Economic Development According to Chandler*

Niklas Engbom — NYU – Stern

Hannes Malmberg — Minnesota

Tommaso Porzio — Columbia

Federico Rossi — Warwick

Todd Schoellman — Minneapolis Fed

June 2026

Abstract

Chandler (1977) shows that large firms require hierarchies of white-collar workers to coordinate complex production. We document that this insight continues to hold globally today, and we show that low educational attainment in developing countries limits the supply of white-collar workers and constrains firm size. We extend the occupational choice model of Lucas (1978) to allow entrepreneurs to reorganize their firms by allocating administrative tasks to hired professionals, thereby bringing their firms closer to constant returns to scale. We calibrate the model to be consistent with cross-sectional microdata and benchmark it using causal evidence on the effects of expansions of schooling. Skills explain half of the reorganization of production into large firms that accompanies economic development, while structural transformation and reductions in barriers are needed to explain the remaining shift.

Keywords: skills, white-collar workers, returns to scale, firm size, endogenous duality

*We thank participants at numerous conferences and seminars for helpful comments and feedback. We are particularly grateful to Jonas Gathen, Jan Grobovšek, Juan Vizcaino, Bryan Seegmiller, Michael Peters, and Paco Buera for insightful discussions. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1 Introduction

In a seminal contribution, Chandler (1977) explores the transformation of American businesses during the Second Industrial Revolution. The defining technologies of this era leveraged economies of scale and scope to achieve productivity gains. As firms adopted these technologies and grew, they encountered new logistical challenges: sourcing a steady supply of inputs, coordinating mass production across product lines and establishments, and marketing and selling large volumes of output. The firms that successfully met these challenges recruited and organized a hierarchy of white-collar workers such as managers, accountants, purchasing agents, and clerks. Firms and countries that failed to invest in this hierarchy did not benefit fully from new technologies and lost ground to competitors that did (Chandler, 1977, 1990).

This paper argues that Chandler's thesis remains central for understanding development today. Using census and labor force survey data from nearly one hundred countries, we document that large firms consistently employ a higher share of white-collar workers. Economic development is also associated with a shift toward white-collar employment, especially in manufacturing and low-skill services. Both findings are consistent with Chandler's historical work.

Our main new empirical finding is that cross-country differences in educational attainment account for nearly all of the gap in white-collar employment shares between developing and developed countries. Globally, the share of white-collar workers rises sharply with education—from about 10 percent of workers with no schooling to about 90 percent of those with tertiary education. Conditional on education, however, white-collar employment rates are similar across countries. Education appears to be a necessary ingredient for building the large white-collar workforce that Chandler showed was essential for scaling up firms and profitably adopting modern technologies.

These facts motivate a model of the endogenous reorganization of production from self-employment to large, white-collar-intensive firms. Following Lucas (1978), the economy contains a continuum of individuals with heterogeneous skills who choose between entrepreneurship and wage work. We enrich this framework in two ways. First, workers can perform either blue-collar production tasks or white-collar administrative tasks. Second, entrepreneurs choose not only how much to produce, but also how to organize production. Production requires a continuum of administrative tasks. For each task, the entrepreneur decides whether to do it herself or hire professionals. Hired professionals are costly but scalable, whereas the entrepreneur can perform only a fixed amount of administrative work herself.

We provide conditions under which the full task assignment problem is equivalent to a simplified one in which the entrepreneur operates a decreasing returns Cobb-Douglas production function with laborers and professionals as inputs.[1] This equivalence allows us to summarize firm organization by a single choice variable: the share of administrative tasks assigned to professionals. Assigning more tasks to professionals increases the white-collar factor share, weakens decreasing returns to scale, and increases firm size.

We impose assumptions on the skill intensity of occupations such that occupational choices can be characterized by two cutoff skill levels. Low-skill workers are indifferent between entrepreneurship and working as laborers; intermediate-skill workers become professionals; and high-skill workers become entrepreneurs. Low-skill and high-skill workers both choose to be entrepreneurs, but they operate very different types of firms. Low-skill entrepreneurs operate small firms and professionalize few tasks, while high-skill entrepreneurs run large firms and professionalize a large share of tasks. Thus, the model generates an endogenous dual economy with traditional and modern firms coexisting.

Under simplifying assumptions, the analytical model yields sharp characterizations of the forces that shape the organization of production. The main comparative static shows that an exogenous increase in the aggregate supply of skills raises the share of white-collar workers through a pure composition effect, consistent with our empirical accounting results. The underlying mechanism is that skilled workers enter both sides of the professional labor market: they supply professional labor, but they also become entrepreneurs who operate modern firms that demand professional labor. In the analytical case, the growth of modern firms acts as *skill-biased organizational change* that offsets the increase in supply and leaves the education wage premium constant. The growth in the modern sector also pulls less-skilled workers from traditional entrepreneurship into large firms as laborers.

We then develop a richer quantitative model to assess the importance of skills, barriers, and structural transformation in the reorganization of production. The model relaxes the simplifying assumptions of the analytical model and incorporates four sectors, each with its own technology and potential productivity gain from professionalizing administrative tasks. We close the model using the structural transformation preferences from Comin, Lashkari and Mestieri (2021).

We calibrate the model to fit a rich set of cross-sectional moments that build on our motivating facts about the relationships among education, occupational choice, sectoral choice, and the organization of production. Although overidentified, the model provides a good fit to these moments. We benchmark the quantitative strength of the main

[1] This equivalence builds on a similar result by Acemoglu and Restrepo (2018), but our aggregation yields endogenously decreasing returns to scale as in Akcigit, Alp and Peters (2021).

mechanism by comparing the model's predictions with evidence from plausibly exogenous expansions of schooling (Duflo, 2001; Cox, 2025). The data support the model's main mechanism of skill-biased organizational change.

We use the model as a laboratory to understand the differences in the organization of production between the average low-income economy, where traditional employment dominates, and the average middle-income economy, where roughly half of workers are employed in medium and large firms. This comparison allows us to address the question of why many developing countries have not yet started the process of reorganization documented by Chandler.

Our counterfactuals yield two main insights. First, reorganization is not a mechanical consequence of structural transformation. Giving the low-income economy the sectoral productivities and distortions of the middle-income economy induces structural transformation but no change in the employment share in medium and large firms. The shortage of skills limits the supply of white-collar labor and prevents a reorganization of production.[2]

Second, increasing skills alone generates one-half of the observed growth in employment in medium and large firms. The bottleneck to further expansion is that higher skills by themselves generate almost no structural transformation. Agriculture remains the dominant sector – but it is also the sector that benefits least from a reorganization of production into large firms staffed by white-collar workers. Matching the full shift therefore requires additional forces that reallocate activity out of agriculture and into sectors where white-collar labor is more productive.

Our paper owes an obvious debt to Chandler's work. We combine his historical, narrative work with detailed cross-country evidence to show the broad importance of skilled, white-collar workers for the reorganization of production. In doing so, we contribute to a literature that links the supply of skills to the organization of production.[3] We also relate to a recent body of work that allows firms to choose their returns to scale, or studies the consequences of that choice.[4]

We are closely related to three recent papers that combine elements of these litera-

[2]This is consistent with the evidence that management quality is lower in developing countries (Bloom and Van Reenen, 2007), that raising management quality in developing countries raises profits (Bloom et al., 2013), and that high-quality management in developing countries is expensive (Hjort, Malmberg and Schoellman, forthcoming).

[3]See Murphy, Shleifer and Vishny (1991), Garicano and Rossi-Hansberg (2006), Porzio (2017), Gomes and Kuehn (2017), Gottlieb, Poschke and Tueting (2025), and Bandiera et al. (2025) for work studying how the supply and allocation of talent across jobs shape firm organization, technology choice, occupational choice, and aggregate output. For related evidence on how the organization of labor changes over the course of development, see also Bandiera et al. (2022).

[4]See Hubmer et al. (2026), Tamkoç (2024), Argente et al. (2025), and Kopytov, Taschereau-Dumouchel and Xu (2026).

tures in the context of development. Akcigit, Alp and Peters (2021) formulate a model in which entrepreneurs choose how many tasks to delegate in response to contract enforcement, generating endogenous returns to scale through a different mechanism than ours. Amaral and Rivera-Padilla (2025) generate a dual economy through an extensive-margin technology choice, whereas in our model duality arises despite all entrepreneurs having access to a common technology. Their empirical focus is also different: they connect the model to data on technology adoption, whereas we focus on the organization of production. Finally, Cox (2025) provides novel empirical evidence on the causal effects of expanding schooling and develops a model with non-homothetic production functions in which non-agricultural sectors use college-educated workers more intensively. He focuses on the consequences of building colleges in Brazil, whereas we bring to bear cross-country data. Our model provides a microfoundation for the non-homothetic production function that also generates endogenous duality, which is key for our results.

Our model's main mechanism is related to the idea of appropriate technology adoption (Basu and Weil, 1998; Acemoglu and Zilibotti, 2001). Caselli and Coleman (2006) first showed that cross-country variation in the relative supply of skilled labor implies enormous, counterfactual movements in the skill premium when viewed through the lens of the standard Katz and Murphy (1992) imperfect substitutes labor aggregator. They propose endogenous adoption of skill-biased technologies as the resolution to this puzzle; Rossi (2022) provides evidence in favor of this mechanism. In our model, endogenous skill-biased organizational change plays the analogous role: as skilled labor becomes more abundant, firms reorganize in ways that raise the demand for skill, muting movements in education wage premia.

2 Motivating Evidence

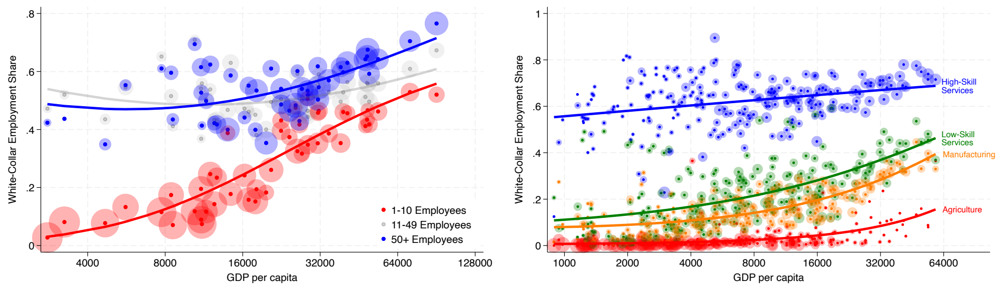
This section documents several facts that motivate our analysis. We use representative data sets drawing on nearly one hundred countries around the world to show the relevance of Chandler's insights today. We then provide new evidence on the role of skills in the reorganization of production into large firms. We summarize the data and main results here and provide additional details in Appendix A.

2.1 White-Collar Labor and Production

As discussed in the introduction, we build on two essential insights of Chandler's historical work. The first is that as firms adopted new technologies and grew, they encountered logistical challenges that required hierarchies of white-collar workers.[5] We document the systematic importance of white-collar workers for large firms using the labor force survey (LFS) database of Donovan, Lu and Schoellman (2023). Occupations are harmonized at the 1-digit level of the International Standard Classification of Occupations (ISCO). We classify codes 1–4 – managers, professionals, technicians and associate professionals, and clerks – as white-collar occupations, and codes 5–9 as blue-collar occupations. The database also groups firms into three size categories: small (1-10 employees), medium (11-49), and large (50+).

FIGURE 1: THE WHITE-COLLAR EMPLOYMENT SHARE AND DEVELOPMENT

(a) By Firm Size (b) By Sector



Notes: Each marker corresponds to a country × firm size group (left panel) or country × year × sector (right panel). The bubbles around the markers are proportional to the employment share of the firm size group within each country (left panel) or the sector within each country × year (right panel). The lines show the fits of multinomial logistic regressions on a quadratic in log GDP per capita.

Figure 1a plots the share of white-collar workers by firm size category in each country against the country's PPP GDP per capita, taken from Penn World Tables 10.0 (Feenstra, Inklaar and Timmer, 2015). Each marker in this figure corresponds to a country × firm size category, with the three firm size categories plotted using different colors. In this and subsequent figures, we scale the size of each marker in proportion to the employment share of the relevant category in the country as a whole and include the fit of a logit regression with a quadratic in log GDP per capita for reference. The central pattern is that white-collar employment shares differ sharply across firm size categories: medium and large firms systematically use more white-collar labor than small firms.

Chandler's second insight is that new technologies did not reorganize all industries equally. Manufacturing, transportation, and wholesale and retail trade were reshaped dramatically, while changes were smaller or nonexistent in other industries.[6] We use the

[5] In his words, administrative coordination "became the central function of modern business enterprise"; without it, firms were little more than "federations of autonomous offices" that "could not lower costs through increased productivity" (Chandler, 1977, pp. 7–8).

[6] Again in Chandler's words, "...modern business enterprise first appeared, grew, and continued to

census microdata from IPUMS International (Ruggles et al., 2025) to measure white-collar employment shares by sector and country. We define white-collar workers as in the labor force survey database. We use industry codes to divide workers into four broad sectors following Herrendorf and Schoellman (2018): agriculture, manufacturing, low-skill services, and high-skill services.

Figure 1b plots the share of white-collar workers by sector in each country against the country's PPP GDP per capita. Each marker in this figure captures a country $\times$ year $\times$ sector, with the four sectors plotted using different colors. Two patterns stand out. First, there are large level differences in the white-collar intensity of the sectors. High-skill services use white-collar labor intensively in all countries, whereas agriculture uses almost no white-collar labor in any country; low-skill services and manufacturing have intermediate shares of white-collar workers. Second, development is associated with a transformation of manufacturing and low-skill services (which includes transportation and wholesale and retail trade) toward more white-collar-intensive production, exactly as Chandler (1977) documented for U.S. history. Results for more detailed industries are available in Appendix A.

2.2 Skills and the Organization of Production

Our perspective on contemporary development differs from Chandler's historical account in one important respect. Chandler emphasizes that new technologies and expanding markets increased firm size, which in turn required hierarchies of white-collar workers.[7] We seek instead to understand why these technologies and large-scale production have not been adopted in developing countries today, more than a century after they were invented. Our view is that low educational attainment limits the potential white-collar workforce and slows the reorganization of firms and adoption of new technologies. Consistent with this, Appendix A.4 shows that the typical developing country has only recently acquired secondary completion rates comparable to what the United States had at the onset of the Second Industrial Revolution. More than half still have secondary completion rates lower than what the United States had at the end of this period.

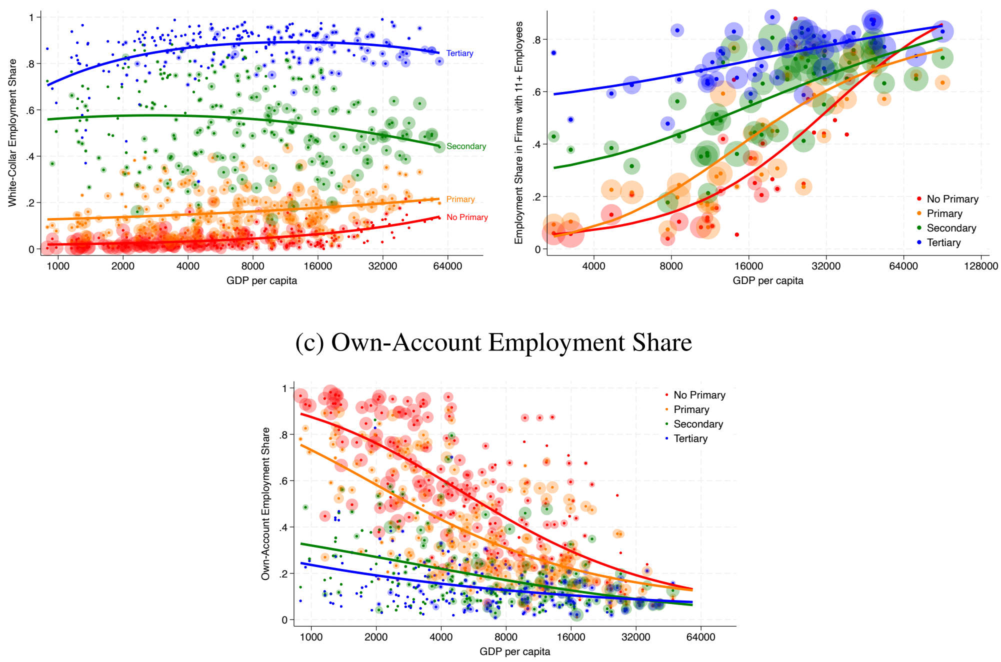
There is substantial variation in the share of white-collar workers, ranging from 10 percent of the workforce in the poorest countries to 60 percent in the richest countries

flourish in those sectors and industries characterized by new and advancing technology and expanding markets." Elsewhere, "administrative coordination was rarely more profitable than market coordination." (Chandler, 1977, p. 8).

[7] Ferraro, Iacopetta and Peretto (2024) offer a theory closer to this spirit, where growing market size induces a switch from owner-managed to professionally managed firms.

(Figure A.1). We show that this variation is almost entirely accounted for by differences in human capital. Using international census data, we measure human capital as educational attainment in four broad bins: less than primary completed, primary completed, secondary completed, and tertiary completed. Figure 2a plots the share of white-collar workers in each country $\times$ year $\times$ education cell against PPP GDP per capita, with the four education levels plotted using different colors.[8] The striking finding is that, conditional on education, white-collar employment shares are essentially uncorrelated with development. For example, 50–60 percent of secondary-educated workers are employed in white-collar occupations in both the poorest and the richest countries.

FIGURE 2: EDUCATION AND THE ORGANIZATION OF PRODUCTION



(a) White-Collar Employment Share

(b) Medium/Large Firm Employment Share

(c) Own-Account Employment Share

Notes: Each marker corresponds to a country $\times$ education ($\times$ year for panels (a) and (c)) observation. The bubbles around the markers are proportional to the employment share of the education within each country ($\times$ year for panels (a) and (c)). Medium/Large firms are those with 11+ employees. The lines show the fits of multinomial logistic regressions on a quadratic in log GDP per capita.

This fact is highly robust. Appendix A.3 shows that this pattern holds in both the time series and the cross section, when we use alternative measures of skills such as test scores, and when we exclude white-collar occupations that may be less relevant to modern businesses. Across all these alternative specifications, an accounting exercise shows that human capital accounts for 68–102 percent of the aggregate cross-country

[8] Gottlieb, Grobovšek and Monge-Naranjo (2025) also use cross-country data to document large differences in occupational choices by educational attainment.

variation in the share of white-collar workers.[9] Overall, the strength and consistency of these results motivate us to model a link between a worker's skills and their occupational choices.

This result previews an important composition effect in our model: countries with higher aggregate educational attainment have a larger pool of potential white-collar workers, which makes it more profitable to reorganize production into large firms. However, this composition effect is not the entire story. We document how the organization of production varies by educational attainment and development. Figure 2b plots the share of workers employed in medium and large firms against PPP GDP per capita, while Figure 2c plots the share of workers who are own-account self-employed against PPP GDP per capita (constructed using labor force survey data and international census data, respectively). Each marker corresponds to a country $\times$ education level, with the four education levels plotted using different colors.

These figures show that education accounts for much less of the cross-country variation in firm organization than it does of the variation in white-collar employment. For example, only about twenty percent of workers with primary education are in white-collar occupations globally. Despite this fact, development is associated with a large change in where they work: in the poorest countries, more than half are engaged in own-account self-employment; in the richest, roughly three-quarters work for medium and large firms. These patterns point to an equilibrium mechanism: as educated workers enable firms to scale, less-educated workers are drawn out of own-account work and into larger firms. We now turn to a model that formalizes this mechanism.

3 Analytical Model

These motivating facts lead us to develop a model in which skills affect development by enabling firms to reorganize production and scale up. The model features a continuum of individuals with heterogeneous skills who make occupational choices. Individuals who become entrepreneurs also choose how to organize production, which we model as a choice over how many administrative tasks to assign to professional white-collar workers.[10] This section presents a simplified, one-sector version of the model. Section 4 characterizes optimal choices and provides analytical results that build intuition for key mechanisms. We enrich the model and take it to the data in Section 5. Proofs for

[9] An implication of our findings is that large firms also use educated workers more intensively around the world, which is consistent with contemporaneous work by Gottlieb, Poschke and Tueting (2025).

[10] Because managers are part of the white-collar professional workforce, we deviate from Lucas by referring to the founder and residual claimant of the firm as the entrepreneur.

this section and the next are in Appendix B.

3.1 Environment

We model the long-run (static) equilibrium of an economy where labor is the only factor of production. The economy is populated by a unit mass of heterogeneous individuals who differ in their skill z, which is continuously distributed on a support $(0, \infty)$ according to a CDF $G(z)$. Individuals maximize their income.

The core element of our model is the entrepreneur's problem. We integrate a task assignment model in the spirit of Acemoglu and Restrepo (2018) and Akcigit, Alp and Peters (2021) into the Lucas (1978) span-of-control model. The production process uses two task inputs. First, production requires a production task that is performed by laborers (e.g., machine operators who work the assembly line). We denote by n_ℓ the efficiency units of laborers hired by the firm.

Second, production also requires a unit continuum of administrative tasks. For each task, the entrepreneur chooses whether to *professionalize* the task – that is, to hire dedicated professionals to perform it. If she professionalizes task i and hires $n_p(i)$ efficiency units of professional labor, then she receives $a(i)n_p(i)$ units of task output. The term $a(i)$ captures the relative productivity of professionalizing task i. If she does not professionalize task i, then she receives a fixed task output of 1, capturing administrative work performed by the entrepreneur in a residual or ad hoc manner rather than directing professionals. For example, Bloom et al. (2013) show that many important administrative functions such as performance monitoring, inventory control, and order sequencing are not performed in any planned or formal way in Indian manufacturing firms. Formally, task output is $\tilde{n}(i) = \max\{1, a(i)n_p(i)\}$, so professionalization raises administrative output only when done at sufficient scale.

Administrative task inputs are combined through an unweighted Cobb-Douglas aggregator. The resulting administrative input is then combined with production task input in a second Cobb-Douglas production function, with output elasticities γ_p and γ_ℓ. Finally, entrepreneurs face a wedge $\tilde{\tau}(\{n_p(i)\})$ that rises with the intensity of professional labor use. We interpret this wedge broadly as capturing legal restrictions, taxes, regulatory barriers, and other non-labor costs that make it difficult to operate large, formal firms. The proceeds are rebated lump sum to individuals.

Formally, an entrepreneur with skill z solves

$$\pi(z) = \max_{\{n_p(i)\}_{i\in[0,1]}, n_\ell} \tilde{\tau}(\{n_p(i)\}) \left\{ zA \exp\left(\int_0^1 \log \tilde{n}(i)^{\gamma_p} di \right) n_\ell^{\gamma_\ell} \right. \tag{1}$$
$$\left. - w_p \int_0^1 n_p(i) di - w_\ell n_\ell \right\}$$

s.t.

$$\tilde{n}(i) = \max\{1, a(i) n_p(i)\}$$
$$n_p(i) \geqslant 0 \quad \text{for} \quad i \in [0,1] \qquad \text{and} \qquad n_\ell \geqslant 0.$$

3.2 Equivalence Result

Without loss of generality, we order tasks in descending order by their relative productivity $a(i)$. We also assume a convenient functional form for the distortion.

ASSUMPTION 1. *The wedge $\tilde{\tau}(\{n_p(i)\})$ takes the following functional form: $\tilde{\tau}(\{n_p(i)\}) = \exp\left(\int_0^1 \log \tilde{n}(i)^{-\tau\gamma_p} di\right)$.*

Under these assumptions, Lemma 1 shows that the multi-dimensional problem (1) can be simplified to the choice of the share q of tasks to professionalize and how much professional and laborer input to hire.

LEMMA 1 (Equivalence Result)**.** *The problem of the entrepreneur (1) is equivalent to the following simplified problem, where q is the share of professionalized tasks and n_p is the professional labor input per task:*

$$\pi(z) = \max_{q\in[0,1], n_p\geqslant 0, n_\ell \geqslant 0} \left\{ z\tilde{A}(q) \left[n_p^{\alpha(q)} n_\ell^{1-\alpha(q)} \right]^{\eta(q)} - q w_p n_p - w_\ell n_\ell \right\}, \tag{2}$$

where

$$\tilde{A}(q) \equiv A \times \left(\exp \frac{1}{q} \int_0^q \log a(i)^{\gamma_p(1-\tau)} di \right)^q,$$
$$\eta(q) \equiv q\gamma_p(1-\tau) + \gamma_\ell,$$
$$\alpha(q) \equiv \frac{q\gamma_p(1-\tau)}{\eta(q)}.$$

The main implication is that the entrepreneur's problem reduces to a standard Cobb-Douglas profit maximization problem over two types of labor, with one additional choice: the entrepreneur also chooses the share of tasks to professionalize. We refer to this choice as determining the organization of the firm because it jointly determines the factor share of professionals $\alpha(q)$ and the returns to scale in production $\eta(q)$. The effect on returns to scale plays a central role in our results. Intuitively, professionalizing a task replaces a fixed entrepreneurial input with a scalable hired input, reducing the severity of decreasing returns to scale. We return to this point when we characterize the optimal choices of q for different types of entrepreneurs in Section 4.1. Let $y(z)$ denote the output implied by the solution to problem (2) for an entrepreneur with skill z.

The expressions for $\alpha(q)$ and $\eta(q)$ clarify the interpretation of Assumption 1. When $\tau = 0$, there is no tax on hiring professionals; when $\tau = 1$, the tax is prohibitive and none are hired. In the latter case the firm has fixed, decreasing returns to scale and the model collapses to the standard Lucas (1978) case. The functional form implies that for $\tau > 0$, taxes affect firm organization for all firms that choose $q > 0$. In this sense our model features a correlated distortion, as in Hopenhayn (2014). A proportional revenue tax, by contrast, distorts organization only for marginal entrepreneurs.[11]

3.3 Equilibrium

Each individual chooses among three occupations: entrepreneurship, professional work, and laborer work. A worker with skill z earns income $\pi(z)$ as an entrepreneur, $w_p z^\rho$ as a professional, and $w_\ell z^\chi$ as a laborer, where w_p and w_ℓ are the equilibrium wages per efficiency unit. The parameters ρ and χ govern how intensively skills are used in professional and laborer work.

Each worker chooses the occupation that maximizes income,

$$\phi(z) = \max \left\{ \underbrace{w_\ell z^\chi}_{\text{Laborer}}, \underbrace{w_p z^\rho}_{\text{Professional}}, \underbrace{\pi(z)}_{\text{Entrepreneur}} \right\}. \tag{3}$$

The occupational choice yields shares of workers with skill level z that choose to be entrepreneurs, professionals, and laborers $\omega_\pi(z), \omega_p(z)$, and $\omega_\ell(z)$, respectively.

We define an equilibrium in our setting, which requires that agents' occupational choices maximize their objectives and that all labor markets clear.

Definition of Competitive Equilibrium *A competitive equilibrium consists of: i. wages per efficiency unit for laborers and professionals, (w_p, w_ℓ); ii. the share of tasks to professionalize, hired labor input of professionals and laborers, and profits for each*

[11] This result follows because the tax lowers wages by the same proportion in general equilibrium.

entrepreneur z, ($q(z)$, $n_p(z)$, $n_\ell(z)$, $\pi(z)$); iii. shares of individuals in each occupation ($\omega_\pi(z)$, $\omega_p(z)$, $\omega_\ell(z)$) such that:

1. *entrepreneurs maximize firm profits by solving* (1)*;*

2. *$\omega_\pi(z)$, $\omega_p(z)$, $\omega_\ell(z)$ satisfy the occupational choice* (3)*, that is,*

$$\begin{aligned}\omega_\pi(z) > 0 \quad &\textit{only if} \quad \phi(z) = \pi(z),\\ \omega_p(z) > 0 \quad &\textit{only if} \quad \phi(z) = w_p z^\rho,\\ \omega_\ell(z) > 0 \quad &\textit{only if} \quad \phi(z) = w_\ell z^\chi;\end{aligned}$$

3. *the markets for professionals and laborers clear,*

$$\begin{aligned}\int q(z) n_p(z)\omega_\pi(z) dG(z) &= \int z^\rho \omega_p(z) dG(z),\\ \int n_\ell(z)\omega_\pi(z) dG(z) &= \int z^\chi \omega_\ell(z) dG(z).\end{aligned}$$

4 Characterization and Analytical Results

We now characterize equilibrium occupational choices and the organization of production. We then provide analytical results on how an increase in skills reorganizes production. For the remainder of the paper we restrict attention to a parametric function for $a(i)$ that yields convenient analytical solutions.

