



# 1 Introduction

A key characteristic of modern economic growth is the systematic transformation of the organization of production (Kuznets, 1973).<sup>1</sup> In poor countries, it is organized along traditional lines: the majority of workers are self-employed or employed in small, slow-growing single-establishment firms whose owners also manage the enterprise on a day-to-day basis.<sup>2</sup> By contrast, most workers in rich countries are employed in modern business enterprises: large, multi-establishment firms with a separation of ownership and management.

This shift in firm organization requires the formation of a class of professional, salaried managers who set strategy, allocate resources, and monitor and coordinate production (Chandler, 1977). In this paper, we document that the price of middle management varies little with development, which implies that its relative price is much higher in less developed countries. We establish that this high relative price deters the adoption and spread of modern business enterprises. We also provide new suggestive evidence as to why management is expensive in poor countries.

We start with the data. An important challenge is that the average quality of management is strongly correlated with development (Bloom *et al.*, 2014).<sup>3</sup> Given this, comparisons of average management earnings across countries confound quality variation with price variation. To address this challenge, we use a proprietary database collected and maintained by a global compensation consulting company (the “Company”). The Company specializes in informing large, modern businesses operating in developing and emerging economies – including many prominent multinational firms – how their salaries and compensation packages compare to local market trends. The Company measures local market trends using data on what past clients pay similar workers. Its database is a cumulative record of actual compensation by leading firms to over 300,000 workers. The database covers mostly workers in middle management roles in 146 countries worldwide.

The information in the database permits two detailed quality adjustments. First, the Company devotes substantial labor resources to standardizing jobs across firms and countries to a common, detailed scheme so that it can provide clients with valuable “apples-to-apples” comparisons of pay.

---

<sup>1</sup>A number of the other, interrelated characteristics cited by Kuznets have generated substantial literatures, including structural transformation among sectors (Herrendorf *et al.*, 2014), urbanization (Lagakos, 2020), openness to trade (Vaugh, 2010), and the rapid adoption of technologies (Comin & Hobijn, 2010).

<sup>2</sup>See Gollin (2008) on the prevalence of self-employment, Bento & Restuccia (2017), Poschke (2018), and Bento & Restuccia (2021) for facts about firm and establishment size, and Hsieh & Klenow (2014) for facts about establishment growth, all relative to development.

<sup>3</sup>Similarly, Chandler (1977) notes that traditional firms do employ managers, but that their activities are equivalent to those of the lowest level of managers in modern firms.

Second, some clients hire the Company to inform them about pay at establishments in multiple countries in the same year, allowing us to estimate pay variation across countries for the same firm and job.

Our main empirical finding is that the price of middle management varies little with development. The elasticity of the average compensation of middle managers in the Company's database with respect to GDP per worker, both adjusted to 2017 international dollars, is just 0.16.<sup>4</sup> We then control for (standardized) job fixed effects and estimate an even lower elasticity of 0.09. Finally, we control for firm-job-year interactions and find an elasticity that is a precisely estimated zero. We explore the heterogeneity of our results by firm type and skill level and validate them against other sources. The elasticity of the quality-adjusted price of management is consistently close to zero.

Our second step is to establish that the high relative price deters the adoption and spread of modern business enterprises. Here, our paper intersects with a large literature that proposes explanations for the relationship between firm size and development, including financial frictions, differences in sectoral composition, or access to reliable electricity (Buera *et al.*, 2011; Buera & Kaboski, 2012; Fried & Lagakos, 2020). We propose an additional, complementary factor, which is the high price of management. We conduct a quantitative evaluation of a model of the optimal firm structure to isolate the importance of this factor.

Our model builds on an appropriate technology adoption framework (Basu & Weil, 1998; Acemoglu & Zilibotti, 2001; Caselli & Coleman, 2006). Inspired by Chandler (1977), we model the relevant technology choice as the organization of the firm and the structure of production itself. Under the traditional structure the owner manages the firm, which necessarily limits its size to a single, small establishment. Alternatively, firms can choose the modern structure, in which case they can grow to have a large establishment or multiple establishments and enjoy economies of scale in production.<sup>5</sup> However, they need to hire middle managers to monitor and coordinate production. The share and size of modern firms depends on the price of middle management as well as the relative productivity of the two technologies and other factors proposed in the literature, which we capture in a generic wedge.

