Jobless Industrialization

Laura Alfaro [*] Marcela Eslava[†] Martí Mestieri[‡] Luis Felipe Sáenz[§]

March 20, 2025

Preliminary Please do not circulate

Abstract

This paper examines how resource misallocation across firms, sectors, and between modern and traditional segments of the economy affects development. Using detailed household and firm microdata from 1970s Colombia, combined with a general equilibrium occupational choice model, we show that low levels of human capital, along with firm-specific distortions, trap labor in low-productivity activities, sustaining low aggregate income and widespread self-employment. This generates "jobless industrialization," where industrial growth increases value-added without proportional employment gains. We leverage our microdata, which is representative of the formal and informal workforce and businesses, to jointly estimate the distribution of human capital- productivity, modern firms' technology parameters, demand elasticity, and firm-specific distortions. These estimates serve as the foundation for our quantification analysis and counterfactual simulations. We find that improving human capital is the most effective way to boost income and reduce inequality. A 25% increase in embodied human capital could more than double income levels and close half the income gap with the United States while narrowing income disparities between informal and formal workers. A 25% contraction in firm-specific and regulatory distortions—such as size-based constraints and minimum wage policies— impacts development to a lesser extent.

Keywords: Informality, (De)industrialization, Value Added, Employment, Distortions.

JEL Codes: E24, E26, O11, O14, O25.

[*]Harvard Business School

[†]Universidad de Los Andes.

[‡]Institut d'Analisi Economica - CSIC, CREI, BSE and CEPR.

[§]University of South Carolina.

We thank the audiences at the ASSA Meetings, CEDE, CEPR-Bread, IEA Conference, LACEA, Oxford-STEG conference, Penn State, RIDGE, SEA, UAB, USC, Tubinguen for their comments and suggestions. We thank Dani Rodrik for the useful discussion and comments and Federico Huneeus for his discussion of an earlier draft. Andrés Soto has provided outstanding research assistance.

1 Introduction

Emerging economies often struggle to allocate resources efficiently to their more productive uses. Compared to richer counterparts, these economies misallocate their resources across sectors and across producers within a sector, leading to significant income losses (Hsieh and Klenow 2009; Restuccia and Rogerson 2008; McMillan and Rodrik 2011). These economies also exhibit a partially dual character, with a substantial portion of their workforce working in informal segments with large presence of self-employment and micro-entrepreneurship, low capital intensity, and outdated production methods (Lewis 1979; La Porta and Shleifer 2014; Eslava et al. 2023). Such segments particularly dominate the bottom end of the income distribution in these societies, which frequently are highly unequal and lack a solid middle class.

We leverage rich historical microdata and a general equilibrium model with occupational choice to quantitatively assess the extent to which three key margins of misallocation have shaped the underwhelming development trajectories of today's emerging economies: (i) misallocation across producers within a sector, (ii) across sectors, and (iii) between modern and traditional segments. Our analysis underscores the limited job creation capacity of modern firms in these economies, driven by the interaction between human capital distribution, distortions across segments and producers, and the uneven concentration of modern industries across sectors. We show how these factors contributed to *jobless industrialization* in emerging markets—where industrial growth increased value-added without a proportional rise in employment. Our findings suggest that enhancing embodied human capital and technology are the most powerful levers for boosting aggregate income while simultaneously reducing modern sectors' joblessness and inequality. Although distortions affecting modern firms—whether firm-specific, size-correlated, or broader regulatory constraints—are also significant, their role is comparatively less pronounced, particularly in environments with high human capital.

We examine the role of different barriers that contribute to jobless modernity, with a focus on four key factors: human capital constraints, labor market distortions, firm-level inefficiencies, and the adoption of labor-saving technologies. To disentangle the role of these mechanisms, we introduce a general equilibrium model of constrained occupational choice, in which modern and traditional firms emerge endogenously across sectors based on the distribution of human capital, institutional distortions, and available technologies. Our model provides a framework for understanding the occupational allocation of workers and the underlying forces shaping employment dynamics in modern and traditional segments of the economy. The mechanisms account for the empirical patterns observed in LDCs, including low average incomes, a large informal sector, and misallocating productive talent within the modern economy. Additionally, our model accounts for heterogeneity in occupational choice, allowing for the possibility that some high-human-capital individuals may choose traditional/informal production due to entrepreneurial preferences or specialized skills.

We inform the quantification of our model with a rich micro-data from Colombia during its peak period

of industrialization in the early 1970s. This historical episode provides a unique empirical setting to study jobless modernity, as Colombia experienced a surge in manufacturing's share of GDP without a proportional increase in employment. Our dataset integrates household labor surveys with the Annual Manufacturing Survey—a comprehensive census of modern manufacturing firms—enabling us to estimate key structural parameters, including firm-level productivity, labor regulations, and employment constraints. Crucially, our framework allows us to disentangle different sources of inefficiency, such as payroll taxes, revenue distortions, and firm-size-dependent constraints, as well as the role played by human capital and labor saving technologies, allowing for a more precise identification of the forces driving jobless modernity. Through counterfactual analysis in our calibrated economy, we analyze the role of these mechanisms in generating limited employment in modern firms, inequality and low-levels of income per capita. We highlight three important takeways from our exercises.

First, our findings indicate that enhancing embodied human capital is the most effective mechanism for increasing aggregate income, reducing joblessness and reducing inequality. A 25% rightward shift of the distribution of embodied human capital closes half the income gap with the US (more than a two-fold increase), reduces the income gap between informal and formal workers from 40% to 10%, and the gap between formal workers and their employers by 16%. It also increases modern employment by 13%, significantly reducing joblessness in modernity. By contrast, a similar 25% shift in the distribution of managerial ability increases income by 40% while also increasing inequality by mainly benefiting modern entrepreneurs.

Second, the adoption of modern labor-saving technologies also has a substantial impact on welfare and joblessness. In particular, assigning a technology that has the labor intensity of Colombian firms in the 2000s (which we also estimate using Colombian firm-level data from) to the 1974 Colombian economy would boost entry into entrepreneurship and, through general equilibrium effects, pool more workers out of self-employment. As a result, the entire economy would become substantially more productive, reducing aggregate prices and boosting real averaging income almost four-fold. While this increases top-bottom inequality substantially and reduces the nominal income of the self-employed, we find that the welfare level of the self-employed would almost double because the drop in prices more than offsets the decline in nominal income.

Finally, distortions affecting modern firms—-whether firm-specific, size-correlated, or linked to a flat minimum wage—-are also significant, but less so than improvements in embodied human capital or labor-saving technologies. Reducing labor distortions (the minimum wage or labor taxes) by 25% increases aggregate income close to 10%. We find a larger effect of close to 15% if the reduction applies to revenue distortions. Reducing the dispersion of distortions across the firm-size distribution has a similar (albeit slightly smaller) effect. These results are driven by the rise in modernity, which is more than proportional in employment relative to value added. The decrease in joblessness is most marked when the regulatory change falls on the minimum wage. However, unlike investments in embodied human capital, reductions

in the minimum wage widen the income gap between formal workers and their employers. In fact, a significant increase in the real minimum wage, such as Colombia has adopted since the 1970s, can erode the gains in inequality reduction of a 25% increase in embodied human capital, in fact leading to larger income gaps between informal workers and those in the formal sector that before the combined policy change. Overall, we find that cross-country differences in structural change depend on differential barriers to the adoption of modern production and the differential concentration of modernity in different sectors.

Related Literature This paper contributes to several strands of literature. First, it builds on the extensive research examining the impact of misallocation on economic development, particularly the role of firm-size distortions in productivity gaps across development levels (Tybout 2000; Restuccia and Rogerson 2008; Hsieh and Klenow 2009; Alfaro et al. 2009; Eslava et al. 2023; Bento and Restuccia 2021). A central mechanism in our analysis is the coexistence of firms with heterogeneous productivity in equilibrium alongside labor and product market institutions that differentially affect firms based on their productivity levels. While much of this literature focuses on firm-level distortions, we extend this framework by emphasizing the misallocation between the economy's modern/formal and traditional/informal segments. Our approach integrates microdata covering employment and value generation in formal and informal businesses, including self-employed individuals—an often-overlooked but crucial component of economic activity in LDCs.

A persistent challenge in the misallocation literature is linking firm-specific distortions, inferred from deviations in firm size, to specific regulatory constraints. The richness of our dataset allows us to disentangle the effects of firm-level payroll taxes, revenue distortions, and other factor-biased inefficiencies. We quantify the relative impact of these distortions, along with the minimum wage, on the growth of modern firms. Our analysis reveals that these distortions are size-dependent: payroll taxes and revenue distortions disproportionately hinder the expansion of the most productive firms, whereas residual capital distortions tend to constrain the growth of lower-productivity businesses. Disentangling labor-market distortions is particularly relevant in our study, not only because of its direct link to jobless modernization but also due to previous findings that labor-related distortions have a more significant impact than revenue distortions in shaping firm-size distributions and productivity gaps in Colombia's modern sector (Eslava et al. (2023)).

Our focus on the formal-informal labor market divide is also rooted in early development literature (Lewis 1979; Harris and Todaro 1970). In many developing economies, the informal sector accounts for nearly half of all employment, with competing views on its role in economic development (Lewis 1954; Harris and Todaro 1970; Rauch 1991; La Porta and Shleifer 2014). We build on these perspectives by framing misallocation away from high-productivity firms as a key explanatory factor. In doing so, we expand the concept of misallocation beyond firm-level distortions to include the broader segmentation of economic activity between modern/formal and traditional/informal sectors, particularly in the context of labor-biased distortions.

Our paper also relates to recent work on the role of labor market institutions in shaping employment patterns across formal and informal jobs. For example, Dix-Carneiro et al. (2021) examine how labor regulations interact with trade dynamics to influence employment allocation. While their study emphasizes trade effects, we isolate the tensions between labor, product, and capital market institutions, and explore how these frictions interact with human capital and technological constraints. A key distinction in our analysis is the introduction of occupational choice as a fundamental mechanism, which is critical for understanding the employment-value added gap in LDCs, given the dominance of self-employment in the informal sector. This aligns with findings from Amodio et al. (2022), who document how self-employment influences monopsony power in labor markets, particularly in Peru.

Our research intersects with the literature on premature deindustrialization (Rodrik 2016; Herrendorf et al. 2014; McMillan and Zeufack 2022). The phenomenon of "hump-shaped" manufacturing shares along the growth path—where industrialization peaks earlier and at lower income levels in late industrializing economies—has been widely documented (Rodrik 2016). Our findings complement this literature by highlighting the role of jobless industrialization, wherein the growth of modern sectors does not translate into proportional employment gains. This aligns with Herrendorf et al. (2022), who show that labor productivity gaps are more significant in manufacturing than in the aggregate economy in poorer countries, contradicting earlier claims of absolute convergence in manufacturing productivity (Rodrik 2013). Their inclusion of small, informal firms in the analysis is particularly relevant to our approach, as we emphasize the role of informality in shaping employment dynamics.

Beyond the development literature, our study contributes to ongoing debates on the structural drivers of premature deindustrialization. Recent work proposes a range of mechanisms to explain this phenomenon, including differences in agricultural productivity (Huneeus and Rogerson 2020), technology adoption lags (Fujiwara and Matsuyama 2020), and trade integration effects (Sposi et al. 2021). While these studies highlight heterogeneous development paths, our analysis focuses specifically on the divergence between employment and value-added along the development trajectory, offering a novel explanation for jobless industrialization and its implications for firm-size distribution.

Methodologically, our paper builds on production function estimation techniques, particularly those developed in Eslava et al. (2023), to identify firm-level distortions and productivity wedges. Our empirical approach leverages a newly constructed historical micro-dataset on manufacturing, allowing us to analyze misallocation over a 70-year period. More broadly, our work aligns with a growing body of literature that uses microdata to inform macroeconomic development patterns, including studies on financial frictions (Buera and Moll 2015) and human capital constraints (Mestieri et al. 2017). Buera et al. (2021) provides a useful survey of these methodologies.

The remainder of the paper is organized as follows. Section 2 presents the theoretical model that underpins our quantitative analysis. Sections 3 and 4 focus on the empirical quantification. In Section 3, we estimate key parameters and moments for modern manufacturing, leveraging detailed historical micro-

data from Colombia. Section 4 calibrates the model and conducts counterfactual exercises to assess the impact of product and labor market distortions on employment generation. The final section concludes.

2 Framework

We develop a rich model that captures the salient features of emerging economies—such as Colombia—discussed so far. Importantly, this model serves as the coherent basis for our measurement (including the production function estimation), calibration and counterfactual analysis.

Our economy has four fundamental sources of heterogeneity. First, in terms of production, we allow the coexistence of self-employment and modern technologies. Modern technologies are run by entrepreneurs and employ labor, capital, and intermediates as inputs. Second, the economy is composed of three sectors that load differently on self-employment and modern technologies. For example, we find that manufacturing is more modern-intensive than agriculture. Third, agents in the economy are heterogeneous in their abilities and, because of this, they self-select into being self-employed, workers of modern firms or entrepreneurs. Fourth, there are distortions that affect production technologies asymetrically. Self-employment escapes these distortions while modern firms face idiosyncratic distortions that vary across the firm types. In particular, modern firms hire workers while facing minimum wage constraints and size-dependent distortions affecting revenue, labor, and capital (which we estimate jointly with the production function estimation).

2.1 Preferences

Consumers' preferences are defined over the consumption of three types of sectoral goods: agriculture (A), industry (I), and services (S):

$$U = \left(\sum_{i \in \{A,I,S\}} \lambda_i^{\frac{1}{\epsilon}} C_i^{\frac{\epsilon-1}{\epsilon}} \right)^{\frac{\epsilon}{\epsilon-1}}$$

with $0 < \epsilon < 1, \lambda_i > 0$ for all i and the normalization that $\lambda_I = 1$. In turn, the production of the sectoral goods is performed by a combination of modern (M) and traditional (T) production, according to:

$$C_i = \left(\phi_{i,M}^{\frac{1}{\rho}} C_{i,M}^{\frac{\rho-1}{\rho}} + (1-\phi_{i,M})^{\frac{1}{\rho}} C_{i,T}^{\frac{\rho-1}{\rho}} \right)^{\frac{\rho}{\rho-1}}$$

with $\phi_{i,M} \in (0,1)$ for all i and $\rho > 0$.

Modern varieties are imperfect substitutes, while traditional production is not differentiated:

$$C_{i,M} = \left(\sum_{j \in N_{i,M}} d_{i,j}^{\frac{1}{\sigma}} c_{i,j}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}} ; C_{i,T} = \sum_{j \in N_{i,T}} q_{i,j}$$

where $\sigma > 1$, j is a subindex for firms and $\{N_{i,M}, N_{i,T}\}$ denote, respectively, the sets of firms in modern and traditional production in sector i. The $d_{i,j}$ vector captures quality/taste differentials across modern varieties in sector i.