ASSUMPTION 2. *The relative productivity of professionalizing task i is a decreasing function of i: $a(i) = \beta^{1/\gamma_p}(1-i)^{\theta/\gamma_p}$.*

Intuitively, β controls the overall level of productivity of professionalizing tasks, while θ controls the dispersion of productivity of professionalizing tasks. This function implies that the log productivity gain from professionalizing a task is a decreasing, concave function of i with $\lim_{i\to 1}\log(a(i)) = -\infty$ when $\theta > 0$. Under this assumption, the endogenous productivity term becomes

$$\begin{aligned}\tilde{A}(q) &= A \times \left(\exp\left(\frac{1-\tau}{q}\int_0^q (\log\beta + \theta\log(1-i))\,di\right)\right)^q\\ &= Ae^{-q\theta(1-\tau)}\beta^{(1-\tau)q}(1-q)^{-\theta(1-\tau)(1-q)}.\end{aligned}$$

Note that $\lim_{q\to 0}\tilde{A}(q) = A$ while $\lim_{q\to 1}\tilde{A}(q) = A\beta^{1-\tau}e^{-\theta(1-\tau)}$.

4.1 Organization of Production

We start by characterizing the choices of an individual with skill z, conditional on becoming an entrepreneur. These decisions and the resulting profits are inputs to the equilibrium occupational choice, which we discuss next. An entrepreneur takes wages as given and chooses the share of administrative tasks to professionalize q and the efficiency units of laborers n_ℓ and professionals n_p to hire to maximize profits. Using the representation of Lemma 1 and the properties of the Cobb-Douglas production function, we can solve for the profits as a function of parameters, the skill z, and the (endogenous) organization of production q,

$$\tilde{\pi}(z;q) \quad = \quad (1-\eta(q))z\tilde{A}(q)\left[\tilde{n}_p(z;q)^{\alpha(q)}\tilde{n}_\ell(z;q)^{1-\alpha(q)}\right]^{\eta(q)}, \tag{4}$$

where $\tilde{n}_p(z;q)$ and $\tilde{n}_\ell(z;q)$ are the optimal labor inputs of entrepreneur z if she uses organization q. We can in turn solve for the composite labor input in the standard way to find

$$\tilde{n}_p(z;q)^{\alpha(q)}\tilde{n}_\ell(z;q)^{1-\alpha(q)} \quad = \quad \left[z\tilde{A}(q)\left(\frac{(1-\tau)\gamma_p}{w_p}\right)^{\alpha(q)}\left(\frac{\gamma_\ell}{w_\ell}\right)^{1-\alpha(q)}\right]^{\frac{1}{1-\eta(q)}}. \tag{5}$$

Equations (4) and (5) show that the expressions for labor utilization and profits are similar to their counterparts in standard span of control models (Lucas, 1978). The main novel feature is that several elements on the right-hand side of these expressions depend on the share of tasks that are professionalized, q. These include the productivity term $\tilde{A}(q)$, the factor share of professionals $\alpha(q)$, the returns to scale $\eta(q)$, and therefore the elasticity of firm size and profits with respect to entrepreneurial skill.

We use equations (4) and (5) to characterize optimal firm organization: the share of tasks professionalized and the resulting scale of production. Lemma 2 establishes that when the gains from professionalizing tasks are sufficiently heterogeneous across tasks (θ is sufficiently large), then the optimal organization of production is a smooth and well-behaved function of the entrepreneur's skill. When θ is sufficiently large, the entrepreneur's objective is quasi-concave in q.[12]

Lemma 2 (Optimal Organization of Production)**.** *Let $\theta > \frac{\gamma_p^2(1-\tau)}{1-\gamma_\ell}$. The entrepreneur's optimal organizational choice, $q(z)$, is governed by a cutoff skill level $\hat{z}_q$: for $z \leqslant \hat{z}_q$, the firm does not professionalize any tasks ($q(z) = 0$), while for $z > \hat{z}_q$, the degree of*

[12]Conversely, if θ is zero, the entrepreneur's problem instead is convex in q and has the feature that entrepreneurs either professionalize no tasks or all of them. We use this feature to help derive analytical results in Section 4.3.

professionalization $q(z)$ is a strictly increasing and differentiable function of skill that converges to full professionalization for large z, $\lim_{z\to\infty} q(z) = 1$. The value of the cutoff is given by

$$\log \hat{z}_q = (1-\gamma_\ell)\left[1 - \underbrace{\log(1-\tau)}_{\textit{Distortions}} + \underbrace{\log\frac{w_p/\gamma_p}{w_\ell/\gamma_\ell}}_{\textit{Skill premium}} - \underbrace{\frac{1}{\gamma_p}\log\beta}_{\textit{Scalability}}\right] + \underbrace{\log\frac{w_\ell/\gamma_\ell}{A}}_{\textit{Labor cost level}}, \qquad (6)$$

and the firm's optimal output as a function of the entrepreneur's skill satisfies

$$\frac{\partial \log y(z)}{\partial \log z} = \begin{cases} \frac{1}{1-\gamma_\ell} & \textit{if } z \leqslant \hat{z}_q, \\ \frac{1+\gamma_p(1-\tau)\frac{dq(z)}{d\log z}}{1-\eta(q(z))} & \textit{if } z > \hat{z}_q. \end{cases}$$

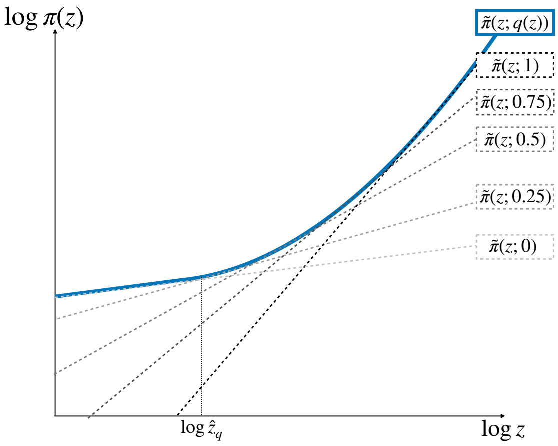
The cutoff $\hat{z}_q$ indexes the attractiveness of professionalization: a higher cutoff means that fewer entrepreneurs professionalize tasks. Equation (6) shows that the cutoff increases with the relative cost of professional labor (w_p/w_ℓ) and the distortion (τ), while it decreases with the scalability parameter (β). The final "Labor cost level" term indicates that professionalization is also less attractive when wages are high relative to productivity A, since this lowers the desired scale of production. In equilibrium, however, this force drops out because wages scale with A.

Figure 3 provides a graphical representation of the results of Lemma 2. Each gray line shows log-profits as a function of the log of the entrepreneur's skill for a given choice of q (e.g., $\tilde{\pi}(z,q)$). A higher q implies a higher elasticity of profits with respect to skill. This reflects the fact that a higher q reduces the degree of diminishing returns, disproportionately benefiting more skilled entrepreneurs. The blue line is the upper envelope of the gray curves. It represents entrepreneurial profits taking into account the optimal choice of the organization of production.

Lemma 2 and Figure 3 identify two very different types of entrepreneurs. Entrepreneurs with sufficiently low z find it optimal to choose $q = 0$ and hire only laborers. They face relatively severe decreasing returns (since $\eta(0) = \gamma_\ell$) and therefore operate small firms in equilibrium. The elasticity of output with respect to skill is $\frac{1}{1-\gamma_\ell}$, which is the familiar expression from Lucas (1978). We interpret these entrepreneurs as representing traditional production—own-account workers or small firms with little labor specialization, as in Bassi et al. (2025). We refer to them as traditional entrepreneurs.

Entrepreneurs with sufficiently high z professionalize at least some tasks. The share of tasks they professionalize rises with their own skill, implying that the white-collar employment share also increases with the entrepreneur's skill. The elasticity of out-

FIGURE 3: ORGANIZATION OF PRODUCTION AND FIRM PROFITS



put with respect to the entrepreneur's skill is larger than the standard $\frac{1}{1-\gamma_\ell}$. It is also increasing in q, which is consistent with recent evidence from Queiró (2022) that the thickness of the firm size distribution tail is increasing in the entrepreneur's education level. We refer to firms that professionalize administrative tasks as modern firms, and to their owners as modern entrepreneurs.[13]

Thus, individuals with different levels of skill z operate very different types of firms if they become entrepreneurs. We now solve for occupational choices, which inform us about who chooses entrepreneurship in equilibrium.

4.2 Occupational Choice

We now characterize who becomes a laborer, professional, traditional entrepreneur, or modern entrepreneur in equilibrium. The occupational choices depend on the equilibrium returns to skills in the various occupations. In the previous section, we characterized equilibrium profits as a function of skills for traditional and modern entrepreneurs and showed that the elasticity of profits with respect to skill is higher for modern entrepreneurs. Assumption 3 completes the ordering of the elasticity of income with respect to skill across all four occupations.

ASSUMPTION 3. *The parameters χ and ρ satisfy*

$$\underbrace{\chi}_{\textit{Laborers}} = \underbrace{\frac{1}{1-\gamma_\ell}}_{\textit{Traditional Entrepreneurs}} < \underbrace{\rho}_{\textit{Professionals}} < \underbrace{\frac{1}{1-\gamma_p(1-\tau)-\gamma_\ell}}_{\textit{Modern Entrepreneurs}}.$$

[13] Banerjee and Newman (1993) also develop a model of the endogenous allocation of workers to different types of firms, although the underlying mechanism is different from ours. Empirically, we distinguish modern and traditional entrepreneurs based on their occupations; see Section 5.2 for details.

Note that the elasticity for modern entrepreneurs applies for a hypothetical entrepreneur who professionalizes all administrative tasks.

The ordering in Assumption 3 implies that low-skill workers have a comparative advantage as laborers or traditional entrepreneurs, while high-skill workers have a comparative advantage as professionals or modern entrepreneurs. This comparative advantage drives occupational sorting, as described in the following lemma.

LEMMA 3 (Occupational Choice)**.** *Given Assumption 3, the equilibrium satisfies the following properties:*

1. *there exist cutoffs $\hat{z}_0 \leqslant \hat{z}_1 < \hat{z}_2$ such that individuals with $z \leqslant \hat{z}_0$ are laborers or traditional entrepreneurs, those with $z \in (\hat{z}_1, \hat{z}_2)$ are professionals, while those with $z \in [\hat{z}_0, \hat{z}_1]$ or $z \geqslant \hat{z}_2$ are modern entrepreneurs;*

2. *the equilibrium incomes satisfy*

   - $w_\ell z^\chi \geqslant \tilde{\pi}(z, 0)$ *with equality on the support of traditional entrepreneurs, i.e., for those $z \leqslant \hat{z}_0$ with $\omega_\pi(z) > 0$.*
   - *if $\hat{z}_0 = \hat{z}_1 : w_\ell \hat{z}_0^\chi = w_p \hat{z}_1^\rho$, $w_p \hat{z}_2^\rho = \pi(\hat{z}_2)$;*
   - *if $\hat{z}_0 < \hat{z}_1 : w_\ell \hat{z}_0^\chi = \pi(\hat{z}_0)$, $\pi(\hat{z}_1) = w_p \hat{z}_1^\rho$, $w_p \hat{z}_2^\rho = \pi(\hat{z}_2)$;*

3. *there are traditional entrepreneurs – i.e. $\int_{\underline{z}}^{\hat{z}_0} \omega_\pi(z) dG(z) > 0$ – if and only if*

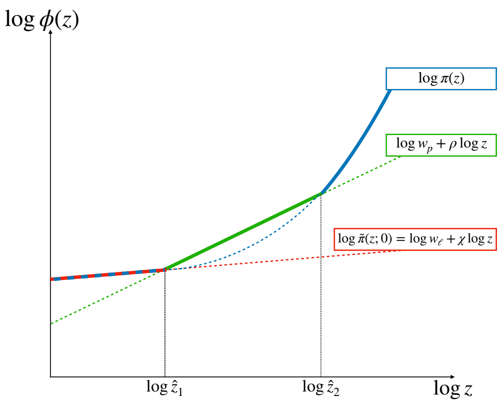
$$\int_{\underline{z}}^{\hat{z}_0} z^\chi dG(z) > \int_{\hat{z}_0}^{\infty} n_\ell(z) \omega_\pi(z) dG(z).$$

Lemma 3 shows that the equilibrium can take one of two possible structures. Figure 4 shows the simpler case where $\hat{z}_0 = \hat{z}_1$, which implies that all modern entrepreneurs are more skilled than all professionals. Appendix Figure B.1 shows the alternative case with $\hat{z}_0 < \hat{z}_1$. Figure 4a shows potential income in each occupation, given equilibrium prices, as a function of skill z. The red line is the wage for laborers, which is identical to the profit of traditional entrepreneurs (in an equilibrium with some traditional entrepreneurs). The green line is the wage of professionals, which is increasing in z, with elasticity modulated by ρ. Finally, the blue line shows the profit of entrepreneurs (both traditional and modern), which takes into account the optimal organization q.

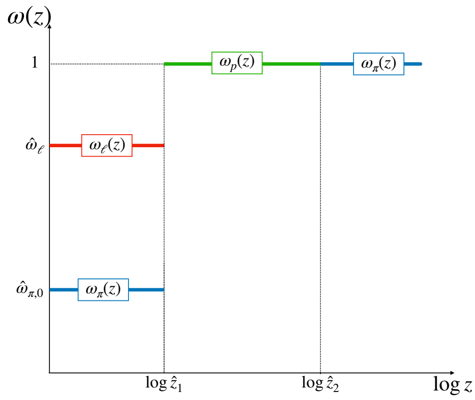
Figure 4b shows the resulting occupational choice. Workers with low skill levels are indifferent between becoming traditional entrepreneurs and working as laborers. Workers with intermediate skill levels earn the most as professionals and consequently choose that occupation. Workers with the highest skill levels choose the most skill-intensive occupation, which is modern entrepreneurship.

FIGURE 4: OCCUPATIONAL CHOICE AND ENDOGENOUS DUALITY

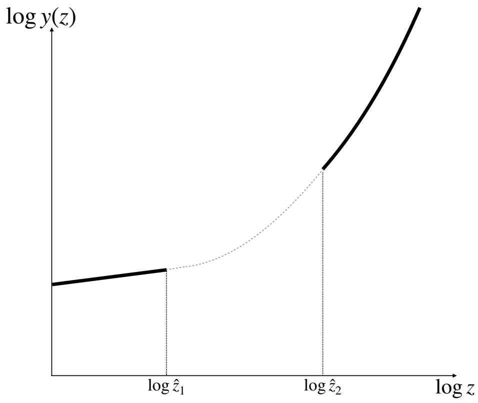
(a) Income



(b) Occupational Choice



(c) Firms' Output in Equilibrium



Finally, Figure 4c illustrates a key implication of Lemma 3: equilibrium in this model can feature duality. If modern entrepreneurs are sufficiently numerous, they absorb all low-skill workers as laborers. If not, the surplus low-skilled workers turn to traditional entrepreneurship. In this equilibrium, both the most- and least-skilled individuals become entrepreneurs – but they operate distinct types of firms with very different productivity levels.

DEFINITION 1. *An equilibrium features* duality *if there is a positive mass of traditional entrepreneurs, that is, if* $\int_{\underline{z}}^{\hat{z}_0} \omega_\pi(x) dG(x) > 0$.

Our empirical results in Section 2 focus on the share of workers in blue-collar versus white-collar occupations. When taking the model to the data, we treat laborers and traditional entrepreneurs as blue-collar workers, and professionals and modern entrepreneurs as white-collar workers.

4.3 Analytical Results: Increase in the Supply of Skills

In this section, we explore how an increase in the aggregate supply of skills can generate a reorganization of production. We introduce a simplified version of the model that can be solved analytically. Proposition 1 builds intuition for the quantitative results in Section 5 and shows that the model is consistent with the motivating facts documented in Section 2.

To obtain closed-form comparative statics, we collapse the professionalization choice to a binary decision and align the skill intensity of professionals and modern entrepreneurs.

ASSUMPTION 4. *The productivity of professionals does not vary across tasks: $\theta = 0$.*

ASSUMPTION 5. *The income of professionals is as skill-sensitive as the profits of a modern entrepreneur, $\rho = \frac{1}{1-\gamma_\ell-\gamma_p(1-\tau)}$.*

ASSUMPTION 6. *The value of $\beta^{1-\tau}(1-\gamma_\ell-(1-\tau)\gamma_p)^{1-\gamma_\ell-(1-\tau)\gamma_p}\left[(1-\tau)\gamma_p\right]^{(1-\tau)\gamma_p}$ is decreasing in τ.*

Assumption 4 implies that the entrepreneurial profit function is convex in q, and thus there are only two types of entrepreneurs: traditional entrepreneurs who choose $q = 0$ and modern entrepreneurs who choose $q = 1$. Assumption 5 implies that high-skilled workers are indifferent between modern entrepreneurship and working as a professional. Finally, Assumption 6 ensures that modern firms become less profitable when the wedge τ increases. This requires β to be sufficiently large, because a higher wedge also raises the profit share $1 - \gamma_\ell - (1-\tau)\gamma_p$.

PROPOSITION 1 (Impact of an Increased Skill Supply)**.** *Start from an equilibrium featuring duality. Consider a uniform increase in the supply of skills that shifts up each individual by a factor of $\kappa > 1$, implying a new distribution function, $\tilde{G}(z) = G(z/\kappa)$.*

There exists a threshold $\hat{\kappa}$ such that for any $\kappa < \hat{\kappa}$, an increase in κ yields:

1. ***Constant wages and cutoff rules.*** *Wages per efficiency unit for both types of workers and the occupational cutoff $\hat{z}$ remain constant. Consequently, there is no change in the set of optimal occupational choices conditional on skill level z.*

2. ***Reorganization of production.*** *The share of white-collar workers and average firm size rise, while the share of traditional entrepreneurs falls.*

The proof of Proposition 1 is provided in the Appendix. The key step is to recognize that, as long as duality persists, wages (per efficiency unit of labor) remain unchanged in response to shifts in the supply of skills. To build intuition for this result, we explain each wage in turn.

An increase in the supply of skills increases the share of individuals who choose to work as professionals or become modern entrepreneurs. The increase in the share of professionals tends to push the wage of professionals down. However, the increase in the share of modern entrepreneurs raises the demand for professionals to staff the growing number of large firms. This *skill-biased organizational change* pushes the wage for professionals up. In general, the net effect on professional wages is ambiguous. In this special case, the fact that skilled workers are indifferent between working as professionals and becoming modern entrepreneurs ensures that the supply and demand forces exactly offset, leaving professional wages unchanged.

A different mechanism keeps the wages of laborers constant. The increase in skills both reduces the supply of laborers and increases the demand for laborers (to work in the larger number of modern firms). However, low-skilled workers are indifferent between working as laborers and becoming traditional entrepreneurs, and we begin in an equilibrium with duality. The traditional entrepreneurs act as a reserve supply of laborers who step in to meet the increased demand. As a result, laborer wages remain fixed until this reserve is exhausted.

The property that wages are invariant to the supply of skills relies on simplifying assumptions specific to this analytical model. However, two key mechanisms carry over to the general framework: first, skilled workers contribute to both the supply of and demand for professional labor; and second, traditional entrepreneurs act as a reserve labor supply for laborers. These features help explain why the model generally produces muted movements in relative wages. They are important when taking the model to the data, as available evidence shows that developing and developed countries exhibit broadly similar relative wages despite vastly different supplies of skilled workers (Caselli and Coleman, 2006; Rossi, 2022).

With wages invariant to the supply of skills, the remaining results arise through composition effects. The increase in the aggregate supply of skills leads to more modern entrepreneurs and an increase in the size of firms. The growth in the modern sector pulls workers from traditional entrepreneurship into laborer jobs in large, modern firms.

These properties are consistent with the motivating facts outlined in Section 2. The fact that occupational choices depend only on the worker's skill level z and not the aggregate skill level is consistent with the findings about occupational choice by education level shown in Figure 2a. The resulting growth in large, white-collar-intensive firms is consistent with Figure 1a. The fact that equilibrium reorganization pulls blue-collar workers out of traditional entrepreneurship and into large firms is consistent with

Figures 2b and 2c.

Increases in skills also raise aggregate output. We show in Appendix B.3 that if $\tau = 0$, the economy is efficient. Then a standard envelope argument implies that, to first order, the output effect from any change in the skill distribution is purely compositional: $dY = \int \phi(z)\, dg(z)$, where $\phi(z)$ is the equilibrium payoff of a worker with skill z and $dg(z)$ denotes the perturbation to the skill distribution induced by the change in κ. If the initial economy features a firm size distortion $\tau > 0$, the output gains are generally larger since the modern sector is inefficiently small (see discussion in Appendix B.3). The same logic applies if externalities, increasing returns, or other amplification mechanisms make modern production inefficiently low. We focus on understanding the reorganization of production and leave these amplification mechanisms to future work.

We conclude the analysis of changes in the supply of skills by considering what happens when an increase in skill supply is large enough to fully exhaust the reserve of traditional entrepreneurs. In this case, the equilibrium no longer features duality and the economy's behavior is very different, as shown in the following corollary.

COROLLARY 1. *An increase in skills to $\kappa = \hat{\kappa}$ implies an end to duality: no entrepreneur chooses $q = 0$. Any increase in skills beyond $\hat{\kappa}$ yields the following:*

1. *the cutoff type $\hat{z}(\kappa)$ satisfies $\log \hat{z}(\kappa) = \log(\kappa/\hat{\kappa}) + \log \hat{z}$, where $\log \hat{z}$ is the cutoff in the economy with duality;*

2. *the share of white-collar workers is constant;*

3. *average firm size is constant;*

4. *the relative wage of professionals to laborers declines;*

5. *the probability of choosing white-collar work conditional on skill level z declines.*

It is worth highlighting the second and third parts of the corollary: both firm size and the labor force composition in terms of white- and blue-collar workers remain constant. Thus, an increase in skills only leads to a reorganization of production if the initial equilibrium features duality. A straightforward but important implication is that an increase in skills does not generate a reorganization of production in the Lucas (1978) model.

Summing up, Proposition 1 shows that an increase in skills can generate a reorganization of production consistent with the data. At the same time, Lemma 2 shows that skills are not the only force shaping firm organization: distortions and relative wages also matter, even in this simple economy. Our next goal is to study the quantitative importance of competing driving forces in a richer calibrated model that relaxes the simplifying assumptions used to obtain these propositions.

5 Quantitative Model

This section develops the quantitative model used for the counterfactuals in Section 6. We extend the model to make it suitable for quantitative analysis; calibrate it to match the organization of production in middle-income countries; construct a low-income counterpart; and benchmark its mechanisms against causal evidence on educational expansions.

5.1 Extensions and Mapping to the Data

The quantitative model enriches the analytical model in three dimensions. First, we extend the model to include multiple sectors, so that it can speak to the relationship between structural transformation and the reorganization of production. Second, we describe the mapping from observed schooling to model skills. Third, we allow for preference shocks that smooth the relationship between comparative advantage and occupational choice, enabling the model to replicate the empirical occupational choice patterns.

Education and Skills. We use the four education groups observed consistently in the cross-country data – less than primary complete, primary complete, secondary complete, and tertiary complete – to proxy for unobserved skills. Within education group i, skills z are lognormally distributed with mean $z_{\mu,i}$ and standard deviation z_σ. We normalize $z_{\mu,\text{No Primary}} = 0$.

Sectors. We extend the analytical model to include the four sectors discussed in Section 2: agriculture, manufacturing, low-skill services, and high-skill services. Each sector j produces a differentiated good that trades at price p_j. Sectors differ in their technologies, which include both physical productivity, A_j, and the curvature parameters in the two types of labor, $\gamma_{p,j}$ and $\gamma_{\ell,j}$. These parameters allow sectors to differ in average firm size, through $\gamma_{\ell,j} + \gamma_{p,j}$, and in the gains from reorganizing production around white-collar workers, through $\gamma_{p,j}$.

Sectors also differ in the intensity with which they use skills. We impose $\chi_j = (1 - \gamma_{\ell,j})^{-1}$ so that traditional entrepreneurship and working as a laborer are equally skill-intensive within each sector, consistent with the case that we analyzed in the analytical model. To further reduce dimensionality, we set the skill intensity of professionals equal to the midpoint between the skill intensity of laborers/traditional entrepreneurs and that of fully professionalized modern entrepreneurs,

$$\rho_j \;=\; \frac{1}{2}\left(\frac{1}{1-\gamma_{\ell,j}} \;+\; \frac{1}{1-\gamma_{p,j}(1-\tau)-\gamma_{\ell,j}}\right).$$

Recall that with $\rho_j < \frac{1}{1-\gamma_{p,j}(1-\tau)-\gamma_{\ell,j}}$, workers with intermediate skill levels have a comparative advantage in working as professionals whereas workers with the highest skill levels have a comparative advantage in becoming modern entrepreneurs, unlike the special case we used in Section 4.3. Finally, the sectors are affected equally by the distortion τ.

Preference Shocks. In our analytical model, workers maximize income, leading to sharp occupational choice cutoffs (see Figure 4b). In the quantitative model, we allow workers to have idiosyncratic taste shocks over both occupations and sectors. Conditional on choosing sector j, workers receive idiosyncratic preference shocks for the three occupations that are i.i.d. draws from a type-I extreme value distribution with shape parameter ξ. Hence, the share of workers in sector j choosing to be laborers is

$$\omega_{\ell,j}(z) \;=\; \frac{\left(w_{\ell,j} z^{\chi_j}\right)^{\xi}}{\left(w_{\ell,j} z^{\chi_j}\right)^{\xi} + \left(w_{p,j} z^{\rho_j}\right)^{\xi} + \left(\pi_j(z)\right)^{\xi}}. \tag{7}$$

Similar expressions give the share who choose to be professionals and entrepreneurs. Because preference shocks smooth choices, indifference no longer defines traditional entrepreneurship. We therefore classify entrepreneurs as traditional if they would earn more as laborers than as professionals.

Workers anticipate these occupational preference draws when choosing sectors. Hence, up to sector-wide factors that we discuss below, the maximum expected income a worker derives from choosing sector j is

$$\phi_j(z) \;\propto\; \left(\left(w_{\ell,j} z^{\chi_j}\right)^{\xi} + \left(w_{p,j} z^{\rho_j}\right)^{\xi} + \left(\pi_j(z)\right)^{\xi}\right)^{\frac{1}{\xi}}. \tag{8}$$

Workers also receive idiosyncratic preference shocks for the four sectors that are i.i.d. draws from a type-I extreme value distribution with shape parameter ν. Similarly, these shocks imply that the share of workers with skill z who choose sector j depends on their expected earnings in j relative to the other sectors. However, a substantial literature documents a divergence between sectoral employment shares and relative sectoral wages. For example, the large literature on the agricultural productivity gap shows that most workers in developing countries work in agriculture despite agriculture offering lower wages than other sectors (Gollin, Lagakos and Waugh, 2014; Herrendorf and Schoellman, 2018). We follow this literature by introducing a sectoral wedge: a worker with skill z in sector j earns $\delta_j z^{\varphi_j}$ times their paid income. The term δ_j captures the distortion that is common to all workers, while z^{φ_j} allows for a correlated distortion that affects skilled workers differentially. As is standard, this wedge stands in for factors such as geography, information frictions, or labor market distortions that reduce the effective income from entering or switching to a sector. We impose that $\varphi_{\text{mfg}} = \varphi_{\text{ls}} = 0$ and $-\varphi_{\text{agr}} = \varphi_{\text{hs}} = \varphi > 0$.

The share of workers with skill z choosing sector j is then given by

$$\sigma_j(z) = \frac{\left(\delta_j z^{\varphi_j} \phi_j(z)\right)^\nu}{\sum_{k\in J}\left(\delta_k z^{\varphi_k}\phi_k(z)\right)^\nu}. \tag{9}$$

Markets. We close the model with a representative consumer who has average worker income and chooses sectoral consumption to maximize a non-homothetic CES utility function as in Comin, Lashkari and Mestieri (2021). Utility U is implicitly defined by

$$\sum_{j\in\{\text{ag,mfg,hs,ls}\}} \Upsilon_j^{\frac{1}{\sigma}} \left(\frac{C_j}{U^{\epsilon_j}}\right)^{\frac{\sigma-1}{\sigma}} = 1. \tag{10}$$

This approach captures consumption responses to changing prices and rising average income without requiring us to solve for expenditures across the full distribution of workers.

5.2 Targeted Moments, Identification, and Model Fit

We calibrate the model to the average middle-income country, defined as countries with GDP per capita between 10 and 50 percent of the U.S. level. This step disciplines the model's occupational, sectoral, and firm-size margins before we use it for counterfactuals. We then recalibrate a limited set of driving forces to construct the average low-income economy, defined as countries with GDP per capita below 10 percent of the U.S. level. Section 5.4 benchmarks the model by comparing its predictions with well-identified causal evidence, and Section 6 uses it to decompose the transition from low- to middle-income economies.

