sector against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015]. The sectoral patterns echo our findings from panel b of Figure 1. On the left, we plot the share of employees in large firms for each sector separately. We find that while agricultural employment is consistently more concentrated in small firms than within the industry or service sector, the share of employment in large firms is increasing for all three sectors with rising GDP per capita. On the right, we plot the share of employees within each sector who are high-skilled. While the service sector typically employs more skilled workers than industry and agriculture, all sectors become increasingly skill-intensive with rising GDP per capita.

Figure C-1: Sectoral patterns by income group



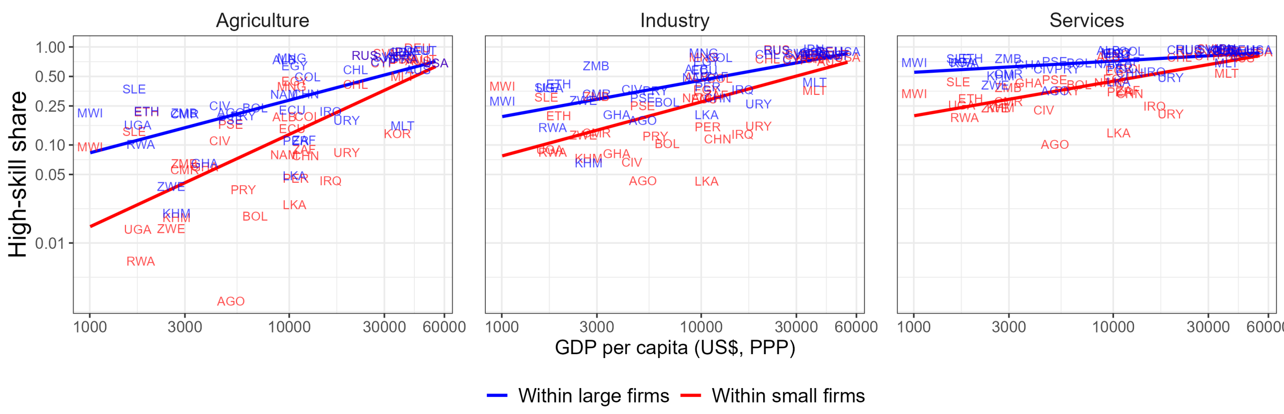
(a) Employment in large firms by sector.

(b) High-skill employment by sector.

Notes. This figure shows the share of employment in large firms (panel (a)) and the share of high-skill workers in total wage employment (panel (b)) by sector against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015].

Our main message is, however, regarding the skill intensity of small- and large firms: Figure C-2 shows the skill intensity of large and small firms by sector against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015]. The skill intensity is defined as the share of high-skill workers in total sectoral wage employment by firm size. As in panel c of Figure 1, large firms are always more skill-intensive than small firms. This pattern also persists within each sector separately. While it is true that agriculture is always less skill-intensive than industry and services, the gap between small and large firms' skill intensity in agriculture in low-income countries is particularly large (factor 9). We observe the same patterns as for the aggregate data: with rising GDP per capita, the gap in skill intensity between small and large firms closes. This finding also holds within sectors.

Figure C-2: Skill intensity and firm size by sector



Notes. This figure shows the skill intensity of large and small firms by sector against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015]. The skill intensity is defined as the share of high-skill workers in total sectoral wage employment by firm size. The sectors are agriculture, manufacturing, and services.

Including the self-employed. As described in the main text, another salient feature when comparing the organization of production between low- and high-income countries is that low-income countries have much larger shares of self-employment. The majority of these self-employed individuals are "own-account" workers, that is, self-employed without employees, who can effectively be considered as operating single-person firms. We report our results for the case where we assign self-employment to our small firm category in Table C-1, Table C-2, and Figure C-3. The main results persist, and the shift out of small firms to large firms along the GDP per capita spectrum becomes even more striking. The share of employment in small firms increases from 57 to 86% of employment in low-income countries. In lower-middle, upper-middle, and high-income countries, the shares increase from 52 to 78%, 33 to 53%, and 17 to 29%, respectively. We also observe a shift in how the self-employed enter employment. In low-income countries, the inclusion primarily inflates the small firm - low skill category, while in high-income countries, it increases the small firm - high skill category. Hence, self-employed individuals in low-income countries are typically less skilled than the average wage-workers in small firms, while they are more skilled than the average wage-worker in small firms in high-income countries. The other findings remain quantitatively similar, with large firms being more skill-intensive than small firms, but this gap closes with increasing GDP per capita, and large firms employ more skilled labor even when the skill premium is high.

Table C-1: Employment statistics: Including the self-employed

Panel A: Employment by skill

	Country Income Group			
Skill category	Low	Lower-middle	Upper-middle	High
High skill	0.15	0.29	0.55	0.86
Low skill	0.85	0.71	0.45	0.14
Number of countries	16	8	13	17

Panel B: Employment by firm size

	Country Income Group			
Firm size category	Low	Lower-middle	Upper-middle	High
Large firm	0.14	0.22	0.47	0.71
Small firm	0.86	0.78	0.53	0.29
Number of countries	16	8	13	17

Notes. This table shows the skill composition and employment by firm size when including self-employed workers in addition to wage-workers. In Panel B, self-employed individuals are counted as employed in small firms.

Table C-2: Skill intensity by firm size (L_h/L) & relative wages (w_h/w_l): Including the self-employed

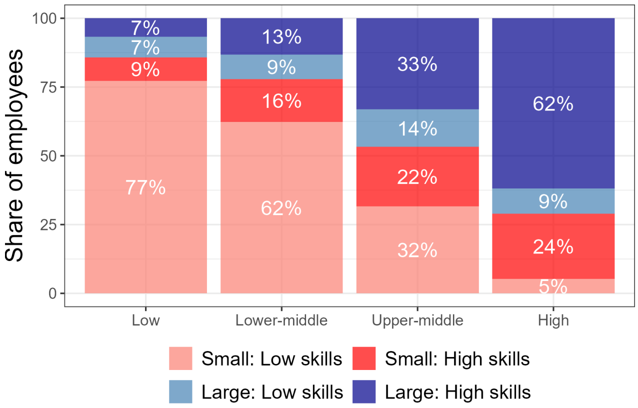
	Country Income Group			
Firm category	Low income	Lower-middle income	Upper-middle income	High income
Large firm	0.45	0.55	0.67	0.87
Small firm	0.1	0.22	0.46	0.83
N (Cross-section)	16	8	13	17
Relative wage (w_h/w_l)	2.54	2.3	1.63	1.5
N (Wage sample)	10	5	9	7

Notes. This table reports the average share of high-skill workers in large and small firms across income groups. The income groups correspond to low- [$0, $3,000], lower-middle- [$3,000, $10,000], upper-middle- [$10,000, $30,000], and high- [$30,000, ∞] income categories. The skill intensity is defined as the share of high-skill workers in total wage employment (L_h/L) and in large (L_h^b/L^b) and small firms (L_h^s/L^s). The relative wage is defined as the ratio of the average wage of high-skill workers to that of low-skill workers (w_h/w_l).