The nested CES structure implies the following demand system for different segments within sectors and for sectoral goods:

$$\begin{aligned} &\frac{E_i}{E} = \lambda_i \left(\frac{P_i}{P} \right)^{1-\epsilon} \qquad \frac{E_{i,M}}{E_i} = \phi_{i,M} \left(\frac{P_{i,M}}{P_i} \right)^{1-\rho} \qquad \frac{E_{i,T}}{E_i} = (1-\phi_{i,M}) \left(\frac{P_{i,T}}{P_i} \right)^{1-\rho} \\ &P_i = \left(\phi_{i,M} P_{i,M}^{1-\rho} + (1-\phi_{i,M}) P_{i,T}^{1-\rho} \right)^{\frac{1}{1-\rho}} \qquad P = \left(\sum_{i \in \{A,I,S\}} \lambda_i P_i^{1-\epsilon} \right)^{\frac{1}{1-\epsilon}} \end{aligned} \tag{1}$$

where $P_{i,T}$ is the price of traditional goods in sector i and final expenditure in the different segments is given by $E_i = P_i C_i$, $E_{i,M} = P_{i,M} C_{i,M}$, and $E_{i,T} = P_{i,T} C_{i,T}$.

A modern firm j in sector i faces demand from consumers:

$$c_{i,j} = \frac{E_{i,M}}{P_{i,M}^{1-\sigma}} d_{i,j} p_{i,j}^{-\sigma} = d_{i,j} \left(\frac{p_{i,j}}{P_{i,M}} \right)^{-\sigma} \left(\frac{P_{i,M}}{P_i} \right)^{-\rho} \left(\frac{P_i}{P} \right)^{-\epsilon} \frac{\phi_{i,M} \lambda_i E}{P} \tag{2}$$

where $P_{i,M} = \left(\sum_{j \in N_{i,M}} d_{i,j} p_{i,j}^{1-\sigma} \right)^{\frac{1}{1-\sigma}}$ is an aggregate price index of the modern segment in sector i, and the last equality comes from replacing $E_{i,M}$ using system 1.

2.2 Technologies

An agent can operate one technology only, modern or traditional (not both) in any sector of their choosing (but only one sector). Entrepreneurs bring their productivity a to their businesses. Those operating the modern technology hire workers (l) and use capital (k) and modern intermediate inputs (m) to produce. In particular, an agent with the ability a_j operating one of these technologies generates output according to:

$$j \in N_{i,M} : q_{i,j} = a_j l_j^{\gamma_l} k_j^{\gamma_k} m_j^{\gamma_m} \tag{3}$$

$$j \in N_{i,T} : q_{i,j} = a_i \tag{4}$$

with $\gamma_l + \gamma_k + \gamma_m = 1$. Within a segment (modern, traditional), technologies are assumed to be the same across sectors. We thus assign all heterogeneity across sectors to preferences, which makes our interpretation cleaner.[1]

Traditional production is only used for final consumption, while modern production is used both for final and intermediate consumption. Total intermediate consumption aggregates the use of intermediary inputs produced by (modern) firms indexed by j across the three sectors,

$$\mathrm{Y}_m = \left(\sum_{i \in \{A,I,S\}} \sum_{j \in N_{i,M}} d_{i,M,j}^{\frac{1}{\sigma}} v_{i,j}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}$$

The resulting demand for intermediates faced by producer j is:

$$v_{i,j} = d_{i,j} \left(\frac{p_{i,j}}{p_m} \right)^{-\sigma} \mathrm{Y}_m \tag{5}$$

where total expenditure in intermediates is $p_m \mathrm{Y}$ and the aggregate price of intermediates is:

$$p_m = \left(\sum_{i \in \{A,I,S\}} \sum_{j \in N_{i,M}} d_{i,j} \, p_{i,j}^{1-\sigma} \right)^{\frac{1}{\sigma-1}} \tag{6}$$

As we argue below, we can normalize the price of the traditional sector to one in all sectors without loss of generality. The value of (gross) total output in the economy is therefore:

$$\sum_{i \in \{A,I,S\}} \left(\sum_{j \in N_{i,T}} q_{i,j} + \sum_{j \in N_{i,M}} p_{i,j} q_{i,j} \right) = \sum_{i \in \{A,I,S\}} \left(\sum_{j \in N_{i,T}} q_{i,j} + \sum_{j \in N_{i,M}} p_{i,j} (c_{i,j} + v_{i,j}) \right) = E + p_m \mathrm{Y}_m$$

where we are allowing for the price of each formal producer j, $p_{i,j}$, to differ since modern producers are heterogeneous. A producer, however, charges the same price for the same good independent of whether it is used for final or intermediate consumption.

2.3 Occupational Choice

The economy is composed of a mass of N agents with heterogeneous abilities, each endowed with one unit of time to be allocated between operating one of the two technologies described above (M, T) or supplying labor. An agent with draws (a', d'', Γ, s'''):

- Supplies a' effective units of labor to a modern firm if working in the labor market, for a wage of w per unit of effective labor, earning wa'.

[1] There is little empirical guidance regarding technological differences across sectors within modern and traditional segments. However, Norris Keiller et al. (2024) find similar production elasticities for sectors in services (restaurants and retail) and manufacturing electronics.

- Produces $a's'''$ units of output if operating a traditional technology with production function 4 in any sector, earning $a's'''$ since the price of traditional goods is the numeraire.
- Operates a modern firm with productive efficiency a' in production function 3 and quality/taste differential d'' in the demand function, in any of the three sectors in the economy, facing taxes Γ and earning after-tax profits.

An agent can only choose one occupation. The natural choice is, thus, the occupation yielding the highest income. Agents attributes (a, d, s) are drawn from independent distributions with densities $f_a(a)$, $f_d(d)$ and $f_s(s)$. We estimate the shapes of these distributions from micro data on individuals and firms. $f_a(a)$ is truncated at a_{min} and a_{max}, while $f_d(d)$ has associated cumulative Pareto distribution F_d, truncated from below at d_{min}. Although abilities are independently drawn, there may be ex-post correlation between a and d among formal firms, given selection into occupations. Taxes Γ are drawn from distributions that may or may not be independent from a and d, also estimated from microdata (on modern firms).

2.4 Markets

We assume that all modern firms, independent of the sector they are in, acquire inputs in non-segmented markets. Thus, there is a single equilibrium price in each of those input markets that all modern firms face. We allow for firm-specific taxes (distortions) to revenue and to capital and labor, but not for materials.

Labor market: The labor market, where only modern firms demand labor, is subject to regulations. Wage payments by modern firms to any given individual worker cannot be less than a minimum earning requirement e_{min} ($wa_j \geq e_{min}$ for any worker with ability a_j). There is also a tax on all hired labor, τ^w, so that the labor cost for modern firms per efficiency unit is $(1 + \tau^L)w$.

Capital markets: There is a fixed capital stock that modern firms can access at a fixed rental rate of r per unit of capital. Conditions to access capital, however, may not be homogeneous across firms. Thus, the effective rental total cost to firm j is $(1 + \tau_j^k)r$.

Market for materials: A modern firm in any sector buys inputs from all other modern firms. We assume away distortions to the use of materials, an assumption that is standard, for instance, in the estimation of production functions, markups, and labor markdowns. Consistent with this assumption, Eslava et al. (2023) find that in the Colombian data used in our calibration, the relevant dispersion in input prices is almost entirely attributable to wages rather than to dispersion in materials prices.

Good markets: there is a revenue tax on the output produced in the modern, i.e. formal, segment. The revenue tax, denoted τ^R, may be firm-specific and may be correlated with firm characteristics a_j, d_j, where j denotes the entrepreneur who owns the firm. By contrast, there is no revenue tax for traditional (informal) producers.

We assume all tax payments from firms are rebated to all agents in the form of lump sum transfers. Thus, a modern firm j faces $\Gamma_j = \{\tau_j^R, \tau_j^l, \tau_j^k\}$. While we refer to τ^R, τ^l, and τ^K as *taxes*, these more generally

represent distortions to revenue and the use of inputs. In our quantification, we measure τ_l from data on employers' contributions to social security and other government programs that are mandatory and paid per worker. We leave τ_k and τ_R to capture, respectively, other factor-biased distortions and factor-unbiased distortions to business size.

Because we assume that only modern firms face minimum wage requirements and taxes, "modern" and "formal" are interchangeable in our setting, as are "traditional", "informal" and "self-employed".[2]

2.5 Equilibrium

A competitive equilibrium is given by a set of prices $(w, \{p_{i,j}\}, P_{i,M})$ such that:

1. Agents' (mutually exclusive) occupational choices maximize each agent's income. That is, the choice of an agent of talent (a, d) between becoming a worker, formal entrepreneur, or self-employed maximizes the agent's income.

2. The total supply of effective (modern) labor equals the total demand of effective labor in each sector i:

$$\sum_{j \in workers_i} a_{i,j} = \sum_{i \in \{A,I,S\}} \sum_{j \in N_{i,M}} l_{i,j}$$

3. The markets for traditional and modern goods clear in each sector. Total income, inclusive of taxes rebated by the government, equals total final expenditure:

$$E = \sum_{i \in A,I,S} \left(\sum_{j \in N_{i,T}} q_{i,j} + \sum_{j \in N_{i,M}} c_{i,j} \right) = \sum_{i \in A,I,S} \left(\sum_{j \in N_{i,T}} a_j + \sum_{j \in N_{i,M}} (\pi_j + wl_j + rk_j) \right) + \hat{T}$$

where $\hat{T} = \sum_{i \in A,I,S} \sum_{j \in N_{i,M}} \left(\tau_j^R p_{i,j} q_{i,j} + \tau_j^l l_{i,j} + \tau_j^k k_{i,j} \right)$.

We now solve for the equilibrium. To characterize the income of a modern entrepreneur, which we need to solve for the optimal occupational choice as a function of the agent's (a, d), we start solving the problem of the modern firm. We then characterize the occupational choice. For ease of exposition, we present the solution to a model where $s = 1$ for all agents but re-introduce s for the quantification.

2.5.1 Factor Demand in the Modern Sector

The profit maximization problem of a modern firm operating in sector i is

[2] While obviously overly simplistic, these assumptions capture the much higher prevalence of different forms of informality in the traditional segments in less developed economies. When taking the model to the data, we label as traditional all businesses with less than 10 employees (including self-employment). Using data from 11 countries in Latin America, (Eslava et al. 2024) find that labor informality, defined as not saving for a pension, had a prevalence of 75% in this segment vs. 12% in the group of modern firms (those with 10+ employees).

$$\max_{\{l_{i,j};k_{i,j};m_{i,j}\}} (1-\tau_j^R)p_{i,j}q_{i,j} - w(1+\tau_j^l)l_{i,j} - r(1+\tau_j^k)k_{i,j} - p_m m_{i,j}$$

subject to the total demand function for modern good j in sector i, given by the sum of final and intermediate demand for j, 2 and 5:

$$p_{i,j} = D_{i,M}^{\frac{1}{\sigma}} d_{i,j}^{\frac{1}{\sigma}} q_{i,j}^{-\frac{1}{\sigma}} = \left(\frac{E_{i,M}}{P_{i,M}^{1-\sigma}} + \frac{p_m \Upsilon}{p_m^{1-\sigma}}\right)^{\frac{1}{\sigma}} d_{i,j}^{\frac{1}{\sigma}} q_{i,j}^{-\frac{1}{\sigma}} \tag{7}$$

where $D_{i,M}$ is an aggregate demand shifter for modern manufacturing in sector i, given by

$$D_{i,M} = \frac{E_{i,M}}{P_{i,M}^{1-\sigma}} + \frac{p_m \Upsilon_m}{p_m^{1-\sigma}} = \left(\frac{1}{P_{i,M}}\right)^{-\sigma} \left(\frac{P_{i,M}}{P_i}\right)^{-\rho} \left(\frac{P_i}{P}\right)^{-\epsilon} \frac{\phi_{i,M}\lambda_i E}{P} + \frac{p_m \Upsilon_m}{p_m^{1-\sigma}}$$

We focus our analysis on the case of monopolistic competition with constant markups.[3] The FOC for labor can be written as:

$$(1-\tau_j^R)\frac{p_{i,j}q_{i,j}}{l_{i,j}}\gamma_l \frac{\sigma-1}{\sigma} = w(1+\tau_j^l) \tag{8}$$

Analogous conditions apply in equilibrium for capital and materials. Plugging these FOCs into profits, we obtain that in equilibrium the firm's profits are proportional to (after tax) revenue:

$$\pi_{i,j} \quad = \quad (1-\tau_j^R)p_{i,j}q_{i,j}\left(1-\frac{1}{\mu}\right) \tag{9}$$

where $\frac{\sigma}{\sigma-1} = \mu$, the markup under monopolistic competition, given isoelastic demand.

The First Order Conditions also imply capital-to-labor and materials-to-labor ratios of $\frac{k}{l} = \frac{\gamma_k}{\gamma_l}\frac{w(1+\tau_j^l)}{r(1+\tau_j^k)}$ and $\frac{m}{l} = \frac{\gamma_m}{\gamma_l}\frac{w(1+\tau_j^l)}{p_m}$. Using these ratios into the FOC for labor 8 after replacing q_{ij} and p_{ij} with 3 and 7, we obtain optimal labor demand for firm j:

$$l_{i,j} = \frac{D_{i,M}}{\mu^\sigma}\left(\frac{(a_j * d_j^{\frac{1}{\sigma-1}})^{\sigma-1}(1-\tau_j^R)}{c^{\sigma-1}\hat{\tau}_j^{\sigma-1}}\right)\left(\frac{\gamma_l}{w(1+\tau_j^l)}\right) \tag{10}$$

where $(a_j * d_j^{\frac{1}{\sigma-1}})$ is quality-adjusted productivity of the firm owned by entrepreneur j, and we have defined as:

[3] We relax this assumption in the appendix and use this more general model in the model estimation.

$$
\begin{aligned}
c &= \left(\frac{w}{\gamma_l}\right)^{\gamma_l}\left(\frac{r}{\gamma_k}\right)^{\gamma_k}\left(\frac{p_m}{\gamma_m}\right)^{\gamma_m} \\
\hat{\tau}_j &= \frac{(1+\tau_j^l)^{\gamma_l}(1+\tau_j^k)^{\gamma_k}}{(1-\tau^R)}
\end{aligned}
$$

such that c is a modernity-wide pre-tax cost index and $\hat{\tau}_j$ is a firm-specific distortion ("tax") index.