Externally Calibrated Parameters. We begin with parameters fixed outside the model, either directly from data or from existing estimates. Table 1 summarizes these parame-

TABLE 1: EXTERNALLY CALIBRATED PARAMETERS

Panel A. Education-specific		**<Primary**	**Primary**	**Secondary**	**Tertiary**	**Source**
υ	Share	0.15	0.39	0.33	0.13	Ruggles et al. (2025)
Panel B. Aggregate						
ξ	Occup. preferences	8.00				Dix-Carneiro (2014); Ashournia (2018)
ν	Sectoral preferences	8.00				Dix-Carneiro (2014); Ashournia (2018)
σ	Aggregate elasticity	0.43				Comin, Lashkari and Mestieri (2021)
$\eta_{\text{mfg}}(1)$	Returns to scale in Mfg	0.80				Buera, Kaboski and Shin (2011)
τ	Firm size wedge	0.00				Normalization
Panel C. Sector-specific		**Agr**	**Mfg**	**Ser (HS)**	**Ser (LS)**	
ϵ	Sector elasticity	0.25	1.00	1.88	1.12	Comin, Lashkari and Mestieri (2021)
$\frac{\gamma_{p,j}}{\eta_j(1)}$	Factor share of prof.	0.27	0.55	0.90	0.66	Own calculation

ters.

Panel A reports the share of workers in each education group, υ_i, in the average middle-income country, which we compute from international census data (Ruggles et al., 2025). Panel B reports five aggregate parameters taken from the literature. The parameters ξ and ν govern the dispersion of taste shocks across occupations and sectors. We discipline these parameters using estimates of long-run labor reallocation from the trade literature. Our central estimate is $\xi = \nu = 8$, consistent with Dix-Carneiro (2014) and Ashournia (2018); we show results for values in the range of 4–16 in Appendix E.[14] We estimate the elasticity of substitution in demand across sectoral outputs to be 0.43 following Comin, Lashkari and Mestieri (2021), using their replication files and code but splitting services into low-skill and high-skill services. We fix the returns to scale of a fully modernized firm ($q = 1$) in manufacturing at 0.80, a value broadly consistent with the literature—for example, Buera, Kaboski and Shin (2011) use 0.79. Finally, we normalize $\tau = 0$ for the middle-income country.[15]

Panel C reports two sector-specific parameters that we calibrate externally. First, we estimate sectoral income elasticities of demand following Comin, Lashkari and Mestieri (2021). Second, the ratio $\gamma_{p,j}/(\gamma_{p,j} + \gamma_{\ell,j})$ determines the compensation share of professionals in fully modernized firms. We calibrate this to match the observed compensation share of professionals in large firms, which we calculate using labor force sur-

[14] Both papers provide estimates of the long-run sectoral elasticity. Revenga (1992) estimates that the five-year sectoral elasticity is 4. Artuç and McLaren (2015) estimate the long-run elasticity across sectors and occupations, finding similar magnitudes and values as large as 20.

[15] Inspecting the entrepreneur's problem in equation (2) shows that only the product $\gamma_p(1-\tau)$ appears, so this normalization does not affect our ability to fit the data. Following the discussion of Proposition 1, the main consequence is that normalizing $\tau = 0$ implies that the middle-income economy is undistorted and so minimizes the output change that results from expanding schooling.

vey data for middle-income countries. Below, we describe how we internally calibrate $\gamma_{p,j} + \gamma_{\ell,j}$; with this value in hand, we can recover the underlying structural parameters $\gamma_{p,j}$ and $\gamma_{\ell,j}$.

Internally Calibrated Parameters. We next turn to the vector of internally calibrated parameters. A useful property of the model is that sectoral choices depend on $\delta_j p_j A_j$, but not on its individual components. Equation (9) implies this because wages, profits, and hence ϕ_j are all homogeneous of degree one in $p_j A_j$. This property implies that it is sufficient to calibrate $\delta_j p_j A_j$ for three sectors in order to match the allocation of employment across sectors, normalizing $p_{\text{mfg}} A_{\text{mfg}} \delta_{\text{mfg}} \equiv 1$. We then back out $p_j A_j$ to be consistent with relative nominal value added by sector from the Productivity Level Database (Inklaar, Marapin and Gräler, 2023). Since sectoral outputs do not have meaningful scales, we normalize $p_j = 1$ for all sectors, which pins down A_j.

We have the following 13 parameters to estimate:

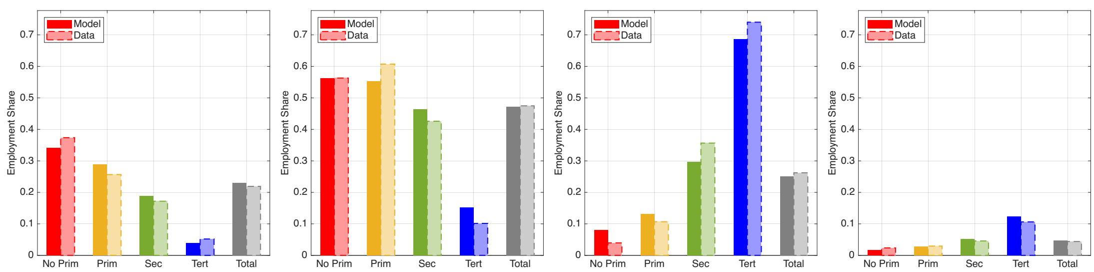
$$\mathbf{p} \quad = \quad \Big[\, z_{\mu,2}\ ,\ z_{\mu,3}\ ,\ z_{\mu,4}\ ,\ z_\sigma\ ,\ \beta\ ,\ \theta\ ,\ \varphi\ ,\ \{\eta_j\}_{j\in\{\text{ag,hs,ls}\}}\ ,\ \{\delta_j A_j p_j\}_{j\in\{\text{ag,hs,ls}\}} \Big].$$

We select the parameter vector to minimize the weighted sum of squared deviations between moments in the model and moments in the data.

Targeted Moments and Model Fit. We summarize the targeted moments and model fit graphically, leaving exact values and weights to Appendix C.1. Broadly, we want the model to replicate the motivating empirical patterns described in Section 2, which cover the distribution of labor across occupations, sectors, and firm sizes, as well as how these distributions vary with educational attainment. We now describe more specifically the moments we target.

FIGURE 5: OCCUPATIONAL SHARES

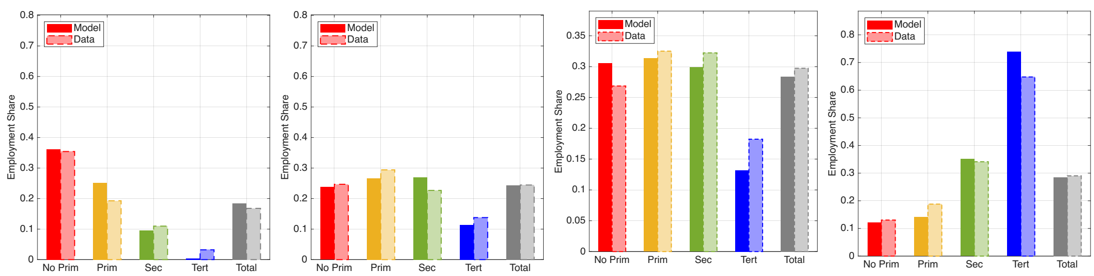
(a) Traditional Entr. (b) Blue-Collar Workers (c) White-Collar Workers (d) Modern Entr.



We first target aggregate occupational choices and occupational choices by education. The model allows for four occupational choices: traditional entrepreneurs, laborers, professionals, and modern entrepreneurs. We map employees in the data to wage workers in the model and self-employed workers to entrepreneurs. We use occupation codes to subdivide each category: workers with blue-collar occupations are laborers, while workers with white-collar occupations are professionals. The self-employed with blue-collar occupations are traditional entrepreneurs, while the self-employed with white-collar occupations are modern entrepreneurs.[16] Figure 5 shows these moments in the data and in the model. We also target the distribution of labor across sectors at the aggregate and within each education level (Figure 6). Finally, we target the distribution of educational shares within sectors. Although these moments are renormalized versions of those in Figure 6, the model is highly overidentified, and including both helps ensure a good fit along the education-sector dimension.

FIGURE 6: SECTORAL SHARES

(a) Agriculture (b) Manufacturing (c) Services (LS) (d) Services (HS)



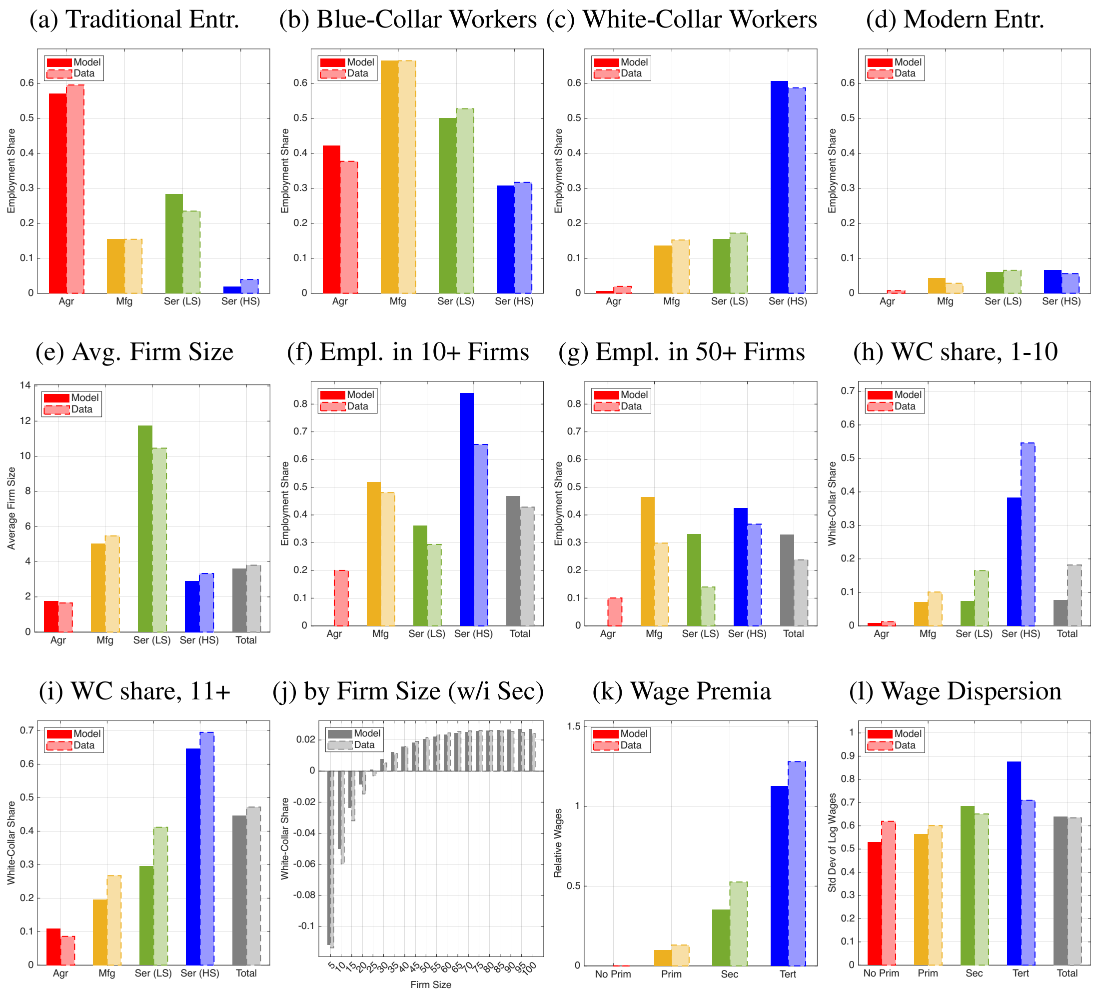
Importantly, we want the model to capture sectoral differences in the organization of production. For example, high-skill services are characterized by many large firms, while roughly half of individuals in agriculture are traditional entrepreneurs. Accordingly, we target occupational shares, average firm size, and the shares of employment in firms with more than 10 and more than 50 employees for each sector (Figure 7).

One of our key motivating facts, consistent with Chandler's narrative, is that larger firms employ a higher share of white-collar workers. We therefore target the share of white-collar employment within each sector for small versus medium/large firms, as well as the relationship between firm size (number of employees) and the relative white-collar share, estimated from the subset of LFS countries with detailed firm-size categories using a cubic in firm size and sector fixed effects (Panels (h)–(j) of Figure 7).

Finally, we target wage outcomes. Specifically, we match the residual wage gaps across education groups and the residual wage dispersion within education groups (Pan-

[16] We do not use data on firm size because information on whether the self-employed are employers and on the number of employees in their firms is less widely available. We include unpaid family workers with the self-employed. Because most have blue-collar occupations, they are then largely counted as traditional entrepreneurs.

FIGURE 7: DIFFERENCES IN WITHIN-SECTOR ORGANIZATION AND WAGES



els (k)–(l) of Figure 7). We residualize wages for the estimated effect of potential experience, gender, and location (*geolevel1*) fixed effects, removing variation associated with factors absent from our model.[17]

In total, we target 125 moments. Although some are collinear by construction (e.g., employment shares sum to one), the model remains overidentified. Nonetheless, it fits the empirical patterns well.

Identification. All parameters are jointly identified. Appendix C.2 provides simulation results showing that the parameters are locally and globally well-identified and documents which moments are most informative for each parameter. Here, we provide an overview of the identification that builds on these results.

[17]These moments are computed using 11 IPUMS cross-sections and 22 LFS countries with available wage information. See Appendices A and C.1 for details on the construction and aggregation of the country-specific estimates.

The parameters z_μ and z_σ govern the distribution of unobserved skill z within and between education groups. In the model, more skilled workers are more likely to choose skill-intensive occupations and also have higher earnings. The parameters z_μ and z_σ can thus be pinned down by heterogeneity in occupational choices and earnings within and between school groups. For example, the model requires a sizable gap $\mu_4 - \mu_2$ relative to σ_z to be consistent with the fact that most college-educated individuals choose white-collar occupations but few primary-educated individuals do so.

The intercept β governs the average productivity of professionalizing tasks. A higher β raises the return to professionalization, increasing the share of white-collar workers and modern entrepreneurs at the expense of blue-collar workers and traditional entrepreneurs.

While β governs the average productivity of professionalizing tasks, the parameter θ modulates heterogeneity in the productivity of professionalizing tasks and thus heterogeneity in the organization of production. As discussed in Section 4.3, a low value of θ implies that entrepreneurs choose to professionalize either no tasks or all of them, which in turn makes the profit schedule highly sensitive to the entrepreneur's skill. Conversely, when θ is high, some tasks remain costly to professionalize even for highly skilled entrepreneurs and profits are less sensitive to the entrepreneur's skill. The extent to which skilled individuals sort into modern entrepreneurship is therefore informative about θ.

The sector-specific returns to scale of a fully modernized firm, $\gamma_{p,j} + \gamma_{\ell,j}$, are disciplined by firm size distributions: all else equal, sectors with higher $\gamma_{p,j} + \gamma_{\ell,j}$ have larger firms in equilibrium.

The correlated sectoral wedge, φ, helps match the strong observed sorting of skilled workers into high-skill services and away from agriculture. While cross-sector wage profiles already generate some sorting—since high-skill services are more skill-intensive than agriculture—matching the magnitude observed in the data requires a positive value of φ.

Finally, the combination of parameters $\delta_j p_j A_j$ are identified from sectoral employment shares: larger values lead to a greater fraction of workers in sector j.

Parameter Estimates. Table 2 summarizes our internally calibrated parameter values. As expected, more educated workers are on average more skilled (Panel A). For instance, $z_{\mu,\text{Tertiary}} = 1.21$ implies that tertiary-educated workers are more than three times as productive as workers with less than primary completed.

Panel B summarizes our estimates of the economy-wide parameters. The standard deviation of skills within education groups is roughly equal to the mean difference

TABLE 2: INTERNALLY CALIBRATED PARAMETERS

Panel A. Education-specific		**<Primary**	**Primary**	**Secondary**	**Tertiary**
z_μ	Average skills relative to no primary		0.14	0.50	1.21
Panel B. Aggregate					
z_σ	St.d. of skills cond. on education	0.51			
β	Productivity of hired professionals, intercept	1.47			
θ	Productivity of hired professionals, slope	1.11			
Panel C. Sector-specific		**Agr**	**Mfg**	**Ser (HS)**	**Ser (LS)**
φ	Non-pecuniary skill-dependent return	-1.53	0.00	1.53	0.00
χ	Skill-sensitivity of laborers	1.86	1.56	1.08	1.37
ρ	Skill-sensitivity of professionals	2.29	3.28	2.38	3.10
γ_ℓ	Curvature in laborers	0.46	0.36	0.07	0.27
γ_p	Curvature in professionals	0.17	0.44	0.66	0.53
A	Sectoral TFPQ	1.52	1.00	0.37	1.13
δ	Sectoral wedge	0.59	1.00	1.00	0.91

in skills between secondary and less than primary completed. That is, we estimate a non-trivial overlap in the distribution of skills across education groups. Combined with our estimates of $\gamma_{p,j}$ and $\gamma_{l,j}$, our estimate $\theta = 1.11$ implies that the restriction $\theta > \gamma_{p,j}^2(1-\tau)/(1-\gamma_{\ell,j})$ holds in all sectors.

Panel C shows the parameters that vary by sector. The non-pecuniary skill-dependent return to going into high-skill services relative to agriculture (φ) is large, which is required to drive the strong sorting of skilled workers away from agriculture and toward high-skill services given the relatively modest differences in returns to skill across sectors. Our assumption that $\chi_j = (1-\gamma_{\ell,j})^{-1}$ implies the highest skill-sensitivity of laborers in agriculture, while the assumption that ρ_j is the midpoint of the skill-sensitivity of traditional and modern entrepreneurship implies the highest skill-sensitivity of professionals in manufacturing.

The curvature in blue-collar labor $\gamma_{\ell,j}$ ranges between 0.27–0.46 across low-skill services, manufacturing and agriculture, but is much lower in high-skill services. The curvature in professional labor, $\gamma_{p,j}$, ranges from 0.17 in agriculture to 0.66 in high-skill services. Since prices are normalized $p_j \equiv 1$, differences in physical productivity A_j correspond to revenue productivity. Agriculture has the highest revenue productivity and high-skill services the lowest. Matching the relatively modest employment share in agriculture therefore requires a large wedge in agriculture relative to manufacturing, $\delta_{\text{agr}} = 0.59$, while the other sectors have much smaller wedges.

5.3 Cross-Country Calibration

We next construct a low-income counterpart to the benchmark middle-income economy. This exercise is useful for two purposes. First, it provides a disciplined way to compare the model with evidence from economies at different stages of development. Second, it allows us to decompose the transition from low-income to middle-income economies into changes in skills, productivity, and wedges.

We augment the model with one additional margin to study low-income economies. We introduce a land wedge, λ, that lowers the effective curvature of agricultural production with respect to laborers from $\gamma_{\ell,agr}$ to $\gamma_{\ell,agr} - \lambda$ in the low-income economy. We normalize $\lambda = 0$ in the middle-income economy. Appendix B.4 shows that missing land rental markets provide one microfoundation for sharper decreasing returns in agriculture. More generally, λ captures the well-documented barriers to the formation of large farms in developing countries (Ayerst, Brandt and Restuccia, forthcoming; Chen, Restuccia and Santaeulàlia-Llopis, 2023; Foster and Rosenzweig, 2022). The land wedge allows the model to fit the reorganization of production within agriculture as countries develop: the transition from a sector dominated by traditional entrepreneurs, or owner-operated farms, to one in which hired labor plays a larger role. This margin is important because agriculture contains much of the labor force that can potentially move into manufacturing and services. At the same time, because this transition does not involve an expansion of white-collar labor, it is distinct from the main reorganization margin emphasized in the paper.

The recalibration allows a limited set of parameters to differ between the two economies: those governing skills, productivities, and wedges. The remaining parameters, including those governing the productivity of professionalizing tasks and the role of laborers and professionals in production in each sector, are held fixed. Thus, the low-income economy differs in endowments, productivities, and wedges, but not in the underlying technology for reorganizing production.

We take the education shares, υ_i, directly from the data (Ruggles et al., 2025).[18] We calibrate $\delta_j^c p_j^c A_j^c$ in country c to match differences in sectoral employment shares between low-income and middle-income economies, maintaining the normalization $\delta_{\text{mfg}}^c p_{\text{mfg}}^c A_{\text{mfg}}^c \equiv 1$. We choose the firm-size wedge, τ^c, to match the difference in the employment share at firms with 11 or more workers, and the land wedge, λ^c, to match the difference in agricultural self-employment.

We decompose $\delta_j^c p_j^c A_j^c$ into the underlying objects using a two-step procedure. First,

[18]This choice is conservative, since education quality and life-cycle human capital formation are lower in developing countries (Schoellman, 2012; Lagakos et al., 2018).

TABLE 3: CALIBRATED CROSS-COUNTRY PARAMETERS AND MOMENTS

	Value			MIE-LIE Gap	
Parameter	LIE	MIE	Target	Data	Model
Panel A. Externally calibrated					
Education shares, υ_i			Education shares		
<Primary	0.501	0.147	<Primary	-0.354	-0.354
Primary	0.312	0.392	Primary	0.080	0.080
Secondary	0.147	0.330	Secondary	0.183	0.183
Tertiary	0.039	0.130	Tertiary	0.091	0.091
Panel B. Internally calibrated					
Sectoral wedge, δ_j			Sector share		
Agriculture	6.948	0.590	Agriculture	-0.316	-0.316
Manufacturing	1.000	1.000	Manufacturing	0.101	0.101
High-skill services	0.794	0.997	High-skill services	0.140	0.140
Low-skill services	1.269	0.914	Low-skill services	0.076	0.076
Firm size wedge, τ	0.020	0.000	Empl. in 11+ firms	0.318	0.318
Land wedge, λ	0.316	0.000	Self-empl. in agr.	-0.296	-0.296
Relative productivity, A_j			Sector value added		
Agriculture	0.297	1.521	Agriculture	0.136	0.136
Manufacturing	1.000	1.000	Manufacturing	0.000	0.000
High-skill services	0.206	0.373	High-skill services	0.040	0.040
Low-skill services	0.652	1.129	Low-skill services	0.177	0.177
Aggregate TFP, $\overline{A}$	0.443	1.000	Log real GDP	1.453	1.453
Panel C. Additional moments (MIE-LIE Gap)					
White collar share	Data	Model	Wage premium	Data	Model
<Primary	0.025	0.011	<Primary	-0.046	-0.104
Primary	0.034	0.016	Primary	-0.039	-0.081
Secondary	0.044	0.035	Secondary	0.000	0.000
Tertiary	0.048	0.075	Tertiary	0.109	0.116

Notes: MIE-LIE gap is the difference between the middle-income and low-income economy.

we use sectoral value added to disentangle the wedge δ_j^c from revenue productivity $p_j^c A_j^c$. Second, we use the demand system to solve for p_j^c and A_j^c such that the observed sectoral output is demanded given prices and income levels, with cross-country aggregate productivity differences $\overline{A}^c$ informed by real manufacturing output.

Table 3 reports the resulting parameters and targeted moments. The right-hand side of the table reports each target as the difference between the middle-income and low-income economies. Because this recalibration is exactly identified, the model replicates the targeted moments by construction.

The calibrated low-income economy features much lower educational attainment:

half of the labor force has less than primary education, as compared to just 15 percent in the middle-income economy. In the aggregate, its TFP is lower and it also faces a firm size wedge of $\tau = 0.020$. At the sectoral level, the largest differences are in agriculture. The land wedge $\lambda = 0.316$ implies that the effective curvature of production with respect to laborers is much lower (approximately 0.14). Agriculture is also much less productive in the low-income economy. Given this, the model requires a large sectoral wedge in agriculture to rationalize the high agricultural employment share.

Panel C shows additional results that help validate the model's mechanism. The left column shows that the white-collar employment share conditional on education rises only modestly between low-income and middle-income economies. Most of the rise in white-collar employment and in employment at large firms is accomplished through composition effects, consistent with the data (Figure 2a) and the analytical model (Proposition 1). The right column shows the change in education wage premia. The two economies have similar education wage premia in the data despite the fact that the middle-income economy has a much more educated workforce; the college wage premium increases by 10.9 log points. The canonical Katz and Murphy (1992) model in which skilled and unskilled workers are treated as imperfect substitutes implies much larger movements in relative wages: a decline of 78 log points for an elasticity of substitution of 1.5.[19] By contrast, our model has small relative wage movements because the model features skill-biased organizational change through the shift toward larger, white-collar-intensive firms. This shift is a form of appropriate technology adoption: it raises the skill intensity of production and offsets the increase in the relative supply of skills, as in Caselli and Coleman (2006).

5.4 Benchmarking with Causal Evidence

Our calibrated model is consistent with a rich set of cross-sectional facts on the relationship among education, occupations, sectors, and the organization of production for low-income and middle-income countries. The key property is skill-biased organizational change. When education expands, the model absorbs the larger white-collar workforce by expanding employment at large firms, without requiring large movements in the skill premium. We now benchmark this mechanism against causal evidence from plausibly exogenous expansions of education. Details of these exercises and similar evidence on the model's consistency with the effects of management training on the reorganization of production are in Appendix D.

[19]Following Rossi (2022), for this calculation we assign tertiary-educated workers to the skilled bin and the rest to the unskilled bin, using relative wages as weights. We then feed in the middle-income education composition, keeping the technology fixed at the low-income level.

TABLE 4: CAUSAL EVIDENCE ON EDUCATIONAL EXPANSIONS

	White Collar	Agri	LS Serv	HS Serv	Self-Emp Share	Emp Share Large Firms	Log College Premium
Panel A. Cox (2025)							
Data	0.029	-0.063	0.068	0.034	-0.050	0.026	-0.124
	(0.009)	(0.011)	(0.022)	(0.009)	(0.011)	(0.009)	(0.048)
Model	0.016	-0.006	-0.016	0.028	-0.011	0.023	0.020
Panel B. Duflo (2001)							
Data	0.021	-0.056	-0.018	0.019	-0.059		
	(0.022)	(0.029)	(0.032)	(0.023)	(0.034)		
Model	0.044	-0.050	0.004	0.038	-0.047		

Notes: The "Data" rows report the estimated coefficients from Cox (2025) and our adaptation of Duflo (2001), with standard errors in parentheses. The estimates from Cox (2025) are re-scaled to represent the effect of an increase of 3.5 p.p. in the college share. The estimates from Duflo (2001) represent the effect of an additional year of schooling. The "Model" rows report the corresponding model-based coefficients. See Appendix D for more details on the implementation of the two exercises.

There is a growing literature on the effects of expansions of educational attainment. Cox (2025) is especially useful for our purposes because he studies how expanded college availability affects college attainment, sectoral employment, and several measures of the organization of production in affected regions.[20] To approximate Cox's experiment in our model, we start from the calibrated middle-income economy. We give it the educational attainment distribution that prevailed in Brazil in 2000. We exogenously shift workers from secondary to tertiary education in a magnitude consistent with Cox's IV estimate, letting wages adjust in equilibrium (but holding goods prices fixed). We compare our model results to Cox's IV estimates for a range of outcomes of interest.

We also benchmark the model against evidence from an exogenous education expansion in a low-income country. We use the individual-level effects of the INPRES primary school construction program in Indonesia (Duflo, 2001). Following Porzio, Rossi and Santangelo (2022), we adopt the empirical specification in Duflo (2001), but focus on sectoral and occupational outcome variables. In the model, we approximate the experiment as a partial equilibrium exercise in which we shift the educational distribution consistent with the first-stage results, keeping all prices and wages fixed.

Table 4 shows the results, with the comparison to Cox in Panel A and the comparison to Duflo in Panel B. The first column shows that the model generates a shift toward white-collar employment that is larger than the Indonesian case but smaller than

[20] See also Porzio, Rossi and Santangelo (2022), Russell, Yu and Andrews (2024), Coelli et al. (2023), Vu (2024), Verma (2025), Nimier-David (2023), and Doxey, Karger and Nencka (2026) for qualitatively similar evidence that is less detailed or harder to map to our model.

the Brazilian case. The next three columns show the sectoral employment effects. The model replicates the rise in high-skill services well. It also generates a decline in agriculture that is close to the Indonesian evidence, but much smaller than the empirical decline in Brazil. This difference is intuitive: shifting workers from secondary to tertiary education has little direct effect on agricultural employment because neither group is likely to work in agriculture. The model can generate an additional decline in agriculture as large firms expand and pull workers out of agriculture, but this force is small relative to Cox's estimates. The model also produces small changes in low-skill service sector employment, whereas there is a large increase in Brazil and a decline in Indonesia.