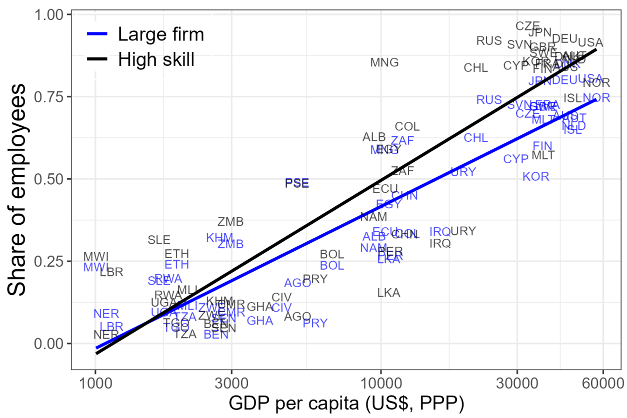
Figure C-3: Size, skill, premia: Including the self-employed

Panel A: Employment by firm size & skill.

(a) Composition of employment.

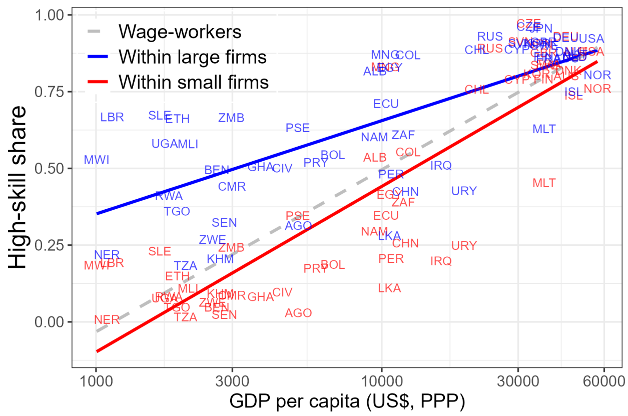


(b) Firm size & skill.

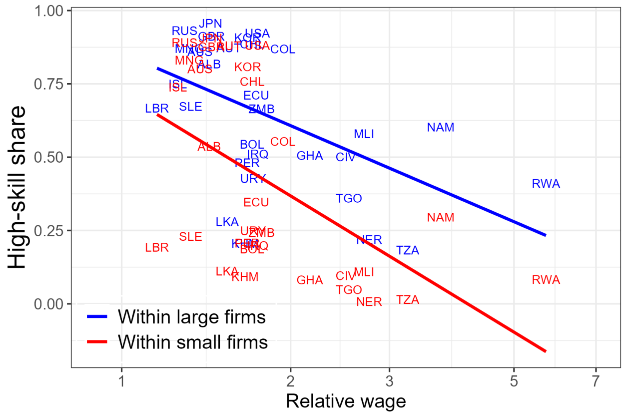


Panel B: Skill intensity by firm size & wage premia.

(c) Skill intensity by firm size (L_h^i/L^i).



(d) Skill intensity against relative wages.



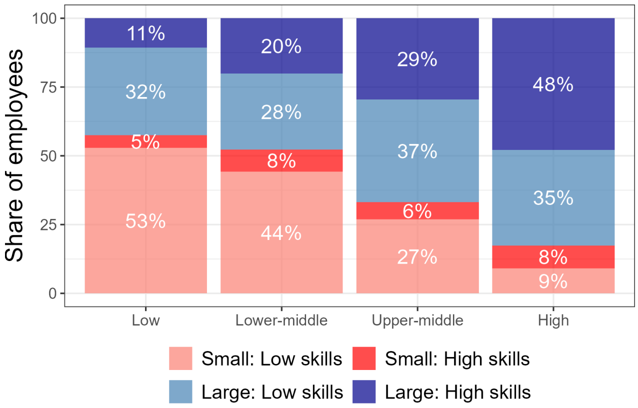
Notes. This figure is the equivalent of Figure 1, whereby we include self-employment as small firm employment. Panel (a) shows the share of low-skill and high-skill workers employed in small and large firms across four income groups. Shares are averaged over the country-year observations in each group. The income groups correspond to low- [\$0, \$3,000], lower-middle- [\$3,000, \$10,000], upper-middle- [\$10,000, \$30,000], and high- [\$30,000, ∞] income categories. Due to rounding, the reported shares may not sum to 100%. Panel (b) plots, against GDP per capita, the share of total wage employment in large firms (blue) and the share of high-skill wage workers in total wage employment (black). Workers are classified as high-skill if they have more than twelve years of schooling. Panel (c) plots, against GDP per capita, high-skill employment: i) as a share of all wage employment (grey dotted line), ii) as a share of employment in large firms (blue line), and iii) as a share of employment in small firms (red line). Panel (d) shows skill intensity (the fraction of high-skill employment in total employment within each firm size) against the relative wage of high- to low-skill workers. Skill intensity is shown for large firms (blue) and small firms (red). GDP per capita is in US dollars at purchasing power parity provided by Feenstra et al. [2015]. The lines represent the fitted values of a linear regression, and country-level observations are denoted by their corresponding ISO codes. The country-year observations and the underlying surveys are reported in Table A-2.

Alternative skill definition. We also show our main results for the alternative skill definition, where we define low-skilled as having 12 or fewer years of education. This corresponds to a high school degree in most countries. Hence, everyone with an education beyond a finished high school degree is high-skilled. We show the results in Table C-3, Table C-4, and Figure C-4. Naturally, the share of high-skilled employment is lower across all categories. The share of total employment by firm size is equal to the reported values in the main text. We find that the overall patterns are, again, robust to this alternative skill definition. As in the main text, large firms always employ more high-skilled workers than small firms, especially when skilled workers are scarce and expensive. As skills are more abundant in high-income countries and the skill premium is lower there, the gap in the skill intensity between small and large firms decreases.