---

<sup>4</sup>Cavallo *et al.* (2019) document a similarly low pay elasticity for workers engaging in freelance work on a popular website. This finding is also related to work on the large firm wage premium or the foreign firm wage premium, although our finding is stronger because we show that it holds even within-firm across-country (see e.g. Oi & Idson, 1999; Aitken *et al.*, 1996; Lipsey & Sjöholm, 2004; Alfaro-Urena *et al.*, 2021).

<sup>5</sup>We focus on economies of scale, but there are other important aspects of firm structure. Becker & Murphy (1992) emphasize specialization, while Garicano & Rossi-Hansberg (2006b) and Garicano & Rossi-Hansberg (2006a) emphasize the organization of the firm into a management hierarchy as a way to economize on knowledge.

We calibrate the model such that it replicates the expansion of middle management and modern business enterprise in the United States from 1900–1960. This expansion is well-documented, was relatively undistorted, and entailed re-organization of production into modern business enterprises within a number of industries that are relevant for developing countries today, including retailing, petroleum refining, meat packing, or flour milling. We isolate the importance of the relative price of middle managers by lowering the price in the poor country to be the same as the rich country in the model. This change raises the revenue share of modern enterprises from 17 to 56 percent. Output expands by 31 percent in the poorest countries, closing 12 percent of the output gap to the rich countries. A decomposition following [Basu & Fernald \(2002\)](#) shows that the results stem primarily from reallocating labor from the traditional to the modern sector. The output effect of this reallocation is amplified because our calibration procedure yields a wedge to the adoption of modern firm structure in developing countries that reflects the importance of non-management impediments to adopting modern firms. Reallocating labor in the face of this wedge has a first-order effect on output as in [Baqae & Farhi \(2020\)](#).

Finally, we examine why management is relatively expensive in poor countries. The zero compensation elasticity is likely at least partially explained by the fact that the labor market for skilled managers is global. It would require a striking coincidence to generate the same result through offsetting supply and demand shifts for countries across a wide range of development. We also provide evidence for migration, cross-border training, and expatriates playing important roles in this market. However, limiting emigration from developing countries would be unlikely to equalize relative prices across countries. We use cross-country data on the quantity and quality of education to show that a much smaller share of the work force in developing countries possesses the literacy skills required to work as a middle manager. Firms may also choose to pay efficiency wages given the greater difficulty of enforcing contracts in developing countries and monitoring managers from afar.

Our work is most closely tied to the new literature demonstrating the importance of management. [Bloom \*et al.\* \(2014\)](#) show that management quality is lower in less developed countries among domestic firms but not among establishments of foreign multinational firms. This finding naturally raises the question of why more firms in poor countries do not choose better management. Our findings point to the relative cost of these managers as an important impediment, which we quantify. The quantitative results are related to recent work that uses quasi-experimental evidence to show that management and firm structure respond to distance and labor supply within a country ([Gumpert](#)

*et al.*, 2019; Feng & Valero, 2020). Finally, we provide some suggestive results on why managers are scarce in poor countries that connect with existing work on their education and high-skill labor markets (Bloom *et al.*, 2013; Guner *et al.*, 2018).

We also contribute to the literature on appropriate technology adoption. Whereas most recent work emphasizes the importance of the skill intensity of technology, our emphasis on firm structure is more in the spirit of Stewart (1977). We show that this has the potential to generate large quantitative results because the relative price of middle managers varies much more across countries than does the relative price of capital or educated labor (Banerjee & Duflo, 2005; Caselli, 2016). Finally, our paper contributes to the literature on cross-country differences in human capital (Caselli, 2005; Hendricks & Schoellman, 2018). Rather than focus on evidence based on inputs such as years of schooling, we provide evidence of scarcity of a particular, important set of skills.

The rest of the paper proceeds as follows. Section 2 describes the key features of the dataset on salaries paid by multinational firms. Section 3 documents the facts about compensation among these firms. Section 4 presents our appropriate technology model. Section 5 provides the calibration and quantitative results. Section 6 provides evidence on why the relative price of managers is higher in developing countries. Finally, Section 7 offers a brief conclusion.

## 2 Data

Our empirical analysis makes use of a proprietary database collected and maintained by a global compensation consulting company (the “Company”). Broadly, compensation consultants provide clients with information on the level and trends in pay for a given job title and region. Relative to its competitors, the Company’s niche is compensation in developing and emerging markets. Firms whose business operations extend into these markets hire the Company to provide information on how their compensation aligns with market conditions.