The next two columns capture the central mechanism of the reorganization of production. The self-employment results closely track the agriculture results: the model generates a decline that is in line with the Indonesian evidence but small relative to the Brazilian evidence. For the employment share at large firms (defined here as firms with five or more employees, following Cox), the model produces a result that is very close to the evidence (0.023 versus 0.026).

The final column reports the college wage premium. In the canonical model with imperfect substitution between unskilled and skilled workers, the expansion in college attainment in Brazil would be expected to produce a 30 log point decline in the college wage premium (based on the calculation discussed in footnote 19). Cox finds a much smaller decline of 12.4 log points. Our model produces a slight increase in relative wages because shifting workers from secondary to tertiary education shifts their comparative advantage from working as professionals to becoming modern entrepreneurs. Our results may differ from Cox's if his estimates capture a transitory decline during adjustment. Consistent with this interpretation, his estimated wage decline is larger than typical estimates in the literature. For example, Duflo (2004) uses a different design to estimate the effect of INPRES on the primary school wage premium in Indonesia and finds no statistically significant effect. Another possible reason is that the marginal workers shifted into tertiary education by his policy may not have a comparative advantage in modern entrepreneurship.

6 Counterfactual Experiments

Now that we have calibrated the model and benchmarked its mechanisms against quasi-experimental evidence, we use it to conduct counterfactual experiments that explore the sources of the reorganization of production. Our experiments focus on the transition from the low-income to the middle-income economy. This transition is associated with a substantial reorganization of production, with agricultural self-employment declining by roughly 30 percentage points and the share of workers employed in medium and large firms rising by a similar amount (Table 3). This reorganization corresponds to the transition the United States experienced in the 19th century and the beginning of the transition documented by Chandler. Our data and model are better suited for this early phase of reorganization than for the later emergence of very large, multi-product, multi-establishment firms. Empirically, establishment size is top-coded at low levels in most labor force surveys. Theoretically, very large firms raise important new margins that we abstract from: entrepreneurs not only professionalize tasks but also organize the hired white-collar workers into a management hierarchy to coordinate knowledge flows within the increasingly large firms, as in Garicano and Rossi-Hansberg (2006).

6.1 Structural Transformation and the Organization of Production

We first ask whether the reorganization of production is a natural consequence of structural transformation. We take the calibrated low-income economy and replace the main parameters governing structural transformation $(\overline{A}, A_j, \delta_j, \lambda)$ with their calibrated values from the middle-income economy. We then solve for the new equilibrium, allowing demand to respond through income and substitution effects using the non-homothetic CES specified in equation (10).

Table 5 reports the results of all the experiments from this section. The first row, labeled "Data", reports the actual difference between the middle-income and the low-income economy for each outcome. The second row, labeled structural transformation, reports the counterfactual difference between the low-income economy with the structural transformation parameters of a middle-income economy and the calibrated low-income economy. Panel A shows the aggregate outcomes, Panel B the outcomes by sector, and Panel C the outcomes by education level.

The main finding is that structural transformation and the reorganization of production are largely distinct phenomena. The results in Panel B show that giving the low-income economy the structural transformation parameters of the middle-income economy generates almost all of the decline in agricultural employment and more than the observed rise in manufacturing and low-skill services employment. The model falls short only for high-skill services.

However, this experiment generates almost no reorganization of production. In the aggregate, the employment share in medium and large firms is essentially unchanged, while the decline in self-employment is 28 percent of what is observed in the data (-0.116 versus -0.417). Panels B and C contain a number of moments that help explain

Table 5: Understanding the Low-to-Middle-Income Transition

Panel A. Aggregate Outcomes	**11+ share**	**Self-empl.**	**White collar**	**Ln VAPW**
Data	0.318	-0.417	0.159	1.453
Structural transformation	0.000	-0.116	0.015	0.870
Schooling	0.173	-0.105	0.066	0.594
Skill-biased technical change	0.173	-0.101	0.069	0.349
Correlated firm-size distortion	0.173	-0.068	0.047	0.143
Panel B. Outcomes by Sector	**Agr**	**Mfg**	**Ser (HS)**	**Ser (LS)**
Employment shares				
Data	-0.316	0.101	0.140	0.076
Structural transformation	-0.308	0.121	0.087	0.101
Schooling	-0.025	-0.009	0.053	-0.019
Skill-biased technical change	0.023	-0.013	0.015	-0.026
Correlated firm-size distortion	0.016	-0.005	0.009	-0.019
Employment share at 11+ firms				
Data	0.184	0.328	0.235	0.243
Structural transformation	0.000	-0.176	-0.121	-0.090
Schooling	0.000	0.402	0.065	0.385
Skill-biased technical change	0.000	0.507	0.152	0.431
Correlated firm-size distortion	0.000	0.416	0.190	0.449
Panel C. Outcomes by Education	**<Primary**	**Primary**	**Secondary**	**Tertiary**
White collar share				
Data	0.025	0.034	0.044	0.048
Structural transformation	0.011	0.015	0.027	0.016
Schooling	-0.013	-0.018	-0.033	-0.007
Skill-biased technical change	0.054	0.071	0.112	0.086
Correlated firm-size distortion	0.031	0.045	0.089	0.115
Wage				
Data	-0.046	-0.039	0.000	0.109
Structural transformation	-0.101	-0.082	0.000	0.096
Schooling	-0.009	0.005	0.000	-0.051
Skill-biased technical change	-0.038	-0.023	0.000	0.011
Correlated firm-size distortion	-0.005	-0.003	0.000	0.010

Notes: Data reports the difference between the middle-income and low-income economy. The remaining rows report the change between a counterfactual economy and the calibrated low-income economy. *Structural transformation* gives the low-income economy the productivities ($\overline{A}A_j$), land wedge (λ), and sectoral wedges (δ_j) of the middle-income economy. *Schooling* gives the low-income economy the education distribution (υ_i) of the middle-income economy. *Skill-biased technical change* and *Correlated firm-size distortion* give the low-income economy a higher β or a lower τ chosen to give an increase in the employment share at medium and large firms of 17.3 percentage points.

why this experiment generates limited reorganization. Agriculture is the sector with the lowest employment share at large firms (Figure 7f). Structural transformation reallocates employment toward sectors with more large firms, which increases demand for white-collar labor and the relative wage of more educated workers. However, workers do not switch much toward white-collar employment within education groups, consistent with the motivating fact in Figure 2a and the model fit in Figure 5. With the supply of skilled labor fixed, the model can accommodate higher relative demand for skilled labor only by *reducing* large-firm employment shares within sectors. This decline affects low-skill services and manufacturing, two sectors that have growth in the large-firm employment share in the data. We conclude that structural transformation alone cannot explain the reorganization of production.

6.2 The Effect of Expanding Education

We now consider the effect of an exogenous expansion of the supply of educated workers. This experiment is the quantitative analogue to Proposition 1 in Section 4.3 and corresponds closely to the model exercises in Section 5.4. We start again from the calibrated low-income economy, replace only its distribution of educational attainment with that of the middle-income economy, and solve for the new equilibrium. The results are reported in the third row of Table 5, labeled schooling.

In contrast to structural transformation alone, schooling generates a substantial reorganization of production. At the aggregate level, it generates 54 percent of the growth in the employment share at medium and large firms (17.3 percentage points versus 31.8 percentage points in the data). Within manufacturing and low-skill services, the increase in medium- and large-firm employment shares exceeds the increase observed in the data. The decline in traditional entrepreneurship and the growth of white-collar employment are both somewhat weaker, 25 and 42 percent of the observed shifts.

The main obstacle that prevents the model from achieving a full reorganization of production is the absence of structural transformation. In Section 4.3 we emphasized that endogenous duality allowed expanding modern firms to draw laborers out of traditional entrepreneurship.[21] The quantitative model adds an important insight: endogenous duality interacts with structural transformation because much of the reserve labor force of traditional entrepreneurs is in agriculture. While structural transformation does not on its own cause reorganization, the model cannot generate a complete reorganization without structural transformation.

Expanding education particularly benefits manufacturing and low-skill services be-

[21] Additional counterfactuals isolating the importance of endogenous duality are in Appendix E.

cause these sectors have the most scope to reorganize production into large firms. However, the estimated elasticity of substitution between the output of different sectors in the structural transformation literature is well below 1. With a low elasticity of substitution, the model generates large offsetting declines in the relative price of manufacturing and low-skill service outputs. This price decline dampens the incentive to expand these sectors. The model requires exogenous shifts in the parameters $(\overline{A}, A_j, \delta_j, \lambda)$ to achieve structural transformation and a complete reorganization of production.

The quantitative model violates the envelope theorem conditions because of the distortions τ, λ, and δ, as well as workers' preferences over occupations and sectors. Even so, the first-order approximation of the change in output is 59.4 log points, closely replicating the actual model-generated change. The approximation works well for two reasons. First, the most quantitatively important distortions in the low-income economy are λ and δ_{agr}, which affect agriculture (Table 3). However, increasing schooling shifts few workers out of agriculture, so this margin matters little here. Second, skill-biased organizational change lets the model absorb large changes in skill supply with small changes in relative wages and marginal products, which implies that second-order effects are small. For example, shutting down endogenous duality yields a smaller reorganization of production (an 11 percentage point rise in the employment share at medium and large firms, versus 17 in the baseline) and a smaller 44 log point increase in output per worker (Appendix E).

6.3 The Effects of Technical Change and Reducing Distortions

We have introduced skills as a new potential driver of the reorganization of production and the growth of large firms. The literature has suggested two other forces. First, technical change may have explicitly favored the growth of large firms through a variety of mechanisms (Garicano and Rossi-Hansberg, 2006; Autor et al., 2020; Reichardt, 2025). Second, reductions in financial constraints or size-dependent distortions may have favored the growth of productive firms relative to unproductive firms (Buera, Kaboski and Shin, 2011; Guner, Ventura and Xu, 2008; Restuccia, Yang and Zhu, 2008; Hsieh and Klenow, 2009). In this section we investigate how these forces shape the organization of production in our model.

We model technical change as an increase in the productivity of professionalizing tasks β and reductions in distortions as a decline in τ. Both of these changes favor the growth of modern entrepreneurship and large firms (e.g., equation (6)). For this counterfactual, we start from the calibrated low-income economy and solve for the rise in β or the reduction in τ that produces an equilibrium rise in the employment share at medium and large firms of 17.3 percentage points, which is the effect of increasing schooling. We choose this strategy to highlight that all three forces can generate a reorganization of production, but that they work differently in the model. The calibrated β rises from 1.47 to 2.14, which does not affect the productivity of a firm with $q = 0$ but raises the productivity of a firm with $q = 1$ by 44 percent. The calibrated τ falls from 0.02 to -0.026, which has no effect on the returns to scale of a firm with $q = 0$ but raises the returns to scale of a manufacturing firm with $q = 1$ from 0.79 to 0.81.

The results of these experiments are shown in the fourth and fifth rows of Table 5. Panel A shows that expansions of education, technical change, and reductions in distortions generate similar reorganizations of production. The rise in the employment share at medium and large firms is 17.3 percentage points in each case by design. Conditional on this normalization, the three forces generate similar decreases in self-employment and increases in white-collar employment shares. Expanding education has the largest effect on value added per worker because it entails a direct productivity effect for workers.

The results in Panel C show that expanding education works through a very different mechanism than increasing β or decreasing τ. Intuitively, expanding education increases the supply of potential white-collar labor, whereas increasing β or decreasing τ increases the demand for white-collar labor. As discussed after Proposition 1, endogenous occupational choice dampens wage responses to shocks through two mechanisms: skill-biased organizational change and a reserve labor force of traditional entrepreneurs. However, Panel C shows that expanding education lowers white-collar employment rates within education groups and reduces education wage premia; raising β or lowering τ has the opposite effect.

These findings are closely related to Buera et al. (2022), who also investigate a model with structural change across sectors that differ in their skill intensity. Whereas they use the canonical model with imperfect substitution between unskilled and skilled workers, we use our model of occupational choice and endogenous organization of production. Nonetheless, many of the results are similar. Our finding that structural transformation is driven mainly by sectoral productivity and sectoral wedges is closely related to their finding that it is largely driven by sector-specific, skill-neutral technical progress. Like them, we find that expansions in education tend to lower the skill premium while structural transformation and technical change tend to increase it. The main difference between our papers is that our model features skill-biased organizational change that dampens movements in relative wages in response to all shocks.

Thus far we have treated the supply of skills as exogenous. This is natural when we consider policies, such as school expansions, that shift educational attainment directly.

However, reductions in distortions or technical change raise education wage premia, which should generate an incentive for workers to acquire more schooling. While the literature provides evidence of a causal link from technologies to educational attainment, we lack well-established evidence on the general strength of this mechanism.[22] Thus, we do not provide a quantitative evaluation of this feedback effect.

These results show that technical change and reductions in distortions can reorganize production, and may also contribute indirectly by raising educational attainment. However, they are unlikely to explain the entirety of the reorganization of production. There are, after all, large differences in educational attainment between low-income and middle-income economies. The literature makes clear that educational attainment is responsive to school construction and expansion, so some of these differences are due to policy. The literature also provides clear evidence of a causal effect running from schooling to a reorganization of production. Taken together, the evidence and counterfactuals imply that low educational attainment is a genuine barrier to the reorganization of production in low-income economies.

7 Conclusion

Chandler (1977) argued that the large firms of the Second Industrial Revolution required hierarchies of white-collar workers to coordinate production at scale. We have shown that this insight remains central for understanding development today.

We develop a theory of occupational choice that extends the classic work of Lucas (1978) by allowing entrepreneurs to decide both how much to produce and how to organize production. In equilibrium, the least and most skilled workers become entrepreneurs, but they organize production very differently. We calibrate the model to rich cross-sectional data and benchmark it against causal evidence on the effects of expanding education. The calibrated model implies that educational attainment is a necessary ingredient for the reorganization of production from traditional self-employment to large firms. Structural transformation and reductions in barriers matter as well, but without a sufficiently educated workforce, they generate only a limited reorganization of production. Low educational attainment therefore remains an important barrier to the emergence of large, white-collar-intensive firms in low-income economies.

Our model predicts that expansions of primary, secondary, and tertiary education have different effects on the organization of production and relative wages.[23] Future

[22] See Jensen (2012) for an example of well-identified evidence.

[23] For example, we find that expanding tertiary education in Section 5.4 raises the college wage premium, but that expanding overall education in Section 6.2 lowers it.

work could investigate this further empirically and theoretically. We also abstract from a number of features to focus on the link from skills to occupational choice and the organization of production. We treat skills as exogenous, but in the long run the supply of skills may itself be endogenous to skill-biased organizational change. We also abstract from physical capital and electrification, which we view as important but already well studied in the literature (e.g., Buera, Kaboski and Shin, 2011; Fried and Lagakos, 2023). Finally, we focus on the role of skills in enabling a reorganization of production in manufacturing and low-skill services, consistent with Chandler's historical work. For today's developed countries, the educated, white-collar workforce is increasingly devoted to the high-skill service sector, which is arguably shaped by different forces. These are all important avenues for future research.

References

Acemoglu, Daron, and Fabrizio Zilibotti. 2001. "Productivity Differences." *Quarterly Journal of Economics*, 116(2): 563–606.

Acemoglu, Daron, and Pascual Restrepo. 2018. "The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment." *American Economic Review*, 108(6): 1488–1542.

Akcigit, Ufuk, Harun Alp, and Michael Peters. 2021. "Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries." *American Economic Review*, 111(1): 231–275.

Amaral, Pedro S., and Alberto Rivera-Padilla. 2025. "Cross-Country Income Dispersion, Human Capital, and Technology Adoption." mimeo, Cal State University – Fullerton.

Argente, David, Sara Moreira, Ezra Oberfield, and Venky Venkateswaran. 2025. "Scalable Expertise: How Standardization Drives Scale and Scope." mimeo, Yale School of Management.

Artuç, Erhan, and John McLaren. 2015. "Trade policy and wage inequality: A structural analysis with occupational and sectoral mobility." *Journal of International Economics*, 97(2): 278–294.

Ashournia, Damoun. 2018. "Labour Market Effects of International Trade When Mobility is Costly." *Economic Journal*, 128(616): 3008–3038.

Autor, David, David Dorn, Lawrence F. Katz, Christina Patterson, and John Van Reenen. 2020. "The Fall of the Labor Share and the Rise of Superstar Firms." *Quarterly Journal of Economics*, 135(2): 645–709.

Ayerst, Stephen, Loren Brandt, and Diego Restuccia. forthcoming. "Distortions, Producer Dynamics, and Aggregate Productivity: A General Equilibrium Analysis." *American Economic Journal: Macroeconomics*.

Bandiera, Oriana, Ahmed Elsayed, Anton Heil, and Andrea Smurra. 2022. "Presidential address 2022: Economic development and the organisation of labour: Evidence from the jobs of the world project." *Journal of the European Economic Association*, 20(6): 2226–2270.

Bandiera, Oriana, Ananya Kotia, Ilse Lindenlaub, Christian Moser, and Andrea Prat. 2025. "Meritocracy Across Countries." Mimeo, Columbia Business School.

Banerjee, Abhijit V., and Andrew F. Newman. 1993. "Occupational Choice and the Process of Development." *Journal of Political Economy*, 101(2): 274–298.

Bassi, Vittorio, Jung Hyuk Lee, Alessandra Peter, Tommaso Porzio, Ritwika Sen, and Esau Tugume. 2025. "Self-Employment Within the Firm." mimeo, Columbia University.

Basu, Susanto, and David N. Weil. 1998. "Appropriate Technology and Growth." *Quarterly Journal of Economics*, 113(4): 1025–1054.

Bloom, Nicholas, and John Van Reenen. 2007. "Measuring and Explaining Management Practices Across Firms and Countries." *Quarterly Journal of Economics*, 122(4): 1351–1408.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Buera, Francisco J, Joseph Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review*, 101(5): 1964–2002.

Buera, Francisco J., Joseph P. Kaboski, Richard Rogerson, and Juan I. Vizcaino. 2022. "Skill-Biased Structural Change." *Review of Economic Studies*, 89(2): 592–625.

Caselli, Francesco, and Wilbur John Coleman, II. 2006. "The World Technology Frontier." *American Economic Review*, 96(3): 499–522.

Chandler, Jr., Alfred D. 1977. *The Visible Hand: The Managerial Revolution in American Business.* Cambridge, Massachusetts:Harvard University Press.

Chandler, Jr., Alfred D. 1990. *Scale and Scope: The Dynamics of Industrial Capitalism.* Cambridge, Massachusetts:Harvard University Press.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis. 2023. "Land Misallocation and Productivity." *American Economic Journal: Macroeconomics*, 15(2): 441–465.

Coelli, Federica, Difei Ouyang, Weidi Yuan, and Yuan Zi. 2023. "Educating Like China." mimeo, University of Zurich.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural Change with Long-Run Income and Price Effects." *Econometrica*, 89(1): 311–374.

Cox, Alvaro. 2025. "Fostering Development Through Higher Education: College Attainment, Firms and Economic Growth." mimeo, Universidad Carlos III de Madrid.

Dix-Carneiro, Rafael. 2014. "Trade Liberalization and Labor Market Dynamics." *Econometrica*, 82(3): 825–885.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman. 2023. "Labor Market Dynamics and Development." *Quarterly Journal of Economics*, 138(4): 2287–2325.

Doxey, Alison, Ezra Karger, and Peter Nencka. 2026. "High Schools and the Uneven Rise in American Opportunity." NBER Working Paper 35068.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.

Duflo, Esther. 2004. "The Medium Run Effects of Educational Expansion: Evidence from a Large School Construction Program in Indonesia." *Journal of Development Economics*, 74(1): 163–197.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

Ferraro, Domenico, Maurizio Iacopetta, and Pietro Peretto. 2024. "The Rise and Evolution of the Innovative Firm: A Tale of Technology, Market Structure, and Managerial Incentives." mimeo, Arizona State University.

Foster, Andrew D, and Mark R Rosenzweig. 2022. "Are there too many farms in the world? Labor market transaction costs, machine capacities, and optimal farm size." *Journal of Political Economy*, 130(3): 636–680.

Fried, Stephie, and David Lagakos. 2023. "Electricity and Firm Productivity: A General-Equilibrium Approach." *American Economic Journal: Macroeconomics*, 15(4): 67–103.

Garicano, Luis, and Esteban Rossi-Hansberg. 2006. "Organization and Inequality in a Knowledge Economy." *Quarterly Journal of Economics*, 121(4): 1383–1435.

Gollin, Douglas, David Lagakos, and Michael E. Waugh. 2014. "The Agricultural Productivity Gap." *Quarterly Journal of Economics*, 129(2): 939–993.

Gomes, Pedro, and Zoë Kuehn. 2017. "Human capital and the size distribution of firms." *Review of Economic Dynamics*, 26: 164–179.

Gottlieb, Charles, Jan Grobovšek, and Alexander Monge-Naranjo. 2025. "From Skills to Occupations: Comparative Advantage and Cross-Country Income Differences." mimeo, Aix-Marseille School of Economics.

Gottlieb, Charles, Markus Poschke, and Michael Tueting. 2025. "Skill Supply, Firm Size, and Economic Development." mimeo, Aix-Marseille School of Economics.

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

Herrendorf, Berthold, and Todd Schoellman. 2018. "Wages, Human Capital, and Barriers to Structural Transformation." *American Economic Journal: Macroeconomics*, 10(2): 1–23.

Hjort, Jonas, Hannes Malmberg, and Todd Schoellman. forthcoming. "The Missing Middle Managers: Labor Costs, Firm Structure, and Development." *American Economic Review: Insights*.

Hopenhayn, Hugo. 2014. "On the Measure of Distortions." NBER Working Paper 20404.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and Manufacturing TFP in China and India." *Quarterly Journal of Economics*, 124(4): 1403–1448.

Hubmer, Joachim, Mons Chan, Sergio Salgado, and Guangbin Hong. 2026. "Scalable versus Productive Technologies." mimeo, University of Pennsylvania.

Inklaar, Robert, Ryan Marapin, and Kaira Gräler. 2023. "Tradability and Sectoral Productivity Differences Across Countries." GGDC Research Memorandum 195.

Jensen, Robert. 2012. "Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India." *Quarterly Journal of Economics*, 127(2): 753–792.

Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in Relative Wages, 1963-1987: Supply and Demand Factors." *Quarterly Journal of Economics*, 107(1): 35–78.

Kopytov, Alexandr, Mathieu Taschereau-Dumouchel, and Zebang Xu. 2026. "Endogenous Returns to Scale." mimeo, University of Rochester.

Lagakos, David, Benjamin Moll, Tommaso Porzio, Nancy Qian, and Todd Schoellman. 2018. "Life Cycle Wage Growth across Countries." *Journal of Political Economy*, 126(2): 797–849.

Lucas, Jr., Robert E. 1978. "On the Size Distribution of Business Firms." *Bell Journal of Economics*, 9(2): 508–523.

Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1991. "The Allocation of Talent: Implications for Growth." *Quarterly Journal of Economics*, 106(2): 503–530.

Nimier-David, Elio. 2023. "Local Human Capital and Firm Creation: Evidence from the Massification of Higher Education in France." mimeo, Cornell.

Porzio, Tommaso. 2017. "Cross-Country Differences in the Optimal Allocation of Talent and Technology." mimeo, Columbia University.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo. 2022. "The Human Side of Structural Transformation." *American Economic Review*, 112(8): 2774–2814.

Queiró, Francisco. 2022. "Entrepreneurial Human Capital and Firm Dynamics." *Review of Economic Studies*, 89(4): 2061–2100.

Reichardt, Hugo. 2025. "Scale-Biased Technical Change and Inequality." mimeo, CREi.

Restuccia, Diego, Dennis T. Yang, and Xiaodong Zhu. 2008. "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis." *Journal of Monetary Economics*, 55(2): 234–250.

Revenga, Ana L. 1992. "Exporting Jobs?: The Impact of Import Competition on Employment and Wages in U.S. Manufacturing." *Quarterly Journal of Economics*, 107(1): 255–284.

Rossi, Federico. 2022. "The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation." *American Economic Review*, 112(1): 235–266.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Quinn Heimann, Jane Lee, and Nate Merrill. 2025. "Integrated Public Use Microdata Series, International: Version 7.6 [dataset]." https://doi.org/10.18128/D020.V7.6.

Russell, Lauren C., Lei Yu, and Michael J. Andrews. 2024. "Higher Education and Local Educational Attainment: Evidence from the Establishment of U.S. Colleges." *Review of Economics and Statistics*, 106(4): 1146–1156.

Schoellman, Todd. 2012. "Education Quality and Development Accounting." *Review of Economic Studies*, 79(1): 388–417.

Tamkoç, Nazım. 2024. "Managers, Talent Misallocation and Productivity." mimeo, World Bank.

Verma, Bipul. 2025. "Tertiary Education and Growth in Services: Evidence from College Expansion in India." mimeo, University of Minnesota.

Vu, Khoa. 2024. "Higher Education Expansion and the Rise of the Skill-Intensive Service Sector." mimeo, University of Minnesota.

Online Appendices

A Data Appendix

This appendix provides details on the data and further results related to the motivating facts established in Section 2.

A.1 Data Construction

IPUMS International. We use all cross-sections with available information on educational attainment, occupation, and sector. This gives us 218 cross-sections from 74 countries, spanning the global income distribution from Mali to the United States. Within each cross-section, we restrict the sample to employed individuals ages 16–65. For the construction of wage moments, we focus on 11 cross-sections with available wage information: Brazil 2010, Canada 2001, India 1999, Indonesia 1995, Jamaica 2001, Mexico 2010, Panama 2010, Trinidad and Tobago 2000, United States 2015, Uruguay 2006, Venezuela 2001. We use hourly wages whenever available and restrict the sample to wage workers with high levels of labor market attachment (35+ weekly hours worked or full-time status, if available). We follow Rossi (2022) in implementing a number of country-specific adjustments, and we refer the reader to Appendix A in that paper for the details.

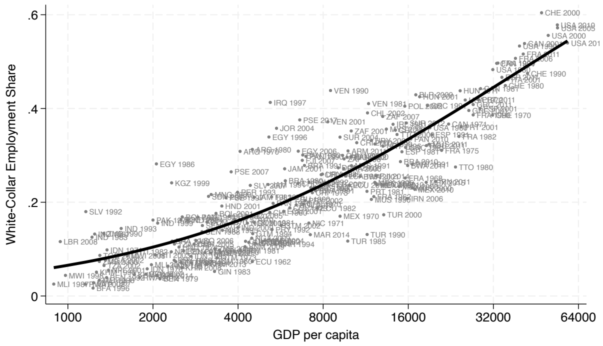
LFS. The main source for labor force surveys is the rotating panel labor force survey database previously used in Donovan, Lu and Schoellman (2023) and Donovan et al. (2026). We augment this with cross-sectional data from Senegal's ENES labor force survey to expand coverage of low-income economies (Agence nationale de la Statistique et de la Démographie, 2017–2024). We focus on data from countries that ask workers about the number of employees at the establishment where they work. We pool all available years within a country and use the country-level average. This gives us 46 countries, spanning the global income distribution from Rwanda to Luxembourg. We again restrict the sample to employed individuals ages 16–65. For the construction of wage moments, we focus on a partially overlapping sample of 22 countries with available wage information. We again use hourly wages and restrict the sample to wage workers with high levels of labor market attachment.

GDP per capita. We construct real GDP per capita in PPP (*rgdpo*/*pop*) from Penn World Tables 10.01 (Feenstra, Inklaar and Timmer, 2015). We extrapolate to years after 2019 by applying the growth rate of GDP per capita in PPP from the World Development Indicators (World Bank, 2025).

A.2 Further Results: White-Collar Employment

An important feature of the data is the large cross-country differences in the share of white-collar workers. Figure A.1 uses census data from Ruggles et al. (2025) to show that this varies from 10 percent in the poorest countries to 60 percent in the richest.

FIGURE A.1: WHITE-COLLAR EMPLOYMENT AND DEVELOPMENT



Notes: Each marker corresponds to a country $\times$ year observation. The line shows the fit of a logistic regression on a quadratic in log GDP per capita.

In Section 2.1 we document two facts that build on Chandler's insight about the role white-collar labor plays in production. The first fact, shown in Figure 1a, is that large firms use white-collar labor more intensively. Although this fact partly reflects structural transformation, we show here that the same pattern holds after controlling for sector or industry.