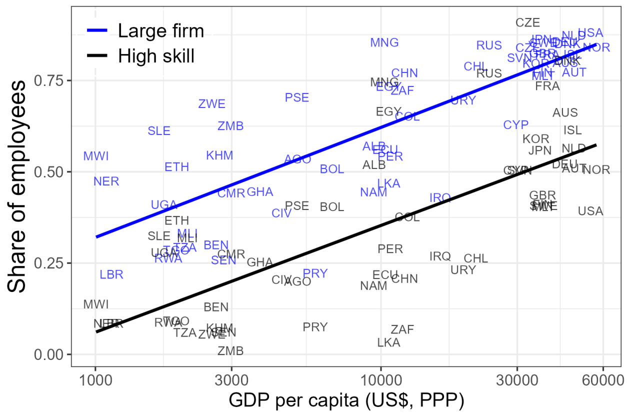
Figure C-4: Size, skill, premia: Alternative skill definition

Panel A: Employment by firm size & skill.

(a) Composition of employment.

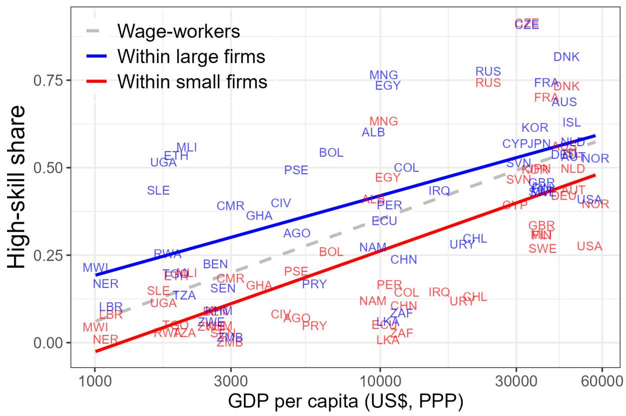


(b) Firm size & skill.

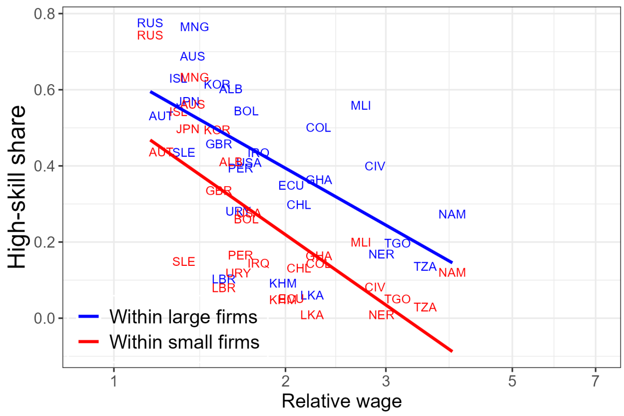


Panel B: Skill intensity by firm size & wage premia.

(c) Skill intensity by firm size (L_h^i/L^i).



(d) Skill intensity against relative wages.



Notes. Panel (a) shows the share of low-skill and high-skill workers employed in small and large firms across four income groups. Shares are averaged over the country-year observations in each group. The income groups correspond to low- [\$0, \$3,000], lower-middle- [\$3,000, \$10,000], upper-middle- [\$10,000, \$30,000], and high- [\$30,000, ∞] income categories. Due to rounding, the reported shares may not sum to 100%. Panel (b) plots, against GDP per capita, the share of total wage employment in large firms (blue) and the share of high-skill wage workers in total wage employment (black). Workers are classified as high-skill if they have more than twelve years of schooling. Panel (c) plots, against GDP per capita, high-skill employment: i) as a share of all wage employment (grey dotted line), ii) as a share of employment in large firms (blue line), and iii) as a share of employment in small firms (red line). Panel (d) shows skill intensity (the fraction of high-skill employment in total employment within each firm size) against the relative wage of high- to low-skill workers. Skill intensity is shown for large firms (blue) and small firms (red). GDP per capita is in US dollars at purchasing power parity provided by Feenstra et al. [2015]. The lines represent the fitted values of a linear regression, and country-level observations are denoted by their corresponding ISO codes. The country-year observations and the underlying surveys are reported in Table A-2.

Table C-3: Employment statistics: Alternative skill definition

Panel A: Employment by skill

	Country Income Group			
Skill category	Low	Lower-middle	Upper-middle	High
High skill	0.15	0.28	0.36	0.56
Low skill	0.85	0.72	0.64	0.44
Number of countries	16	8	13	17

Panel B: Employment by firm size

	Country Income Group			
Firm size category	Low	Lower-middle	Upper-middle	High
Large firm	0.43	0.48	0.67	0.83
Small firm	0.57	0.52	0.33	0.17
Number of countries	16	8	13	17

Notes. This table shows the skill composition and employment by firm size for our alternative skill definition.

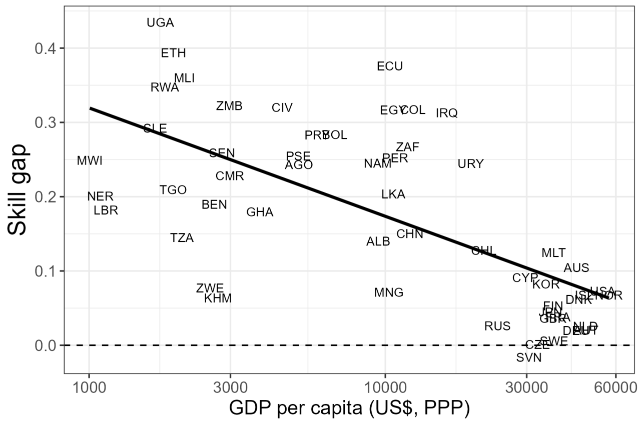
Table C-4: Skill intensity by firm size (L_h/L) & relative wages (w_h/w_l): Alternative skill definition

	Country Income Group			
Firm category	Low income	Lower-middle income	Upper-middle income	High income
Large firm	0.25	0.39	0.42	0.58
Small firm	0.08	0.17	0.24	0.48
N (Cross-section)	16	8	13	17
Relative wage (w_h/w_l)	3.12	2.48	1.81	1.43
N (Wage sample)	10	5	9	7

Notes. This table reports the average share of high-skill workers in large and small firms across income groups. The income groups correspond to low- [$0, $3,000], lower-middle- [$3,000, $10,000], upper-middle- [$10,000, $30,000], and high- [$30,000, ∞] income categories. The skill intensity is defined as the share of high-skill workers in total wage employment (L_h/L) and in large (L_h^b/L^b) and small firms (L_h^s/L^s). The relative wage is defined as the ratio of the average wage of high-skill workers to that of low-skill workers (w_h/w_l).