As we discuss further below, the typical client for the Company is a modern, multi-establishment firm. Clients that hire the Company thus begin by selecting which establishment or establishments will participate in the market comparison. For each establishment, human resources personnel report the positions that are present and the average compensation by position.

The Company’s central business proposition is to return to the client select moments of the distribution of compensation for each position in the local market. For these figures to be meaningful, it is

essential that the Company provide “apples-to-apples” comparisons. To this end, the Company does not take the position titles reported by the client at face value. Instead, it employs professional jobs analysts who conduct interviews to learn about the tasks, responsibilities, and skills associated with each position. They use this information to translate each position into their own internal, globally standardized job classification scheme. This step ensures that workers the Company analysts deem “accountants” in any firm or country perform similar tasks and have similar responsibilities and so makes the compensation comparisons meaningful. This work is invaluable for our purposes because it means that the data on compensation for the same job across countries is much more comparable than that produced by the standard method, which involves economists or national accountants applying crosswalks to data that include workers’ self-reported occupations.

The database we have access to only records the harmonized job title, not the original title provided by the client. However, we have access to select reports the Company has provided to clients for establishments in developing countries that list both the original position title and the standardized job title. These reports indicate that Company analysts systematically downgrade job titles in developing countries. For example, the client may have a position that it calls senior accountant, but after interviews the Company analysts would deem it to be equivalent only to accountant or junior accountant by global standards.

After providing the market comparison to the client, the Company adds the client’s data to its database for future use. Thus, the Company’s definition of market compensation is based on the compensation actually paid by previous clients in the same labor market; the market compensation data provided to future clients in the same labor market will be based in part on the current client’s pay. The Company defines a labor market at the city level. However, there are only data for one city per country (generally the capital city, sometimes the business hub if that is different) and so we use country and city interchangeably. The Company’s standardized job classification scheme includes more than 200 titles ranked both by the broad occupation (accounting) and skill level (junior accountant, senior accountant, etc.).

We have access to the database as of late 2015, which in turn reflects compensation reported by clients spanning the years 2005–2015. Each observation reports the firm name, city/country, year, standardized job classification, the average compensation of workers in the position in the establishment, and in many cases also the total number of such workers. All observations pertain to local workers; expatriates are reserved to a separate database, which unfortunately we cannot access.

While there is no other information in the database, we use the firm name to merge on the firm's industry, profit/non-profit status, and headquarters location. For our analysis, we restrict attention to for-profit firms and remove all NGOs, educational institutions, embassies, international organizations, and public firms. The remaining firms come from a wide variety of sectors, including banking, consulting, health care, mining and other natural resources, technology, telecommunications, and transport (using Standard Industrial Classification codes). In the end, this leaves us with data on pay for more than 300,000 workers from 1,219 firms in 146 countries.

Table 1 provides statistics on how our sample is distributed across countries and firms. For Panel A we aggregate the sample to the country level and merge on GDP per worker, measured in 2017 international dollars from [World Bank \(2018\)](#). This panel shows that we cover a wide range of the income distribution, with a 90-10 ratio of more than a factor of 16. It also shows that the database covers hundreds or thousands of workers in most countries.

**TABLE 1: SAMPLE DISTRIBUTION**

	Percentiles				
	10th	25th	50th	75th	90th
<i>Panel A: Countries (146)</i>					
GDP p.w., 2017 intl \$	4,774	12,224	27,872	49,069	77,792
Workers	388	895	1,511	2,944	4,858
<i>Panel B: Firms (1,219)</i>					
Countries	1	1	1	1	3
Unique Jobs	9	13	18	31	46
Workers	12	26	64	159	460

Table shows the distribution of the sample when aggregated to the level of the country or firm. Percentiles are computed separately for each moment, so the country with the median GDP per worker is different from that with median number of workers. All statistics refer to the final sample of 172,582 country-year-firm-job observations representing 316,452 total workers after imposing sample restrictions discussed in the text.

For Panel B we aggregate the sample to the (parent) firm level. We have 1,219 firms in the database. The first row shows that the majority of firms contribute observations for a single country. However, about twenty percent of firms appear in the database for multiple countries. The top ten percent of firms appear in three or more countries; the top firm contributes observations for 81 different countries. The remaining rows show that the median firm contributes 18 different jobs and provides data on pay for 64 workers.