Solving equation (8) for pre-tax revenue and plugging optimal labor demand 10, we obtain optimal levels of revenue and profits for firm j in sector i :

$$
p_{i,j}q_{i,j} = \left[D_{i,M}^{\frac{1}{\sigma-1}}\frac{a_j d_j^{\frac{1}{\sigma-1}}}{\mu c \hat{\tau}_j}\right]^{(\sigma-1)} \tag{11}
$$

$$
\pi_{i,j} = \left(1-\frac{1}{\mu}\right)(1-\tau_j^R)D_{i,M}\left(\frac{a_j d_j^{\frac{1}{\sigma-1}}}{\mu c \hat{\tau}_j}\right)^{(\sigma-1)} \tag{12}
$$

Besides aggregate conditions in the modern sector where the firm is ($D_{i,M}$), the markup and input prices, profits depend on the firm's quality-adjusted productivity $a_j d^{\frac{1}{\sigma-1}}$ (inherited from the firm's operator), and the set of idiosyncratic distortions to revenue, capital and labor faced by the operator.

$$
\pi_{i,j} = \left(1-\frac{1}{\mu}\right)\left(\underbrace{\frac{E_{i,M}}{P_{i,M}^{1-\sigma}}}_{D_{i,M}=D_{i',M}}\right)\left[\frac{\overbrace{a_j d_j^{\frac{1}{\sigma-1}}}^{\Delta=\text{qual.-adj productivity}}}{c\mu}\underbrace{\frac{(1-\tau_j^R)^{1+\frac{1}{\sigma-1}}}{(1+\tau_j^l)^{\gamma_l}(1+\tau_j^k)^{\gamma_k}}}_{\hat{\tau}=\text{idio.dist.}}\right]^{\sigma-1}
$$

Under monopolistic competition and isoelastic demand, the optimal price is a constant markup times the marginal (distortion-adjusted) cost of production:

$$
p_{ij} = \frac{\mu c \hat{\tau}_j}{a_j} \tag{13}
$$

2.5.2 Occupational Choice

The agent's problem is solving:

$$
\max\{a, wa, \pi(a,d,T)\}
$$

Therefore, i) working is (weakly) preferred to being self-employed if $w \geq 1$; ii) conditional on the

agent's d, modern entrepreneurship is preferred if $a > a^E(d) = max\{a_S^E(d), a_W^E(d)\}$ where:

$$a_S^E = \left[D_{i,M} d_{i,M} \left(1 - \frac{1}{\mu}\right) (1 - \tau_j^R) \left(\frac{1}{\mu c \hat{\tau}_j}\right)^{(\sigma-1)} \right]^{\frac{1}{2-\sigma}}$$

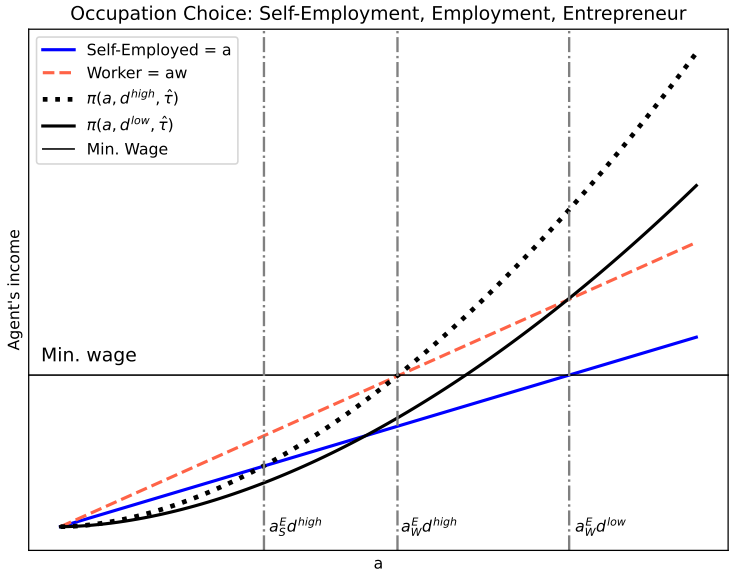
$$a_W^E = \left[\frac{D_{i,M} d_{i,M}}{w} \left(1 - \frac{1}{\mu}\right) (1 - \tau_j^R) \left(\frac{1}{\mu c \hat{\tau}_j}\right)^{(\sigma-1)} \right]^{\frac{1}{2-\sigma}}$$

As will become clear, in equilibrium $w \geq 1$, so $a^E = a_W^E$, and $D_{i,M} = D_M$ (i.e., an agent in any given occupation is indifferent between sectors).

Because profits also depend on the agent's entrepreneurial ability d, cutoff 14 implies a "contract curve" of sorts between a and d. When profits are convex in quality-adjusted productivity Δ (i.e., when $\sigma > 2$, which holds in our calibration), the a^E cutoff is strictly decreasing in d. That is, relatively low a agents could prefer becoming modern entrepreneurs conditional on sufficiently high d's.

Figure 1 depicts agents' incomes under the different occupations as a function of the agent's (a, d) for the case where $w > 1$, where the two black lines correspond to profits for different levels of d:

Figure 1: Income as a function of an agent's occupation, a and d holding s constant.



Unconstrained Occupational Choice: No Minimum Earnings. Agents with $a \geq a_W^E(d)$ and $a \geq a_S^E(d)$ become modern entrepreneurs or employers. Since Inada conditions in preferences guarantee a positive demand for both traditional and modern production, in equilibrium $w = 1$. As a result, $a^E(d) = a_W^E(d) = a_S^E(d)$, all non-employers (agents with $a < a^E(d)$) are indifferent between being self-employed or workers, and there is a positive supply of both activities. We denote the equilibrium wage in this unconstrained case

as $w^U(=1)$. In equilibrium, total labor, in this case, is simply determined by labor demand by modern entrepreneurs:

$$L^U = \sum_d \sum_{a>a^E(d)} l(a,d,w=1)$$

The supply of self-employment corresponds to all non-employers minus labor demand. We label it N_T^U

Constrained Occupational Choice Under the Minimum Earnings Regime. Consider now an environment in which regulation establishes a lower bound e_{min} for modern workers' earnings. Denote by L_T^C and N_T^C the numbers of salaried workers and self-employed workers in this constrained equilibrium. In particular, the constraint $wa \geq e_{min}$ is imposed on modern firms' payments to a worker characterized by talent a:[4] Workers with $a < \frac{e_{min}}{w}$ are thus not hired as employees and become self-employed.

That is, non-employers (workers with $a < a^E(d)$) sort between self-employment and salaried employment according to the following rule: Be self-employed if $a < \frac{e_{min}}{w}$; Work as a modern sector employee if $\frac{e_{min}}{w} \leq a$ and $w > 1$. Because in equilibrium $w \geq 1$, the only other possible case is that of unconstrained workers ($a \geq \frac{e_{min}}{w}$) facing a wage $w = 1$. Workers who are, in this case, indifferent between self-employment and labor.

Non-binding minimum wage: The equilibrium wage would still indeed be $w = 1$, $N_T^C = N_T^U$ and $L_T^C = L_T^U$, if the number of workers with $a < \frac{e_{min}}{w}$ falls below N_T^U. Although the fractions of agents in the different occupations are not affected by the minimum earnings requirement in this case, the regulation does imply a type of sorting that is absent in the unconstrained case. In the unconstrained equilibrium, workers who are indifferent between self-employment and labor are randomly assigned, while in the non-binding minimum wage case, all workers endowed with low abilities $a < \frac{e_{min}}{w}$ will be assigned to self-employment.

Binding minimum wage If, instead, the minimum wage is high enough that the number of workers with $a < \frac{e_{min}}{w}$ is larger than N_T^U, then the number of workers able to work in the modern sector will not be enough to cover the labor demand that would prevail at $w = 1$. Thus, in equilibrium, $w > 1$, and aggregate salaried employment will be equal to the mass of effective units not assigned to self-employment.

The model taken to the data for our quantification assigns each agent a random draw of s so that some high a agents may choose to become self-employed. This is done to capture the fact that, in the data, income segmentation between self-employed workers and others, or informal workers and others, is imperfect. In particular, although workers at the low end of the income distribution are mostly self-

[4]In the context of our application (Colombia at the peak of industrialization), indeed, there is a monthly (rather than hourly) minimum wage. Of course, the context is also characterized by high informality, with informal workers in general earning below the minimum monthly salary and self-employment capturing a significant fraction of informal employment. In our model, modern firms hire formally and are thus subject to the minimum earnings constraint, while the traditional sector is equated to self-employment and assumed informal, i.e., not subject to the minimum earnings limit.

employed and informal, informality and self-employment are also present (with much less intensity) in other parts of the income distribution.

2.5.3 General Equilibrium

Inada conditions impose positive consumption of all segments, as long as $1 > \phi_{i,M} > 0$ and $\lambda_i > 0$. Thus, an agent (a, d, Γ) who chooses to be in an occupation must be indifferent between this occupation in agriculture, industry or services. Such indifference renders a symmetry condition that is key in the solution of our model:

$$\alpha_{i,M} \equiv \frac{E_{i,M}}{E_M} = \frac{P_{i,M}^{1-\sigma}}{\sum_{i=A,S,I} P_{i,M}^{1-\sigma}} = \frac{L_{i,M}}{L_M} = \frac{N_{i,M}}{N_M} \tag{14}$$

where $L_{i,M}$ is the total modern demand for labor in sector i and $N_{i,M}$ is the mass of modern entrepreneurs in the same sector (in effective units of labor), and $\{E_M, L_M, N_M\}$ aggregate modernity over the three sectors.

To see that the symmetry condition 14 holds, notice first that because entrepreneurs are indifferent between the three sectors and we do not ad-hoc assign specific types of entrepreneurs to each sector, the distribution of entrepreneurs' (a, d, Γ) types is the same across sectors. Therefore, $p_{i,j} = \frac{\mu c \hat{\tau}_j}{a_j} = p_j$ and $P_{i,M}^{1-\sigma} = \sum_{N_{i,M}} d_j p_j^{1-\sigma} = \frac{N_{i,M}}{N_M} \sum_{i=A,S,I} \sum_{N_{i,M}} d_j p_j^{1-\sigma} = \frac{N_{i,M}}{N_M} \sum_{i=A,S,I} P_{i,M}^{1-\sigma}$. The fact that $\frac{L_{i,M}}{L_M} = \frac{N_{i,M}}{N_M} \left(= \frac{P_{i,M}^{1-\sigma}}{\sum_{i=A,S,I} P_{i,M}^{1-\sigma}} \right)$ follows from a similar logic that a firm's labor demand depends on the firm's type, which does not systematically vary over sectors. Finally, indifference also requires profits from modern entrepreneurship to equate across sectors for a given type of entrepreneur. Thus, $D_{i,M} = \frac{E_{i,M}}{P_{i,M}^{1-\sigma}} + \frac{p_m Y}{p_m^{1-\sigma}}$ must be independent of i (equation 12). From this, it follows that $\frac{P_{i,M}^{1-\sigma}}{\sum_{i=A,S,I} P_{i,M}^{1-\sigma}} = \frac{E_{i,M}}{E_{A,M}+E_{S,M}+E_{I,M}} = \alpha_{i,M}$.[5]

The symmetry property allows us to solve the model sequentially. Individuals are first allocated across occupations. This determines not only the total masses of salaried workers, modern entrepreneurs and traditional entrepreneurs, but also total income, and therefore final expenditure. These allocations are then assigned across sectors according to expenditure fractions that depend on relative demand.

Algorithm For given E, E_M and P_M, $\alpha_{i,M}$ depends only on preferences and can thus be solved for (at least numerically). This can be seen using $E_{i,M} = \phi_{i,M} \left(\frac{P_{i,M}}{P_i} \right)^{1-\rho} \lambda_i \left(\frac{P_i}{P} \right)^{1-\epsilon} E$ from 1 and replacing $P_{i,M} = \alpha_{i,M}^{\frac{1}{1-\sigma}} P_M$, $P_i = \left(\phi_{i,M} P_{i,M}^{1-\rho} + (1 - \phi_{i,M}) \right)^{\frac{1}{1-\rho}}$ and $P^{1-\epsilon} = \sum_i \lambda_i P_i^{1-\epsilon}$, such that

$$\alpha_{i,M} = \frac{E_{i,M}}{E_M} = \frac{\phi_{i,M} \lambda_i}{E_M / E} \alpha_{i,M}^{\frac{1-\rho}{1-\sigma}} \left(\frac{P_M}{P_i(P_M, \alpha_{i,M})} \right)^{1-\rho} \left(\frac{P_i(P_M, \alpha_{i,M})}{P(P_M, \alpha_{i,M})} \right)^{1-\epsilon}$$

[5] $\pi_{A,j} = \pi_{I,j} = \pi_{S,j}$ implies $D_{i,M} = \frac{E_{i,M}}{P_{i,M}^{1-\sigma}} + \frac{p_m Y}{p_m^{1-\sigma}}$ to be independent of i, so that $P_{I,M}^{1-\sigma} = E_{I,M} \frac{P_{A,M}^{1-\sigma}}{E_{A,M}}$ and $P_{S,M}^{1-\sigma} = E_{S.M} \frac{P_{A,M}^{1-\sigma}}{E_{A,M}}$. Denote $E_M = E_{A,M} + E_{I,M} + E_{S,M}$ and $\alpha_{i,M} = \frac{E_{i,M}}{E_M}$, notice $p_m^{1-\sigma} = P_{A,M}^{1-\sigma} + P_{I,M}^{1-\sigma} + P_{S,M}^{1-\sigma}$, and replace the $P_{i,M}^{1-\sigma}$, to obtain $\frac{p_m^{1-\sigma}}{P_{A,M}^{1-\sigma}} = \frac{E_{A,M}+E_{S,M}+E_{I,M}}{E_{A,M}} = \frac{1}{\alpha_{A,M}}$.

Moreover, aggregating over all modern firms' demand for materials $p_m Y_m = \frac{\gamma_m}{\mu} R^{net}$ where we have defined $R^{net} \equiv \sum_{j \in i} \left((1 - \tau_j^R) p_{i,j} q_{i,j} \right) = \sum_{j \in i} \left(p_{i,j} q_{i,j} \right) - G_R = E_M + p_m Y - G_R$ and G_R is total τ^R collections by the government. Then

$$R^{net} \left(1 - \frac{\gamma_m}{\mu} \right) = E_M - G_R \tag{15}$$

$$\sum_{j \in M} p_{i,j} q_{i,j} = \frac{E_M}{1 - \frac{\gamma_m}{\mu}} - \frac{\frac{\gamma_m}{\mu} G_R}{1 - \frac{\gamma_m}{\mu}}$$

3 Micro estimation

Detailed firm- and individual-level microdata is necessary for our quantification for at least three reasons: Our model relies on heterogeneous productivity across agents and across the firms they run. We need to estimate the distributions of talents $f_a(a)$ and $f_d(d)$. We do so using individual-level income information representative of the whole workforce, and firm-level information for modern firms. Our model also depends on firm-specific distortions in the modern segment, $\tau_j^l, \tau_j^k, \tau_j^R$. We estimate the distributions of such distortions and their potential dependence from $f_a(a)$ and $f_d(d)$ using firm-level data. Finally, to quantify the model, we require technology parameters $\gamma_l, \gamma_k, \gamma_m$ and the elasticity of substitution across firms.

We take the firm production and demand functions to the data on modern firms to estimate these parameters. Re-estimating these functions for later periods in time also allows us to gain insight into what are plausible magnitudes for the relative changes in γ_l as technologies become more labor-saving. We later use these insights in counterfactual exercises designed to gauge the importance of labor-saving technological change in explaining jobless modernity.

3.1 Historical Micro Data

We leverage long-time series of individual and firm microdata from Colombia. The Colombian data, generally unavailable for other countries, is unparalleled in scope and detail in most of developing countries and provides microdata level information for over 50 years. In particular, we have recovered and use microdata from two official historical sources produced by the Colombian Statistics Bureau, DANE: The Annual Manufacturing Survey available and the Household Survey.

The Household Surveys are the source of official statistics on employment and the distribution of income and are thus representative of the whole economy, formal and informal, in all sectors. For all respondents that take part in production, the survey records whether they are salaried workers, own-account workers without employees or employers, and the size of the business where their activity takes place in broad categories. This provides a comprehensive picture of income and employment in the whole economy, their allocation across sectors and different types of firms, and the distribution of earnings in the economy and within each of these segments. Although there have been methodological changes over

time, the Household Surveys go back in time several decades. Unlike other waves for years close to the peak of industrialization, the digital records of the 1971 wave contain the individual responses on the size of the business to which the worker is affiliated, besides income and occupational category. For this reason, we use the household survey information from 1971.