Aggregate vs. per-country patterns. As discussed in the main text, the fitted lines shown in Figure 1 may be driven by outliers and falsely indicate that the skill intensity between small and large firms is lower in low-income settings. To address this concern, we calculate the difference in the skill intensity of large and small firms for each country and plot this resulting skill gap, rather than the individual shares. In Figure C-5, we show that the results are not driven by some outliers, but that the skill gap is positive in almost all countries. On a per-country basis, large firms consistently have a skill mix that uses more skilled workers than small firms. This gap in the skill intensity is particularly large for countries with low GDP per capita.

Figure C-5: Gap in skill intensity (large vs. small firms)

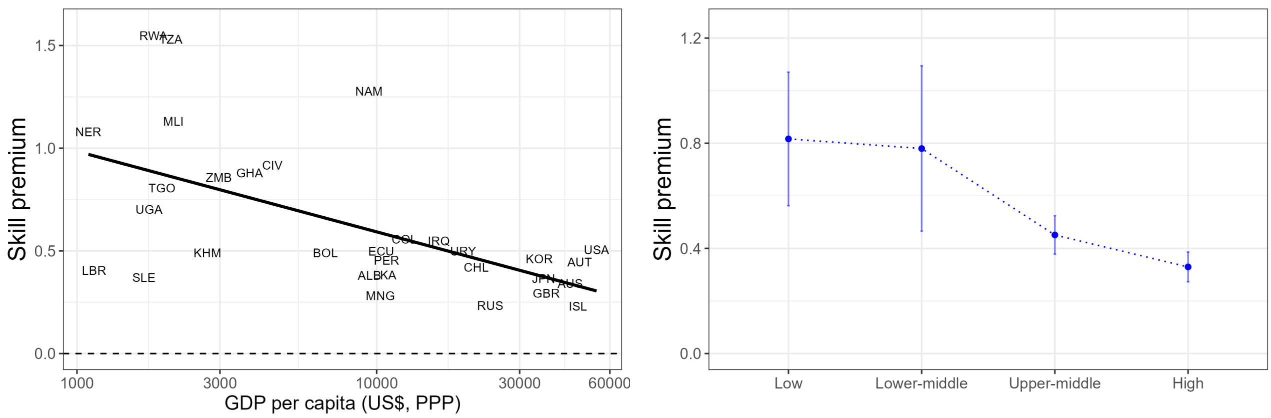


Notes. This figure shows the gap in the share of high-skilled employment between large and small firms. The skill gap is defined as: $L^b_h/L^b - L^s_h/L^s$. We show the skill gap against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015].

Mincer regression. In the main text, we have shown the "raw" relative average wages of high- vs. low-skilled workers, and Figure 2 has shown that the relative wages are higher in low-income countries. One concern may be that there are confounding factors, such that we are falsely attributing personal characteristics that influence wages to skill. To address this issue, we run a Mincer wage regression for each country separately.

The left panel of Figure C-6 plots the estimated β coefficients for each country as well as a fitted line through the estimated country coefficients. The right panel of Figure C-6 plots the average coefficient and its standard error within each of our income categories.

Figure C-6: Skill premium by income



(a) Country-level coefficient of being high-skilled from Mincer regression.

(b) Average of country-level skill premia estimated from Mincer regressions.

Notes. This figure shows the estimated coefficient of being high-skilled from running a Mincer regression for each country separately. Panel a plots the estimated coefficients against GDP per capita (in US dollars at purchasing power parity) as provided by Feenstra et al. [2015]. Panel b plots the average estimate within each income group. Error bars represent the 95% confidence intervals of the group means.

D Model appendix

D-1 Equilibrium definition

For a given set of parameters $\alpha, z_m, \gamma_i, \mu_i, rho_i, A, \nu, c^e, c^f$, an equilibrium consists of $\{w_h, w_l, L_h, L_l, M, M_e \hat{z}, z^*, m^b\}$ such that:

1. Skill content of employment for each technology $i \in \{s, b\}$:

$$\frac{L_l^i}{L_h^i} = \Big(A^{\frac{\rho^i-1}{\rho^i}} \, \frac{1-\mu^i}{\mu^i} \, \frac{w_l}{w_h}\Big)^{-\rho^i} \;=\; \Omega^i.$$

2. Labor Demand for skilled and unskilled workers i, $L_h^i(z)$:

$$L_h^i(z) \;=\; \Big(\frac{z^{\,1-\nu}\,\Theta^i\,\gamma^i}{\tilde{\Omega}^i\, w_h}\Big)^{\frac{1}{1-\gamma^i}}.$$

3. Productivity cutoffs for entry and technology choice

$$\pi^s(z^*) \;=\; \pi^b(z^*), \quad z^* = w_h^{\frac{1}{1-\nu}} \Big(\frac{\Pi^s}{\Pi^b}\Big)^{\frac{1}{1-\nu}} / \Big(\tfrac{1}{1-\gamma^b} - \tfrac{1}{1-\gamma^s}\Big),$$

$$\pi^s(\hat{z}) = c^f\,\Lambda\, w_h, \quad \hat{z} = \Big(\frac{c^f\,\Lambda}{\Pi^s}\Big)^{\frac{1-\gamma^s}{1-\nu}} \; w_h^{\frac{1}{1-\nu}}.$$

4. Free Entry condition when $\hat{z} < z^*$:

$$w_h = \Big(1/[c^e\,\Lambda]\Big)^{\frac{1-\nu}{\alpha}} \Big[A^s\,\Pi^s\,\Pi_2 + A^b\,\Pi^b\,\Pi_1 - z_m^{\alpha}\,c^f\,\Lambda\,\big(\tfrac{c^f\,\Lambda}{\Pi^s}\big)^{-\alpha\,\frac{1-\gamma^s}{1-\nu}}\Big]^{\frac{1-\nu}{\alpha}}.$$

   when $\hat{z} > z^*$ adjust integrals accordingly.