It is important to emphasize at the outset that these firms are not representative employers in their labor markets. Indeed, given the prevalence of small, traditional firms in developing countries,

a representative sample of firms would be of little use in characterizing the price or compensation of high-quality middle management. Instead, our sample consists almost entirely of modern business enterprises. The firms that hire the Company tend to be large, multi-establishment firms; three-fourths of our earnings observations come from foreign affiliates of multinational firms. The multinational firms are based primarily in North America (predominantly the United States), followed by Africa and Europe. Many firms in the database are large, well-known, publicly listed companies. To this point, the publicly listed U.S. firms in the database account for 32 percent of all revenue and 44 percent of all R&D investment in Compustat North America.

The database consists primarily of workers in middle management roles, with some associated support workers (cleaners, guards, and the like).<sup>6</sup> There are few production workers. To help visualize the occupational distribution in our database, we construct a crosswalk to match every job in the Company database to the closest 1-digit International Standard Classification of Occupations (ISCO)-08 occupation group. We then compute the distribution of employment across these ten bins in the Company database among poor countries, which we define as those having income less than or equal to Bolivia (roughly \$18,500 in 2017 international dollars).

We compare this distribution to one constructed from nationally representative data sets for countries with similar income levels. Details of the data sets are available in Appendix A.1. Figure 1a shows that the two distributions are quite different. Representative samples show that the typical worker in developing countries is engaged in sales, farming, trades work, or elementary occupations. By contrast, the workers in the Company's database are focused in management, as well as the business subsets of professional, technical, and clerical occupations.

This occupational distribution is quite similar to the one that prevails among workers employed in the business service sector in the United States, as shown in Figure 1b. The high degree of similarity leads us to infer that the establishments in the Company's database are primarily local headquarters that coordinate sales and marketing for large firms from the country's capital city or business hub. We can verify that some firms also have production or sales establishments in the same country, but these establishments are not in the database.

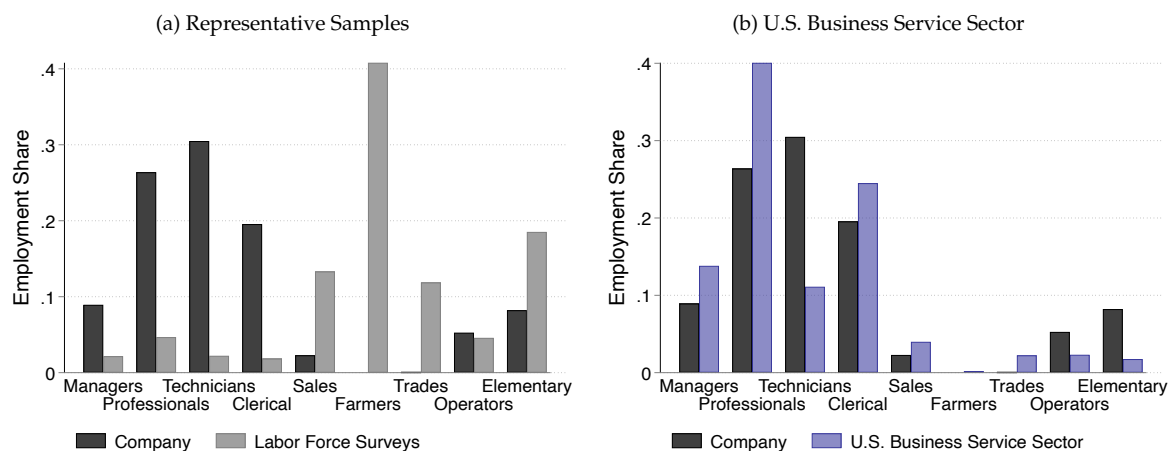
The database reports gross and net compensation for all positions in three categories: base wage, bonus, and other income. Our preferred measure of compensation is total gross pay: gross wage, gross bonus, and other gross income. All amounts are reported to us in contemporaneous U.S. dollars;

---

<sup>6</sup>We conjecture that this selection represents the firms and establishments with the greatest demand for the Company's services: foreign-owned firms who are uncertain about the market for specialized, uncommon, highly compensated workers.