The Annual Manufacturing Survey is a census of all manufacturing establishments with at least 10 employees. We use data for 1974-1989.[6] It records detailed information at the establishment level on the value of output and input use (capital, labor, intermediates consumed), including quantities and values for all individual products and all individual material inputs produced by the establishments plus payroll. A firm in the model corresponds to an establishment (plant) in the data. The within-plant product level information is used to estimate establishment-level prices that are comparable across establishments of the manufacturing sector, even for multiproduct plants. Plant prices are then used to measure plant output for uni- and multi-product plants with differentiated products and to subsequently estimate the plant production and demand functions 3 and 7 with the methodology proposed by (Eslava et al. (2023)).

In the data, we define the modern segment to correspond to establishments with 10 or more employees. The traditional segment is the complementary set of all establishments with less than 10 employees, including own-account individuals. Since the household surveys classify workers by sector and firm size category and into self-employed/salaried workers/employers, we can construct the share of employment and income (i.e., value-added) represented by modernity in each sector. We also have the detailed distribution of income overall, by sector, and by segment (modern/traditional) within each sector. A common concern when aggregating income data from household surveys is that individuals at the very top of the income distribution are likely under-represented, implying that value-added may be underestimated. We address this concern by scaling income in each sector to reflect the aggregate reported in GGDC data.

The Annual Manufacturing Survey then provides us with a detailed picture of the distribution of productivity and distortions across firms of the modern segment and, consequently, for modern entrepreneurs. Modern manufacturing is the main modern activity during the episode we study. For this reason, the lack of equally detailed data on modern services or modern agriculture, and the difficulties inherent to measuring firm productivity and distortions in services and agriculture in a comparable way to manufacturing, we assume that the distributions we observe in manufacturing also represent the distributions in the modern segments of services and agriculture.

The historical microdata for Colombia is crucial in our quantification. In light of the model, we use the income distributions from the household surveys to estimate the ex-post distribution of a for salaried and own-account workers. We use the Annual Manufacturing survey to estimate the parameters of technology

[6]The data has been collected annually since 1974, but there is a gap in the electronic records in 1977-1982: for this set of years, the information on individual products and material inputs is missing. Thus, although the peak occurred at the end of the 1970s, we use the available data for the full 1974-1989 period to represent the peak of industrialization. We do not use solely 1974-1976 because we have less assurance of the quality of the data for the earliest years, especially at the detailed product-input level. Since the methods we use to separately measure quantities and prices rely on chained firm-level prices, we take advantage of the robust data for the 1980s to improve the quality of our price measurement in the early years.

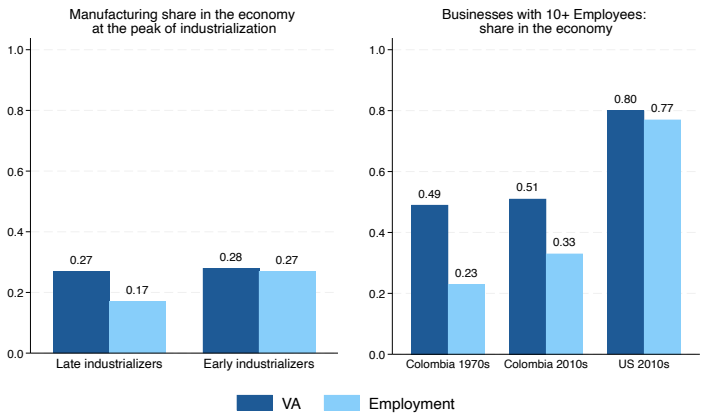
and demand $\{\gamma_l, \gamma_k, \gamma_m, \sigma\}$, and subsequently obtain measures of $a, d, \tau_l, \tau_k, \tau_r$ for each manufacturing establishment. These correspond to the ex-post distribution of each of these dimensions of talent among those who self-select as entrepreneurs. We now explain the estimation in detail.

3.2 Stylized Facts

The inclusion of multiple mechanisms in our model not only seeks to contrast possible explanations for joblessness but is also necessary to replicate important features of the data that are frequently neglected. We seek to capture:

Thin and Jobless Modernity Figure 2 shows that today the modern segment (all businesses with 10+ employees) carries low weight in the economy in Colombia compared to the US, using 2019 data from Colombia's household data and the US Current Population Survey. During our period of study, the weight was even lower, by about 10 percentage points. Moreover, modernity has been jobless since the 1970s in Colombia. While in 1971 the modern share of value added was 40%, the analogous share in employment was only 33%. No similar joblessness is observed in the US today, where the weight of the modern segment is similar in employment and value-added. [7] Different features in our model are key to reproducing jobless modernity: labor distortions biased against the growth of the largest businesses, a talent distribution with a very fat bottom tail.

Figure 2: **Jobless modernity**



Left panel: Data from GGDC 10-Sector Database. *Early Starters*: Great Britain, United States, Denmark, France, The Netherlands, and Sweden. *Late Starters*: Argentina, Brazil, Chile, Colombia, Costa Rica, Egypt, Spain, Indonesia, Italy, Japan, South Korea, Mexico, Mauritius, Malaysia, Nigeria, Philippines, Thailand, Taiwan, and South Africa. Sample considers only countries that reach a manufacturing value-added share over 20% of GDP at some point in their history. Right panel: Data from 1971 and 2019 official household survey for Colombia, and from 2019 Current Population Survey for USA, for each of Agriculture, Industry and Services. Weighted averages across sectors using the weight of the sector in a respective variable from GGDC 10-Sector Database.

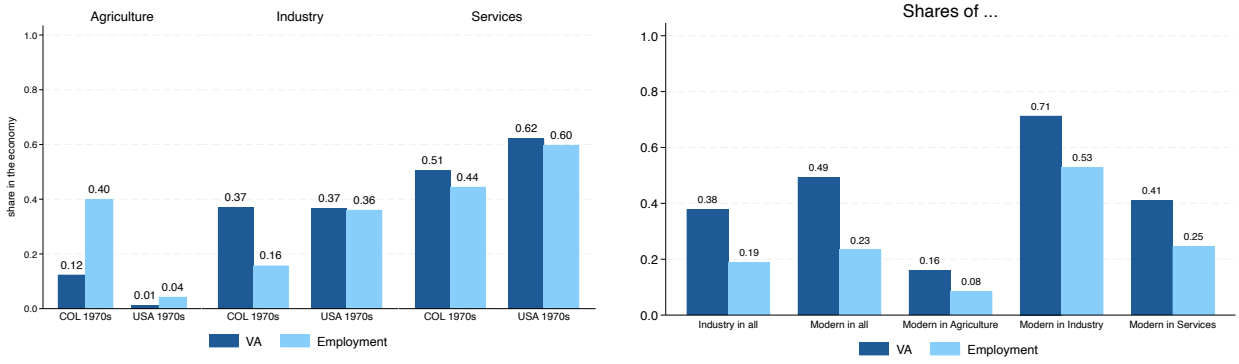
[7] We don't find evidence either that modernity was much more jobless in the 1970s than today in the US. The 1979 US CPS reports firm size in different categories than those used in 2019, but the comparable measurement is possible at the 25-employee threshold. Businesses with 25+ workers represented similar shares of employment and value-added in the US in the 1970s, while in Colombia, their weight in value-added more than doubled that in employment.

Jobless Industrialization One face of jobless modernity in emerging economies is the joblessness of sectors where modernity carries a larger weight. During the process of industrialization, this phenomenon manifested in jobless industrialization–a much larger weight of industry in an economy's value-added than its employment–in the countries that industrialized late but in contrast to those that industrialized early with job-full industrial industrialization. Figure 2 shows this fact for the peak of industrialization.

Colombia is a typical case of jobless industrialization: at the peak of the industrialization process, the industry's weight in the economy's aggregate value added more than doubled that in aggregate employment (Figure 3a). Meanwhile, agriculture was job-full, contributing much more to employment than value-added, while employment and value-added shares were much closer in services. There is no evidence of marked divorce between employment and value-added shares in different sectors in the US over time.

The differential preference for modernity across sectors in the model helps reproduce jobless industrialization, together with differential preferences for consuming agricultural vs. manufactured goods vs. services. The (calibrated) weights of modern/traditional within sector-level preferences and of different sectors in aggregate preferences interact with distortions measured in the data to generate differential joblessness across sectors. The weight of modernity indeed differs across sectors in the 1970s Colombian data: modern value added has a much larger weight in industry and services than in agriculture, and also a larger weight in the industry than services (see Figure 3b).

Figure 3: **Joblessness and Modernity Shares**



(a) Jobless vs. job-full sectors

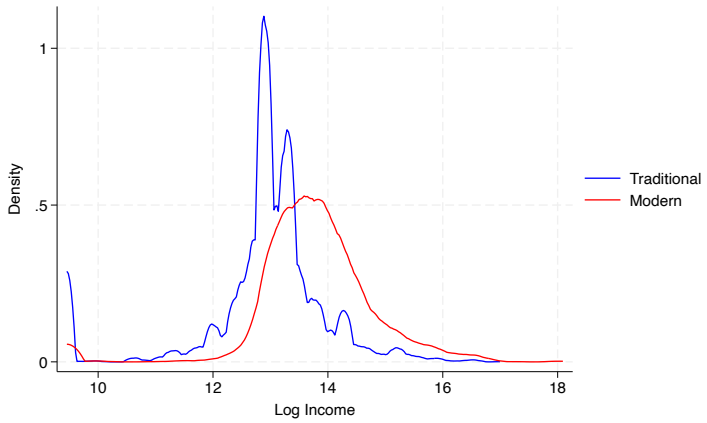
(b) Modernity share across all sectors

Left panel: Data from the GGDC 10-Sector Database. Sector shares are for the year 1975. Right panel: Industry shares are taken from the GGDC 10-Sector Database for 1971. Modern economy shares are based on 1971 data combined with Colombia's official national household survey. Economy-wide modern shares are calculated as weighted averages across sectors, using sector weights from the relevant GGDC variable. Within-sector modern shares are estimated from household survey data.

Modern Modernity Figure 4 displays the distribution of Colombian workers' earnings in the segments that we have defined as modern and traditional. It indicates that businesses with 10+ employees indeed pay much higher earnings per worker than those in the segment that we label as traditional. The modern

earnings distribution is significantly shifted to the right compared with that in traditional businesses. At the same time, the segmentation is not perfect: the ranges of the two distributions overlap, and a non-negligible fraction of workers in traditional businesses make relatively large earnings.

Figure 4: **Earnings distribution: modern vs. traditional**



In the model, the presence of a regulatory minimum earnings constraint selects low-ability workers into self-employment. In contrast, the convexity of profits on entrepreneurial talent implied by modern product diversification selects individuals with high productivity and the ability for product differentiation into modern entrepreneurship individuals. These features explain higher earnings in modernity relative to traditional businesses. An additional self-employment shock, s, is crucial to capture the concurrent presence of high-earnings individuals in the traditional segment. At the time, the minimum wage was sector- and region-specific. Our calculations indicate that about 17% of workers made earnings below the minimum wage level of their respective sector-region pair. This relatively low level of the minimum wage at the time (compared to one close to the 50th income percentile today) is taken into account in our calibration.

3.3 Firm level estimation

The manufacturing data directly report, for each unit, revenue (pq), total salary payments (wl), capital stocks (k), value of material inputs used ($p_m m$). We build, based on our theory for product demand, plant level deflators for output, p, and materials, p_m, which we use to deflate revenue and value of materials to obtain y and m. Effective labor l is measured from the idiosyncratic variation in total salary payments, given an equilibrium wage w homogeneous across production units. In the data, the empirical analog of our theoretical modern firms is the manufacturing establishment, but we will continue to use the language "firm" for consistency.[8] The data allows us to accurately measure, for the estimation period, technological parameters that are crucial for simulations in the model. We can also re-estimate them for later periods,

[8]Over 90% of firms are uni-establishment in the data.

Table 1: **Factor and demand elasticities**

Year	γ_l	γ_k	γ_m	σ
1974 - 1989	0.374	0.105	0.542	2.391
2000 - 2012	0.227	0.113	0.674	2.793

shedding light on reasonable patterns of evolution over relatively long periods that we later use to assess the role of factor-biased technology changes.

We let both technology and demand parameters vary across 29 manufacturing sectors and for the period of study and the first decade of the XXI century. Details of that estimation are presented in the appendix. Table 1 presents results for the average manufacturing sector. There is a noticeable decrease in the elasticity of labor in production from the peak of industrialization to the post-2000 decade, from 0.374 to 0.227, while the coefficient for capital has remained fairly stable, and that for materials has increased by almost 0.12 points. Interestingly, σ is lower than usually estimated for advanced economies and increases over time from 2.4 to 2.8.

Having estimated the parameters of the firm problem, we obtain a for each firm as the idiosyncratic residual of the production function 3 and d as the idiosyncratic residual of demand function 7. We work only with the idiosyncratic part of firm-level a and d by taking the log difference vs. the mean of logs in the respective sector and year.

We subsequently estimate idiosyncratic firm-level distortions. We directly measure $\tau^l = \frac{\text{payroll tax}}{wl}$ for each establishment, taking advantage of reports in the data about the payroll taxes paid by each individual establishment. Using only the idiosyncratic component of τ_l and of k and l in the capital-labor ratio, we obtain (the idiosyncratic component of) $1 + \tau^k = \frac{l}{k}(1 + \tau^l)$. We then solve for $(1 - \tau^R)$ re-writing equation 11 as:

$$(1 - \tau^R) = \frac{\mu(1 + \tau^l)^{\gamma_l}(1 + \tau^k)^{\gamma_k}(pq)^{\frac{1}{(\sigma-1)}}}{D_{i,M} a d^{\frac{1}{\sigma-1}}} \tag{16}$$

Since all firm-level variables and parameters in the right hand side of this equation have been measured at this point, we can estimate the idiosyncratic variation in $(1 - \tau^R)$ up to $D_{i,M}$.

Table 2 presents correlation coefficients across these moments. As in the original work of Eslava et al. (2023) for 1982-2012, revenue positively correlates with quality-adjusted productivity $ad^{\frac{1}{\sigma-1}}$ and each component. At the same time, quality-adjusted productivity is strongly negatively correlated with revenue distortions and with a function of the payroll tax. In other words, the most highly productive firms in the survey have substantially smaller revenue than their productivity would predict, partly because taxes on labor and revenue distortions affect them more harshly. Interestingly, the negative correlation between productivity and τ^l disappears in the 2000s.

Table 2: **Correlation matrix for establishment attributes**

	Δ	a	d	pq	τ^R	τ^k	τ^l	$\frac{1-\tau^R}{(1+\tau^k)^{\gamma_k}(1+\tau^l)^{\gamma_l}}$
	1974-1989							
Δ	1,00							
a	0,17	1,00						
d	0,96	-0,09	1,00					
R	0,98	0,09	0,97	1,00				
τ^R	0,22	0,36	0,12	0,09	1,00			
τ^k	-0,01	0,08	-0,04	-0,03	-0,02	1,00		
τ^l	0,48	0.08	0,47	0,49	-0,07	0,00	1,00	
$\frac{1-\tau^R}{(1+\tau^k)^{\gamma_k}(1+\tau^l)^{\gamma_l}}$	-0,26	-0,48	-0,14	-0,10	-0,84	-0,11	-0,04	1,00
	2010-2019							
Δ	1,00							
a	0,25	1,00						
d	0,80	-0,24	1,00					
R	0,84	0,00	0,90	1,00				
τ^R	0,28	0,42	0,07	0,02	1,00			
τ^k	-0,04	0,05	-0,08	-0,08	-0,01	1,00		
τ^l	0,01	0.01	0,01	0,01	0,00	-0,01	1,00	
$\frac{1-\tau^R}{(1+\tau^k)^{\gamma_k}(1+\tau^l)^{\gamma_l}}$	-0,32	-0,55	-0,04	0,02	-0,79	-0,09	-0,01	1,00

3.4 Estimation of underlying type distributions f_a and f_d

To estimate the distributions f_a and f_d used for calibrating the model, we use the estimates of a and d for manufacturing model firms, the overall income distribution reported in the household survey, and insights from the model. The distribution of income for workers and the self-employed directly gives us the distribution of a for this set of individuals. For modern entrepreneurs, we have the idiosyncratic distribution of a.