Both IPUMS and labor force survey data are harmonized to a common industry variable with fifteen codes. For most of the paper, we further aggregate industries into four broad sectors. High-skill services consist of industries whose workers average at least 13 years of schooling in the United States, which includes education; financial services and insurance; health; public administration; other services; and real estate and business services. Low-skill services consist of hotels and restaurants; private household services; communication and transportation; and wholesale and retail trade. Manufacturing also includes construction, mining, and utilities. Agriculture is its own code.

We use the microdata from the labor force survey database to estimate the probability that a worker has a white-collar occupation as a function of firm size. We use a linear probability model for ease of interpretation; results from a logit are similar. Table A.1, column (1) reports estimates controlling for country fixed effects. In this case we

TABLE A.1: WHITE-COLLAR EMPLOYMENT AND FIRM SIZE

	White-Collar Employment Share		
	(1)	(2)	(3)
Medium	0.200***	0.135***	0.106***
	(0.000)	(0.000)	(0.000)
Large	0.248***	0.178***	0.146***
	(0.000)	(0.000)	(0.000)
Fixed Effects	Country	Country + Sector	Country + Industry
R-squared	0.10	0.23	0.26
Sample Size (m)	29.8	29.8	29.8

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

find that workers in medium and large firms are 20–25 percentage points more likely to have white-collar occupations, consistent with Figure 1a. In column (2) we control also for broad sector fixed effects, while in column (3) we control instead for industry fixed effects. Doing so reduces the point estimates by about half. However, even within the same sector or industry it remains the case that workers in medium and large firms are 10–18 percentage points more likely to have white-collar occupations.

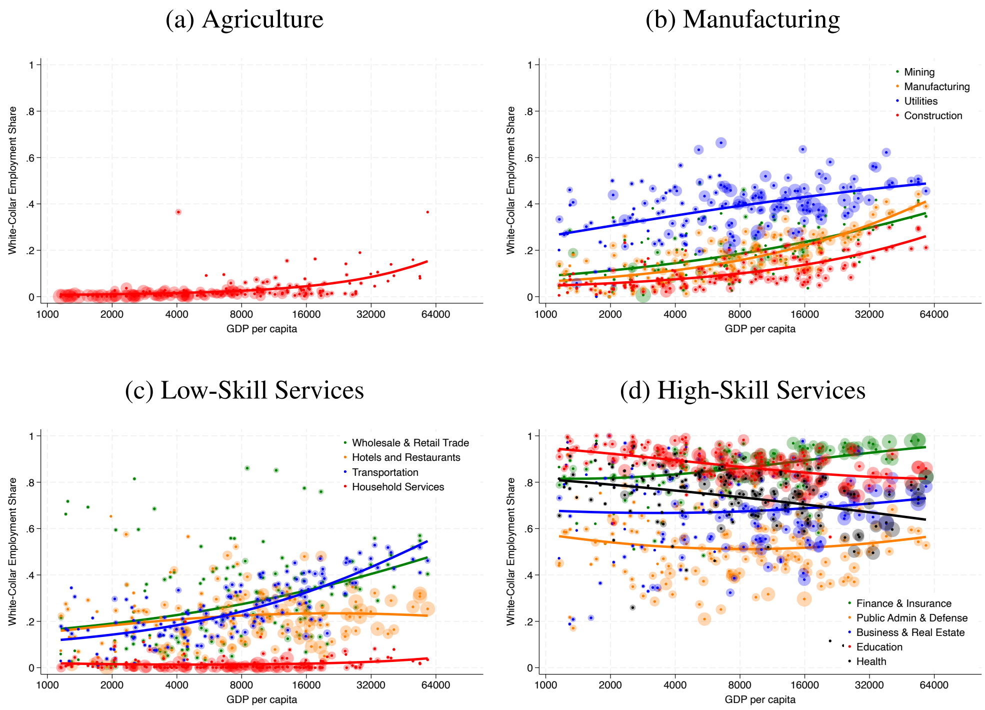
The second fact, shown in Figure 1b, is that development is associated with rising white-collar employment shares in some sectors. We now decompose broad sectors into the underlying detailed industry codes (where such codes are available) in Figure A.2. Each panel corresponds to a broad sector and plots its component detailed industries. The observations are at the country $\times$ year $\times$ detailed industry level.[24] The figure shows that the main detailed industries affected by a reorganization of production are manufacturing, wholesale and retail trade, and transportation, whose fitted white-collar shares rise by approximately 30, 35, and 40 percentage points, respectively, from the poorest to the richest economies. These are precisely the industries emphasized most by Chandler (1977).

A.3 Further Results: Skills and White-Collar Employment

Our main new empirical result is that most of the cross-country differences in the employment share of white-collar workers can be accounted for by differences in skills. In the main text we use educational attainment as our baseline proxy for human capital

[24] We focus on a subsample of 54 countries (153 cross-sections) where we observe separately all the 14 industries listed in Figure A.2. We do not use the "Other Services" category that might include different industries across countries.

Figure A.2: Detailed Sectors and White-Collar Employment



Notes: Each marker corresponds to a country × year × sector observation. The bubbles around the markers are proportional to the employment share of the sector within each country × year. The lines show the fits of multinomial logistic regressions on quadratic polynomials in log GDP per capita.

and show the results using cross-country data. Here we show that the result is robust to alternative measures, samples, and decompositions.

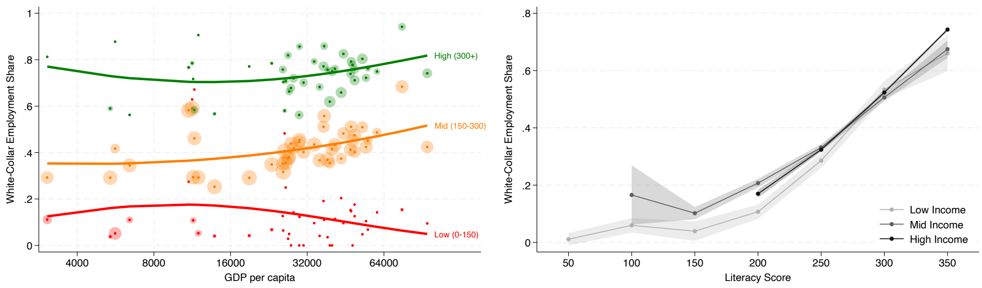
A.3.1 Alternative Measures of Skills: Adult Test Scores

As an alternative to educational attainment, we study white-collar employment shares as a function of adult test scores for a large set of countries. For this analysis we use data from the Organisation for Economic Co-operation and Development (OECD)'s PIAAC Survey of Adult Skills and the World Bank's STEP Skills Measurement Program. The OECD PIAAC surveyed roughly 5,000 adults aged 15-65 in more than 40 countries. Its tests measure skills in literacy, numeracy, and problem solving. The World Bank STEP program builds on and expands the scope of PIAAC by surveying 2,000-4,000 adults aged 16-65 in 12 lower-income countries/regions. The surveys measure literacy and socioemotional skills. We combine the two datasets and focus on literacy, which is measured in both surveys, following Caunedo, Keller and Shin (2023). Our final sample includes 43 countries spanning the income distribution from Kenya to Norway.

Figure A.3 presents results using adult literacy scores instead of education. The same patterns apply: workers with higher test scores are much more likely to work in white-collar occupations; cross-country differences in white-collar employment shares conditional on test scores are small.

FIGURE A.3: WHITE-COLLAR EMPLOYMENT SHARE AND LITERACY SCORES

(a) Share by Literacy Score Group (b) Propensities by Income Group



Notes: Each marker in Panel (a) corresponds to a country × year × literacy score bin observation. The bubbles around the markers are proportional to the employment share of the literacy score bin within each country × year. The lines show the fits of multinomial logistic regressions on quadratic polynomials in log GDP per capita. Panel (b) displays estimates (and shaded 95% confidence intervals) from individual-level regressions of white-collar status on 7 literacy score bin fixed effects (0-49, 50-99, ..., 250-299, 300+), controlling for age group (16-20, 21-25,...,61-65) and gender fixed effects. Any country with GDP p.c. lower than 6000 USD is reclassified as low-income to increase the sample size of this group. Observations below the 1st or above the 99th percentiles of the literacy score distribution within each income group are dropped. Observations are reweighted so that each country contributes equally to the regressions.

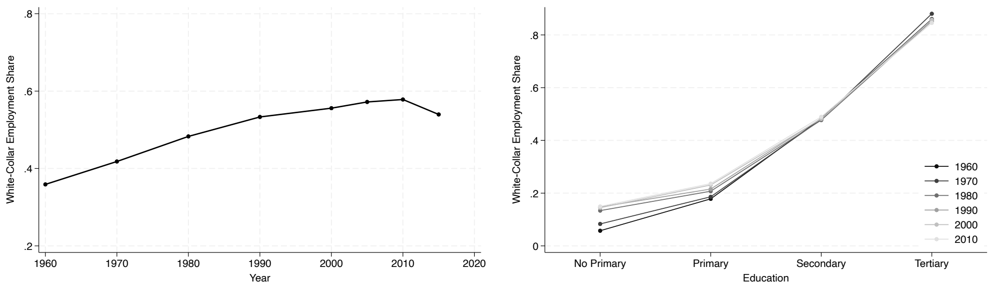
A.3.2 Time-Series Results

The analysis in Section 2 combines the cross-sectional and time-series variation by pooling all available surveys. This subsection shows that the same relationship holds within countries over time. Figure A.4 starts by focusing on the United States, the country with the longest available time series. Figure A.4a shows that the white-collar share of employment increased by more than 20 percentage points between 1960 and 2015. Figure A.4b shows that the share of workers employed in white-collar occupations conditional on education is remarkably constant across decades, implying that virtually all the aggregate increase in Figure A.4a can be accounted for by changes in the educational composition over time.

Figure A.5 shows the share of workers working in a white-collar occupation conditional on education for all countries in the sample. Figures A.5a, A.5b, and A.5c show results for low-income, middle-income, and high-income countries, while the lines within each figure capture the estimated share for different time periods.

FIGURE A.4: WHITE-COLLAR SHARE OVER TIME – UNITED STATES

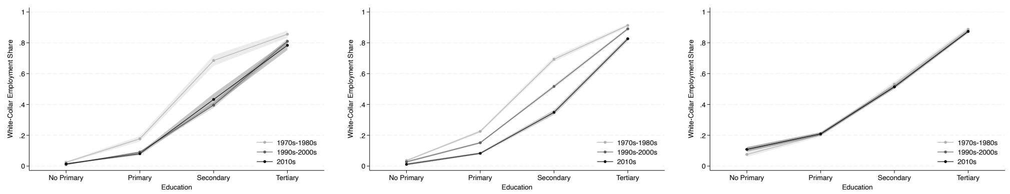
(a) White-Collar Share Over Time　　(b) Propensities by Decade



Notes: Panel (a) plots the average white-collar employment share for each year. Panel (b) displays estimates from individual-level regressions of white-collar status on the 4 education dummies, controlling for age group (16-20, 21-25,...,61-65) and gender fixed effects (confidence intervals omitted for visual clarity).

FIGURE A.5: WHITE-COLLAR SHARE OVER TIME – ALL COUNTRIES

(a) Low Income　　(b) Middle Income　　(c) High Income



Notes: The panels display estimates (and shaded 95% confidence intervals) from individual-level regressions of white-collar status on the 4 education dummies, controlling for age group (16-20, 21-25,...,61-65) and gender fixed effects. Observations are reweighted so that each country contributes equally to the regressions.

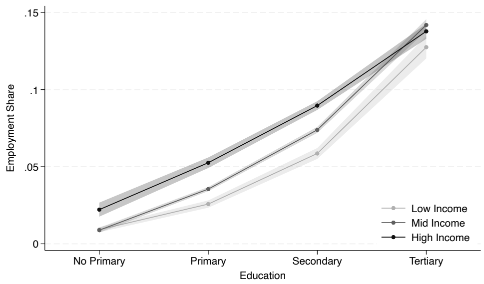
The share of workers choosing white-collar occupations is very stable in high-income countries. For low- and middle-income countries there is a decline in the white-collar share of primary- and secondary-educated workers. One possible explanation for this declining share is that the years 1970–2010 correspond to a period of massive educational expansion in these countries. Recent work suggests that this expansion may have lowered education quality, which would imply that educational attainment does not map into skills in a consistent way over time (Le Nestour, Moscoviz and Sandefur, 2023). Nevertheless, differences across education groups remain large in all periods, and changes in the educational composition can account for most of the variation in the white-collar employment share over time.

A.3.3 Detailed Occupations

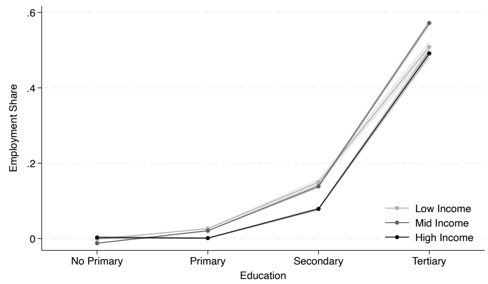
Figure A.6 shows the results separately for the four white-collar occupations. The share of managers and professionals monotonically increases with education, with profiles quite comparable across countries. For associate professionals and clerks, the gradient is quite strong at lower levels of education, though the educational profile flattens out or even decreases between the secondary and tertiary levels.

FIGURE A.6: DETAILED WHITE-COLLAR OCCUPATIONS AND EDUCATION

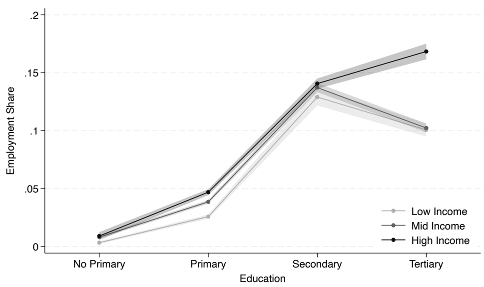
(a) Managers



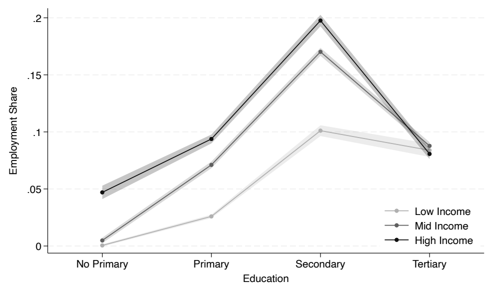
(b) Professionals



(c) Technicians and Associate Professionals



(d) Clerks



Notes: The panels display estimates (and shaded 95% confidence intervals) from individual-level regressions of occupational dummies on the 4 education dummies, controlling for age group (16-20, 21-25,...,61-65) and gender fixed effects. Observations are reweighted so that each country contributes equally to the regressions.

A.3.4 Summary Accounting Results

In Section 2.2 we document that human capital accounts for a substantial share of cross-country differences in the share of white-collar workers. This appendix formalizes that idea with a simple accounting exercise. We collapse the Ruggles et al. (2025) data to the white-collar employment share at the country × year × education × 5-year age group × gender cell. We run a regression of the white-collar employment share on log GDP per capita, weighting cells by their employment shares within each cross-section (so that all cross-sections are weighted equally). We include dummies for gender and age groups.

We refer to the estimated coefficient on log GDP per capita as the unconditional semi-elasticity of white-collar employment with respect to development. We then re-estimate the specification adding education fixed effects. We refer to the estimated coefficient on log GDP per capita in this case as the conditional semi-elasticity.

We measure the share of the white-collar-development relationship accounted for by skills as

$$\text{Accounting Share} = 1 - \frac{\text{Conditional Semi-Elasticity}}{\text{Unconditional Semi-Elasticity}}.$$

TABLE A.2: ACCOUNTING RESULTS: ROBUSTNESS

	Unconditional Semi-Elasticity	Conditional Semi-Elasticity	Accounting Share
(1) Baseline	0.112	0.015	0.865
	(0.001)	(0.001)	
(2) Within Sector	0.050	0.001	0.985
	(0.001)	(0.001)	
(3) Country and Decade FE	0.039	0.001	0.968
	(0.005)	(0.003)	
(4) Men	0.082	0.003	0.965
	(0.001)	(0.001)	
(5) Women	0.160	0.029	0.816
	(0.002)	(0.001)	
(6) Literacy Score	0.118	-0.002	1.021
	(0.002)	(0.002)	
(7) No Education, Health, Public Admin	0.102	0.033	0.678
	(0.001)	(0.001)	
(8) Private Sector Only	0.114	0.027	0.763
	(0.001)	(0.001)	

Notes: The Table shows the results of the accounting exercises described in the text. Rows 1-5 and 7-8 use data from IPUMS International, while Row 6 uses data from PIAAC and STEP.

Table A.2 displays the results. In the baseline case, the unconditional semi-elasticity (shown in Figure A.1) is 0.112, while the semi-elasticity conditional on education is 0.015. This implies that variation in the aggregate supply of skills accounts for 87 percent of the cross-country correlation between white-collar employment share and development. Rows (2)–(6) show that human capital retains a large accounting role within sectors, within countries over time, separately by gender, and when skills are measured using literacy scores (as discussed in Appendix A.3.1).

Some of the shift toward white-collar workers is due to the growth of occupations such as doctors or teachers that are not involved in the administrative coordination of firms. Our baseline analysis deals with this by grouping most such workers into a high-skill service sector that plays little role in our analysis. As a complementary analysis, we also explore excluding such workers entirely. Rows (7)–(8) show that excluding sectors such as education or health or focusing only on the private sector does not materially change our main result.

A.4 Educational Attainment in Developing Countries

This appendix shows that educational attainment in many developing countries today is low relative to the levels that prevailed in the United States during the Second Industrial Revolution. This comparison helps explain why low educational attainment may remain an important impediment to the growth of medium and large firms in developing countries today.

The U.S. Second Industrial Revolution is conventionally dated to 1870–1914. We measure the educational attainment of workers by birth cohort using the full count 1940 U.S. Census, the first to collect such data nationally (Ruggles et al., 2025*b*). For cohorts who would have been ages 15–24 in 1870, the average secondary completion rate was 11 percent; for cohorts who would have been ages 15–24 in 1914, the secondary completion rate had risen to 22 percent.[25]

We compare this to contemporary educational attainment data from Barro and Lee (2013). We focus on educational attainment of people ages 15–24; older workers generally have lower educational attainment. We also focus on countries with a 2019 GDP per capita (rgdpe/pop) of less than $5,000 in the Penn World Tables 10.01 (Feenstra, Inklaar and Timmer, 2015). Among the 33 countries below this threshold with education data, 8 have secondary completion rates among young workers below 11 percent and 20 have secondary completion rates below 22 percent. The median country in this set achieved an 11 percent secondary completion rate only in 2010, and had not yet achieved a secondary completion rate of 22 percent as of the most recent data.

B Model Appendix

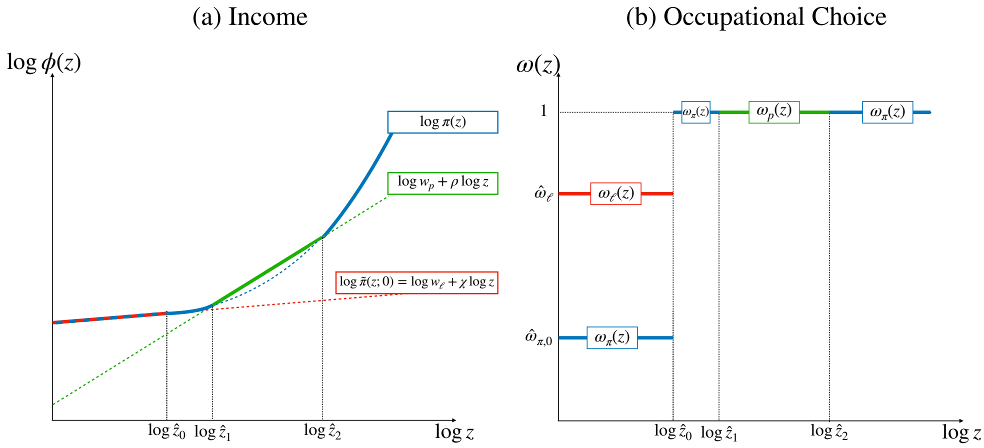
This section contains additional results referred to in the text as well as proofs of selected results.

[25] Using responses from older workers in the census introduces some survivorship bias, but is unavoidable given the lack of earlier data on educational attainment.

B.1 Visualizations of Alternative Occupational Choice Rules

Lemma 3 provides a general characterization of occupational choices. Figure 4b visualizes one possible income schedule and choice rule with the feature that $\hat{z}_0 = \hat{z}_1$. In this case, all modern entrepreneurs are more skilled than all professionals. An alternative case that can arise in equilibrium features $\hat{z}_0 < \hat{z}_1$, such that some modern entrepreneurs are less skilled than all professionals. Figure B.1 shows the income schedule and occupational choices for this case.

FIGURE B.1: ALTERNATIVE OCCUPATIONAL CHOICE RULES



B.2 Proofs of Selected Results

Proof of Lemma 1.

Substituting in the expression for $\tilde{\tau}$ from Assumption 1, the profit maximization problem of an entrepreneur with skill z is given by

$$\max_{\{n_p(i)\}_{i\in[0,1]},\, n_\ell \geqslant 0} \; zA \exp\Big(\int_0^1 \log \tilde{n}(i)^{(1-\tau)\gamma_p}\, di\Big) n_\ell^{\gamma_\ell} \;-\; w_p \int_0^1 n_p(i)\, di \;-\; w_\ell n_\ell. \tag{B.1}$$

We note that the optimal allocation $\{n_p(i)\}$ must satisfy a cutoff rule, with a uniform choice $n_p(i) = n_p$ for $i \leqslant q$ and $n_p(i) = 0$ otherwise. The equal allocation follows from the concavity of the aggregator, and since $a(i)$ is decreasing in i, the firm optimally chooses to professionalize the highest-productivity tasks, which corresponds to the interval $[0, q]$.

Substituting this expression for $n(i)$ into the objective function means that we obtain a choice over q, n_p, and n_ℓ,

$$\max_{q\in[0,1],\, n_p\geqslant 0,\, n_\ell\geqslant 0} \; zA\exp\Big(\int_0^q \log a(i)^{(1-\tau)\gamma_p}\, di\Big)\, n_p^{(1-\tau)\gamma_p q}\, n_\ell^{\gamma_\ell} \;-\; qw_p n_p \;-\; w_\ell n_\ell.$$

This can be reparameterized into

$$\max_{q\in[0,1],\, n_p\geqslant 0,\, n_\ell\geqslant 0} \; z\tilde{A}(q)\Big[n_p^{\alpha(q)} n_\ell^{1-\alpha(q)}\Big]^{\eta(q)} \;-\; qw_p n_p \;-\; w_\ell n_\ell,$$

with $\tilde{A}(q) = A\exp\Big(\int_0^q \log a(i)^{(1-\tau)\gamma_p}\, di\Big)$, $\alpha(q) = \frac{(1-\tau)\gamma_p q}{\eta(q)}$ and $\eta(q) = (1-\tau)\gamma_p q + \gamma_\ell$.

Proof of Lemma 2.

The expressions (4) and (5) imply that total profits satisfy

$$\pi(q,z) = (1-\eta(q))\left[z\tilde{A}(q)\right]^{\frac{1}{1-\eta(q)}}\left[\left(\frac{\gamma_p(1-\tau)}{w_p}\right)^{\alpha(q)}\left(\frac{\gamma_\ell}{w_\ell}\right)^{1-\alpha(q)}\right]^{\frac{\eta(q)}{1-\eta(q)}}.$$

For each z, the entrepreneur chooses q to maximize this function, and we write

$$f(q,z) \equiv \log\pi(q,z)$$

for the log profit function, which is twice-differentiable whenever $q < 1$.[26] The properties of the solution can be derived from properties of the first and second derivatives of $f(q,z)$. We start by establishing these derivatives, and then use our findings to establish the properties of the optimal solution.

1. Properties of derivatives of f. The profit expression can be written as

$$f(q,z) \equiv \log\pi(q,z) = \log[1-\eta(q)] + \frac{1}{1-\eta(q)}\left[\log z + \Phi(q)\right],$$

where $\eta(q) = \gamma_\ell + q(1-\tau)\gamma_p$ and

$$\Phi(q) = \log\tilde{A}(q) - \eta(q)\log(w_\ell/\gamma_\ell) - q(1-\tau)\gamma_p\log\frac{w_p/[(1-\tau)\gamma_p]}{w_\ell/\gamma_\ell}.$$

[26]For $q = 0$, we define the derivatives of f as right-derivatives.

Using that we have $\eta''(q) = 0$, we can derive

$$f_q(q, z) = \frac{1}{1-\eta(q)} \left[-\eta'(q) + \frac{\eta'(q)}{1-\eta(q)} [\log z + \Phi(q)] + \Phi'(q) \right], \tag{B.2}$$

$$f_{qq}(q, z) = \frac{1}{1-\eta(q)} \left[\Phi''(q) + \frac{\eta'(q)^2}{1-\eta(q)} + 2\eta'(q) f_q(q, z) \right], \tag{B.3}$$

$$f_{qz}(q, z) = \frac{\eta'(q)}{(1-\eta(q))^2} \frac{1}{z}. \tag{B.4}$$

In addition, one critical property is that given our assumption on θ, f does not have any local minima. In particular, for all z and $q \in [0, 1)$,

$$f_q(q, z) = 0 \quad \Longrightarrow \quad f_{qq}(q, z) < 0. \tag{B.5}$$

That is, if there is any stationary point of f, it has to be a local maximum.[27]

2. Existence of a cutoff $\hat{z}_q$. To establish the existence of a cutoff, we first note from (B.2) that for $q = 0$, the partial derivative $f_q(0, z)$ is strictly increasing in z, going from $-\infty$ to $+\infty$, which means that there is a unique value $\hat{z}_q$ such that

$$f_q(0, \hat{z}_q) = 0.$$

The value $\hat{z}_q$ is our candidate cutoff point. To verify this, consider the firm's choice of q.

For any $z \leqslant \hat{z}_q$, we have that $f_q(0, z) \leqslant 0$. This means that at $q = 0$, the profit function is either decreasing (for $z < \hat{z}_q$) or is momentarily flat before decreasing (for $z = \hat{z}_q$, since the second derivative is negative by (B.5)). Since (B.5) also implies that there are no interior local minima where the function could turn back to achieve a higher value, the profit $f(q, z)$ for any $q > 0$ cannot exceed its value at the origin. Therefore, the optimal choice must be the corner solution $q(z) = 0$.

Conversely, for any $z > \hat{z}_q$, we have $f_q(0, z) > 0$. A strictly positive derivative at $q = 0$ means that profits can be increased by choosing a small positive q. Thus, $q = 0$ is no longer optimal, and the firm will choose an interior solution $q(z) > 0$. Hence, $\hat{z}_q$

[27] To derive this, we note that

$$f_{qq}(q^*, z) < 0 \iff \frac{(1-\tau)^2\gamma_p^2}{1 - q(1-\tau)\gamma_p - \gamma_\ell} < \frac{\theta(1-\tau)}{1-q} \iff \frac{1-q}{1 - q\frac{(1-\tau)\gamma_p}{1-\gamma_\ell}} \frac{(1-\tau)\gamma_p^2}{1-\gamma_\ell} < \theta,$$

where we use that $\Phi''(q) = (\log \tilde{A})''(q) = -\frac{\theta(1-\tau)}{1-q}$. Since $(1-\tau)\gamma_p + \gamma_\ell < 1$, the factor $(1-\tau)\gamma_p/(1-\gamma_\ell) < 1$, and the left-hand side is maximized at $q = 0$. Hence, $f_{qq} < 0$ under our assumption that $\theta > (1-\tau)\gamma_p^2/(1-\gamma_\ell)$.

has the desired properties.

3. Organizational choice $q(z)$ is monotonic in z. We now show that the optimal choice $q(z)$ is increasing in skill z. The result is immediate when comparing any $z_1 \leqslant \hat{z}_q$ with any $z_2 > \hat{z}_q$. The non-trivial case is to show that $q(z)$ is strictly increasing for all $z > \hat{z}_q$.

To establish this, we first use the fact that (B.2) implies an Inada condition at $q = 1$: the marginal profit tends to minus infinity as we approach the upper bound of professionalization: $\lim_{q \to 1} f_q(q, z) = -\infty$. In particular, in that expression, we have

$$\Phi'(q) = (1 - \tau)\theta \log(1 - q) + C_\Phi,$$

where C_Φ is a constant independent of q. Since $\log(1 - q) \to -\infty$ as $q \to 1$ and all other terms in the derivative are bounded, the overall derivative tends to $-\infty$.