**FIGURE 1: OCCUPATIONAL DISTRIBUTION OF COMPANY DATA (DEVELOPING COUNTRIES)**



original data were either reported in U.S. dollars or were converted to dollars using market exchange rates. We make several adjustments to make sure that these amounts can be averaged and compared across countries and years, which is complicated by the fact that some emerging markets grow rapidly and hence experience rapid wage increases.

Our approach is to first convert all earnings back into local currency units using contemporaneous market exchange rates. We then adjust all amounts to year 2017 local currency units by adjusting for the average rate of nominal wage growth between year  $t$  and year 2017, inferred from the growth rate of nominal GDP per worker. This adjustment makes salaries comparable over time by assuming that each occupation would have experienced the aggregate average wage growth; it misses any occupation-specific wage growth over this decade. Finally, we convert year 2017 wages in local currency units to year 2017 international dollars using the PPP exchange rate.<sup>7</sup> Our next goal is to study how the PPP-adjusted compensation of middle managers varies across countries.

### 3 Empirical Results

Now that we understand the nature of the database, we use it to address our main question of interest: how does the price of middle management vary with development? We estimate regressions of the

<sup>7</sup>All data for the adjustments from [World Bank \(2018\)](#). PPP exchange rate inferred from the ratio of GDP per capita reported in local currency units and international dollars in year 2017.

form

$$\log(w_{c,t,f,j}) = \gamma + \eta \log(y_c) + \beta X_{c,t,f,j} + \varepsilon_{c,t,f,j} \quad (1)$$

where  $w_{c,t,f,j}$  is the total gross compensation for workers in country  $c$  and year  $t$  working for firm  $f$  in standardized job  $j$ ,  $y_c$  is the GDP per worker in country  $c$ , and  $X$  is a potential vector of controls. The main parameter of interest is  $\eta$ , the elasticity of compensation with respect to GDP per worker.

This compensation elasticity captures how much the price of middle management varies with development. Two simple benchmarks can help build intuition. The first is a standard neoclassical growth model with homogeneous labor. A representative firm in each country takes input prices as given and produces output using a Cobb-Douglas production function with country-specific total factor productivity. In this model, compensation per employee is the labor share times GDP per worker, which implies that the compensation elasticity is one. The second benchmark is a simple application of the law of one price with heterogeneous labor. If a given type of worker earns the same compensation in all countries, then the compensation elasticity is zero.

Table 2 shows the results from estimating equation (1). Recall that each observation in our database includes the number of workers and average compensation per country-year-firm-job; we weight the regression by the number of workers and report robust standard errors. Column (1) shows the simplest specification, which includes no controls at all. In this case, the estimated elasticity is 0.16. In column (2) we add year and job fixed effects. Job fixed effects control for cross-country differences in the skill composition of workers. We can see that the implied compensation elasticity is cut in half in this specification, to 0.088.<sup>8</sup> In column (3) we allow for job-year interactions, which produces nearly identical results.

In columns (4) and (5) we include the identity of the firm as a control. Column (4) includes a full set of year, job, and firm fixed effects, which reduces the compensation elasticity to nearly zero. Finally, we showed in Table 1 that about twenty percent of the firms in our sample engage the Company to benchmark multiple affiliates, usually in the same year. This fact offers us the opportunity to study pay variation across countries within a fixed firm, job, and year. This result is useful for alleviating any residual concern about the comparability of job classifications. The results are shown in column (5). We lose about 40 percent of the sample. The compensation elasticity in this specification is a

---

<sup>8</sup>These results are consistent with the recent work of Bayer & Kuhn (2019), who show that job levels account for a substantial share of multiple measures of wage dispersion within the United States and Germany. Here, we show that the same is true even between countries.

**TABLE 2: ESTIMATED ELASTICITY OF COMPENSATION WITH RESPECT TO GDP PER WORKER**

	(1)	(2)	(3)	(4)	(5)
Log GDP p.w.	0.162*** (0.0345)	0.0881*** (0.00829)	0.0848*** (0.00716)	0.0150* (0.00585)	-0.00863 (0.00644)
Fixed Effects	None	Year + Job	Year × Job	Year + Job + Firm	Year × Job × Firm
R-squared	0.017	0.543	0.550	0.685	0.738
N	173,105	173,094	172,912	173,091	115,152

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

precisely estimated zero: there is no variation across countries in the PPP-adjusted compensation within a given firm, job, and year.