These two distributions are then put together to obtain overall f_a. This poses two challenges: 1) the techniques employed in our firm estimation only allow us to estimate idiosyncratic attributes of firms so that by construction, our firm estimates of a are centered around $a = 1$, while the income reports in the household survey are in monetary units; 2) there is apparent measurement error in the left tail of income reports in the household surveys, with a mass of zeroes for a fraction of, mainly, the self-employed.

We deal with these problems by taking advantage of the optimal occupational choice, which implies that the range of a that the ex-post selects into entrepreneurship is the same range of the ex-ante distribution f_a. We thus take a_{max} to correspond to the maximum income of workers in the Household Survey and obtain a_{min} by requiring the range to replicate the range of a in the manufacturing firm data.[9] Our calibration uses the non-parametric distribution of $f(a)$ thus obtained.

The optimal occupational choice also implies that modern entrepreneurs (firms) with large enough a should display a distribution of d similar to that in the population. With this in mind, we estimate parameters for a Pareto distribution for d using the d distribution observed among the top decile of the a distribution among manufacturing firms.

[9] We estimate the range and distribution in log terms.

4 Quantification: Calibration Strategy and Results

Having estimated key parameters and moments in the microdata of modern manufacturing, we proceed to calibrate the model by leveraging our estimates to account for the parameterization that defines the equilibrium prices and allocations. The model calibration seeks to account quantitatively for the employment and value-added shares of modern production within manufacturing and to shed light on the several mechanisms discussed potentially driving the "divorce" between employment and value-added shares within a general equilibrium framework. The calibration is based on the occupational choice outlined in the previous section. We simulate an economy populated with agents endowed by a triplet a, d, and s. We currently feed the calibration with data moments for the peak of industrialization in the 1970s and target moments for the same period; in particular, the information from the household survey is for 1971 and that from the Annual Manufacturing Survey for 1974-1982. We also use the production function estimates for the period 1974 - 1982.

4.1 Calibration Strategy

Production function parameters The modern sector's production function parameters are estimated directly from the historical firm-level data. In particular, we use the estimates from 1974 through 1982to parameterize the production function parameters of the modern gross output production function. We note that the estimation yields a production function for modern sectors that is constant returns to scale up to the first decimal position, that is, $\gamma_k + \gamma_l + \gamma_m = 1.021$, even though our estimation does not constrain the sum to be one. In the calibration we adjust the materials elasticity to ensure exact constant returns to scale, which we assumed in solving the model. Estimating the firm-level data also gives us a distribution of firm-level TFP (a) and product appeal d. We fit this joint distribution non-parametrically (through splines) and feed it to the population (see the discussion below).

Demand Parameters Our estimation method also yields an estimate of the elasticity of substitution between modern varieties, $\sigma = 2.9$. We use the estimates for the 1974-1982 period. The other elasticity of substitution parameters of the demand functions are externally calibrated. Following the literature in structural change, we calibrate the elasticity of substitution between broad sectors to be complements, $\epsilon = 0.2$, along the lines of Herrendorf et al. (2014). There is little guidance in the literature regarding the elasticity of substitution between modern and traditional manufacturing. Feng et al. (2024) calibrate an elasticity of 3.9, which is larger than the elasticity of substitution between modern varieties that we estimate. Since we deem modern varieties to be at least equally substitutable with each other as with traditional varieties, we require $\rho \leq \sigma$. Given our estimate of $\sigma = 2.9$, we set a nudge below this value, $\rho = 2.5$. The remaining taste parameters of the utility functions are calibrated to match the sectoral shares of value added and the modernity shares within each sector, as described in Table 4.

Table 3: **Summary of Model Parameters Estimated in the Microdata or Exogenously set**

Parameter	Value	Target or Empirical Counterpart
	Parameters Estimated from Firm-level Data	
γ_k	0.105	Capital intensity
γ_l	0.374	Labor intensity
γ_m	0.521	Materials intensity
σ	2.391	Elasticity Substitution modern varieties
Average τ_l	0.47	Avg. $\frac{\text{Contributions on workers}}{\text{Wage Payments}}$
	Distributions Estimated from Firm-level Data	
Density (a, d)	$[\cdot]_{122\times30}$	Talent a and product appeal d joint distribution
$\tau^R(a, d)$	$[\cdot]_{122\times30}$	Revenue distortions matrix, Non-parametric fit
$\tau^l(a, d)$	$[\cdot]_{122\times30}$	Labor distortion matrix, Non-parametric fit
$\tau^k(a, d)$	$[\cdot]_{122\times30}$	Capital distortion matrix, Non-parametric fit
	Parameters Set Exogenously from the Model	
ρ	2.5	Elasticity substitution Modern-Traditional
ϵ	0.15	Elasticity substitution broad sectors (Herrendorf et al., 2014)
Average τ^R	0.20	Corporate income tax rate
Average τ^k	0	Dividends tax rate

Distortions As explained above, our firm-level estimation delivers a firm-specific estimate of idiosyncratic revenue, capital, and labor distortions (see section 3.3). Given a type (a, d) we compute the average distortion for these types $\tau^R(a, d)$, $\tau^k(a, d)$, and $\tau^l(a, d)$ and fit each distortion non-parametrically through splines. This gives us the profile of distortions faced by a type (a, d) in the model. Because, at this point, we have only considered idiosyncratic variation across firms, we scale each of the distortions (τ^R, τ^l, τ^k) by an intercept. For τ_l, directly measured in the data, the intercept is simply the average level of contributions paid per unit of wage, which stands at $\tau_l = 0.47$. For τ^R and τ_k, we fix intercepts that approximate tax rates of corporate income and dividends, respectively. We set $\tau^R = 0.2$ and $\tau^k = 0$, taken from historical accounts on Colombian tax reforms (Cardenas and Perry, 1986).

Lastly, we calibrate the minimum earnings requirement e_{min} and the distribution of s to match the modern sectors' employment share and the likelihood of agents above and below median income to be self-employed.[10] Tables 3 and 4 summarize our calibration strategy.

Model Solution Given the draws in (a, d, s), our calibration strategy computes the occupational choice for each agent conditional on its draw and the distortions should the agent choose to become an entrepreneur. We conjecture values for the aggregate demand shifter D_{im}, the price of the intermediates, and the wage in modern manufacturing labor market w and compute a *tâtonnement* process until we find market clearance.

[10]Specifically, we assume that s follows a truncated Pareto, with a minimum value equal to 1. We then calibrate the tail parameter and the upper bound of the distribution.

Table 4: **Calibrated Model Parameters**

Parameter	Value	Model	Target	Description
λ_I	0.44	38%	39%	Industry Exp. Sh.
λ_S	0.48	50%	50%	Serv. Exp. Sh.
$\phi_{A,M}$	0.31	4%	4%	Sh. Agr. in Mod.
$\phi_{I,M}$	0.60	52%	55%	Sh. Ind. in Mod.
$\phi_{S,M}$	0.35	44%	41%	Sh. Serv. in Mod
e_{min}/income p_{50}	0.79	50%	51%	Value added in Self Emp.
Pareto tail s	1.05	17%	17%	Sh. earning below min. wage
Upper bound s	1.75	79%	72%	Sh. self empl. above min. wage

4.2 Calibration Results and Counterfactual Exercises

Table 4 presents calibrated model parameters. The model matches close to precisely the shares in value added of modernity in the economy and in all sectors, and of each sector in the economy, as well as the share of people who earn below the minimum wage. It is also close to reproducing the fraction of workers who earn more than the minimum wage among the self-employed, with a larger margin of error than other targeted moments.

Other baseline results are reported in column (2) of Tables 5 and 6, including untargeted moments, together with the respective moments from the data (column 1). The baseline calibration reproduces the fact that the self-employed earn, on average, less than modern workers and modern entrepreneurs, with this last group displaying the highest income. The model also shows the importance of manufacturing in the patterns of the time: it is the sector that loads more heavily in modernity. It is, therefore, the sector most subject to distortions and plays a central role in creating jobless modernity.

Understanding the Role of Distortions We conduct different sets of counterfactual exercises. A first group, presented in columns $3, 5, 6, 7$ of tables 5 and 6, examines the effect of reducing by 25% respectively the tax rates on modern firms' use of labor, revenue, or both, and the minimum wage. These institutions distort modernity relative to the traditional segment bu making the former more costly. We discuss the two tables jointly; Table 5 records results for the allocations while Table 6 presents the distributional consequences of our exercises.

We compare all these experiments to the baseline calibration. Each of these policies expands the size of the modern sector relative to that benchmark and does so more in terms of employment than value-added, reducing the joblessness of modernity. Aggregate income increases non-negligibly by close to 10 p.p. in most cases and by more than 15 p.p. in the experiments that reduce revenue taxes. Interestingly, reducing the dispersion of τ^L by 25% results in an allocation and aggregate income that are close to reducing its intercept in the same proportion. That is, lowering distortions within modernity by 25% has an effect similar to reducing distortions between modernity and informality by 25%. As reported in table 2, the

Table 5: **Baseline calibration and regulatory experiments: allocations**

	Data	Baseline	↓ 25% avg. τ^L	↓ 25% disp. τ^L	↓ 25% avg. τ^R	↓ 25% avg. all τ	↓ 25% min earn	↑ 25% min earn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Modern VA	49.2	50.0	54.0	53.8	54.2	56.4	53.7	50.6
Modern Emp	23.3	17.1	19.5	19.5	20.0	22.1	23.5	14.5
Modern VA / Employment	2.1	2.9	2.8	2.8	2.7	2.6	2.3	3.5
U/Ubaseline		1.000	1.094	1.086	1.16	1.19	1.09	1.02
Entrepreneurs	0.4	5.7	6.4	6.4	7.0	7.3	7.4	5.4
Salaried	23.0	11.4	13.2	13.2	13.0	14.7	16.1	9.1
Self employed	76.6	82.9	80.5	80.5	80.0	77.9	76.5	85.5
Modern share within sector								
Agriculture	15.6	16.0	17.9	17.9	20.4	21.1	17.9	15.9
Industry	71.4	71.9	73.4	73.3	75.0	75.4	73.3	71.9
Services	40.8	45.7	48.0	47.9	50.6	51.3	47.9	45.6
Total Exp. Shares								
Agriculture	12.6	12.7	13.0	13.0	13.4	13.4	13.0	12.7
Industry	37.9	37.5	37.1	37.1	36.8	36.7	37.1	37.5
Services	49.5	49.8	49.9	49.9	49.9	49.9	49.9	49.8

Table 6: **Baseline calibration and regulatory experiments: distributional consequences**

	Data	Baseline	↓ 25% avg. τ^L	↓ 25% disp. τ^L	↓ 25% avg. τ^R	↓ 25% avg. all τ	↓ 25% min earn	↑ 25% min earn
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
U/Ubaseline		1.00	1.09	1.09	1.16	1.19	1.09	1.02
Avg. real profits entrep	8.29	10.81	11.44	11.32	12.29	12.42	9.74	11.84
Avg. real inc. formal work	2.14	1.41	1.52	1.52	1.60	1.59	1.19	1.59
Avg. real inc. self-emp	1.00	1.00	1.04	1.03	1.10	1.14	1.09	0.98
Avg. profits entrep / Avg. inc. self-emp	8.29	10.81	11.05	10.97	11.19	10.90	8.92	12.12
Avg. inc. formal work / Avg. inc. self-emp	2.14	1.41	1.47	1.47	1.45	1.40	1.08	1.63
Avg. profits entrep / Avg. inc. formal work	3.87	7.66	7.50	7.46	7.70	7.79	8.22	7.44
Share people below min earn	17.00	16.78	16.78	16.78	16.78	16.78	8.51	26.70
Share of self-emp above min earn	72.00	79.40	76.50	76.50	76.00	73.50	74.30	80.20
Share of self-emp above median inc.	60.00	70.20	66.10	66.10	65.40	61.40	67.60	70.80
Earnings p50-p10	2.70	2.36	2.36	2.36	2.36	2.37	2.42	2.33
Earnings p95-p10	14.60	14.05	14.05	14.05	14.42	14.84	14.24	13.85

most significant τ^l fall on the highest productivity firms. Remember also that all government revenue collections are rebated.

The creation of modern employment in this group of policies is highest with the reduction of the minimum wage, as is the closure of the gap between modern value added and employment. Modern entrepreneurship also experiences a large increase in this experiment, although not as much as modern salaried employment. Reducing the minimum wage is also the only among these policy experiments that closes the income gap between the self-employed and employers (the lowest- and highest-earning occupations in this economy), as well as between the self-employed and formal workers (the middle "class"). This last group expands but also sees the gap with their employers' earnings widen.

Interestingly, *increasing* the minimum wage by 25% has effects that are not the mere opposite of re-

ducing it. Increasing the minimum wage does contract modern employment and increases joblessness in modernity. But it does so without reducing aggregate income. Although the group of formal salaried workers experiences a large contraction, its income relative to other occupations rises, making the occupation attractive for better-qualified workers. Modern average productivity thus increases. As expected, an increase in the minimum wage makes workers better off than their employers, but it benefits only salaried workers in the modern segment. Informal workers with low wages become more numerous and lose income relative to formal workers and their employers.

The increase in aggregate income is largest in the experiments that reduce taxes on modern firms. The gap between income in modernity and self-employment also grows in these experiments. Reducing taxes on modern firms' revenue, moreover, also widens the income gap between modern employers and their employees. Modern entrepreneurship is the occupation that expands the most in size in all tax-reducing experiments.

The Role of Human Talent We also conduct policy experiments that affect the distributions of human talents (a, d). These distributions underlie the businesses' productivity distribution in our model. They are also crucial for the effects of any given level of the minimum wage, which is more or less binding depending on the mass of individuals for whom $a < e_{min}$. In the model, a corresponds to embodied human capital: a component of productivity that the individual can use to run a modern firm producing a output units per unit of input, to provide a effective units of labor in a modern firm, or to produce a units of output in the traditional segment. Meanwhile, d is a talent specific to the production of differentiated goods that can only be used if the individual runs a modern firm.

Although, in applied terms, policies that affect a most evidently education, it is also the case that a clearly captures abilities that go beyond what can be delivered by increasing years of education. a covers transversal talents that also depend on the quality of education at all levels and the cultivation of soft and basic reasoning skills. Other factors affect the general productivity of the workforce, from health to transportation. Policies that affect d are probably closest to interventions to improve managerial capabilities and talent for innovation.