5. Labor market for high-skilled workers clears such that:

$$L_h = \frac{M}{1-G(\hat{z})}\Big[\Big(\frac{\Theta^s\,\gamma^s}{\tilde{\Omega}^s\,w_h}\Big)^{\frac{1}{1-\gamma^s}}\;\bar{z}^s + \Big(\frac{\Theta^b\,\gamma^b}{\tilde{\Omega}^b\,w_h}\Big)^{\frac{1}{1-\gamma^b}}\;\bar{z}^b\Big].$$

6. Labor market for low-skilled workers clears such that:

$$L_l = \frac{M}{1-G(\hat{z})}\Big[\Omega^s\,\Big(\frac{\Theta^s\,\gamma^s}{\tilde{\Omega}^s\,w_h}\Big)^{\frac{1}{1-\gamma^s}}\;\bar{z}^s + \Omega^b\,\Big(\frac{\Theta^b\,\gamma^b}{\tilde{\Omega}^b\,w_h}\Big)^{\frac{1}{1-\gamma^b}}\;\bar{z}^b\Big].$$

D-2 Derivations

In this section, we show the full derivations of the model equations.

Two types of firms exist (small and large). We refer to these using superscript s(mall) and b(ig) to avoid letter clashes, with generic superscript i.[9]

Firms differ in their productivity z. We abstract from physical capital. Each firm produces a final good using skilled and unskilled labor, L_h and L_l. These are combined in a CES production function with weight μ^i on the unskilled and elasticity of substitution ρ^i. These two parameters differ between small and large firms. Technology may be skill-biased. We denote the relative productivity of high-skilled workers by A. Production has decreasing returns to scale, with parameter $0 < \gamma^s < \gamma^b < 1$.

We also allow for an output tax τ, which may vary with a firm's productivity and can capture distortions à la Restuccia and Rogerson [2008] and others.

[9] Note that whereas in the empirical analysis above, large (small) referred to firms with at least (less than) ten employees, here large and small refer to technology choices. These will correlate with firm size, but the size threshold need not be at ten. In our quantitative analysis below, we compute model statistics both by size (as in the empirical analysis) and by technology type (the model object).

Output of a firm of size i with productivity z is then given by

$$y^i(z) = z\left[\mu^i L_l^{i\,\frac{\rho^i-1}{\rho^i}} + (1-\mu^i)(AL_h^i)^{\frac{\rho^i-1}{\rho^i}}\right]^{\frac{\rho^i}{\rho^i-1}\gamma^i}.$$

The firm chooses skilled and unskilled labor inputs to maximize profits. Dropping firm-type superscripts i for conciseness, the problem is to maximize

$$\pi(z) = \max_{L_l, L_h}(1-\tau(z))z\left[\mu L_l^{\frac{\rho-1}{\rho}} + (1-\mu)(AL_h)^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}\gamma} - w_l L_l - w_h L_h$$

The first-order conditions for this problem are

$$(1-\tau(z))\gamma\mu z\left(\frac{y(z)}{z}\right)^{1-\frac{\rho-1}{\rho\gamma}} L_l^{-\frac{1}{\rho}} = w_l$$

$$(1-\tau(z))\gamma(1-\mu)A^{\frac{\rho-1}{\rho}} z\left(\frac{y(z)}{z}\right)^{1-\frac{\rho-1}{\rho\gamma}} L_h^{-\frac{1}{\rho}} = w_h.$$

The optimal ratio of skilled to unskilled workers is thus common for all firms of a given size type, regardless of productivity z, and is given by

$$\left(\frac{L_l}{L_h}\right)^i = \left(A^{\frac{\rho-1}{\rho}}\frac{1-\mu^i}{\mu^i}\frac{w_l}{w_h}\right)^{-\rho^i}.$$

Denote this by Ω^i. This implies $L_l = \Omega^i L_h$, and

$$y(z) = zL_h^{\gamma}\underbrace{\left[\mu\Omega^{\frac{\rho-1}{\rho}} + (1-\mu)A^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}\gamma}}_{\equiv\Theta}.$$

From this, the first order condition for L_h is

$$(1-\tau(z))z\gamma\Theta L_h^{\gamma-1} = w_h\underbrace{\left(1+\frac{w_l}{w_h}\Omega\right)}_{\tilde{\Omega}}.$$

It follows that the optimal demand for skilled labor is

$$L_h(z) = \left(\frac{(1-\tau(z))z\Theta\gamma}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma}}.$$

Optimal overall employment in the firm is

$$L(z) = L_h(z) + L_l(z) = (1+\Omega)\left(\frac{(1-\tau(z))z\Theta\gamma}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma}}.$$

Optimal output (net of distortions) is

$$\begin{aligned} y^i(z) &= (1-\tau(z))z\left[\left(\frac{(1-\tau(z))z\Theta\gamma}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma}}\right]^{\gamma}\underbrace{\left[\mu\Omega^{\frac{\rho-1}{\rho}} + (1-\mu)A^{\frac{\rho-1}{\rho}}\right]^{\frac{\rho}{\rho-1}\gamma}}_{\equiv\Theta} \\ &= ((1-\tau(z))z\Theta)^{\frac{1}{1-\gamma}}\left(\frac{\gamma}{\tilde{\Omega}w_h}\right)^{\frac{\gamma}{1-\gamma}}. \end{aligned}$$

From this, it follows that variable profits of a firm with productivity z are

$$\begin{aligned}\pi(z) &= ((1-\tau(z))z\Theta)^{\frac{1}{1-\gamma}}\left[\left(\frac{\gamma}{\tilde{\Omega}w_h}\right)^{\frac{\gamma}{1-\gamma}} - (w_l\Omega + w_h)\left(\frac{\gamma}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma}}\right] \\ &= ((1-\tau(z))z\Theta)^{\frac{1}{1-\gamma}}\left(\tilde{\Omega}w_h\right)^{-\frac{\gamma}{1-\gamma}}\left[\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}}\right] \\ &\equiv \Pi((1-\tau(z))z)^{\frac{1}{1-\gamma}} w_h^{-\frac{\gamma}{1-\gamma}},\end{aligned}$$

where

$$\Pi = \Theta^{\frac{1}{1-\gamma}}\tilde{\Omega}^{-\frac{\gamma}{1-\gamma}}\left[\gamma^{\frac{\gamma}{1-\gamma}} - \gamma^{\frac{1}{1-\gamma}}\right]. \tag{D-1}$$

Profits increase monotonically in z, from 0 for z of 0 to infinity as z goes to infinity. Note that both Π and γ differ by firm type.