We investigate the heterogeneity of this result along two dimensions. First, we consider whether it differs much between foreign affiliates of multinational firms and domestic establishments, inferred from whether an establishment is in the same country as the firm’s headquarters. The results are shown in Table 3. We cannot include firm fixed effects when investigating domestic establishments, so we control for job-year interactions as in column 3 of Table 2. The first column repeats those results for comparison.

**TABLE 3: ESTIMATED ELASTICITY OF COMPENSATION BY ESTABLISHMENT TYPE**

	All	By Firm Type	
		Foreign	Domestic
Log GDP p.w.	0.0848*** (0.00716)	0.0813*** (0.00677)	0.0826*** (0.0142)
Fixed Effects	Year × Job	Year × Job	Year × Job
R-squared	0.550	0.537	0.631
N	172,912	138,705	33,997

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$ 

The remaining two columns show the results for foreign affiliates and domestic establishments. Note again that the majority of our sample is foreign affiliates (here,  $137,705/172,912 \approx 80$  percent). However, the estimated compensation elasticity for the two groups is actually quite similar. This implies that our findings are not particular to affiliates of multinational firms.

We also investigate how our results vary by skill level. Like most compensation consulting firms, the Company’s job classification scheme includes a measure of skill that crosses occupation borders, so that some human resource officers and some accountants can be deemed to be at the same skill

level. We aggregate skill levels into four broad groups to avoid disclosing the Company’s business information. The bottom skill level includes workers who are not in middle management roles. These workers are the cleaners, guards, drivers, and so on that are captured in the data simply because they work in the local headquarters. The remaining three skill groups capture different skill levels of middle managers. The low skill level includes workers with clerical jobs, such as secretaries. The medium skill level includes workers with business associate and business professional jobs, such as accountant. The high skill level includes those with upper management role, such as senior executive.

**TABLE 4: ESTIMATED ELASTICITY OF COMPENSATION BY SKILL LEVEL**

	All	By Skill Level			
		Non-Management	Low	Medium	High
Log GDP p.w.	0.0848*** (0.00716)	0.242*** (0.0192)	0.0977*** (0.0131)	0.0276*** (0.00618)	-0.00343 (0.00555)
Fixed Effects	Year × Job	Year × Job	Year × Job	Year × Job	Year × Job
R-squared	0.550	0.269	0.218	0.100	0.153
N	172,912	10,882	75,165	50,906	35,949
Example Job		Driver	Secretary	Accountant	Senior Executive

Standard errors in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Table 4 shows the implied compensation elasticity for these different skill groups, each estimated with job-year interactions (which control for heterogeneity across countries in the mix of jobs within each skill level group). The first column again shows that the elasticity in the aggregate is 0.08. Turning to the results by skill level, there is a very clear pattern: the elasticity is lower for workers with higher skill levels. While the elasticity is 0.24 for the non-management workers, it falls to 0.10 for the least-skilled managers, 0.03 for the medium-skilled managers, and is actually negative, albeit statistically insignificant, for the highest-skilled managers.

The low compensation elasticity for middle managers – equivalently, higher relative compensation for middle managers in developing countries – is the central empirical finding of our paper. In Sections 4 and 5 we take these relative price facts as given and investigate their consequences for the adoption of middle management and modern business enterprises. In Section 6 we provide preliminary evidence on why compensation might be high. But first, it is worthwhile to validate our results against other data sources covering this market.

### 3.1 Alternative Data Sources for Middle Manager Compensation

We start by considering whether alternative data sources show evidence of similar trends for the compensation of middle managers. At least one data source clearly does not. As we describe in Appendix A.2, some nationally representative datasets include data on labor earnings. We use these datasets to estimate the relative earnings of middle managers (as compared to other workers or GDP per worker) for these countries. Representative datasets also show evidence of higher relative earnings of middle managers in poorer countries, but to a much smaller extent. The implied compensation elasticity for this dataset is 0.57–0.63; alternatively, the relative earnings of middle managers is 1.7 times that of other workers in poor countries.

However, these results reflect compensation among a representative sample of middle managers. Given differences in management quality and the share of modern firms, we interpret this comparison as not getting at the true quality-adjusted price of middle management. To provide further evidence of high-quality management, we turn to data from an alternative source: recruiting firms. Recruiting firms play a complementary role to compensation consulting firms. Whereas compensation consulting firms' pay advice can be used to help with worker retention, recruiting firms help with vacancy fulfillment. Our specific data comes from Robert Walters, a self-described "global, specialist professional recruitment consultancy."<sup>9</sup> Robert Walters provides recruiting services for many of the same types of positions and in many of the same countries as the Company.