Column (2) of Table 7 and 8 reports the results from a counterfactual exercise that shifts the distribution of a to the right by 25%. The effects of this exercise are massive on aggregate income and modern employment. Real income more than doubled, closing almost half the income gap with the US at the time. This is an impressive multiplication of the direct effect of this 25% increase in hicks-neutral productivity. The magnification is due to general equilibrium effects, including amplification through the IO matrix and the fact that any given level of the minimum wage becomes a less binding distortion.

Moreover, the share of employment that is modern grows by over 13 p.p., almost doubling. The income gaps between the self-employed and all others shrink, as does the gap between formal workers and their employers. Although many of the distributional and allocative consequences are similar to the experiment

Table 7: **Human talent and technology experiments: allocations**

	Baseline	↑ 25% a	↑ 25% d	↑ 25% a < p50	↑ 25% a > p50	↑ 100% a	↑ 25% a ↑ e_{min} ≈e_{p50}	Modern γ's
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Modern VA	50.0	64.3	58.3	60.1	55.2	79.4	60.0	78.3
Modern Emp	17.1	30.6	22.0	28.1	19.1	52.4	20.0	38.2
Modern VA / Employment	2.9	2.1	2.7	2.1	2.9	1.5	3.0	2.0
U/Ubaseline	1.00	2.16	1.37	1.68	1.29	8.96	1.86	4.01
Entrepreneurs	5.7	10.1	7.2	10.1	5.6	16.0	6.8	15.9
Salaried	11.4	20.5	14.7	18.0	13.6	36.3	13.2	22.3
Self employed	82.9	69.4	78.0	71.9	80.9	47.6	80.0	61.8
Modern share within sector								
Agriculture	16.0	30.7	24.5	24.5	21.3	56.9	26.8	53.7
Industry	71.9	80.2	77.3	77.3	75.5	89.3	78.4	88.4
Services	45.7	59.8	54.6	54.6	51.5	77.4	56.6	75.5
Total Exp. Shares								
Agriculture	12.7	14.4	13.8	13.8	13.5	16.2	14.1	16.0
Industry	37.5	35.6	36.2	36.2	36.6	33.9	36.0	34.1
Services	49.8	49.9	49.9	49.9	49.9	49.9	49.9	49.9

Table 8: **Human talent and technology experiments: distributional consequences**

	Baseline	↑ 25% a	↑ 25% d	↑ 25% a < p50	↑ 25% a > p50	↑ 100% a	↑ 25% a ↑ e_{min} ≈e_{p50}	Modern γ's
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
U/Ubaseline	1.00	2.16	1.37	1.68	1.29	8.96	1.86	4.01
Avg. real profits entrep	10.81	17.00	13.67	12.38	15.76	54.78	20.48	28.53
Avg. real inc. formal work	1.41	2.03	1.80	1.71	1.72	6.07	2.69	3.31
Avg. real inc. self-emp	1.00	1.84	1.21	1.55	1.18	6.41	1.54	2.33
Avg. profits entrep / Avg. inc. self-emp	10.81	9.23	11.28	7.99	13.36	8.55	13.28	12.23
Avg. inc. formal work / Avg. inc. self-emp	1.41	1.10	1.48	1.11	1.46	0.95	1.74	1.42
Avg. profits entrep / Avg. inc. formal work	7.66	8.36	7.61	7.22	9.16	9.03	7.62	8.62
Share people below min earn	16.78	9.70	16.78	9.70	16.78	4.21	30.55	16.78
Share of self-emp above min earn	79.40	66.10	73.60	68.90	77.00	45.30	71.20	54.10
Share of self-emp above median inc.	70.20	57.90	62.20	60.80	66.90	40.50	59.80	41.60
Earnings p50-p10	2.36	2.46	2.36	2.46	2.36	2.38	2.33	2.46
Earnings p95-p10	14.05	15.81	14.68	15.99	14.05	15.76	15.30	19.05

where the minimum wage is reduced, there are two very significant differences: there is a much larger increase in aggregate income, and functional inequality is reduced in favor not only of the self-employed relative to those in modernity but also in favor of workers vis-a-vis employers.

Column 7 shows that because regulations have a more modest effect on aggregate income, the increase in a leads to a large increase in real income due to expended value in modernity, even if accompanied by a rise in the (real) level of the minimum wage. However, the positive distributional consequences suffer an important reversion. The experiment in column 7 simultaneously shifts the distribution of a to the right by 25% while also increasing the level of the minimum wage to the 50th percentile of the scenario reported in column 2. We attempt to replicate in this way the increase of the real minimum wage observed in Colombia, where today it stands close to the median income. Column 7 shows that the gains vs. baseline originally reported in column 2, where a is also increased by 25%, but the minimum wage is not changed,

Table 9: **The role of intermediates: allocations**

	Baseline	↓ 25% γ_m	↓ 25% γ_m + avg. all τ	↓ 25% γ_m + disp. τ^L
Modern VA	50.0	42.4	45.6	43.9
Modern Emp	17.1	13.8	16.4	16.0
Modern VA / Employment	2.9	3.1	2.8	2.7
U/Ubaseline	1.00	0.72	0.79	0.73
Entrepreneurs	5.7	4.0	4.6	4.1
Salaried	11.4	9.8	11.8	11.9
Self employed	82.9	86.2	83.6	84.0
Modern share within sector				
Agriculture	16.0	8.4	10.7	8.9
Industry	71.9	64.2	67.1	65.1
Services	45.7	34.1	38.2	35.3
Total Exp. Shares				
Agriculture	12.7	11.2	11.8	11.4
Industry	37.5	39.4	38.6	39.1
Services	49.8	49.5	49.6	49.5

are eroded. But while income gains continue to be substantial (186 p.p. in column 7 vs. 1 compared to 216 p.p. in column 2 vs. 1), the increase in modern jobs becomes 2.9 p.p. rather than the 13.5 p.p. of column 2. Moreover, the much larger fraction of workers who remain self-employed are also considerably worse off: compared to the baseline, their income gap expands (rather than contracting as it did in column 2) from 1.41 to 1.71 vis-a-vis formal workers, and from 10.81 to 13.28 vis-a-vis employers.

An experiment that shifts the d distribution to the right by 25% also expands modernity and aggregate income. Although it is also more effective in doing so than the experiments reducing regulatory burdens from the previous section, the comparison with the experiment of shifting a is much less favorable. Compared to column (2), results in column (3) show less effectiveness in increasing modern value added and employment, much less effectiveness in delivering aggregate income increases, and counterproductive effects if reducing functional inequality is a goal. The increase in aggregate income is close to 40%, and that of modern employment is less than 5 p.p, compared to 116% and 13 p.p. in column 2, respectively. And it is the profits of modern employers that expand relatively more, widening the gap with both modern workers and those in the traditional sector. The gap between modern workers and the self-employed also grows.

We also study experiments that change a focusing on specific parts of the distribution and show the results in columns (4) and (5). Targeting the bottom tail of the a distribution with a 25% increase (column 4) is much more effective than the opposite experiment in column 5, that is, increasing the upper half of the a distribution. Improvements in a in the bottom half of the distribution imply increases in real income of 60% and of modern employment by more than 10%. It yields the most potent reductions in the gap in income for workers vis-a-vis modern employers, both for workers in the modern and the informal segment. By contrast, increasing a in the upper tail substantially increases inequality. It shrinks the modern segment of the economy: self-employment increases by 9 percentage points, and the modern value-added

Table 10: **The role of intermediates: distribution**

	Baseline	↓ 25% γ_m	↓ 25% γ_m + avg. all τ	↓ 25% γ_m + disp. τ^L
U/Ubaseline	1.00	0.72	0.79	0.73
Avg. real profits entrep	10.81	8.47	9.28	8.62
Avg. real inc. formal work	1.41	1.08	1.22	1.13
Avg. real inc. self-emp	1.00	0.80	0.86	0.81
Avg. profits entrep / Avg. inc. self-emp	10.81	10.59	10.85	10.61
Avg. inc. formal work / Avg. inc. self-emp	1.41	1.35	1.42	1.39
Avg. profits entrep / Avg. inc. formal work	7.66	7.85	7.63	7.66
Share people below min earn	16.78	16.78	16.78	16.78
Share of self-emp above min earn	79.40	83.50	80.30	80.80
Share of self-emp above median inc.	70.20	76.00	71.50	72.30
Earnings p50-p10	2.36	2.36	2.36	2.36
Earnings p95-p10	14.05	13.07	13.66	13.07

share of the economy decreases by 5 pp.

The role of intermediates The usage of intermediates in production plays a central role in amplifying the production distortions faced by the modern segment of the economy. To illustrate this point, Tables 9 and 10 report the allocative and distributional consequences of reducing the intermediate share in production by 25%. We then compare the effect of reducing distortions in this environment relative to our baseline environment.

In particular, and to avoid a taxonomical analysis, we focus on a reduction on the average level of all distortions by 25% and a reduction in the dispersion of 25% of τ_l. Comparing the second column, which reports the effect of simply reducing the intermediate shares, with columns 3 and 4, which also reduce distortions, we find that removing modern production distortions affects relatively less the allocation when comparing it to the baseline results. This is a natural outcome of the roundabout nature of the model: since intermediate production requires modern inputs, all distortions that affect the production of these intermediates appear in the production of intermediates. Since intermediates themselves are an input to the production of modern goods, the effect of distortions are amplified. The more so, the more modern production relies in intermediates. For example, focusing on columnn 4, we see that a reduction on labor wedges increases modern value added by 1.5 pp and modern employment by 2.2 pp. By contrast, in our baseline exercise, the same reduction increase the modern value added share by 3.8 pp and modern employment by 2.4 pp.

5 Conclusions

Emerging economies often struggle to allocate resources efficiently, preventing labor and capital from reaching their most productive uses. Compared to wealthier nations, these economies experience widespread misallocation—within sectors, across sectors, and between modern and traditional segments of the economy—resulting in substantial income losses. A significant portion of the workforce remains concentrated in informal businesses, which are characterized by low capital intensity, outdated production methods, and high levels of self-employment. This results in a significant fraction of the population earning very low incomes. Using historical microdata from Colombia and a general equilibrium model of occupational choice, we analyze how these inefficiencies contribute to "jobless industrialization"—a phenomenon in which industrial growth raises value-added without a corresponding increase in employment. Our findings indicate that the most effective way to address this challenge is by improving embodied human capital. A 25% rightward shift in its distribution could more than double income levels, close half the income gap with the United States, and significantly reduce inequality between informal and formal workers. While distortions affecting modern firms—such as size-based regulations and minimum wage constraints—also contribute to jobless industrialization, their impact remains limited unless coupled with human capital improvements.

We examine a range of structural distortions constraining the expansion of modern production and the creation of formal jobs in emerging economies. These barriers operate at multiple levels: within-sector misallocation prevents high-productivity firms from scaling up, cross-sector distortions hinder the efficient movement of labor and capital, and segmentation between modern and traditional sectors reinforces informality and underemployment. Firm-specific constraints—including payroll taxes, revenue distortions, and financing restrictions—disproportionately burden the most productive businesses, limiting their ability to absorb labor. At the same time, rigid labor market regulations, such as minimum wage policies and hiring restrictions, further constrain job creation in the modern sector. These frictions, combined with human capital deficits, exacerbate the challenges of structural transformation and economic growth.

Our counterfactual results shed light on the importance and incidence of these barriers in preventing late-industrializing economies from achieving employment-rich growth. Through counterfactual simulations, we demonstrate that improving human capital is a more effective strategy for reducing joblessness than merely addressing firm-level distortions. However, reforms targeting labor and revenue distortions must be carefully designed to avoid deepening income disparities between formal and informal workers. Ultimately, our findings suggest that differences in structural transformation across countries stem from varying degrees of barriers to adopting modern production techniques and hiring practices. By addressing these constraints, policymakers can promote a more inclusive and sustainable industrialization process—one that fosters both economic growth and widespread employment opportunities.

References

ALFARO, L., A. CHARLTON, AND F. KANCZUK (2009): *Plant Size Distribution and Cross-Country Income Differences*, University of Chicago Press, 243–272.

AMODIO, F., P. MEDINA, AND M. MORLACCO (2022): "Labor Market Power, Self-Employment, and Development," IZA Discussion Papers 15477, Institute of Labor Economics (IZA).

BENTO, P. AND D. RESTUCCIA (2021): "On average establishment size across sectors and countries," *Journal of Monetary Economics*, 117, 220–242.

BUERA, F. J., J. P. KABOSKI, AND R. M. TOWNSEND (2021): "From Micro to Macro Development," NBER Working Papers 28423, National Bureau of Economic Research, Inc.

BUERA, F. J. AND B. MOLL (2015): "Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity," *American Economic Journal: Macroeconomics*, 7, 1–42.

CARDENAS, M. AND G. PERRY (1986): "Diez años de reformas tributarias en Colombia," Tech. rep., Fedesarrollo.

DIX-CARNEIRO, R., P. K. GOLDBERG, C. MEGHIR, AND G. ULYSSEA (2021): "Trade and Informality in the Presence of Labor Market Frictions and Regulations," Working Paper 28391, National Bureau of Economic Research.

ESLAVA, M., J. C. HALTIWANGER, AND N. URDANETA (2023): "The Size and Life-cycle Growth of Plants: The Role of Productivity, Demand and Wedges," *The Review of Economic Studies*, Forthcoming.

FENG, Y., D. LAGAKOS, AND J. E. RAUCH (2024): "Unemployment and Development," *The Economic Journal*, 134, 614–647.

FUJIWARA, I. AND K. MATSUYAMA (2020): "A Technology-Gap Model of Premature Deindustrialization," CEPR Discussion Papers 15530, C.E.P.R. Discussion Papers.

HARRIS, J. R. AND M. P. TODARO (1970): "Migration, Unemployment and Development: A Two-Sector Analysis," *The American Economic Review*, 60, 126–142.

HERRENDORF, B., R. ROGERSON, AND VALENTINYI (2014): "Growth and Structural Transformation," in *Handbook of Economic Growth*, ed. by P. Aghion and S. Durlauf, Elsevier, vol. 2 of *Handbook of Economic Growth*, chap. 6, 855–941.

HERRENDORF, B., R. ROGERSON, AND ÁKOS VALENTINYI (2022): "New Evidence on Sectoral Labor Productivity: Implications for Industrialization and Development," NBER Working Papers 29834, National Bureau of Economic Research, Inc.

HSIEH, C.-T. AND P. J. KLENOW (2009): "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 124, 1403–1448.

HUNEEUS, F. AND R. ROGERSON (2020): "Heterogeneous Paths of Industrialization," NBER Working Papers 27580, National Bureau of Economic Research, Inc.

LA PORTA, R. AND A. SHLEIFER (2014): "Informality and Development," *Journal of Economic Perspectives*, 28, 109–26.

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

——— (1979): "The Dual Economy Revisited," *The Manchester School*, 47, 211–229.

MCMILLAN, M. AND A. ZEUFACK (2022): "Labor Productivity Growth and Industrialization in Africa," *Journal of Economic Perspectives*, 36, 3–32.

MCMILLAN, M. S. AND D. RODRIK (2011): "Globalization, Structural Change, and Productivity Growth," *NBER Working Paper No. 17143*.

MESTIERI, M., J. SCHAUER, AND R. TOWNSEND (2017): "Human Capital Acquisition and Occupational Choice: Implications for Economic Development," *Review of Economic Dynamics*, 25, 151–186.