Size-dependent distortions. Following Buera and Fattal-Jaef [2018] and others, we model the output tax τ as

$$1-\tau(z) = z^{-\nu}.$$

This implies that for $\nu = 0$, $1-\tau = 1$ for all values of z, and there is no tax. For $\nu > 0$, after-tax revenue falls with productivity, so there are productivity-dependent distortions. With this functional form assumption, the profit function for type i is

$$\pi(z) = \Pi w_h^{-\frac{\gamma}{1-\gamma}} z^{\frac{1-\nu}{1-\gamma}}.$$

Technology choice Because $\gamma^b > \gamma^s$, $\pi^b(z)$ is less than $\pi^s(z)$ for small z, and is larger for large z. Hence, low-productivity firms prefer the small-firm technology, and high-productivity firms the large-firm technology. Denote the cutoff where $\pi_j^s(z) = \pi_j^b(z)$ by z_j^*. At this value,

$$\pi_j^s(z_j^*) = \pi_j^b(z_j^*)$$

This implies

$$\Pi^s(z^*)^{\frac{1-\nu}{1-\gamma^s}} w_h^{-\frac{\gamma^s}{1-\gamma^s}} = \Pi^b(z^*)^{\frac{1-\nu}{1-\gamma^b}} w_h^{-\frac{\gamma^b}{1-\gamma^b}}$$

Hence,

$$z^* = w_h^{\frac{1}{1-\nu}}\left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{1}{1-\nu}/\left(\frac{1}{1-\gamma^b}-\frac{1}{1-\gamma^s}\right)}$$

Because large firms are more skill-intensive, the optimal cutoff for being large, z^*, increases with the wage of skilled workers.

Entry and operational choice. There is an unlimited mass of potential entrants. To start a firm, an entrant pays an entry cost $c_j^e \cdot \Lambda w_h$, and then draws a productivity z from a distribution with *cdf* $G(z)$. We assume that G is a Pareto distribution with parameter α, so its *cdf* is $1-(z_m/z)^\alpha$. Active firms pay a fixed operating cost $c^f \Lambda w_h$.

Both entry costs and fixed costs are in units of labor, as in Klenow and Li [2025]. Given the difference in the skill composition of the population, we assume $\Lambda = w_l/w_h L_l + L_h$, which implies that these costs scale with the aggregate wage bill.

Due to the presence of fixed costs, not all firms are profitable. That is, a firm that drew a productivity z only operates if this yields positive profits. This occurs if productivity exceeds a threshold $\hat{z}_j$ at which

$$\max(\pi_j^s(\hat{z}_j), \pi_j^b(\hat{z}_j)) = 0.$$

Suppose for here that at $\hat{z}_j$, it is optimal to run a small firm, so that we will observe firms of both sizes run by entrepreneurs of both skill types in equilibrium. This implies

$$\Pi^s(\hat{z})^{\frac{1-\nu}{1-\gamma^s}} w_h^{-\frac{\gamma^s}{1-\gamma^s}} = c^f \Lambda w_h.$$

As a consequence,

$$\hat{z} = \left(\frac{c^f \Lambda}{\Pi^s}\right)^{\frac{1-\gamma^s}{1-\nu}} w_h^{\frac{1}{1-\nu}}.$$

If there are no productivity-dependent distortions ($\nu = 0$), the threshold is proportional to the high-skilled wage. It is lower when firm profitability is higher, and higher when fixed costs are higher. If fixed costs are sufficiently low, the cutoff productivity level $\hat{z}$ may fall below the minimum productivity level in the economy, z_m. In that case, $\hat{z} = z_m$.

Note that although the "natural case" is that $\hat{z} < z^*$, so that entrepreneurs choose to run both large and small firms, it is also possible that $\hat{z} > z^*$. Then z^* is not relevant, and the conditions for $\hat{z}$ feature γ^b and Π^b instead of γ^s and Π^s.

In the "natural case", entrants who draw $z < \hat{z}_j$ do not operate, those with $z > z_j^*$ operate a large firm, and those with $z \in (\hat{z}_j, z_j^*)$ operate a small firm.

Comparing thresholds The ratio of thresholds is

$$\frac{z^*}{\hat{z}} = \left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{1}{1-\nu}/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)} \left(\frac{\Pi^s}{c^f}\right)^{\frac{1-\gamma_s}{1-\nu}}.$$

if the constraint $\hat{z} \geq z_m$ is not binding.

Since entrants with $z \geq (<) z^*$ choose the large (small-) firm technology, the share of large firms is

$$m^b \equiv \frac{M^b}{M} = \frac{1-G(z^*)}{1-G(\hat{z})} = (\hat{z}/z^*)^\alpha = \left(\frac{\Pi^b}{\Pi^s}\right)^{\frac{\alpha}{1-\nu}/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)} \left(\frac{c^f}{\Pi^s}\right)^{\alpha\frac{1-\gamma_s}{1-\nu}},$$

if $z^* > \hat{z}$, and 1 otherwise.

Free entry. Firms enter until the expected value of entry, net of the entry cost, is zero. This implies

$$\begin{aligned} c^e \Lambda w_h &= \int_{\hat{z}}^{z^*} \pi^s(z) \mathrm{d}G(z) + \int_{z^*}^{\infty} \pi^b(z) \mathrm{d}G(z) \\ &= \Pi^s w_h^{-\frac{\gamma^s}{1-\gamma^s}} \alpha z_m^\alpha \int_{\hat{z}}^{z^*} z^{\frac{1-\nu}{1-\gamma^s} - \alpha - 1} \mathrm{d}z \\ &\quad + \Pi^b w_h^{-\frac{\gamma^b}{1-\gamma^b}} \alpha z_m^\alpha \int_{z^*}^{\infty} z^{\frac{1-\nu}{1-\gamma^b} - \alpha - 1} \mathrm{d}z - \left(\frac{z_m}{\hat{z}}\right)^\alpha c^f \Lambda w_h. \end{aligned}$$

$\bar{z}$. Define

$$\bar{z}^s \equiv \alpha z_m^\alpha \int_{\hat{z}}^{z^*} z^{\frac{1-\nu}{1-\gamma^s} - \alpha - 1} \mathrm{d}z \quad \text{and} \quad \bar{z}^b \equiv \alpha z_m^\alpha \int_{z^*}^{\infty} z^{\frac{1-\nu}{1-\gamma^b} - \alpha - 1} \mathrm{d}z.$$

With Pareto distributed z, these are

$$\bar{z}^b = \frac{\alpha z_m^\alpha z^{*\frac{1-\nu}{1-\gamma^b} - \alpha}}{\alpha - \frac{1-\nu}{1-\gamma^b}}$$

and

$$\bar{z}^s = \alpha z_m^\alpha \frac{\hat{z}^{\frac{1-\nu}{1-\gamma^s} - \alpha} - z^{*\frac{1-\nu}{1-\gamma^s} - \alpha}}{\alpha - \frac{1-\nu}{1-\gamma^s}}$$