Robert Walters uses its experience in vacancy fulfillment to produce an annual Salary Survey, which lists for select countries/regions and jobs the typical salary range in the current and previous year. The data in the Salary Survey differ from the Company's database in three main ways. First, it is much less detailed. In developing countries it generally aggregates countries into regions (such as East Africa) and focuses on a small set of the most commonly filled jobs. Second, the data reflect Robert Walters' experience placing new workers, including expatriates, rather than payments to all local workers. Finally, they report salaries exclusive of bonuses and other benefits.

We focus on their data for Africa exclusive of South Africa, which contains most of the poorest countries in the Company's sample. The geographic detail in the Salary Survey increases over time; we collect data from the 2017 survey, which was the first to decompose Africa into four geographic regions: North Africa, East Africa, West Africa, and Central-South Africa (Robert Walters, 2017). The Salary Survey includes earnings for 65 roles spread across these four regions along with a salary range

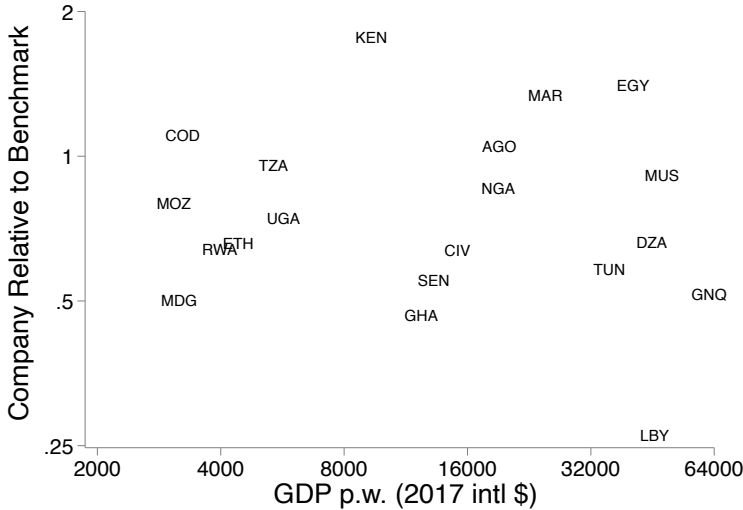
---

<sup>9</sup><https://www.robertwalters.com/news/about-us.html>.

for each. Broadly, the survey supports high salaries. For example, the midpoint of the salary range for a General Manager in Central Africa is \$90,000; for a Head of Supply Chain in East Africa, \$67,500; for an HR Manager in West Africa, \$80,000.

For a more thorough comparison, we match the Robert Walters survey responses to the Company’s database. We map regions to countries by using commentary from the last four years of Salary Surveys to infer the set of countries in each region where Robert Walters is active. We merge occupations using several examples showing actual mappings from common job titles to the Company’s standardized job scheme in developing countries. We replace the salary range with the midpoint and adjust to 2017 international dollars using the same algorithm as we applied to the Company’s database. We compare Robert Walters’ salary figures to the gross salary in the Company database (rather than total gross compensation). This procedure allows us to compare gross salary for 12,000 observations in 19 countries in Africa to equivalent reports from Robert Walters.

**FIGURE 2: RESIDUAL EARNINGS GAPS: COMPARISON WITH RECRUITMENT DATA**



We find that on average the Company compensation is actually 22 percent *lower* than that in Robert Walters. The gap is plausibly accounted for by the fact that Robert Walters includes expatriates in its database. Figure 2 plots against GDP per worker the ratio of Company to Robert Walters gross salary at the country level. There is no strong evidence that the ratio varies systematically with development. In short, these results suggest that other firms with experience in the market for high-quality middle managers in developing countries agree on the level of earnings needed to attract or retain those

workers. We turn now to quantifying the implications of this finding for how firms are organized around the world.

## 4 Model: Appropriate Technology and the Organization of the Firm

This section formulates a model of appropriate technology. Inspired by [Chandler \(1977\)](#), the technology available for adoption is re-organization as a modern business enterprise. When making this choice, firms trade off the benefits of economies of scale against the costs of hiring managers to coordinate the high-velocity, high-volume production ([Coase, 1937](#); [Becker & Murphy, 1992](#)). This decision depends in part on the cost of management, which is the margin we quantify.