RAUCH, J. E. (1991): "Modelling the informal sector formally," *Journal of Development Economics*, 35, 33–47.

RESTUCCIA, D. AND R. ROGERSON (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11, 707–720.

RODRIK, D. (2013): "Unconditional Convergence in Manufacturing," *The Quarterly Journal of Economics*, 128, 165–204.

——— (2016): "Premature deindustrialization," *Journal of Economic Growth*, 21, 1–33.

SPOSI, M., K.-M. YI, AND J. ZHANG (2021): "Deindustrialization and Industry Polarization," Working Paper Series WP 2022-44, Federal Reserve Bank of Chicago.

TYBOUT, J. (2000): "Manufacturing Firms in Developing Countries: How Well Do They Do, and Why?" *Journal of Economic Literature*, 38, 11–44.

APPENDIX

A Data Construction

A.1 Modern firm's problem

The modern firm solves

$$Max_{\{k,l\}} \left(1-\tau^R\right) py - \left(1+\tau^l\right) wl - \left(1+\tau^K\right) rk - p_m m \tag{17}$$

subject to **??**. Revenue $R = py$ is given by,

$$R = D^{\frac{1}{\sigma}} d^{\frac{1}{\sigma}} y^{1-\frac{1}{\sigma}} \tag{18}$$

The first order condition for any factor $x = \{k, l\}$ with price p^x is thus given by

$$\begin{aligned} \left(1+\tau^x\right) p^x &= \left(1-\tau^R\right)\left[p\frac{\partial y}{\partial x} + y\left(\frac{\partial p}{\partial y} + \frac{\partial p}{\partial P}\frac{\partial P}{\partial y}\right)\frac{\partial y}{\partial x}\right] \\ &= \left(1-\tau^R\right) p\left(\frac{\sigma-1}{\sigma}\right)(1-s)\frac{\partial y}{\partial x} \end{aligned} \tag{19}$$

where s is the firm's revenue share in the market, and the last line uses Shepard's lemma ($\frac{\partial P}{\partial y}\frac{y}{P} = -s$), and the facts that $\frac{\partial p}{\partial y}\frac{y}{p} = -\frac{1}{\sigma}$ and $\frac{\partial p}{\partial P}\frac{P}{p} = \frac{\sigma-1}{\sigma}$.

Moreover, the markup is given by:[1]

$$\mu^{-1} = \left(\frac{\sigma-1}{\sigma}\right)(1-s) \tag{20}$$

Thus, optimal k is given by $\left(1+\tau^k\right) rk = \left(1-\tau^R\right)\mu^{-1}\alpha\gamma R$, and the optimal capital-labor ratio is given by

$$l = \frac{(1-\alpha-\phi)\, r\left(1+\tau^k\right)}{\alpha w\left(1+\tau^l\right)} k \equiv \theta k \tag{21}$$

whilst the optimal capital-materials ratio is

$$m = \frac{\phi r\left(1+\tau^k\right)}{\alpha p_m} k \equiv \theta_m k \tag{22}$$

Using these FOCs, profits are given by:

[1] See Hottman, Redding and Weinstein, 2016, and Eslava, Haltiwanger and Urdaneta, 2022.

$$\begin{aligned}\pi &= \left(1-\tau^R\right)R-\left[\left(1+\tau^l\right)w\theta k+\left(1+\tau^k\right)rk+p_m\theta_m k\right]\\ &= \left(1-\tau^R\right)R\left(1-\mu^{-1}\gamma\right)\end{aligned} \tag{23}$$

Further, plugging (21), (22) and (20) into the FOC for capital and replacing p using (**??**), we obtain optimal capital

$$k^{1-\gamma\left(1-\frac{1}{\sigma}\right)} = \frac{\left(1-\tau^R\right)\mu^{-1}\Delta^{\frac{\sigma-1}{\sigma}}\alpha\gamma\theta^{(1-\alpha-\phi)\gamma\frac{\sigma-1}{\sigma}}\theta_m^{\phi\gamma\frac{\sigma-1}{\sigma}}}{\left(1+\tau^k\right)r} \tag{24}$$

$$= \left(1-\tau^R\right)\mu^{-1}\Delta^{\frac{\sigma-1}{\sigma}}\gamma\left(\frac{\alpha}{r(1+\tau^k)}\right)^{1-(1-\alpha)\gamma\frac{\sigma-1}{\sigma}}\left(\frac{1-\alpha-\phi}{w(1+\tau^l)}\right)^{(1-\alpha-\phi)\gamma\frac{\sigma-1}{\sigma}}\left(\frac{\phi}{p_m}\right)^{\phi\gamma\frac{\sigma-1}{\sigma}} \tag{25}$$

where we have denoted quality-adjusted physical productivity by $\Delta = D^{\frac{1}{\sigma-1}}d^{\frac{1}{\sigma-1}}a$.

For future reference, plugging (24) into (21), labor demand is

$$l^{1-\gamma\left(1-\frac{1}{\sigma}\right)} = \left(1-\tau^R\right)\mu^{-1}\Delta^{\frac{\sigma-1}{\sigma}}\gamma\left(\frac{\alpha}{r\left(1+\tau^k\right)}\right)^{\alpha\gamma\left(\frac{\sigma-1}{\sigma}\right)}\left(\frac{1-\alpha-\phi}{w\left(1+\tau^l\right)}\right)^{\left(1-(\alpha+\phi)\gamma\left(\frac{\sigma-1}{\sigma}\right)\right)}\left(\frac{\phi}{p_m}\right)^{\phi\gamma\frac{\sigma-1}{\sigma}} \tag{26}$$

Because $y = a(k\theta^{(1-\alpha-\phi)}\theta_m^{\phi})^{\gamma}$ and $R = py = D^{\frac{1}{\sigma}}d^{\frac{1}{\sigma}}y^{\frac{\sigma-1}{\sigma}}$, we obtain

$$R = \left[\Delta^{\frac{1}{\gamma}}\frac{\left(1-\tau^R\right)}{\left(1+\tau^k\right)^{\alpha}\left(1+\tau^l\right)^{1-\alpha-\phi}}\mu^{-1}\gamma\left(\frac{\alpha}{r}\right)^{\alpha}\left(\frac{1-\alpha-\phi}{w}\right)^{1-\alpha-\phi}\left(\frac{\phi}{p_m}\right)^{\phi}\right]^{\frac{\gamma(\sigma-1)}{\sigma-\gamma(\sigma-1)}} \tag{27}$$

One could thus write revenue (and profits, using (23)) as a the function of quality-adjusted quantity productivity, distortions in product and labor markets (τ^R, τ^l and τ^k), markups and a set of parameters and equilibrium input prices that are constant across firms (of the same sector).

$$R = \left(\Delta\mu^{-\gamma}\frac{\left(1-\tau^R\right)^{\gamma}}{\left(1+\tau^k\right)^{\alpha\gamma}\left(1+\tau^l\right)^{\gamma(1-\alpha-\phi)}}\Theta\right)^{\frac{(\sigma-1)}{\sigma-\gamma(\sigma-1)}} \tag{28}$$

$$\pi = \left(\Delta\left(1-\mu^{-1}\gamma\right)^{\frac{\sigma-\gamma(\sigma-1)}{(\sigma-1)}}\mu^{-\gamma}\frac{\left(1-\tau^R\right)^{\frac{\sigma}{\sigma-1}}}{\left(1+\tau^k\right)^{\alpha\gamma}\left(1+\tau^l\right)^{\gamma(1-\alpha-\phi)}}\Theta\right)^{\frac{(\sigma-1)}{\sigma-\gamma(\sigma-1)}} \tag{29}$$

where $\Theta = \gamma^{\gamma}(\frac{\alpha}{r})^{\gamma\alpha}\left(\frac{1-\alpha-\phi}{w}\right)^{\gamma(1-\alpha-\phi)}(\frac{\phi}{p_m})^{\phi\gamma}$. Revenue and profits are convex in a and in quality-adjusted quantity productivity Δ whenever $\frac{(\sigma-1)}{\sigma-\gamma(\sigma-1)} > 1$ and concave otherwise.

Aggregation of modern sector As shown in Appendix C, modern manufacturing productivity is given

by

$$TFP_{MF} = \frac{C_{MF}}{K^{\alpha}L^{(1-\alpha)}} = D^{-\frac{1}{\sigma-1}} \left(\sum_{I} \left[\left(\frac{\Delta}{tfpr^{\gamma}} \right)^{\frac{1}{\sigma\left(1-\gamma\left(1-\frac{1}{\sigma}\right)\right)}} TFPR_{MF} \right]^{\sigma-1} \right)^{\frac{1}{\sigma-1}} \tag{30}$$

where $tfpr \equiv \frac{py}{k^{\alpha}l^{(1-\alpha)}} = \frac{\left(\frac{\alpha}{\left(1+\tau^{k}\right)r}\right)^{\alpha}\left(\frac{1-\alpha-\phi}{w\left(1+\tau^{l}\right)}\right)^{1-\alpha-\phi}(\frac{\phi}{p_m})^{\phi}\mu}{\gamma\left(1-\tau^{R}\right)}$ is a modern firm's revenue productivity (see Appendix C) and $TFPR_{MF} \equiv \frac{E_{MF}}{K^{\alpha}L^{1-\alpha}}$ is aggregate revenue productivity in manufacturing.

B Estimation of parameters of production and demand functions in modern manufacturing

We follow the structural value-added approach to the estimation of the establishment's production function (Ackerberg et al., 2015, Gandhi et al. 2017). An establishment indexed by i produces y_{it} using labor l_{it} and capital k_{it}, with technology

$$y_{it} = a_{it} l_{it}^{1-\alpha} k_{it}^{\alpha} \tag{31}$$

The underlying assumption is that production is Leontieff, $y_{it} = \min(l_{it}^{1-\alpha} k_{it}^{\alpha} a_{it}, m_{it})$. At the optimum, thus, (31) is satisfied and $m_{it} = l_{it}^{1-\alpha} k_{it}^{\alpha} a_{it}$.

Following Ackerberg, Caves, and Frazer (2015, ACF henceforth), we use material inputs to proxy for productivity. The Leontieff assumption implies the optimal choice of material inputs can be written as:

$$\ln m_{it} = h\left(\ln a_{it}, \ln k_{it}, \ln l_{it}\right)$$

Inverting, we obtain

$$\ln a_{it} = h^{-1}\left(\ln m_{it}, \ln l_{it}, \ln k_{it}\right) \equiv h^{-1}\left(\overrightarrow{W_{it}}\right)$$

Incorporating this expression, recognizing that y_{it} is subject to measurement error and other shocks not observed by either the econometrician or the firm at the time of making input choices, and denoting by $\widehat{y_{it}} = y_{it} e^{\varepsilon_{it}}$ measured y_{ft}, we write:

$$\ln \widehat{y_{it}} = \alpha\gamma \ln k_{it} + (1-\alpha)\gamma \ln l_{it} + h^{-1}\left(\overrightarrow{W_{it}}\right) + \varepsilon_{it} \tag{32}$$

y_{it} is a CES composite of varieties produced by i. We follow Eslava et al. (2022) to jointly estimate 32 and the demand function implied by preferences in our setting:

$$R_{it} = D_t + d_{it} + p_{it}^{1-\sigma} \tag{33}$$

where $R_{it} = p_{it}y_{it}$ is $i's$ revenue. CES preferences across varieties produced within i imply (see Eslava et al. (2023)) $p_{it} = \overline{P^*_{ft}} P_{fB} \left(\Lambda^Q_{ft}\right)^{\frac{1}{\sigma-1}}$ where $s_{ijt} = \frac{p_{ijt}y_{ijt}}{R_{it}}$; $\Omega^i_{t,t-1}$ is the set of products that i produces in both t and $t-1$ and s^*_{ijt} is the share that product j represents in the revenue that i obtains in t from products in $\Omega^i_{t,t-1}$; $\lambda^{QF}_{it} = \frac{\sum_{\Omega^i_{t,t-1}} s_{ijt}}{\sum_{\Omega^i_{t,t-1}} s_{ijt-1}}$; and $\lambda^{QRW}_{it} = \prod_{\Omega^i_{t,t-1}} \left(\frac{s^*_{ijt}}{s^*_{ijt-1,\Omega^i_{t,t-1}}}\right)^{\frac{1}{\|\Omega^i_{t,t-1}\|}}$.

Since $y_{it} = \frac{R_{it}}{p_{it}}$ and $p_{it} = \overline{P^*_{it}} P_{iB} \left(\Lambda^Q_{it}\right)^{\frac{1}{\sigma-1}}$, σ is necessary to measure prices and thus y_{it}. An estimate of σ is not available at the stage of estimating 32 and 33, so we estimate a modified version of 32:

$$\ln \widehat{y^*_{it}} = \ln\left(\frac{R_{it}}{\overline{P^*_{ft}}}\right) = \alpha\gamma \ln k_{it} + (1-\alpha)\,\gamma \ln l_{it} + h^{-1}\left(\overrightarrow{W_{it}}\right) + \frac{1}{1-\sigma}\Lambda^Q_{ft} + \widehat{\varepsilon}_{it} \tag{34}$$

Following Eslava et al. (2023) and the proxy-methods literature on production function estimation, we estimate 34 in two stages. The first stage purges measurement error $\widehat{\varepsilon}_{it}$, while the second stage, jointly estimated with 33 via GMM, estimates (α, γ, σ). Assuming that a_{it} and d_{it} follow flexible AR1 processes:

$$\begin{aligned}
\ln a_{it} &= \pi^a_0 + \pi^a_1 \ln a_{it-1} + \pi^A_2 \ln a^2_{it-1} + \pi^A_3 \ln a^3_{it-1} + \xi^A_{it} \\
\ln a_{it} &= \pi^D_0 + \pi^D_1 \ln D_{it-1} + \pi^A_2 \ln D^2_{it-1} + \pi^A_3 \ln D^3_{it-1} + \xi^D_{it}
\end{aligned}$$

we use the following moment conditions for identification based on those proposed by **?**:

$$E\begin{bmatrix} \ln L_{it} \times \xi^A_{it} \\ \ln K_{it} \times \xi^A_{it} \\ \ln D_{it-1} \times \xi^A_{it} \\ \ln A_{it-1} \times \xi^D_{it} \\ \ln A_{it} \\ \ln D_{it} \end{bmatrix} = 0 \tag{35}$$

Table 11 presents the detailed results.