This Inada condition ensures that the solution is never at the corner $q = 1$. Therefore, for any $z > \hat{z}_q$, we have a strictly interior solution, $q(z) \in (0, 1)$. This interior optimum must satisfy the first-order condition $f_q(q(z), z) = 0$. Since equation (B.5) rules out interior local minima, any point satisfying the first-order condition is the unique global maximum. Totally differentiating the first-order condition with respect to z, we obtain

$$f_{qq}(q, z) \cdot q'(z) + f_{qz}(q, z) = 0 \implies q'(z) = \frac{f_{qz}(q, z)}{-f_{qq}(q, z)}.$$

We know from (B.5) that $f_{qq} < 0$. Together with the complementarity between q and z, $f_{qz} > 0$, from (B.4), we have $q'(z) > 0$.[28] Hence $q(z)$ is strictly monotonic for $z > \hat{z}_q$.

4. Solution $q(z)$ tends to $q = 1$ as $z \to \infty$. To establish the limiting behavior of the organizational choice, we note the following property of the derivative $f_q(q, z)$ in (B.2):

$$\lim_{z \to \infty} f_q(q, z) = \infty. \tag{B.6}$$

This property says that for a given organizational structure $q < 1$, the marginal profit of increasing q becomes arbitrarily large as skill goes to infinity.

To see why this implies $\lim_{z \to \infty} q(z) = 1$, consider any small $\epsilon > 0$. From (B.6), there must exist some skill level $\bar{z}$ such that for all $z > \bar{z}$, the marginal profit is positive even at $q = 1 - \epsilon$; that is, $f_q(1 - \epsilon, z) > 0$. Since we established that the profit function has a single peak, a positive slope at $1 - \epsilon$ implies the optimal choice $q(z)$ must lie to

[28] We assume here that $q(z)$ is differentiable. The result that $q(z)$ is non-decreasing follows more generally from the supermodularity of f, which does not require differentiability.

the right of this point. Hence, for any arbitrarily small $\epsilon > 0$, we can find a $\bar{z}$ such that for all $z > \bar{z}$, we have $q(z) > 1 - \epsilon$, which is the formal definition of $\lim_{z\to\infty} q(z) = 1$.

5. Value of cut-off point $\hat{z}_q$. The cut-off point $\hat{z}_q$ solves

$$f_q(0, \hat{z}_q) = 0 \iff 0 = \left[-\eta'(0) + \frac{\eta'(0)}{1-\eta(0)} \left[\log \hat{z}_q + \Phi(0)\right] + \Phi'(0) \right].$$

Using that $\eta(0) = \gamma_\ell$ and $\eta'(0) = (1-\tau)\gamma_p$, as well as

$$\begin{aligned} \Phi(0) &= \log A - \gamma_\ell \log(w_\ell/\gamma_\ell), \\ \Phi'(0) &= (1-\tau)\log\beta - (1-\tau)\gamma_p \log w_\ell/\gamma_\ell - (1-\tau)\gamma_p \log\left(\frac{w_p/[(1-\tau)\gamma_p]}{w_\ell/\gamma_\ell}\right), \end{aligned}$$

we can manipulate the equation to arrive at

$$\log \hat{z}_q = (1-\gamma_\ell)\left[1 - \log(1-\tau) + \log\frac{w_p/\gamma_p}{w_\ell/\gamma_\ell} - \frac{1}{\gamma_p}\log\beta\right] + \log\frac{w_\ell/\gamma_\ell}{A}$$

as desired.

6. Elasticity of output with respect to z. For $z \leqslant \hat{z}_q$, we can explicitly solve

$$y(z) = z^{\frac{1}{1-\gamma_\ell}} A^{\frac{1}{1-\gamma_\ell}} \left(\frac{w_\ell}{\gamma_\ell}\right)^{-\gamma_\ell/(1-\gamma_\ell)},$$

which yields $\partial \log y(z)/\partial \log z = 1/(1-\gamma_\ell)$ as in the Lemma.

For $z > \hat{z}_q$, we have an interior solution, and the Cobb-Douglas production function implies that

$$\pi(q(z), z) = (1 - \eta(q(z)))y(q(z), z). \tag{B.7}$$

Furthermore, the envelope theorem and the optimality of $q(z)$ implies[29]

$$\frac{\partial \pi(q(z), z)}{\partial z} = \frac{y(q(z), z)}{z}, \qquad \frac{\partial \pi(q(z), z)}{\partial q} = 0.$$

Thus, differentiating (B.7) with respect to z yields

$$\frac{y}{z} = -(1-\tau)\gamma_p y \frac{dq}{dz} + (1-\eta)\frac{dy}{dz} \quad \Longrightarrow \quad \frac{d\log y}{d\log z} = \frac{1 + (1-\tau)\gamma_p \frac{dq}{d\log z}}{1 - \gamma_\ell - \gamma_p(1-\tau)q(z)},$$

[29]For the envelope theorem, note that we can write $\pi(z, q) = \max_{n_l, n_p}[z f(n_l, n_p, q) - w_\ell n_\ell - q w_p n_p]$, and use that the first-order effect of changing z is simply $f(n_l^*, n_p^*, q^*) = y/z$.

where we use $1 - \eta(q) = 1 - \gamma_\ell - q\gamma_p(1-\tau)$, divide both sides with y/z, and rearrange.

Proof of Lemma 3

I. Cutoff Result. We normalize the payoffs for professionals and entrepreneurs relative to the blue-collar worker's payoff. The relative payoffs are

$$\tilde{P}(\log z) := (\alpha_p - \alpha_w) + (\rho - \chi)\log z,$$
$$\tilde{E}(\log z) := \log \pi(z) - (\alpha_w + \chi \log z).$$

An agent with ability z chooses the occupation corresponding to $\max\{0, \tilde{P}(\log z), \tilde{E}(\log z)\}$.

Below, we document a number of facts about the curves $\tilde{P}$ and $\tilde{E}$ and the restrictions that need to hold in equilibrium, which we then use to derive the result.

1. **Shape of the curves $\tilde{P}$ and $\tilde{E}$.** The professional payoff is simply a linear curve with slope $\rho - \chi$. For $\tilde{E}$, 2 implies a slope $1/(1-\gamma_\ell-(1-\tau)\gamma_p q) - \chi$. Since $1/(1-\gamma_\ell) = \chi$, the slope is zero for $z < \hat{z}_q$ where $q = 0$. For $z > \hat{z}_q$, the slope is positive and strictly increasing, with an asymptotic slope of $1/(1-\gamma_\ell-(1-\tau)\gamma_p) - \chi$ which is steeper than the slope of the professional curve $\rho - \chi$. We write $\hat{z}_{e,\ell}$ for the largest z for which $\tilde{E}(\log z) \leqslant 0$.

2. **Positive supply of both laborers and professionals.** In equilibrium, there is a positive measure of modern entrepreneurs, since for sufficiently large z, we are above the cutoff $\hat{z}_q$ and $\tilde{E}$ grows faster than $\tilde{P}$, implying that modern entrepreneurship is the preferred choice. Since modern firms demand both professionals and laborers, labor market clearing requires that these are in positive supply in equilibrium.

3. **No traditional firms with $\tilde{E}(\log z^*) > 0$.** Since the profit elasticity with respect to ability is the same for traditional entrepreneurs and laborers, the payoff of traditional entrepreneurship relative to blue-collar work is constant. If this constant were positive, all individuals would strictly prefer traditional entrepreneurship to being a laborer, implying zero supply of laborers, which violates market clearing.

4. **Almost all modern firms have $\tilde{E}(\log z) > 0$.** First, there are no modern firms where $E(\log z) < 0$, since this would be dominated by being a blue-collar worker. Second, a modern firm needs to have $\log z > \log \hat{z}_q$, and since the entrepreneur's payoff function $\tilde{E}$ is steeper than χ when $\log z > \log \hat{z}_q$, there can at most be a point where $\tilde{E}(\log z) = 0$, so almost all modern firms have $\tilde{E}(\log z) > 0$.

5. **Single Intersection $\hat{z}_{p,\ell}$ of Professional and Laborer Payoffs.** The professional's relative payoff, $\tilde{P}(\log z)$, is an upward-sloping line, as the condition $\rho > \chi$ ensures that professional wages grow faster with ability than those of laborers. We write $\hat{z}_{p,\ell}$ for the unique intersection of $\tilde{P}$ with the horizontal line, giving the point of indifference between being a professional and a laborer.

6. **Exactly Two Intersections $\hat{z}^1_{p,e}, \hat{z}^2_{p,e}$ of Professional and Entrepreneurial Payoffs.** The professional line $\tilde{P}$ starts out steeper than the entrepreneurial curve $\tilde{E}$ for small $\log z$, since $\rho > 1/(1-\gamma_\ell)$, but ends up less steep at large $\log z$, since $\rho < \frac{1}{1-\gamma_\ell-\gamma_p}$. Hence, the professional payoff line begins and ends below the entrepreneurial payoff curve. This configuration allows for zero, one (tangency), or two intersections.[30] However, for the professional labor market to clear, a positive measure of individuals must choose this occupation. This equilibrium condition rules out the zero- and one-crossing cases, as they would leave no region where becoming a professional is the optimal choice. We therefore conclude that the two curves intersect exactly twice, at points we denote $\hat{z}^1_{p,e} < \hat{z}^2_{p,e}$.

7. **Payoff at Second Intersection $\hat{z}^2_{p,e}$ Strictly Dominates Laborer Payoff 0.** The payoff at the second intersection of the professional and entrepreneurial curves has to satisfy $\tilde{E}(\log \hat{z}^2_{p,e}) = \tilde{P}(\log \hat{z}^2_{p,e}) > 0$ to ensure that some z choose to be professionals. Otherwise, being professional would be dominated by laborers for $z \leqslant \hat{z}^2_{p,e}$, and would be dominated by entrepreneurship for $z > \hat{z}^2_{p,e}$.

The sorting cutoffs can now be constructed. We set the highest cutoff at $\hat{z}_2 = \hat{z}^2_{p,e}$. For all $z \geqslant \hat{z}_2$, entrepreneurship is the optimal choice, and it needs to be modern since payoffs exceed 0. The other cutoffs, $\hat{z}_0$ and $\hat{z}_1$, depend on the ordering of the remaining intersection points. Two cases arise:

1. **Case 1:** $\hat{z}^1_{p,e} \leqslant \hat{z}_{p,\ell}$. The professional line intersects the entrepreneur line at or before it intersects the zero line. We set $\hat{z}_0 = \hat{z}_1 = \hat{z}_{p,\ell}$. In this case, individuals with $z \in (\hat{z}_1, \hat{z}_2)$ choose to be professionals, while individuals with $z \leqslant \hat{z}_0$ choose to be either laborers or traditional entrepreneurs.[31]

2. **Case 2:** $\hat{z}_{p,\ell} < \hat{z}^1_{p,e}$. The professional line crosses the zero line strictly before its first intersection with the entrepreneur's line. Since the professional line has

[30] To see that there can be at most two crossings, we note that the difference between the two functions, $\tilde{P}(\log z) - \tilde{E}(\log z)$, is strictly concave. To have more than two crossings would require a local minimum, which is not possible for a strictly concave function.

[31] Modern entrepreneurship is precluded since $\tilde{E}(\log z) \leqslant 0$ in this interval. This follows since $\tilde{E}(\hat{z}_0) \leqslant 0$ and $\tilde{E}$ is weakly increasing.

to be strictly below the entrepreneur's line when it crosses zero, the entrepreneur line has to be strictly above 0 at $\hat{z}_{p,\ell}$, which implies that the entrepreneur crossed 0 at $\hat{z}_{e,\ell} < \hat{z}_{p,\ell}$. We set $\hat{z}_0 = \hat{z}_{e,\ell}$ and $\hat{z}_1 = \hat{z}^1_{p,e}$. As before, those with $z \leqslant \hat{z}_0$ are laborers or traditional entrepreneurs. However, for $z \in (\hat{z}_0, \hat{z}_1]$, individuals choose to be modern entrepreneurs, since $\tilde{E}(\log z) > 0$ and $\tilde{E}(\log z) > \tilde{P}(\log z)$ in this interval.

This construction confirms that there exist cutoffs $\hat{z}_0 \leqslant \hat{z}_1 < \hat{z}_2$ such that individuals with $z \leqslant \hat{z}_0$ are laborers or traditional entrepreneurs, those with $z \in (\hat{z}_1, \hat{z}_2)$ are professionals, and those with $z \in (\hat{z}_0, \hat{z}_1]$ or $z \geqslant \hat{z}_2$ are modern entrepreneurs, proving the result.

II. Equilibrium Incomes

- To ensure there are any blue-collar workers, traditional entrepreneurship cannot have a higher payoff than blue-collar work. Hence, $\pi(z, 0) \leqslant w_\ell z^\chi$. Furthermore, if $\omega_\pi(z)\mathbb{I}_{z \leqslant \hat{z}_q} > 0$, at least one z chooses traditional entrepreneurship, in which case we get equality.

- **Case 1:** $\hat{z}_0 = \hat{z}_1$**.**
  This case corresponds to the scenario where the set of modern entrepreneurs is fully above the set of professionals. The first occupational transition, occurring at $\hat{z}_0$, is directly from worker to professional. At this margin, an agent must be indifferent between the two, yielding the condition

  $$w_\ell \hat{z}_0^\chi = w_p \hat{z}_0^\rho.$$

  The second cutoff, $\hat{z}_2$, is defined by the point where professionals and modern entrepreneurs yield the same payoff, hence

  $$w_p \hat{z}_2^\rho = \pi(\hat{z}_2).$$

- **Case 2:** $\hat{z}_0 < \hat{z}_1$**.**
  In this case, the optimal choice as z increases is to first switch from blue-collar work to modern entrepreneurship, then to professional, and subsequently back to modern entrepreneurship.

  – At the first cutoff, $\hat{z}_0$, individuals are indifferent between working and becoming a modern entrepreneur. This gives the equality: $w_\ell \hat{z}_0^\chi = \pi(\hat{z}_0)$.

- At the second cutoff, $\hat{z}_1$, the marginal individual is indifferent between being an entrepreneur and becoming a professional, leading to $\pi(\hat{z}_1) = w_p \hat{z}_1^\rho$.
- Finally, at the highest cutoff, $\hat{z}_2$, the choice is again between professional and entrepreneur roles, with high-ability entrepreneurs dominating. The indifference condition is: $w_p \hat{z}_2^\rho = \pi(\hat{z}_2)$.

III: Existence of traditional entrepreneurs. The market-clearing condition for blue-collar workers is

$$\int_{\underline{z}}^{\hat{z}_0} \omega_L(z) z^\chi dG(z) = \int_{\underline{z}}^{\infty} n_\ell(z)\omega_\pi(z) dG(z),$$

where $n_\ell(z)$ is the labor demand from a firm of type z. The integral for labor supply (left-hand side) is capped at $\hat{z}_0$ because we have that $\omega_L(z) = 0$ for $z \geqslant \hat{z}_0$. By substituting $\omega_L(z) = 1 - \omega_\pi(z)$ for $z \leqslant \hat{z}_0$, we can rearrange the market-clearing condition as

$$\int_{\underline{z}}^{\hat{z}_0} z^\chi dG(z) = \int_{\underline{z}}^{\hat{z}_0} \omega_\pi(z)[z^\chi + n_\ell(z)] dG(z) + \int_{\hat{z}_0}^{\infty} n_\ell(z)\omega_\pi(z) dG(z),$$

where the left-hand side is the total supply of efficiency units for $z \leqslant \hat{z}_0$, and the right-hand side is the total demand for such individuals, which comes from traditional entrepreneurs and their labor demand, as well as the labor demand from modern entrepreneurs.

From this, we observe that $\omega_\pi(z) > 0$ on the interval $[\underline{z}, \hat{z}_0)$ is equivalent to

$$\int_{\underline{z}}^{\hat{z}_0} z^\chi dG(z) > \int_{[\hat{z}_0,\hat{z}_1)\cup[\hat{z}_2,\infty)} n_\ell(z)\omega_\pi(z) dG(z),$$

where we use that, for $z \geqslant \hat{z}_0$, $\omega_\pi(z)$ is only positive on $[\hat{z}_0, \hat{z}_1) \cup [\hat{z}_2, \infty)$, which is the result.

Proof of Proposition 1

Our proof strategy is to do a guess-and-verify approach where we propose an equilibrium for a given κ, where cutoffs and wages are independent of κ, and then verify that it is an equilibrium if κ is sufficiently small. We then show that comparative statics with respect to κ have the desired form, and that there exists a $\hat{\kappa}$ such that duality disappears.

Production and profit functions. With $\theta = 0$, the optimal organizational choice is either fully traditional ($q = 0$) or fully modern ($q = 1$), since the productivity from professionalizing all tasks is the same. Substituting in $q = 0$ and $q = 1$ into the expressions for $\alpha(q), \eta(q)$, and $\tilde{A}(q)$, we obtain the following two production functions for modern and traditional entrepreneurs:

$$y_0(n_\ell; z) = Azn_\ell l^{\gamma_\ell}, \qquad y_1(n_\ell, n_p; z) = A\beta^{1-\tau} z n_\ell^{\gamma_\ell} n_p^{\gamma_p(1-\tau)}.$$

Standard profit maximization with Cobb-Douglas production functions yields:

$$\pi(z, q = 0) = z^{\frac{1}{1-\gamma_\ell}}(1-\gamma_\ell)A^{\frac{1}{1-\gamma_\ell}}\left(\frac{w_\ell}{\gamma_\ell}\right)^{-\gamma_\ell/(1-\gamma_\ell)},$$

$$\pi(z; q = 1) = (1-\gamma_\ell-(1-\tau)\gamma_p)\left[zA\beta^{1-\tau}\left(\frac{w_\ell}{\gamma_\ell}\right)^{-\gamma_\ell}\left(\frac{w_p}{(1-\tau)\gamma_p}\right)^{-(1-\tau)\gamma_p}\right]^{\frac{1}{1-\gamma_\ell-(1-\tau)\gamma_p}}.$$

Indifference conditions. In our posited equilibrium, there is a single skill cutoff, $\hat{z}$. This candidate equilibrium requires that for all individuals with $z \leqslant \hat{z}$, the returns to being a laborer and a traditional entrepreneur are equal, and for all individuals with $z > \hat{z}$, the returns to being a professional and a modern entrepreneur are equal. At the cutoff, individuals are indifferent between all four occupations. This yields three indifference conditions:

$$\begin{aligned} w_\ell z^\chi &= \pi(z, q = 0) \quad z \leqslant \hat{z}. \\ w_p z^\rho &= \pi(z, q = 1) \quad z > \hat{z}, \\ w_\ell \hat{z}^\chi &= w_p \hat{z}^\rho. \end{aligned}$$

Substituting in the profit expressions, using that $\chi = 1/(1-\gamma_\ell)$ and $\rho = 1/(1-\gamma_\ell-(1-\tau)\gamma_p)$, we obtain the following expressions for the wages and the productivity cut-off

$$w_\ell = A(1-\gamma_\ell)^{1-\gamma_\ell}\gamma_\ell^{\gamma_\ell}, \tag{B.8}$$

$$w_p^{1-\gamma_\ell} w_\ell^{\gamma_\ell} = A\beta^{1-\tau}(1-\gamma_\ell-(1-\tau)\gamma_p)^{1-\gamma_\ell-(1-\tau)\gamma_p}\gamma_\ell^{\gamma_\ell}[\gamma_p(1-\tau)]^{\gamma_p(1-\tau)}, \tag{B.9}$$

$$\hat{z} = \left(\frac{w_p}{w_\ell}\right)^{-\frac{1}{\rho-\chi}}. \tag{B.10}$$

By substituting (B.8) into (B.9), we can solve for w_p and thus w_p/w_ℓ purely in terms of the production parameters $A, \gamma_\ell, \gamma_p, \beta, \tau$, which also lets us calculate $\hat{z}$ from (B.10). Note that the skill distribution does not show up in any of these expressions.

Market clearing. The previous derivations established wages and a cutoff, $\hat{z}$, such that individuals optimally choose their occupations. To finalize the equilibrium, we must allocate individuals on either side of this cutoff to wage work or entrepreneurship and verify that markets clear.

To this end, let Z_{occ} denote the aggregate skill-weighted units in each occupation (e.g., $Z_{e,m}$ for modern entrepreneurs, Z_p for professionals). An individual's contribution is z^ρ if they are in a high-skill role ($z > \hat{z}$) and z^χ if they are in a low-skill role ($z \leqslant \hat{z}$). The market clearing conditions for the total supply of these units are

$$Z_{e,m} + Z_p = \int_{\hat{z}}^{\infty} z^\rho dG_\kappa(z), \tag{B.11}$$

$$Z_{\ell,m} + (Z_{e,t} + Z_{\ell,t}) = \int_0^{\hat{z}} z^\chi dG_\kappa(z), \tag{B.12}$$

where $Z_{\ell,m}$ and $Z_{\ell,t}$ denote blue-collar workers in modern and traditional firms respectively, and $Z_{e,t}$ denotes traditional entrepreneurs.

Since profits in modern entrepreneurship scale with z^ρ, we define profits per efficiency unit as $\tilde{\pi}_m \equiv \pi(z)/z^\rho$. Similarly, for traditional entrepreneurship, $\tilde{\pi}_t \equiv \pi(z)/z^\chi$ for $z \leqslant \hat{z}$. Both $\tilde{\pi}_m$ and $\tilde{\pi}_t$ are constant for all z within their respective domains. Given the Cobb-Douglas production functions, the following ratios of total payments hold:

$$\begin{aligned}
\frac{Z_{e,m}\tilde{\pi}_m}{Z_p w_p} &= \frac{1-\gamma_\ell-(1-\tau)\gamma_p}{\gamma_p(1-\tau)}, \\
\frac{Z_{e,m}\tilde{\pi}_m}{Z_{\ell,m} w_\ell} &= \frac{1-\gamma_\ell-(1-\tau)\gamma_p}{\gamma_\ell}, \\
\frac{Z_{e,t}\tilde{\pi}_t}{Z_{\ell,t} w_\ell} &= \frac{1-\gamma_\ell}{\gamma_\ell}.
\end{aligned}$$

The indifference conditions ($\tilde{\pi}_m = w_p$ and $\tilde{\pi}_t = w_\ell$) simplify these expressions. For the high-skilled group, this directly determines the allocation of efficiency units,

$$Z_{e,m} = \left(\frac{1-\gamma_\ell-(1-\tau)\gamma_p}{1-\gamma_\ell}\right)\int_{\hat{z}}^{\infty} z^\rho dG_\kappa(z), \tag{B.13}$$

$$Z_p = \left(\frac{\gamma_p(1-\tau)}{1-\gamma_\ell}\right)\int_{\hat{z}}^{\infty} z^\rho dG_\kappa(z), \tag{B.14}$$

showing that the aggregate units in modern entrepreneurship and professional work are fixed shares of the total high-skill supply. For the low-skilled individuals, the demand for laborers from modern firms, $Z_{\ell,m}$, is pinned down by $Z_{e,m}$,

$$Z_{\ell,m} = Z_{e,m} \frac{\gamma_\ell}{1 - \gamma_\ell - (1-\tau)\gamma_p} \left(\frac{w_p}{w_\ell} \right),$$

where we recall that the wage ratio w_p/w_ℓ is fully pinned down by primitives in (B.8)-(B.9). The remaining low-skilled units form the traditional sector, which exists as long as the residual supply of low-skilled workers is positive. These units are split between traditional entrepreneurs and laborers in proportion to their income shares,

$$Z_{e,t} = (1 - \gamma_\ell) \left[\int_0^{\hat{z}} z^\chi dG_\kappa(z) - Z_{\ell,m} \right],$$

$$Z_{\ell,t} = \gamma_\ell \left[\int_0^{\hat{z}} z^\chi dG_\kappa(z) - Z_{\ell,m} \right].$$

Duality exists if and only if the term in the brackets is positive, which happens for a sufficiently low κ, which we assume holds in our case. This concludes the equilibrium construction.

Effect of increasing $\hat{\kappa}$. The skill distribution parameter κ only enters through the integrals that determine the total supply of skill-weighted units; it does not affect wages or the cutoff $\hat{z}$ provided duality still exists. An increase in κ instead raises the total units of modern entrepreneurs ($Z_{e,m}$) and thus the laborers they demand ($Z_{\ell,m}$). This shrinks the residual supply for the traditional sector, implying there is a threshold $\hat{\kappa}$ above which duality disappears.

Proof of corollary 1

In the regime where $\kappa > \hat{\kappa}$ so that there is no duality, there are two changes in the equilibrium conditions. First, we remove the requirement (B.8) that low-skill individuals are indifferent between traditional entrepreneurship and blue-collar work. Second, we require that the blue-collar employment of modern firms equal the supply of unskilled workers, i.e., setting $Z_{\ell,t} + Z_{e,t} = 0$ in (B.12).

New equilibrium conditions. To pin down the skilled wage premium $\frac{w_p}{w_\ell}(\kappa)$ and the cut-off $\hat{z}(\kappa)$ as functions of κ, we use the following equilibrium equations:

$$\frac{w_p}{w_\ell}(\kappa) = \hat{z}(\kappa)^{-(\rho-\chi)}, \tag{B.15}$$

$$Z_{e,m} = \left(\frac{1-\gamma_\ell-(1-\tau)\gamma_p}{1-\gamma_\ell}\right)\int_{\hat{z}(\kappa)}^{\infty} z^\rho dG_\kappa(z), \tag{B.16}$$

$$\int_0^{\hat{z}(\kappa)} z^\chi dG_\kappa(z) = Z_{e,m}\left[\frac{\gamma_\ell}{1-\gamma_\ell-(1-\tau)\gamma_p}\frac{w_p}{w_\ell}(\kappa)\right]. \tag{B.17}$$

The first two equations are the same as in the dual economy equilibrium: indifference between blue-collar work and professional work and an expression of the measure of modern entrepreneurs in terms of primitives and the supply of high-skilled workers. The last equation is new, stating that blue-collar employment in modern firms equal the low-skilled supply.

Constructing and verifying equilibrium. To solve the system, we first note that at the boundary $\kappa = \hat{\kappa}$, the solution must be the one from the dual economy, which we denote $\hat{z}_{dual}$ and $(w_p/w_\ell)_{dual}$. This follows from the definition of $\hat{\kappa}$ as the precise skill level at which the blue-collar labor market clears at the dual-economy prices without a traditional sector.

For $\kappa > \hat{\kappa}$, we guess and verify that the equilibrium is a simple scaling of this solution,

$$\log \hat{z}(\kappa) = \log \hat{z}_{dual} + \log(\kappa/\hat{\kappa}),$$
$$\log(w_p/w_\ell)(\kappa) = \log(w_p/w_\ell)_{dual} - (\rho-\chi)\log(\kappa/\hat{\kappa}).$$

It is immediate that this solution respects the indifference between blue-collar work and professional work at the cut-off. For the market clearing condition, we use the structure of the skill distribution, $G_\kappa(z) = G(z/\kappa)$. Specifically, a change of variables ($u = z/\kappa$) shows that for any power ξ, the aggregate skill supply integral scales directly with κ,

$$\int_a^\infty z^\xi dG_\kappa(z) = \int_a^\infty z^\xi \frac{1}{\kappa} g\left(\frac{z}{\kappa}\right) dz = \kappa^\xi \int_{a/\kappa}^\infty u^\xi dG(u).$$

Under our proposed solution, the rescaled cutoff $\hat{z}(\kappa)/\kappa$ is constant and equal to $\hat{z}_{dual}/\hat{\kappa}$. Consequently, the high-skill supply integral scales with κ^ρ, the low-skill supply integral scales with κ^χ. When these terms are substituted into the new market-clearing condition (B.17), we obtain a power κ^χ on the left-hand side and a power κ^ρ on the right-hand side, which perfectly cancels the $-(\rho-\chi)$ power coming from the wage premium.

Last, given solutions for the cut-off and the skilled wage premium, we can recover the wage levels w_p, w_ℓ from the indifference condition between professional work and entrepreneurship (B.9) and the measure of professionals Z_p from (B.14), since both of these conditions hold in the economy without duality.

Properties of equilibrium. From the definition of the equilibrium, we immediately see that the cut-off type satisfies $\log \hat{z}(\kappa/\hat{\kappa}) = \log(\kappa/\hat{\kappa}) + \log \hat{z}$ as stated in the corollary, as well as the skilled wage premium being decreasing with κ. Furthermore, since the cut-off scales proportionally with the probability distribution, the share of workers above the threshold is constant, meaning that the share of white-collar workers is constant even though the conditional probability declines conditional on a skill level z. Last, the share of individuals who are entrepreneurs stays constant, since this is only pinned down by the share of individuals above the threshold and production parameters, both of which are fixed. Thus, average firm size stays the same.