With this definition, the free entry condition becomes

$$c^e \Lambda w_h = \Pi^s w_h^{-\frac{\gamma^s}{1-\gamma^s}} \bar{z}^s + \Pi^b w_h^{-\frac{\gamma^b}{1-\gamma^b}} \bar{z}^b - \left(\frac{z_m}{\hat{z}}\right)^\alpha c^f \Lambda w_h.$$

Using the expressions for z_j^* obtained above, the thresholds become

$$\bar{z}^b = \frac{\alpha z_m^\alpha}{\alpha - \frac{1-\nu}{1-\gamma^b}} w_h^{\frac{1}{1-\gamma^b} - \frac{\alpha}{1-\nu}} \left[\left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{1}{1-\nu}/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)}\right]^{\frac{1-\nu}{1-\gamma^b} - \alpha}$$

and

$$\bar{z}^s = \frac{\alpha z_m^\alpha}{\alpha - \frac{1-\nu}{1-\gamma^s}} w_h^{\frac{1}{1-\gamma^s} - \frac{\alpha}{1-\nu}} \left\{\left(\frac{c^f}{\Pi^s}\right)^{1-\alpha\frac{1-\gamma^s}{1-\nu}} - \left(\frac{\Pi^s}{\Pi^b}\right)^{\left(\frac{1}{1-\gamma^s} - \frac{\alpha}{1-\nu}\right)/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)}\right\}$$

if $\hat{z} > z_m$. Let

$$\Pi_1 \equiv \left[\left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{1}{1-\nu}/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)}\right]^{\frac{1-\nu}{1-\gamma^b} - \alpha}$$

$$\Pi_2 \equiv \left(\frac{c^f}{\Pi^s}\right)^{1-\alpha\frac{1-\gamma^s}{1-\nu}} - \left(\frac{\Pi^s}{\Pi^b}\right)^{\left(\frac{1}{1-\gamma^s} - \frac{\alpha}{1-\nu}\right)/\left(\frac{1}{1-\gamma^b} - \frac{1}{1-\gamma^s}\right)},$$

and

$$A^i \equiv \frac{\alpha z_m^\alpha}{\alpha - \frac{1-\nu}{1-\gamma^i}},$$

so that

$$\bar{z}^b = A^b w_h^{\frac{1}{1-\gamma^b} - \frac{\alpha}{1-\nu}} \Pi_1$$
$$\bar{z}^s = A^s w_h^{\frac{1}{1-\gamma^s} - \frac{\alpha}{1-\nu}} \Pi_2.$$

As a result, the free entry condition becomes

$$c^e \Lambda w_h + z_m^\alpha c^f \Lambda \left(\frac{c^f \Lambda}{\Pi^s}\right)^{-\alpha\frac{1-\gamma^s}{1-\nu}} w_h^{1-\frac{\alpha}{1-\nu}} = w_h^{1-\frac{\alpha}{1-\nu}} \left(A^s \Pi^s \Pi_2 + A^b \Pi^b \Pi_1\right)$$

This can be solved for

$$w_h = \left(\frac{1}{c^e \Lambda}\right)^{\frac{1-\nu}{\alpha}} \left[A^s \Pi^s \Pi_2 + A^b \Pi^b \Pi_1 - z_m^\alpha c^f \Lambda \left(\frac{c^f \Lambda}{\Pi^s}\right)^{-\alpha\frac{1-\gamma^s}{1-\nu}}\right]^{\frac{1-\nu}{\alpha}}$$

The wage increases in the firm profitability. The effect of c^f is more complicated because of selection.

Instead, if $\hat{z} = z_m$,

$$\bar{z}^s \bar{z}^s = A^s \left\{z_m^{\frac{1-\nu}{1-\gamma^s} - \alpha} - w_h^{\frac{1}{1-\gamma_s} - \frac{\alpha}{1-\nu}} \Pi_1\right\}.$$

Here, the free entry condition becomes

$$c^e \Lambda w_h = A^s \Pi^s w_h^{-\frac{\gamma^s}{1-\gamma^s}} z_m^{\frac{1-\nu}{1-\gamma^s} - \alpha} + \Pi_1 (A^b \Pi^b - A^s \Pi^s) w_h^{1-\frac{\alpha}{1-\nu}} - c^f \Lambda w_h.$$

This is a non-linear equation that determines w_h.

Labor market clearing. First, note that the number of firms M and the number of entrants M^e are related by

$$M = (1 - G(\hat{z}))M^e.$$

Then, for high-skilled workers,

$$\begin{aligned} L_h &= M^e \left[\int_{\hat{z}}^{z^*} L_h^s(z) \mathrm{d}G(z) + \int_{z^*}^{\infty} L_h^b(z) \mathrm{d}G(z) \right] \\ &= \frac{M}{1 - G(\hat{z})} \left[\left(\frac{\Theta^s \gamma^s}{\tilde{\Omega} w_h} \right)^{\frac{1}{1-\gamma^s}} \bar{z}^s + \left(\frac{\Theta^b \gamma^b}{\tilde{\Omega} w_h} \right)^{\frac{1}{1-\gamma^b}} \bar{z}^b . \right] \end{aligned}$$

This pins down M. It increases in the number of workers and decreases in $\bar{z}$. Note that $\bar{z}^s$ and $\bar{z}^b$ already contain the relative proportions of large and small firms. When $\hat{z} > z_m$, we can further use the expressions for $\bar{z}$ and $\hat{z}$ to obtain

$$L_h = M z_m^{-\alpha} \left(\frac{c^f}{\Pi^s} \right)^{\alpha \frac{1-\gamma^s}{1-\nu}} \left[\left(\frac{\Theta^s \gamma^s}{\tilde{\Omega}^s} \right)^{\frac{1}{1-\gamma^s}} A^s \Pi_2 + \left(\frac{\Theta^b \gamma^b}{\tilde{\Omega}^b} \right)^{\frac{1}{1-\gamma^b}} A^b \Pi_1 . \right]$$

For low-skilled workers,

$$L_l = \frac{M}{1 - G(\hat{z})} \left[\Omega^s \left(\frac{\Theta^s \gamma^s}{\tilde{\Omega} w_h} \right)^{\frac{1}{1-\gamma^s}} \bar{z}^s + \Omega^b \left(\frac{\Theta^b \gamma^b}{\tilde{\Omega} w_h} \right)^{\frac{1}{1-\gamma^b}} \bar{z}^b \right].$$