Table 11: Factor and demand elasticities by sector (pooling years)

Sector	$1-\alpha$	α	γ	σ_w	σ	$\gamma\left(\frac{1}{1-\sigma}\right)$
311	0,646	0,354	1,052	8,110	4,073	0,794
313	0,674	0,326	1,042	5,531	4,303	0,799
321	0,546	0,454	0,885	4,776	4,036	0,666
322	0,594	0,406	0,679	6,287	4,019	0,510
323	0,539	0,461	0,946	5,268	4,081	0,714
324	0,640	0,360	0,865	5,975	3,957	0,647
331	0,758	0,242	0,887	4,785	3,828	0,655
332	0,580	0,420	0,884	3,616	4,003	0,663
341	0,588	0,412	0,961	2,570	4,058	0,724
342	0,775	0,225	1,053	2,012	4,057	0,794
351	0,657	0,343	1,127	8,042	4,162	0,857
352	0,563	0,437	0,979	6,171	4,174	0,744
355	0,736	0,264	0,975	4,028	4,176	0,741
356	0,703	0,297	1,050	3,467	4,022	0,789
362	0,685	0,315	1,092	3,745	4,031	0,821
369	0,603	0,397	1,019	6,880	4,078	0,770
371	0,685	0,315	1,249	4,856	4,155	0,949
381	0,659	0,341	0,946	1,912	3,955	0,707
382	0,709	0,291	0,959	3,401	4,011	0,720
383	0,688	0,312	1,056	4,447	4,051	0,795
384	0,748	0,252	1,132	7,526	3,990	0,848
385	0,640	0,360	1,002	4,260	3,901	0,745
390	0,741	0,259	0,971	2,009	4,004	0,729
Average	0,659	0,341	0,992	4,768	4,049	0,747
Min	0,539	0,225	0,679	1,912	3,828	0,510
Max	0,775	0,461	1,249	8,110	4,303	0,949

C TFP

Modern manufacturing productivity is given by

$$TFP_{MF} = \frac{C_{MF}}{K^{\alpha}L^{(1-\alpha)}} = \frac{E_{MF}}{K^{\alpha}L(1-\alpha)}\frac{1}{P_{MF}} = \frac{TFPR_{MF}}{P_{MF}} \tag{36}$$

where E_{MF} is total expenditure in manufacturing, and $TFPR_{MF} \equiv \frac{E_{MF}}{K^{\alpha}L^{1-\alpha}}$ is aggregate revenue productivity in manufacturing.

Define also $tfpr = \frac{py}{k^{\alpha}l^{(1-\alpha)}} = \frac{D^{\frac{1}{\sigma}}d^{\frac{1}{\sigma}}y^{1-\frac{1}{\sigma}}}{\left(k^{\alpha}l^{1-\alpha}\right)} = \frac{D^{\frac{1}{\sigma}}d^{\frac{1}{\sigma}}a^{1-\frac{1}{\sigma}}}{\left(k^{\alpha}l^{1-\alpha}\right)^{1-\gamma\left(1-\frac{1}{\sigma}\right)}}$. Moreover, from **??** it is apparent that, at the optimum:

$$k^{\alpha}l^{(1-\alpha)} = \left(\frac{\Delta^{\frac{\sigma-1}{\sigma}}\gamma}{\mu c\left(1-\tau^{R}\right)^{-1}}\right)^{\frac{1}{1-\gamma\left(1-\frac{1}{\sigma}\right)}}. \tag{37}$$

Replacing this into $tfpr$, we obtain:

$$tfpr = \frac{c\mu}{\gamma\left(1-\tau^{R}\right)} \tag{38}$$

where $c = \left(\frac{\alpha}{(1+\tau^{k})r}\right)^{\alpha}\left(\frac{1-\alpha}{w(1+\tau^{l})}\right)^{1-\alpha}$. As a result, we can also write $k^{\alpha}l^{(1-\alpha)} = \left(\frac{\Delta^{\frac{\sigma-1}{\sigma}}}{tfpr}\right)^{\frac{1}{1-\gamma\left(1-\frac{1}{\sigma}\right)}}$. Replacing this into $tfpr$ we obtain $tfpr = pa\left(k^{\alpha}l^{1-\alpha}\right)^{\gamma-1} = pa\left(\frac{\Delta^{\frac{\sigma-1}{\sigma}}}{tfpr}\right)^{\frac{\gamma-1}{1-\gamma\left(1-\frac{1}{\sigma}\right)}}$, so that after some manipulation

$$tfpr^{\frac{\gamma}{\sigma(1-\gamma\frac{\sigma-1}{\sigma})}} = pa\left(\Delta^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\gamma-1}{1-\gamma\left(1-\frac{1}{\sigma}\right)}} \tag{39}$$

Using 39 in 36

$$\begin{aligned}
TFP_{MF} &= TFPR_{MF}\left(\sum_{I} dp^{1-\sigma}\right)^{\frac{1}{\sigma-1}} \\
&= TFPR_{MF}\left(\sum_{I}\left(d^{\frac{1}{\sigma-1}}\frac{a\left(\Delta^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\gamma-1}{1-\gamma\left(1-\frac{1}{\sigma}\right)}}}{tfpr^{\frac{\gamma}{\sigma(1-\gamma\frac{\sigma-1}{\sigma})}}}\right)^{\sigma-1}\right)^{\frac{1}{\sigma-1}} \\
&= D^{-\frac{1}{\sigma-1}}\left(\sum_{I}\left[\left(\frac{\Delta}{tfpr^{\gamma}}\right)^{\frac{1}{\sigma\left(1-\gamma\left(1-\frac{1}{\sigma}\right)\right)}} TFPR_{MF}\right]^{\sigma-1}\right)^{\frac{1}{\sigma-1}}
\end{aligned} \tag{40}$$

where we have used the fact that $\Delta = D^{\frac{1}{\sigma-1}}d^{\frac{1}{\sigma-1}}a$ and $1+\left(\frac{\sigma-1}{\sigma}\right)\frac{(\gamma-1)}{\left(1-\gamma\left(\frac{\sigma-1}{\sigma}\right)\right)} = \frac{1}{\sigma\left(1-\gamma\left(\frac{\sigma-1}{\sigma}\right)\right)}$.

D Data and Additional Results

D.1 Detailed Data Description

The primary data source for the motivating macro facts comes from the Groningen Growth and Development Center's 10-Sector Database, which provides internationally comparable measures for employment (number of workers) and value-added (nominal) for the 10 broad sectors (ISIC rev 3.1) of the economy since 1950. The sectors included are: 1) Agriculture, hunting, forestry, and fishing (AtB); 2) Mining and quarrying (C); 3) Manufacturing (D); 4) Electricity, gas and water supply (E); 5) Construction (F); 6) Wholesale and retail trade, hotels and restaurants (GtH); 7) Transport, storage, and communication (I); 8) Finance, insurance, real estate and business services (JtK); 9) Government services (LtN); and 10) Community, social and personal services (OtP). These data cover advanced and developing economies all over the world. We updated these numbers with World KLEMS for the United States and several European Countries, with the Economic Transformation Database for countries in Africa, Asia, and Latin America, and with the Montevideo-Oxford Latin American Economic History Data Base to complete the value-added series from Brazil since 1950.

In addition, we use the INDSTAT2 from the United Nations Industrial Development Organization (UNIDO), which contains information from industrial surveys from 1963. Important for our purposes, these data do not include the smallest manufacturing firms, which also tend to be informal. Thus, this data is useful to measure the number of workers in manufacturing establishments outside the micro enterprise segment (typically also informal) worldwide.

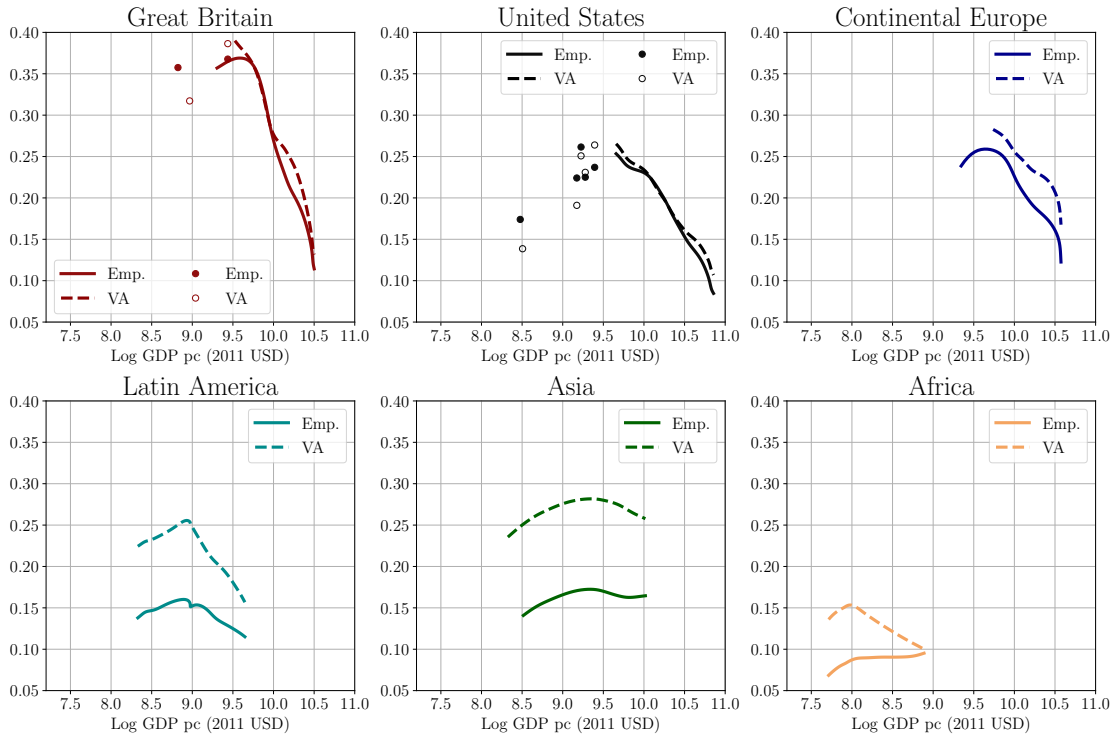
Table 12: **Manufacturing: share of employment vs. share of the value generated by employers of different sizes**

Employer size	Ratio of category's weight in employment to category's weight in ...			Individual-level Survey	
	value added (Manufacturing Survey)	worker earnings (Colombia)	worker earnings (CPS USA)	Weight of category in manufacturing employment (Colombia)	Weight of category in manufacturing employment (USA)
Panel A: 1997					
$\leq$ 10 workers (inc. Self-employed)		1.456	1.237	0.594	0.080
10-50 workers	2.127	0.963	1.222	0.129	0.225
51-100 workers	1.591	0.601	0.925	0.276	0.695
101+ workers	0.836				
Panel B: 2018					
$\leq$ 10 workers (inc. Self-employed)		1.431	1.294	0.564	0.100
10-50 workers	1.772	0.887	1.174	0.125	0.138
51-100 workers	1.537	0.750	1.138	0.036	0.093
101+ workers	0.856	0.662	0.930	0.276	0.669

The Manufacturing Survey (first column) covers all manufacturing establishments with at least 10 employees. The last four columns are based on surveys of individuals representative of all occupied individuals. This is the Annual Social and Economic Supplement of the Consumer Population Survey (ASEC-CPC) for the US. For Colombia, it is the Households Survey (the official source of labor market statistics) in 2018 and the Life Quality Survey in 1997. The Life Quality Survey was also collected in 2018 and yielded results similar to those of the households' survey.

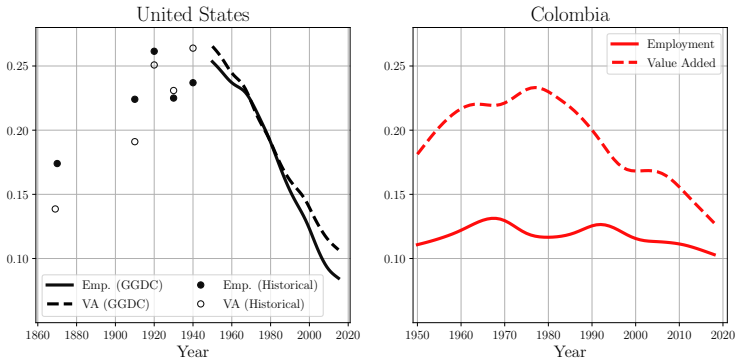
Using data on establishments (i.e., salaried employment) and focusing on a subset of late industrializers

Figure 5: **Manufacturing as a Share of Overall Employment and Value-Added Around the World**



Notes: Data sources: GGDC 10-Sector Database; Economic Transformation Database; World KLEMS. The time series are trended using a Hodrick–Prescott filter with a smoothing parameter of $\lambda = 100$. Regional series are simple yearly averages. The countries included in each region are: i) Europe: Denmark, Spain, France, Great Britain, Italy, The Netherlands, and Sweden; ii) Latin America: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica and Mexico; Asia: China, India, Indonesia, Japan, Korea, Malaysia, Philippines, Thailand and Taiwan; iii) Africa: Botswana, Egypt, Ethiopia, Ghana, Kenya, Morocco, Malawi, Mauritius, Nigeria, Senegal, South Africa, Tanzania and Zambia. Peru and Venezuela were excluded due to a lack of data during the 1950s. The data on Value Added for Brazil was updated backward using the Montevideo-Oxford Latin American Economic History Data Base. Hong Kong and Singapore were excluded from the Asian average, as these economies are city-states.

Figure 6: **Manufacturing as a share of Employment and Value-Added: Colombia vs. US**



Notes: Data sources: GGDC 10-Sector Database; Economic Transformation Database; World KLEMS; MOxLAD Database.

in Latin America, Table 13 shows that micro establishments are much more prevalent in these countries as compared to advanced economies, dominated by early industrializers, especially in terms of their weight

in employment. Although this holds in all sectors, the weight of micro establishments is lesser in manufacturing than in basic services (including retail). Moreover, there is a tight link between the prevalence of microestablishments and that of informality in these late industrializers. Even among salaried workers, the vast majority of informal employment, i.e., employment that does not abide by labor regulations, concentrates on micro-employers (Table 14. The proportion grows to almost 100% of informal workers if self-employment is included. The weight of microestablishments in formal employment is much lower, especially in manufacturing, where it does not reach above 10%.

Table 13: **Establishments under 10 Employees as a % of all (Economic Censuses, circa 2005)**

	Unweighted			Employment-weighted		
	Advanced Economies	Latin America	Colombia	Advanced Economies	Latin America	Colombia
Manufacturing	71,1%	91,6%	87,2%	11,5%	30,3%	31,7%
Wholesale-Retail	85,0%	96,9%	96,0%	27,4%	65,8%	61,4%
Accom-Restaur.			95,7%			69,9%
Other Services	88,1%	91,0%	87,8%	25,7%	34,3%	34,0%

Table 14: **Workers in Establishments of at most 10 Employees as a % of all workers in the category: Formal or Informal. Excludes self-employment**

	Latin America		Colombia		US
	Formal workers	Informal workers	Formal workers	Informal workers	
Manufacturing	10%	69%	8%	71%	4%
Wholesale-Retail	28%	82%	24%	88%	11%
Services	15%	60%	11%	73%	13%
Accom-Restaur.	29%	78%	33%	88%	31%

Latin America: average of Argentina, Bolivia, Colombia, El Salvador, Guatemala, Honduras, Mexico, Paraguay, Peru, and Uruguay. Formal workers: contributing to a pension. Pooled data from 2011 to 2015. Latin America from household surveys, the US from Business Dynamics Statistics. Source: Eslava (2018).

Table 15: **Distribution of Colombian workers across employer size classes by sector and formality status: 2012**

Employer size	Manufacturing		Non-manufacturing		All
	Formal workers (45%)	Informal workers (55%)	Formal workers (35%)	Informal workers (65%)	
$\leq$ 10 workers (incl. Self-employed)	0.9%	7.5%	9.1%	53.7%	71.1%
11-50 workers	1.2%	0.6%	4.1%	1.8%	7.6%
51-100 workers	0.6	0.1%	1.5%	0.2%	2.3%
101+ workers	3.3%	0.2%	14.6%	0.9%	19.0%
Total	6.0%	8.2%	29.2%	56.6%	100%

Source: Colombian Household's Survey. Urban and rural coverage.