B.3 Aggregate output effect

Aggregate output and efficiency. Consider the planner's problem in the economy without the distortion τ. Let $\omega_\pi(z)$, $\omega_\ell(z)$, and $\omega_p(z)$ denote the shares of type-z individuals who become entrepreneurs, laborers, and professionals, respectively. Aggregate output is

$$Y = \int_0^\infty \omega_\pi(z) f\big(n_\ell(z), q(z)n_p(z); z, q(z)\big) g(z)\, dz$$

subject to

$$\omega_\pi(z) + \omega_\ell(z) + \omega_p(z) \leqslant 1 \qquad \text{for all } z, \tag{B.18}$$

and the two labor-market resource constraints

$$\int_0^\infty \omega_\pi(z) n_\ell(z) g(z)\, dz \leqslant \int_0^\infty \omega_\ell(z) z^\chi g(z)\, dz,$$

$$\int_0^\infty \omega_\pi(z) q(z) n_p(z) g(z)\, dz \leqslant \int_0^\infty \omega_p(z) z^\rho g(z)\, dz.$$

Absent the distortion τ, any decentralized equilibrium allocation solves the planner's problem. If λ_l and λ_p are the Lagrangian multipliers on the constraints of laborers and professionals, the planner and the firms' first order condition implies

$$w_l = \lambda_l \qquad w_p = \lambda_p,$$

and the envelope theorem yields that the first-order change in output from a perturbation in density $g \to g + \partial g$ is given by

$$\begin{aligned} dY &= \int_0^\infty [\omega_\pi f + \lambda_l(\omega_\ell z^\chi - \omega_\pi n_l) + \lambda_p(\omega_p(z) z^\rho - \omega_\pi q(z) n_p(z))](\partial g(z))dz = \\ &= \int_0^\infty [\omega_\pi(z)\pi(z) + \omega_\ell(z) w_\ell z^\chi + \omega_p(z) w_p z^\rho](\partial g(z))dz \\ &= \int_0^\infty \phi(z)(\partial g(z))dz, \end{aligned}$$

where the second row uses $\lambda_l = w_l$ and $\lambda_p = w_p$, and the last row uses that workers only have $\omega > 0$ on the choices that maximize $\omega_\pi(z)\pi(z) + \omega_\ell(z) w_\ell z^\chi + \omega_p(z) w_p z^\rho$.

Effect of distortions. When there are distortions $\tau > 0$, the formula above is modified in two ways. First, a positive wedge means that marginal products are generically larger than factor payments, which means that $\phi(z)$ understates the marginal product of increasing the supply of a type s. Second, if wedges are heterogeneous across different producers, the allocation of workers across activities with different wedges has a first-order effect on output.

To illustrate this, consider the model version in Section 4.3 with $q \in \{0, 1\}$, and assume that the wedge takes a simplified form of a fixed tax $\hat{\tau}$ on modern firms (rather than depending on the exact level of n_p).[32] Then, the first-order effect of changing human capital on output can be shown to satisfy

$$\begin{aligned} dY = &\int_{\underline{z}}^{\hat{z}} z^\chi w_L \left[1 + \hat{\tau}\, \omega_{l,\text{mod}}\right] \partial g(z) dz + \int_{\hat{z}}^\infty z^\rho w_P \left[1 + \hat{\tau}\right] \partial g(z) dz \\ &+ \hat{\tau}\, d\omega_{l,\text{mod}} \int_{\underline{z}}^{\hat{z}} z^\chi w_L\, g(z) dz, \end{aligned} \tag{B.19}$$

where $\omega_{l,\text{mod}}$ is the share of efficiency units below the low types that are allocated to blue-collar work in modern firms.

In equation (B.19), the first two terms are the composition effect of the change in $g(z)$, holding fixed the allocation of low-type efficiency units across traditional and modern firms. For low types, the average output value of an efficiency unit is $w_L(1 + \hat{\tau}\, \omega_{\ell,\text{mod}})$, since a share $\omega_{\ell,\text{mod}}$ works in distorted modern firms and the remainder works in undistorted traditional production. For high types, the corresponding value is $w_P(1 + \hat{\tau})$, since high types work only in the modern sector in this simplified economy.

[32] When the wedge depends on n_p, the distortion on professional workers is higher than on entrepreneurs and blue-collar workers, since the private marginal product of professionals is depressed by the wedge increasing with n_p.

The last term is the contribution from reallocation of low-type efficiency units. When $\hat{\tau} > 0$, a low-type efficiency unit has a higher social marginal product in modern firms than in traditional production, so output rises when $\omega_{\ell,\text{mod}}$ increases. There is no analogous wedge term from reallocating low types between traditional entrepreneurship and blue-collar work in traditional firms, since both activities are undistorted. Nor is there an additional wedge term from reallocating high types between professional work and modern entrepreneurship, since both activities take place within the same distorted modern sector.

Proof of efficiency for the decentralized economy. Let (w_ℓ^d, w_p^d) denote the decentralized equilibrium wages, and define the profit of a type-z entrepreneur evaluated at arbitrary firm choices as

$$\Pi(n_\ell, n_p, q; z, w_\ell^d, w_p^d) \equiv f(n_\ell, qn_p; z, q) - w_\ell^d n_\ell - w_p^d q n_p.$$

Let

$$\Pi^d(z) \equiv \max_{q, n_\ell, n_p} \Pi(n_\ell, n_p, q; z, w_\ell^d, w_p^d)$$

be the maximized profit of a type-z entrepreneur at decentralized prices. For any feasible allocation $\{\widehat{\omega}_\pi, \widehat{\omega}_\ell, \widehat{\omega}_p, \widehat{n}_\ell, \widehat{n}_p, \widehat{q}\}$, output satisfies

$$\begin{aligned}
\widehat{Y} &= \int_0^\infty \widehat{\omega}_\pi(z) \left[\Pi(\widehat{n}_\ell(z), \widehat{n}_p(z), \widehat{q}(z); z, w_\ell^d, w_p^d) + w_\ell^d \widehat{n}_\ell(z) + q(z) w_p^d \widehat{n}_p(z)\right] g(z)\, dz \\
&\leqslant \int_0^\infty \widehat{\omega}_\pi(z) \Pi^d(z) g(z)\, dz + w_\ell^d \int_0^\infty \widehat{\omega}_\ell(z) z^\chi g(z)\, dz + w_p^d \int_0^\infty \widehat{\omega}_p(z) z^\rho g(z)\, dz \\
&= \int_0^\infty \left[\widehat{\omega}_\pi(z)\Pi^d(z) + \widehat{\omega}_\ell(z) w_\ell^d z^\chi + \widehat{\omega}_p(z) w_p^d z^\rho\right] g(z)\, dz \\
&\leqslant \int_0^\infty \phi^d(z) g(z)\, dz \\
&= \int_0^\infty \left[\omega_\pi^d(z)\Pi^d(z) + \omega_\ell^d(z) w_\ell^d z^\chi + \omega_p^d(z) w_p^d z^\rho\right] g(z)\, dz \\
&= Y^d,
\end{aligned}$$

where

$$\phi^d(z) = \max\left\{\Pi^d(z),\, w_\ell^d z^\chi,\, w_p^d z^\rho\right\}$$

is the equilibrium payoff of a type-z individual. The first inequality uses profit maximization by firms and the feasibility of the labor allocations. The second inequality uses the fact that each type can earn at most $\phi^d(z)$, while the decentralized allocation assigns each type to an activity attaining this maximum. The final equality follows from labor-market clearing: aggregate output equals entrepreneurial profits plus wage payments.

B.4 Land Market Interpretation of Agricultural Distortion

Our model features more sharply decreasing returns in agriculture in poor countries. A substantial literature has emphasized that frictions in land markets are a key constraint on the ability of productive farmers to scale up (Adamopoulos and Restuccia, 2020; Adamopoulos et al., 2022; Chen, Restuccia and Santaeulàlia-Llopis, 2023). Here, we show how sharper decreasing returns arise naturally in models where land markets are closed.

Formally, consider a setup in which farmers produce output using hired labor and land,

$$f_a(n_\ell, m; z) = (A_a z n_\ell^{\gamma_l})^{(1-\gamma_m)} m^{\gamma_m}, \tag{B.20}$$

where m denotes land input. Further, consider two alternative setups for determining land use and land market clearing: (i) an economy with land rental markets, and (ii) an economy with communal land allocation. In the economy with land rental, we assume that land rents are rebated to farmers and workers in agriculture in proportion to their profit and labor income. The farm problems and market clearing conditions in the two setups are

Rental Markets

$$\max_{n_\ell, m} \tilde{\pi}(z) = p_a f_a(n_\ell, m) - \tilde{w}_\ell n_\ell - rm$$

$$M = \int_z \omega_{\pi,ag}(z) m(z) dG(z)$$

Communal Land Allocation

$$\max_{n_\ell} \pi(z) = p_a f_a(n_\ell, \bar{m}) - w_\ell n_\ell$$

$$M = \bar{m} \times \int_z \omega_{\pi,ag}(z) dG(z). \tag{B.21}$$

In the rental market case, farms are free to rent land at a price r which is determined by market clearing, while in the case with communal land allocation, farms are allocated a fixed amount of land $\bar{m}$ with the size determined by an equal allocation rule. In the rental case, we write $\tilde{\pi}(z)$ and $\tilde{w}_l$ for profits and wages not including rebates. The income received by farmers and workers satisfy

$$\pi(z) = \tilde{\pi}(z) + rM \times \frac{\tilde{\pi}(z)}{\int_z \omega_{pi,ag}(z)[\tilde{\pi}(z) + \tilde{w}_l n_l(z)]dG(z)}, \tag{B.22}$$

$$w_l = \tilde{w}_l + rM \times \frac{\tilde{w}_l}{\int_z \omega_{pi,ag}(z)[\tilde{\pi}(z) + \tilde{w}_l n_l(z)]dG(z)}, \tag{B.23}$$

capturing that aggregate land rents are redistributed in proportion to factor incomes.

The following proposition shows that the reduced form of closing land markets is to make decreasing returns to scale sharper.

PROPOSITION B.1. *In the model with a rental market, farm output, labor input, and profits $y(z)$, $n_l(z)$, and $\pi(z)$ satisfy the following equations*

$$\begin{aligned} n_l(z), y(z) &= \arg\max_{n_l, y} \{\pi(z) = py(z) - w_l n_l\}, \qquad s.t. \quad y = \tilde{A}_a z n_l^{\gamma_l} \\ \tilde{A}_a &= A_a (M/Y_a)^{\gamma_m/(1-\gamma_m)}, \end{aligned}$$

where Y_a is aggregate agricultural output. Under communal land ownership, the corresponding equations are

$$\begin{aligned} n_l(z), y(z) &= \arg\max_{n_l, y} \{\pi(z) = py(z) - w_l n_l\} \qquad s.t. \quad y = (\tilde{A}_a z n_l^{\gamma_l})^{1-\gamma_m} \\ \tilde{A}_a &= A(M/\Omega_\pi)^{\frac{\gamma_m}{1-\gamma_m}}, \end{aligned}$$

where $\Omega_a = \int_z \omega_{ag,z}(z) dG(z)$ is the aggregate measure of farmers.

We provide a proof at the end of this section. The implication of the proposition is that farm outcomes in the model with land can be represented by an equivalent formulation with no explicit land choice, where the role of land is embedded in productivity through a term $\tilde{A}_A$. This term reflects a congestion effect arising from decreasing returns at the aggregate level. In the land rental market case, this congestion operates through higher equilibrium rental prices, which reduce optimal land holdings. Since rental rates are proportional to output, congestion is a function of aggregate output Y_a. In the communal regime, the congestion effect arises because more farmers share the land. Since this effect depends on the measure of farmers, it is a function of $\Omega_a = \int_z \omega_{ag,z}(z) dG(z)$.

This proposition implies that even though our baseline model does not explicitly include land, calibration can be viewed as recovering a composite agricultural productivity term. This interpretation is exact for the simplified land economy above — in which case our calibrated A_A would correspond to the adjusted productivities $\tilde{A}_A$ – and motivates our reduced-form implementation of λ as a barrier to farm scale.

The main inconsistency with a literal land interpretation arises in our counterfactual exercises, since we abstract from how changes in land availability induced by structural transformation affect effective productivity $\tilde{A}_A$. This effect is present, because in a model with land, $\tilde{A}_A$ is not a primitive productivity parameter. Instead, it depends on equilibrium objects such as total agricultural output and the measure of farmers in agriculture. For calibration, this is not problematic: the procedure recovers the composite $\tilde{A}_A$, which can be inverted to obtain primitive productivities given equilibrium allocations. However, in counterfactuals, these equilibrium objects change, implying that $\tilde{A}_A$ should also be adjusted under a land-based interpretation.

Proof of Proposition B.1. The expression for communal ownership is obtained by substituting in the market clearing condition for land in the household problem.

For the case with rental markets, we first note that standard first-order conditions imply that

$$rm(z) = \gamma_m py(z), \quad \tilde{w}_l n_l(z) = \gamma_l(1-\gamma_m)py(z), \quad \tilde{\pi}(z) = (1-\gamma_m)(1-\gamma_l)py(z).$$

Substituting these into (B.22) and (B.23) implies

$$w_l = \frac{\tilde{w}_l}{1-\gamma_m}, \tag{B.24}$$

$$\pi(z) = (1-\gamma_l)py(z). \tag{B.25}$$

Furthermore, using two-stage optimization of the farmer problem and substituting in $m(z) = \gamma_m py(z)/r$ we have that the optimal choice of hired labor and output is

$$n_l(z), y(z) \in \arg\max_{n_l, y}(1-\gamma_m)py - \tilde{w}_l n_l \quad s.t. \quad y = (A_a z n_l^{\gamma_l})^{1-\gamma_m}(\gamma py/r)^{\gamma}_m$$

Furthermore, the market clearing condition for land implies $r = \gamma_m pY_a/M$. Substituting in this expression, and noting that dividing the objective with $1-\gamma_m$ does not change the optimal choice, we have that farmers' choices are given by

$$n_l(z), y(z) \in \arg\max_{n_l, y} py - w_l n_l \quad s.t. \quad y = A_a(M/Y_a)^{\frac{\gamma_m}{1-\gamma_m}} z n_l^{\gamma_l}.$$

Furthermore, together with the result above that profits are given by $\pi(y) = (1-\gamma_l)py(z)$, we obtain the result in the proposition that farm output, labor choices, and profits are given by the solution of a profit maximization problem over only output and labor, given a production function $\tilde{A}_a z n_l^{\gamma_l}$ with $\tilde{A}_a = A_a(M/Y_a)^{\frac{\gamma_m}{1-\gamma_m}}$.

C Estimation Appendix

This appendix documents the quantitative calibration underlying Section 5. Appendix C.1 reports the full set of targeted moments, model counterparts, deflators, and weights.

Appendix C.2 provides evidence that the internally calibrated parameters are well disciplined by these moments.

C.1 Moments, Weighting, and Model Fit

In Section 5.2 we use figures to display the targets and model fit. Here, we provide further details on the construction of these moments. We also describe how we weight the various moments.[33]

At a high level, we target 125 moments in total: 20 moments on occupational shares by education and in the aggregate (Table C.1); 20 moments on sectoral shares by education and in the aggregate (Table C.2); 16 moments on educational shares by sector (Table C.3); 16 moments on occupational shares within sector (Panels A-D in Table C.4); 15 moments on the distribution of firm size by sector (Panels E-G in Table C.4); 30 moments on white-collar employment share by firm size (Panels H-J in Table C.4) and 8 moments on wage gaps between and within education groups (Panels K-L in Table C.4).

Sources and Samples. The moments are constructed using a combination of the census data from IPUMS International and the labor force survey database. All moments that do not involve firm size are constructed using the census data, taking the latest available year for each country. All moments involving firm size are constructed using the LFS data. Since LFS samples are generally smaller, we pool all available years within a country and study the country-level average. Wage moments are based on the combination of 11 cross-sections from IPUMS International and 22 countries from the LFS with the necessary wage data (we use IPUMS International for countries covered in both).

Aggregation. We construct an internally consistent set of moments for a fictional "average" middle-income country. The basic idea is to estimate each moment as a function of log GDP per capita across countries, and then evaluate the fitted relationship at the average GDP per capita of middle-income countries in the IPUMS International sample.

For the moments based on IPUMS International, we first estimate multinomial logistic regressions for the employment shares in each education $\times$ white-collar status $\times$ sector cell. We also estimate the self-employment share conditional on each education

[33] The tables follow the structure and ordering of the figures in the main text, and refer to the corresponding panels. The one exception is Table C.3, which does not have a corresponding figure, since it is simply a different normalization of Table C.2.

$\times$ white-collar status $\times$ sector cell. In both cases, the regressors are a quadratic polynomial in log GDP per capita. We evaluate the fitted values at the average middle-income GDP per capita and aggregate the resulting cells to construct the moments in Tables C.1–C.4, as well as the education shares in Table 1. The wage moments are constructed in the same way, using quadratic fits in log GDP per capita evaluated at the same GDP level.

For the firm-size moments in Panels E–G of Table C.4, we combine IPUMS International and LFS data. Average firm size is inferred from IPUMS as the inverse of the sector-level self-employment rate, and is fitted as a function of log GDP per capita as above. The shares of workers in medium and large firms are computed from the LFS as multinomial-logit fits on a quadratic polynomial in log GDP per capita within each education $\times$ white-collar status $\times$ sector cell, evaluated at the same average middle-income GDP per capita and aggregated using the IPUMS International cell distribution.

Finally, Panels H–J of Table C.4 use LFS data to construct moments on white-collar employment by firm size. For Panels H and I, we collapse firm size into two groups: firms with 1–10 workers and firms with 11 or more workers. We then regress white-collar employment shares on indicators for these firm-size groups interacted with sector. We use the estimated coefficients to predict white-collar shares by sector and firm-size group, and then combine these predictions using employment weights to obtain the moments reported in the table. For Panel J, we use the subset of LFS countries with more detailed firm-size categories. We pool these countries, regress the white-collar share within each firm-size group on a cubic polynomial in log firm size, controlling for sector and country fixed effects, and predict white-collar shares for 20 firm-size groups, from 5 to 100 workers in increments of 5.

Weighting and Fit. As noted above, Tables C.1-C.4 display the exact moments, data value, and model value for the moments used in our calibration. As is standard, we minimize a weighted sum of squared errors. Conceptually, we would like our measure of the error to be the percentage error in fitting each moment. However, we face the usual practical problem that for moments whose data value is close to 0, this percentage error is not well-defined. To avoid this, we instead use a moment-specific deflator, which is generally the average of all data values within a category. For example, in Table C.1, Panel A, we assign the deflator of 0.214 to all of the first four moments; this is simply the average of the data for these first four moments, which avoids assigning excess weight to the (low) value of the share of traditional entrepreneurs with tertiary education.

We also weight moments to ensure that different economic concepts contribute comparably to the overall fit. For instance, we want our estimation to give similar importance to matching the relationship between white-collar employment and firm size (Panel J of Table C.4) and to matching the wage premium for tertiary-educated workers (row 3 of Panel K of Table C.4). Yet the former concept is represented by about twenty separate moments, while the latter corresponds to just one. To balance this, we assign each moment in Panel C a weight of 0.05, so that the entire group sums to one. This logic guides our broader weighting scheme: we group moments by economic concept and allocate equal total weight across groups, based on our assessment of their relevance for the model's overall fit.

While some arbitrariness is unavoidable—since the model is overidentified and not all moments can be matched perfectly—most choices are straightforward. The only departure from equal weighting across groups is that we assign greater overall weight to patterns involving tertiary-educated individuals. For example, in Table C.2, we treat "employment shares across sectors by education" as one concept (implying a per-moment weight of 0.062, since 16×0.062=1), and we treat "employment shares across sectors for tertiary-educated individuals" as an additional concept, which raises the weight for the tertiary-educated rows to 0.312 (= 0.250 + 0.062). The final two columns of each table report the deflator and the corresponding weight applied to each moment.

TABLE C.1: OCCUPATIONAL SHARES

Panel A. Traditional Entrepreneur					**Panel B. Blue-Collar Workers**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
Trad. entr. (No prim.)	0.315	0.374	0.214	0.250	Blue-collar (No prim.)	0.570	0.563	0.424	0.250
Trad. entr. (Prim.)	0.299	0.257	0.214	0.250	Blue-collar (Prim.)	0.566	0.607	0.424	0.250
Trad. entr. (Sec.)	0.174	0.172	0.214	0.250	Blue-collar (Sec.)	0.449	0.426	0.424	0.250
Trad. entr. (Tert.)	0.039	0.052	0.214	1.250	Blue-collar (Tert.)	0.147	0.101	0.424	1.250
Trad. entr. (Overall)	0.226	0.219	0.219	1.000	Blue-collar (Overall)	0.473	0.475	0.475	1.000
Panel C. White-Collar Workers					**Panel D. Modern Entrepreneur**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
White-collar (No prim.)	0.095	0.040	0.311	0.250	Modern entr. (No prim.)	0.019	0.024	0.051	0.250
White-collar (Prim.)	0.113	0.107	0.311	0.250	Modern entr. (Prim.)	0.022	0.029	0.051	0.250
White-collar (Sec.)	0.322	0.356	0.311	0.250	Modern entr. (Sec.)	0.055	0.046	0.051	0.250
White-collar (Tert.)	0.689	0.740	0.311	1.250	Modern entr. (Tert.)	0.125	0.106	0.051	1.250
White-collar (Overall)	0.255	0.262	0.262	1.000	Modern entr. (Overall)	0.046	0.044	0.044	1.000

Notes: Def. is the moment deflator and Wt. is the weight in the minimum-distance objective.

TABLE C.2: SECTORAL SHARES

Panel A. Agriculture					**Panel B. Manufacturing**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
Agr (No prim.)	0.338	0.354	0.250	0.062	Mfg (No prim.)	0.235	0.247	0.250	0.062
Agr (Prim.)	0.301	0.193	0.250	0.062	Mfg (Prim.)	0.245	0.294	0.250	0.062
Agr (Sec.)	0.094	0.110	0.250	0.062	Mfg (Sec.)	0.256	0.227	0.250	0.062
Agr (Tert.)	0.006	0.032	0.250	0.312	Mfg (Tert.)	0.124	0.138	0.250	0.312
Agr (Overall)	0.199	0.168	0.250	1.000	Mfg (Overall)	0.231	0.244	0.250	1.000
Panel C. Services (Low-Skill)					**Panel D. Services (High-Skill)**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.

Sectoral Shares (continued)

Ser. LS (No prim.)	0.305	0.269	0.250	0.062	Ser. HS (No prim.)	0.121	0.130	0.250	0.062
Ser. LS (Prim.)	0.313	0.325	0.250	0.062	Ser. HS (Prim.)	0.141	0.188	0.250	0.062
Ser. LS (Sec.)	0.299	0.322	0.250	0.062	Ser. HS (Sec.)	0.351	0.341	0.250	0.062
Ser. LS (Tert.)	0.132	0.182	0.250	0.312	Ser. HS (Tert.)	0.738	0.648	0.250	0.312
Ser. LS (Overall)	0.284	0.297	0.250	1.000	Ser. HS (Overall)	0.285	0.290	0.250	1.000

Notes: Def. is the moment deflator and Wt. is the weight in the minimum-distance objective.

TABLE C.3: EDUCATIONAL SHARES, WITHIN SECTOR

Panel A. Agriculture					**Panel B. Manufacturing**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
No prim. in Agr	0.248	0.308	0.250	0.062	No prim. in Mfg	0.149	0.148	0.250	0.062
Prim. in Agr	0.592	0.450	0.250	0.062	Prim. in Mfg	0.415	0.472	0.250	0.062
Sec. in Agr	0.156	0.216	0.250	0.062	Sec. in Mfg	0.366	0.307	0.250	0.062
Tert. in Agr	0.004	0.025	0.250	0.062	Tert. in Mfg	0.070	0.073	0.250	0.062
Panel C. Services (Low-Skill)					**Panel D. Services (High-Skill)**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
No prim. in Ser. LS	0.157	0.132	0.250	0.062	No prim. in Ser. HS	0.062	0.066	0.250	0.062
Prim. in Ser. LS	0.433	0.429	0.250	0.062	Prim. in Ser. HS	0.194	0.254	0.250	0.062
Sec. in Ser. LS	0.349	0.359	0.250	0.062	Sec. in Ser. HS	0.407	0.389	0.250	0.062
Tert. in Ser. LS	0.061	0.080	0.250	0.062	Tert. in Ser. HS	0.337	0.291	0.250	0.062

Notes: Def. is the moment deflator and Wt. is the weight in the minimum-distance objective.

TABLE C.4: DIFFERENCES IN WITHIN-SECTOR ORGANIZATION AND WAGES

Panel A. Traditional Entrepreneur					**Panel B. Blue-Collar Workers**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
Trad. entr. in Agr	0.545	0.596	0.256	0.250	Blue-collar in Agr	0.441	0.377	0.472	0.250
Trad. entr. in Mfg	0.137	0.154	0.256	0.250	Blue-collar in Mfg	0.677	0.665	0.472	0.250
Trad. entr. in Ser. LS	0.279	0.235	0.256	0.250	Blue-collar in Ser. LS	0.498	0.528	0.472	0.250
Trad. entr. in Ser. HS	0.023	0.039	0.256	0.250	Blue-collar in Ser. HS	0.306	0.317	0.472	0.250
Panel C. White-Collar Workers					**Panel D. Modern Entrepreneur**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
White-collar in Agr	0.014	0.020	0.233	0.250	Modern entr. in Agr	0.000	0.008	0.039	0.250
White-collar in Mfg	0.142	0.152	0.233	0.250	Modern entr. in Mfg	0.045	0.029	0.039	0.250
White-collar in Ser. LS	0.162	0.172	0.233	0.250	Modern entr. in Ser. LS	0.062	0.065	0.039	0.250
White-collar in Ser. HS	0.608	0.587	0.233	0.250	Modern entr. in Ser. HS	0.064	0.056	0.039	0.250
Panel E. Average Firm Size					**Panel F. Employment in 10+ Firms**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
Avg. size (Agr)	1.835	1.658	5.230	1.000	11+ Firms (Agr)	0.000	0.034	0.084	0.125
Avg. size (Mfg)	5.502	5.472	5.230	1.000	11+ Firms (Mfg)	0.133	0.119	0.084	0.125
Avg. size (Ser. LS)	2.940	3.329	5.230	1.000	11+ Firms (Ser. LS)	0.104	0.089	0.084	0.125
Avg. size (Ser. HS)	11.542	10.460	5.230	1.000	11+ Firms (Ser. HS)	0.242	0.190	0.084	0.125
Avg. size (Overall)	3.677	3.802	3.802	1.000	11+ Firms (Overall)	0.480	0.432	0.336	0.500
Panel G. Employment in 51+ Firms					**Panel H. WC Share, 1–10 Firms**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
51+ Firms (Agr)	0.000	0.017	0.084	0.125	WC % 1–10 (Agr)	0.014	0.013	0.206	0.250
51+ Firms (Mfg)	0.118	0.074	0.084	0.125	WC % 1–10 (Mfg)	0.068	0.100	0.206	0.250
51+ Firms (Ser. LS)	0.095	0.042	0.084	0.125	WC % 1–10 (Ser. LS)	0.079	0.164	0.206	0.250
51+ Firms (Ser. HS)	0.116	0.107	0.084	0.125	WC % 1–10 (Ser. HS)	0.368	0.546	0.206	0.250
51+ Firms (Overall)	0.328	0.240	0.336	0.500	WC % 1–10 (Overall)	0.076	0.182	0.326	0.500
Panel I. WC Share, 11+ Firms					**Panel L. Wage Dispersion**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
WC % 11+ (Agr)	0.107	0.084	0.364	0.250	SD log wages (No prim.)	0.521	0.619	0.619	0.250
WC % 11+ (Mfg)	0.196	0.266	0.364	0.250	SD log wages (Prim.)	0.532	0.601	0.601	0.250
WC % 11+ (Ser. LS)	0.303	0.408	0.364	0.250	SD log wages (Sec.)	0.684	0.651	0.651	0.250
WC % 11+ (Ser. HS)	0.651	0.695	0.364	0.250	SD log wages (Tert.)	0.857	0.710	0.710	0.250
WC % 11+ (Overall)	0.449	0.470	0.326	0.500	SD log wages (Overall)	0.623	0.635	0.635	1.000
Panel K. Wage Premia					**Panel J. WC Shares by firm size**				

Differences in Within-Sector Organization and Wages (continued)

Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
Wage premium (Prim.)	0.035	0.131	0.646	1.000	WC % (firm size 5)	-0.090	-0.116	0.028	0.050
Wage premium (Sec.)	0.366	0.526	0.646	1.000	WC % (firm size 10)	-0.057	-0.062	0.028	0.050
Wage premium (Tert.)	1.122	1.280	0.646	1.000	WC % (firm size 15)	-0.038	-0.034	0.028	0.050
Panel J. WC Shares by firm size					**Panel J. WC Shares by firm size**				
Moment	Model	Data	Def.	Wt.	Moment	Model	Data	Def.	Wt.
WC % (firm size 20)	-0.025	-0.017	0.028	0.050	WC % (firm size 65)	0.023	0.026	0.028	0.050
WC % (firm size 25)	-0.015	-0.005	0.028	0.050	WC % (firm size 70)	0.026	0.027	0.028	0.050
WC % (firm size 30)	-0.007	0.004	0.028	0.050	WC % (firm size 75)	0.028	0.027	0.028	0.050
WC % (firm size 35)	-0.001	0.010	0.028	0.050	WC % (firm size 80)	0.030	0.028	0.028	0.050
WC % (firm size 40)	0.005	0.015	0.028	0.050	WC % (firm size 85)	0.032	0.027	0.028	0.050
WC % (firm size 45)	0.009	0.019	0.028	0.050	WC % (firm size 90)	0.034	0.027	0.028	0.050
WC % (firm size 50)	0.013	0.022	0.028	0.050	WC % (firm size 95)	0.036	0.027	0.028	0.050
WC % (firm size 55)	0.017	0.024	0.028	0.050	WC % (firm size 100)	0.037	0.026	0.028	0.050
WC % (firm size 60)	0.020	0.025	0.028	0.050					

Notes: Def. is the moment deflator and Wt. is the weight in the minimum-distance objective.