This pins down w_l (which, as w_l/w_h, enters Ω^i, and thus Θ^i and $\tilde{\Omega}^i$, and thus Π^i in the equation). In the computation of the equilibrium, it serves to verify the initial guess of w_l/w_h. When $\hat{z} > z_m$, we can again use the expressions for $\bar{z}$ and $\hat{z}$ to obtain

$$L_l = M z_m^{-\alpha} \left(\frac{c^f}{\Pi^s} \right)^{\alpha \frac{1-\gamma^s}{1-\nu}} \left[\Omega^s \left(\frac{\Theta^s \gamma^s}{\tilde{\Omega}^s} \right)^{\frac{1}{1-\gamma^s}} A^s \Pi_2 + \Omega^b \left(\frac{\Theta^b \gamma^b}{\tilde{\Omega}^b} \right)^{\frac{1}{1-\gamma^b}} A^b \Pi_1 . \right]$$

Average firm size equals $(L_l + L_h)/M$.

Aggregate output. Aggregate output net of distortions is:

$$\begin{aligned} Y &= \frac{M}{1 - G(\hat{z})} \left[\int_{\hat{z}}^{z^*} y^s(z) \mathrm{d}G(z) + \int_{z^*}^{\infty} y^b(z) \mathrm{d}G(z) \right] \\ &= \frac{M}{1 - G(\hat{z})} \left[(\Theta^s)^{\frac{1}{1-\gamma^s}} \left(\frac{\gamma^s}{\tilde{\Omega}^s w_h} \right)^{\frac{\gamma^s}{1-\gamma^s}} \bar{z}^s + (\Theta^b)^{\frac{1}{1-\gamma^b}} \left(\frac{\gamma^b}{\tilde{\Omega}^b w_h} \right)^{\frac{\gamma^b}{1-\gamma^b}} \bar{z}^b \right]. \end{aligned}$$

Equilibrium. Equilibrium variables: $w_h, w_l, L_h^s, L_l^s, L_h^b, L_l^b, M^s, M^b, z^*, \hat{z}, \bar{z}^b, \bar{z}^s$ s.t.

1. Skill mix, for each firm type:

$$\left(\frac{L_l}{L_h} \right)^i = \left(A^{\frac{\rho-1}{\rho}} \frac{1-\mu^i}{\mu^i} \frac{w_l}{w_h} \right)^{-\rho^i}.$$

2. Labor demand, for each firm type:

$$L_h^i(z) = \left(\frac{z^{1-\nu} \Theta^i \gamma^i}{\tilde{\Omega}^i w_h} \right)^{\frac{1}{1-\gamma^i}}.$$

3. Labor market clearing, high skill:

$$L_h = \frac{M}{1-G(\hat{z})}\left[\left(\frac{\Theta^s\gamma^s}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma^s}}\bar{z}^s + \left(\frac{\Theta^b\gamma^b}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma^b}}\bar{z}^b.\right]$$

4. Labor market clearing, low skill:

$$L_l = \frac{M}{1-G(\hat{z})}\left[\Omega^s\left(\frac{\Theta^s\gamma^s}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma^s}}\bar{z}^s + \Omega^b\left(\frac{\Theta^b\gamma^b}{\tilde{\Omega}w_h}\right)^{\frac{1}{1-\gamma^b}}\bar{z}^b\right]$$

5. Free entry:

$$w_h = \left(\frac{1}{c^e\Lambda}\right)^{\frac{1-\nu}{\alpha}}\left[A^s\Pi^s\Pi_2 + A^b\Pi^b\Pi_1 - z_m^{\alpha}c^f\Lambda\left(\frac{c^f\Lambda}{\Pi^s}\right)^{-\alpha\frac{1-\gamma^s}{1-\nu}}\right]^{\frac{1-\nu}{\alpha}}$$

if $\hat{z} > z_m$, and

$$c^e\Lambda w_h = A^s\Pi^s w_h^{-\frac{\gamma^s}{1-\gamma^s}} z_m^{\frac{1-\nu}{1-\gamma^s}-\alpha} + \Pi_1(A^b\Pi^b - A^s\Pi^s){w_h}^{1-\frac{\alpha}{1-\nu}} - c^f\Lambda w_h.$$

otherwise

6. Optimal continuation:

$$\pi(\hat{z}) = 0 \Leftrightarrow \hat{z} = \max\left(z_m, \left(\frac{c^f\Lambda}{\Pi^s}\right)^{\frac{1-\gamma^s}{1-\nu}} w_h^{\frac{1}{1-\nu}}\right).$$

7. Firm size choice:

$$\pi^s(z^*) = \pi^b(z^*) \Leftrightarrow z^* = {w_h}^{\frac{1}{1-\nu}}\left(\frac{\Pi^s}{\Pi^b}\right)^{\frac{1}{1-\nu}/\left(\frac{1}{1-\gamma^b}-\frac{1}{1-\gamma^s}\right)}$$

Other definitions and useful objects:

- $\bar{z}$:

$$\begin{aligned}\bar{z}^b &= A^b{w_h}^{\frac{1}{1-\gamma^b}-\frac{\alpha}{1-\nu}}\Pi_1\\ \bar{z}^s &= A^s w_h^{\frac{1}{1-\gamma^s}-\frac{\alpha}{1-\nu}}\Pi_2.\end{aligned}$$

  if $\hat{z} > z_m$ or

$$\bar{z}^s = A^s\left\{z_m^{\frac{1-\nu}{1-\gamma^s}-\alpha} - {w_h}^{\frac{1}{1-\gamma_s}-\frac{\alpha}{1-\nu}}\Pi_1\right\}$$

  otherwise.

- Share large firms:

$$m^b \equiv \frac{M^b}{M} = \frac{1-G(z^*)}{1-G(\hat{z})} = (\hat{z}/z^*)^{\alpha} = \left(\frac{\Pi^b}{\Pi^s}\right)^{\frac{\alpha}{1-\nu}/\left(\frac{1}{1-\gamma^b}-\frac{1}{1-\gamma^s}\right)}\left(\frac{c^f}{\Pi^s}\right)^{\alpha\frac{1-\gamma_s}{1-\nu}},$$

  if $\hat{z} > z_m$.

Note that the auxiliary parameters Ω, Θ, and Π all depend on the wage ratio w_l/w_h and vary by firm type.