C.2 Identification

This section reports two exercises that provide evidence on the identification of the model and the moments most informative for each parameter.

Single-Peaked Minimum Distance Function. We first examine the shape of the minimum distance function to verify that our calibration corresponds to both a local and global minimum. We vary one parameter at a time while holding all others fixed to assess local identification, then vary all parameters jointly over a broad support to assess global identification. In practice, for this second exercise, we draw 8,000,000 parameter vectors spanning a wide range around our estimates.[34]

Figure C.2 summarizes the results. Each panel corresponds to one parameter. The x-axis reports alternative values of that parameter, and the y-axis reports the value of the minimum distance objective. The weights in the objective function are normalized so that vertical distances can be interpreted as increases in the average squared percentage deviation of the moments from their targets. For example, a value of one corresponds to a 1% increase in the squared deviation, or a 10% increase in the average percentage deviation.

Within each panel, the blue line shows the model fit when only the chosen parameter is varied, and thus tests for a local optimum. The red line shows the model fit when all parameters are allowed to vary simultaneously, and thus tests for a global optimum. The y-axes are kept on the same scale across panels to make clear which parameters have larger or smaller effects on the model fit.[35] The figure shows that the parameters

[34] Given the high dimensionality of the parameter space, even 8,000,000 draws cannot fully cover it. To increase precision, we draw more densely near the estimated values.

[35] As a result, the lines are cropped in some panels.

are well identified both locally and globally: each curve has a clear minimum at the calibrated value, and the fit deteriorates as we move away from it. As expected, the global fit is always better, that is, closer to the best fit, than the local fit, since allowing all parameters to adjust improves the overall fit.

Jacobian Matrix. While Figure C.2 confirms that parameters are well identified, it does not reveal which moments are most informative for each parameter. To explore this mapping, Figure C.1 presents a normalized Jacobian matrix summarizing how each parameter affects each targeted economic concept. For each parameter, we select the moment that, according to our heuristic mapping discussed in the main text, should be most directly related to it.

Moments and parameters are ordered so that the moment in the first row corresponds to the parameter in the first column, the second to the second, and so on.

Most moments are self-explanatory. The only two requiring clarification are the sorting variables. “Sorting into modern entrepreneurship" compares the propensity of more-educated workers to enter modern entrepreneurship with the propensity of less-educated workers to enter traditional entrepreneurship.[36] “Sorting into high-skill services" is defined analogously, focusing on employment in high-skill services rather than agriculture.

We use the 8,000,000 model evaluations described above to compute all targeted moments for each parameter vector. For each pair of moment and parameter, we then run a simple univariate regression and store the resulting slope, which captures the sensitivity of that moment to the parameter. All slopes are normalized so that the absolute values sum to one along both rows and columns, using a Sinkhorn balancing algorithm. The resulting normalized matrix is displayed in Figure C.1.

In interpreting the magnitudes in Figure C.1, note that in a hypothetical case where each parameter affects only its corresponding moment, the diagonal elements of the matrix would all equal one. Conversely, if all parameters affected all moments equally, every cell would equal $1/13 \approx 0.077$. In practice, we find that diagonal entries are substantially higher than off-diagonal ones and reach the maximum for each parameter, indicating that the targeted moment is indeed the one most closely associated with that parameter. A similar pattern holds across rows: for nearly every moment, the parameter identified heuristically as its key determinant is also the most influential empirically.

[36] Formally, we compute the difference between the share of individuals with secondary or higher education who are modern entrepreneurs and the share of individuals with primary or less education who are modern entrepreneurs, plus the difference between the share of individuals with secondary or higher education who are traditional entrepreneurs and the share of those with less than primary education who are traditional entrepreneurs.

The only notable exception is sorting into high-skill services, which is also strongly affected—unsurprisingly—by the relative utility parameters for agriculture and services.

Overall, this exercise validates the heuristic mapping between parameters and moments described in the main text.

FIGURE C.1: JACOBIAN MATRIX TO VERIFY IDENTIFICATION ARGUMENT

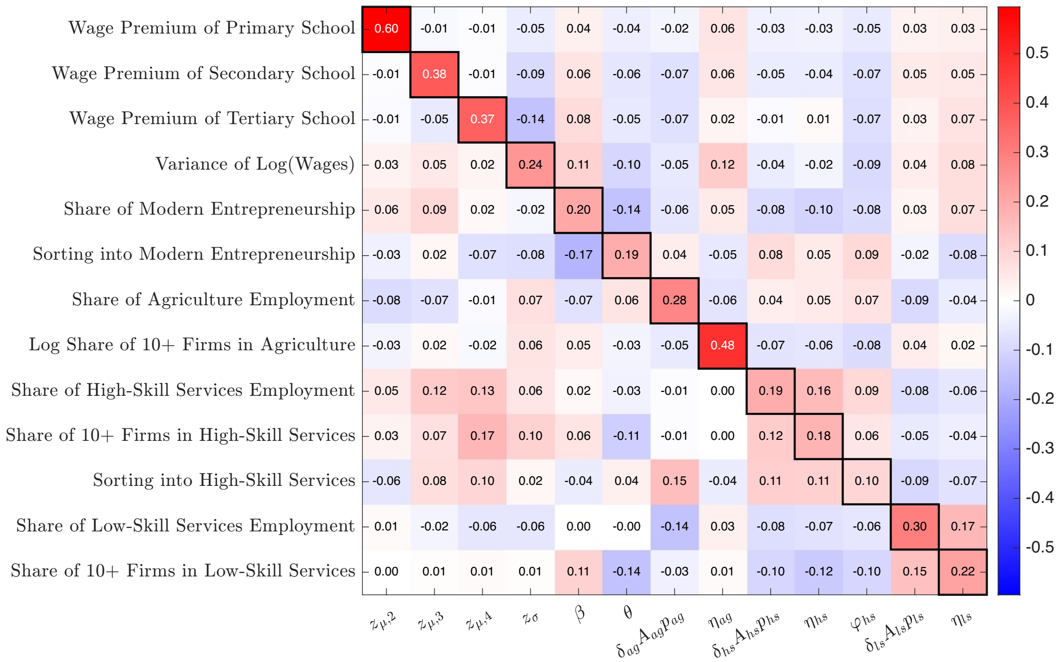
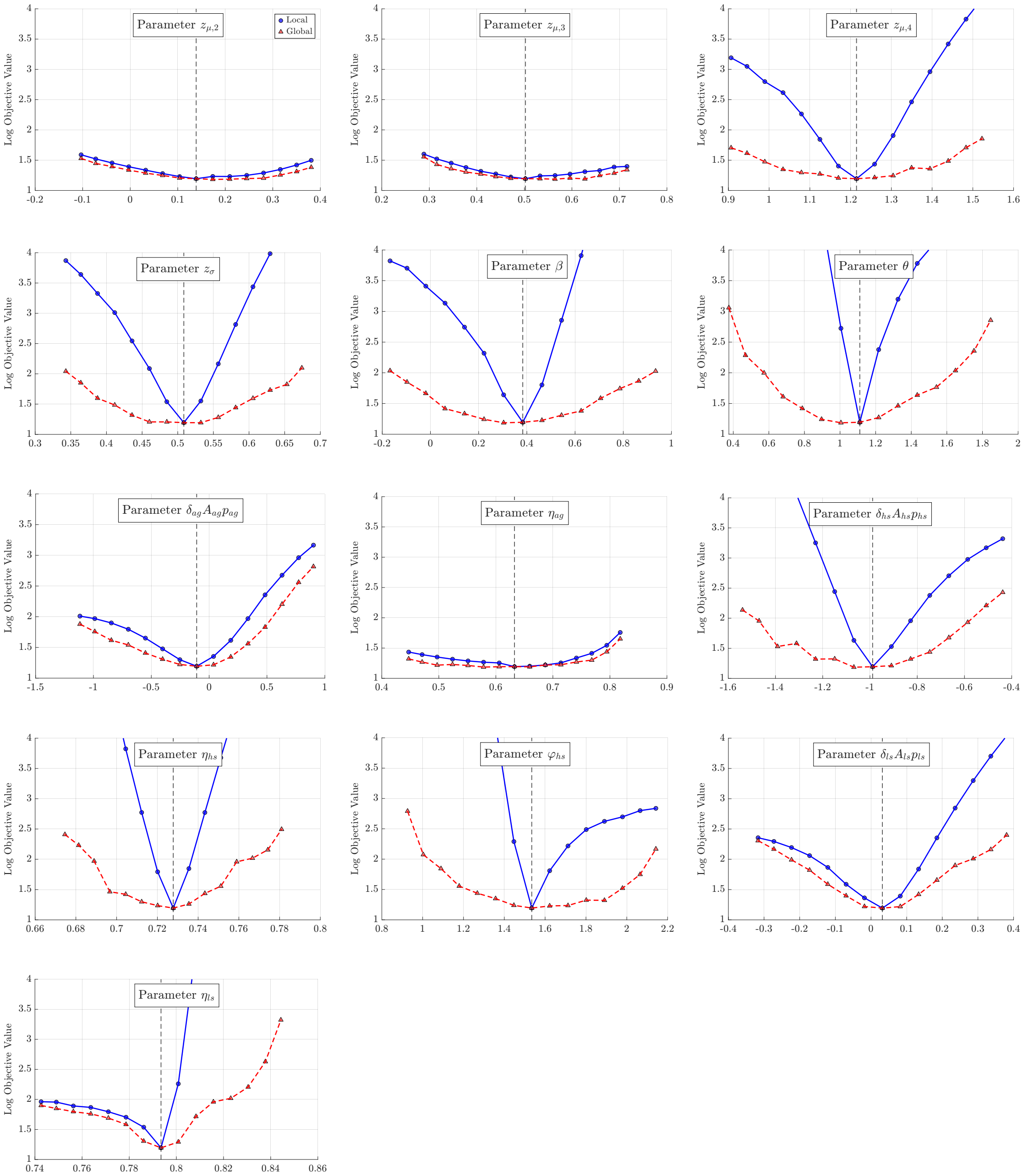


FIGURE C.2: IDENTIFICATION CHECKS



D Benchmarking Against Causal Evidence

This appendix provides details on how we simulate the expansions of education in the model. It also contains the results of a simulation of the effects of management training interventions.

D.1 Brazilian College Expansion

To approximate the experiment in Cox (2025) in the model, we decrease the share of workers with secondary education by 3.5 percentage points and increase the share with tertiary education by the same amount. Given that Cox's design uses regional variation, we hold all sectoral prices fixed, effectively treating Brazilian regions as small open economies. Cox reports the effects of changes in the college share among the 25-34 age group on outcomes measured either for the same age group or the overall population. We use the former when available, and rescale the model-based coefficient by multiplying it by the 25-34 share in Brazil 2000 within the 25-59 sample when only the aggregate results are reported (this applies to the self-employment and large firms' employment share outcomes in Table 4).

D.2 INPRES School Construction Program

We use data from the 1995 Intercensus Population Survey from Ruggles et al. (2025), and impose the same sample restrictions as in the rest of the paper. Following Duflo (2001) and Porzio, Rossi and Santangelo (2022), we take individuals born in 1958-1962 as the control group and individuals born in 1968-1972 as the treatment group. We consider the specification

$$y_{icd} = \alpha_c + \alpha_d + \beta \text{Sch}_{icd} + \sum_k \left(X_d I_i^k \right) \Gamma_k + \varepsilon_{icd}$$

where y_{icd} denotes outcomes for individual i in cohort c and district d, α_c and α_d are cohort and district fixed effects, Sch_{icd} is years of schooling, and $\sum_k \left(X_d I_i^k \right) \Gamma_k$ denotes interactions between cohort fixed effects and a vector of district-level outcomes in 1971 (enrollment rate and the age 5-14 population). We then instrument Sch_{icd} with the interaction between a treated group dummy and the intensity of the program (number of schools built per 1000 children) in district d. We refer the reader to Duflo (2001) for more details on the context and the identification strategy.

Column (1) of Table D.1 shows the first-stage results – an additional school per 1000 students leads to 0.16 years of schooling for the treated cohorts. Columns (2)-(5)

show that this increase in schooling is due to a decline in the share without primary education and an increase across all other education groups, secondary in particular. To compare the model's predictions with the IV results in Table 4, we consider changes in the educational shares given by the coefficients in columns (2)-(5) divided by the coefficient in column (1), so that the variation corresponds to an additional year of schooling, and compute the resulting changes across the different outcomes. Given that the empirical exercise consists of a cross-cohort comparison within a district, we keep all prices and wages fixed.

TABLE D.1: FIRST STAGE RESULTS

	(1) Yrs School	(2) No Primary	(3) Primary	(4) Secondary	(5) Tertiary
Treated x Intensity	0.162	-0.017	0.003	0.010	0.004
	(0.040)	(0.004)	(0.006)	(0.005)	(0.002)
N	44160	44160	44160	44160	44160
F Stat	16.26	14.57	0.29	3.93	3.57
Cohort FE	Yes	Yes	Yes	Yes	Yes
District FE	Yes	Yes	Yes	Yes	Yes

Notes: The Table shows estimates from regressions of either years of schooling (column 1) or dummies for educational attainment (columns 2-5) on the interaction between a treated group dummy (born in 1968-1972) and the number of schools built per 1000 children in the district. All specifications control for cohort fixed effects, district fixed effects and interactions between cohort fixed effects and a vector of district-level outcomes in 1971 (enrollment rate and the age 5-14 population). Robust standard errors in parentheses.

D.3 Benchmarking to the Effects of Management Training

For further evidence on the model mechanism, we turn to evidence from studies that offer management training to firms in low-income and middle-income countries. Two recent papers show that training affects not only firm profitability and sales, but also the organization of treated firms.

Giorcelli (2019) provides evidence from a management training intervention that exposed managers of randomly chosen Italian firms to American-style management practices during the recovery from World War II. She evaluates the effect of this treatment on firm employment and productivity, but also on the ratio of managers per employee. We use the calibrated middle-income economy to simulate this intervention. We focus on firms with at least 10 employees, consistent with the design of the original training program. We simulate a drop in τ that induces an increase in TFPR consistent with Giorcelli (2019)'s estimates. To compute TFPR, we map efficiency units in the model into employment in the data emp by dividing efficiency units by the average efficiency units per worker. We then compute $TFPR = \log y(z) - 0.6 \log emp(z)$. Our main outcome in this case is managers per worker, which we measure by dividing the efficiency units of each type of labor in each firm by the economy-wide average efficiency units per worker of each type of labor.

Bloom et al. (2013) evaluate a randomized intervention that uses a consulting firm to provide management training to the owners of textile plants in India. Again, their intervention not only raised productivity, but also led the firm owners to professionalize a wider range of management tasks. For example, the intervention increased the share of firms that perform routine maintenance on machinery, track inventory, or set clear job descriptions and performance incentives for workers.

We use the calibrated low-income economy to simulate this intervention. We focus on firms with at least 100 employees, consistent with the design of the experiment. We simulate a drop in τ that induces an increase in TFPR consistent with Bloom et al. (2013)'s estimates. We measure TFPR as we did for the Italian intervention; the two studies use very similar labor shares.

The main outcome of interest is the causal effect of the management training intervention on the unweighted share of a range of 38 management practices that a firm had adopted. Conceptually, we compare this to the change in the share of tasks professionalized q in order to induce the necessary TFPR increase in the model.

The main challenge is that the share of management practices and q do not necessarily have a comparable underlying scale or distribution. Our approach is to report each effect in terms of how far it moves treated firms within the baseline (pre-treatment or calibrated) CDF. While Bloom et al. (2013) do not report the full baseline distribution, they do report the min, median, and max (Table I). We fit a triangular distribution to these statistics. We then evaluate the treatment effect in terms of how far in the CDF it moves the treated firms. In the model, we also report how far the change in q moves firms in the distribution, in this case among firms with more than 100 workers in the baseline, calibrated low-income economy.

We replicate these experiments in the calibrated middle-income and low-income economy, respectively. We think of these experiments as information interventions that shared knowledge about a valuable technology. This approach is consistent with the historical evidence that the diffusion of management best practices started within the United States during World War II and diffused internationally afterwards, including to Italy. Bloom et al. (2013) also find information to be an important barrier in India: managers often either were unaware of important practices, or believed that they would not be profitable for their firm.

TABLE D.2: CAUSAL EVIDENCE ON MANAGEMENT IMPROVEMENTS

		TFPR		Manager/worker		Log size		Management	
Study	τ	Data	Model	Data	Model	Data	Model	Data	Model
Giorcelli (2019)	-0.103	0.401	0.401	0.099	0.035	0.300	1.008		
Bloom et al. (2013)	-0.028	0.154	0.154					50.0	9.9

Notes: Each intervention is modeled as a reduction in the firm-size wedge τ (reported in the second column), chosen so the model TFPR increase matches the data TFPR increase. The Giorcelli (2019) intervention is simulated in the calibrated middle-income economy and the Bloom et al. (2013) intervention in the calibrated low-income economy. TFPR computed as $\log(y) - 0.6 \log(\text{employment})$ in the model, where employment is constructed by dividing efficiency units hired by the firm by average efficiency units per worker. Managers/workers ratio and log employment constructed in a similar fashion. Management is computed as the shift in the share of management practices adopted in the data or q in the model, reported as $100\times$ the change in the pre-treatment CDF.

Our baseline approach to modeling this is to exogenously reduce τ for treated firms until the model produces the same change in TFPR as is reported in the original studies. We study the implied effects of this change on other outcomes of interest. Giorcelli (2019) reports the change in employment and managers per worker. The first row in Table D.2 shows that the model's predictions are qualitatively consistent with her estimates. In terms of magnitudes, the model produces a growth in overall employment that is three times larger than her estimates but a growth in managers per worker that is one-third of what she finds. Bloom et al. (2013) report the change in the adoption of a range of management practices. Our model generates a smaller reorganization of production equivalent to a 10 percentage point change in the baseline distribution of q. Again, the model's reorganization response is conservative relative to the best evidence.

It is not obvious that management interventions map into reductions in τ. Following the discussion in Section 6.3, we could also think of management interventions as a diffusion of a new technology that raises β, the gains from professionalizing tasks. Doing so produces roughly half the effect on managers/worker, a similar effect on firm size, and roughly twice the effect on management.

E Endogenous Duality as a Model Mechanism

We stressed in our analytical section the importance of endogenous duality. This feature is important because it creates a reserve labor force of traditional entrepreneurs who can be pulled into wage work when modern firms expand. In this section we conduct two sets of counterfactuals to show the quantitative importance of this mechanism in our model. Each counterfactual considers again the effects of giving the low-income economy the educational attainment of the middle-income economy.

First, we consider the effect of increasing educational attainment in economies that vary in the value of ξ, which governs the dispersion of taste shocks across occupations and hence the elasticity of labor supply across occupations. Larger values of ξ imply that workers are more responsive to wage differentials (equation (7)). As ξ becomes larger, the quantitative model approaches the analytical model in the sense that workers become more likely to choose the highest-paying occupation. On the other hand, smaller values of ξ imply that workers are less responsive to wages and hence less willing to switch from traditional entrepreneurship to working as a laborer (among other margins).

Table E.1 shows the results of expanding schooling with values of ξ ranging from 4 to 16, motivated by the range of parameter estimates in the trade literature. We organize the rows from the least elastic to the most elastic labor-supply specification. Because the schooling expansion is the same in all three cases, the rows can be compared to get a sense of the importance of ξ and endogenous duality.

In many ways these three counterfactuals look similar. All three generate a muted structural transformation. The decline in the self-employment share is one-quarter to one-third of what we observe in the data. The main difference is in the employment share in medium and large firms. When $\xi = 8$ (the benchmark), the model explains 17.3/31.8 = 54 percent of the overall growth in the employment share in medium and large firms in response to a change in skills. In the more elastic case ($\xi = 16$), this share would be larger still, at 84 percent, because the higher labor supply elasticity makes workers more willing to switch occupations and so accommodates the growth in large firms. On the other hand, in the less elastic case, the corresponding share is just 44 percent.

An alternative approach to exploring the role of endogenous duality is to shut down the reserve labor force effect entirely. To accomplish this, we introduce a mean preference shifter for being a laborer. We again give the low-income economy the skill level of the middle-income economy, but now we adjust the preference shifter such that the aggregate traditional entrepreneurship rate remains unchanged. Conceptually, the difference between these economies shows how structural transformation and reorganization would proceed without endogenous duality and the reserve labor force of traditional entrepreneurs.

The results of this experiment are shown in the last row of Table E.1. Increasing skills generates only about three-fifths of the growth in the employment share in medium and large firms that it did in the original counterfactual (10.3 versus 17.3 percentage points). The main reason for this smaller reorganization is the scarcity of la-

TABLE E.1: ROBUSTNESS: THE ROLE OF SCHOOLING

Panel A. Aggregate Outcomes	11+ share	Self-empl.	White collar	Ln VAPW
Data	0.318	-0.417	0.159	1.453
Schooling only				
Inelastic Labor (ξ = 4)	0.139	-0.093	0.066	0.600
Baseline (ξ = 8)	0.173	-0.105	0.066	0.594
Elastic Labor (ξ = 16)	0.266	-0.144	0.093	0.609
No Reserve Labor	0.103	0.000	0.094	0.440
Panel B. Outcomes by Sector	**Agr**	**Mfg**	**Ser (HS)**	**Ser (LS)**
Employment shares				
Data	-0.316	0.101	0.140	0.076
Schooling only				
Inelastic Labor (ξ = 4)	-0.043	-0.007	0.061	-0.011
Baseline (ξ = 8)	-0.025	-0.009	0.053	-0.019
Elastic Labor (ξ = 16)	-0.029	-0.011	0.065	-0.025
No Reserve Labor	-0.012	-0.018	0.043	-0.013
Employment share at 11+ firms				
Data	0.184	0.328	0.235	0.243
Schooling only				
Inelastic Labor (ξ = 4)	0.000	0.280	0.033	0.268
Baseline (ξ = 8)	0.000	0.402	0.065	0.385
Elastic Labor (ξ = 16)	0.000	0.621	0.143	0.649
No Reserve Labor	0.000	0.230	0.042	0.215
Panel C. Outcomes by Education	**<Primary**	**Primary**	**Secondary**	**Tertiary**
White collar share				
Data	0.025	0.034	0.044	0.048
Schooling only				
Inelastic Labor (ξ = 4)	-0.017	-0.030	-0.047	-0.026
Baseline (ξ = 8)	-0.013	-0.018	-0.033	-0.007
Elastic Labor (ξ = 16)	-0.005	-0.006	-0.009	0.040
No Reserve Labor	0.002	0.003	0.007	0.031
Wage				
Data	-0.046	-0.039	0.000	0.109
Schooling only				
Inelastic Labor (ξ = 4)	-0.020	0.009	0.000	-0.061
Baseline (ξ = 8)	-0.009	0.005	0.000	-0.051
Elastic Labor (ξ = 16)	-0.011	0.004	0.000	-0.021
No Reserve Labor	0.036	0.045	0.000	-0.206

Notes: Data reports the difference between the middle-income and low-income economy. The remaining rows isolate the impact of giving the low-income economy the education distribution (υ_i) of the middle-income economy, holding fixed everything else. *Inelastic Labor* ($\xi = 4$), *Baseline* ($\xi = 8$), and *Elastic Labor* ($\xi = 16$) recalibrate the cross-country parameters under three different values of the elasticity of occupational choice, with $\xi = 8$ as the baseline. *No Reserve Labor* augments the baseline schooling shift with a mean preference shifter on being a laborer chosen such that the aggregate self-employment share is fixed in the low-income country (calibrated under $\xi = 8$).

borers; the white-collar employment actually rises by more in this counterfactual than in the baseline.

References for Online Appendices

Adamopoulos, Tasso, and Diego Restuccia. 2020. "Land Reform and Productivity: A Quantitative Analysis with Micro Data." *American Economic Journal: Macroeconomics*, 12(3): 1–39.

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia. 2022. "Misallocation, Selection, and Productivity: A Quantitative Analysis With Panel Data From China." *Econometrica*, 90(3): 1261–1282.

Agence nationale de la Statistique et de la Démographie. 2017–2024. "Enquête Nationale sur l'Emploi (ENES)." Accessed August 1, 2025. https://anads.ansd.sn/index.php/home.

Barro, Robert J., and Jong Wha Lee. 2013. "A New Data Set of Educational Attainment in the World, 1950–2010." *Journal of Development Economics*, 104: 184–198.

Bloom, Nicholas, Benn Eifert, Aprajit Mahajan, David McKenzie, and John Roberts. 2013. "Does Management Matter? Evidence from India." *Quarterly Journal of Economics*, 128(1): 1–51.

Caunedo, Julieta, Elisa Keller, and Yongseok Shin. 2023. "Technology and the Task Content of Jobs Across the Development Spectrum." *World Bank Economic Review*, 37(3): 479–493.

Chandler, Jr., Alfred D. 1977. *The Visible Hand: The Managerial Revolution in American Business.* Cambridge, Massachusetts:Harvard University Press.

Chen, Chaoran, Diego Restuccia, and Raül Santaeulàlia-Llopis. 2023. "Land Misallocation and Productivity." *American Economic Journal: Macroeconomics*, 15(2): 441–465.

Cox, Alvaro. 2025. "Fostering Development Through Higher Education: College Attainment, Firms and Economic Growth." mimeo, Universidad Carlos III de Madrid.

Donovan, Kevin, Will Jianyu Lu, and Todd Schoellman. 2023. "Labor Market Dynamics and Development." *Quarterly Journal of Economics*, 138(4): 2287–2325.

Donovan, Kevin, Will Jianyu Lu, Joseph Pedtke, and Todd Schoellman. 2026. "Labor Market Anatomy of a Macroeconomic Crisis." mimeo, Federal Reserve Bank of Minneapolis.

Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review*, 91(4): 795–813.

Feenstra, Robert C., Robert Inklaar, and Marcel P. Timmer. 2015. "The Next Generation of the Penn World Table." *American Economic Review*, 105(10): 3150–3182.

Giorcelli, Michela. 2019. "The Long-Term Effects of Management and Technology Transfers." *American Economic Review*, 109(1): 121–152.

Le Nestour, Alexis, Laura Moscoviz, and Justin Sandefur. 2023. "The Long-Run Decline of Education Quality in the Developing World." Center for Global Development Working Papers 608.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo. 2022. "The Human Side of Structural Transformation." *American Economic Review*, 112(8): 2774–2814.

Rossi, Federico. 2022. "The Relative Efficiency of Skilled Labor across Countries: Measurement and Interpretation." *American Economic Review*, 112(1): 235–266.

Ruggles, Steven, Lara Cleveland, Rodrigo Lovaton, Sula Sarkar, Matthew Sobek, Derek Burk, Dan Ehrlich, Quinn Heimann, Jane Lee, and Nate Merrill. 2025*a*. "Integrated Public Use Microdata Series, International: Version 7.6 [dataset]." *https://doi.org/10.18128/D020.V7.6*.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025*b*. "IPUMS USA: Version 16.0 [dataset]." *Minneapolis, MN: IPUMS*. Available online at https://doi.org/10.18128/D010.V16.0.

World Bank. 2025. *World Development Indicators.* Washington DC:World Bank.