We consider a static model of a country with a continuum of industries that produce differentiated goods. Goods vary exogenously in how suitable they are for modern production, measured as relative productivity when organized along modern versus traditional lines. We first describe the technology adoption problem for firms within a single industry to highlight the essential forces. Most of our analysis focuses on a parametric model that allows for tractable aggregation, but in [Section 4.2](#) we show that the main qualitative results hold as long as there are increasing returns to scale in production and decreasing returns to scale in coordination. We then describe the aggregate economy, including the range of industries, the households, and the government.

### 4.1 Industry model

Each industry is populated by a large number of ex ante identical firms. There is free entry, with each entrant producing the same homogeneous output by choosing one of two firm organization technologies. Below, we present the two technologies, the firm's choice between them, and the resulting firm size and industry production function.

**Traditional technology.** The traditional technology captures self-employment and small, single-establishment, owner-managed firms. There are no economies of scale, so the production function is linear,

$$F^T(\ell_p) = z_T \ell_p, \quad \ell \leq \bar{\ell}_p, \quad (2)$$

with the constraint reflecting that the owner-manager has a limited span of control.

**Modern technology.** The modern technology features economies of scale in the use of production workers. As a function of the number of production workers, output is

$$y = \kappa z_M \ell_p^{1+\eta}, \quad (3)$$

where  $\eta > 0$  regulates the strength of scale economies, and  $\kappa$  is a constant to facilitate derivations further on. Firms need managers to coordinate the production workers, with the required number of managers growing as a convex function of the number of production workers:

$$\ell_m = \ell_p^{1+\gamma}. \quad (4)$$

The parameter  $\gamma$  introduces a convexity, reflecting that coordination requirements typically grow more than linearly with the number of workers.<sup>10</sup> For example, the number of bilateral interactions between workers grows with the square of the number of workers, and the number of mappings from workers to distinct tasks grows with the factorial of the number of workers. We restrict  $\gamma > \eta$  to ensure that firms choose a finite size for every set of wages, since coordination requirements eventually grow faster than scale economies.

We assume that there is no gain in output from hiring managers beyond the managerial requirement or from hiring production workers beyond the managerial capacity. These assumptions imply that equations (3) and (4) can be captured in the production function

$$F^M(\ell_p, \ell_m) = \kappa z_M \min\{\ell_p, \ell_m^{\frac{1}{1+\gamma}}\}^{1+\eta}. \quad (5)$$

**Technology selection.** To analyze how firms select technologies, it is helpful to express the production functions (2) and (5) in terms of average cost functions. Let  $w_p$  and  $w_m$  be the wages of production workers and managers. To capture other factors impeding the adoption of modern technologies, we further allow that modern firms face a wedge that is equivalent to a proportional  $e^\tau - 1$  tax on their input costs. The average cost functions are:

$$c^T(y) = \frac{w_p}{z_T}, \quad y \leq z_T \bar{\ell}_p, \quad (6)$$

$$c^M(y) = \frac{e^\tau}{z_M \kappa} \left( w_p (y/z_M \kappa)^{-\frac{\eta}{1+\eta}} + w_m (y/z_M \kappa)^{\frac{\gamma-\eta}{1+\eta}} \right). \quad (7)$$

---

<sup>10</sup>Similar results obtain if we require management to scale with output given the simple production function in (3).



With free entry, firms operate on the bottom of their average cost curves. For the traditional technology, average costs are constant at  $c^T \equiv \frac{w_p}{z_T}$ . For the modern technology, differentiating average costs yields the output level  $y^*$  that minimizes average costs. Substituting this output level into the average cost function yields the attained minimum cost, which is

$$c^M(y^*) = \frac{e^\tau}{z_M} w_p^{1-\alpha} w_m^\alpha, \quad \alpha \equiv \frac{\eta}{\gamma}$$

The parameter  $\alpha$  is the managerial compensation share of the modern technology, and we choose  $\kappa$  to cancel out constants from the calculation. Note that the managerial share is simply the ratio of the economies of scale ( $\eta$ ) to the convexity of the coordination costs ( $\gamma$ ). Our restriction that  $\gamma > \eta$  implies that  $0 < \alpha < 1$ .

Firms adopt the modern technology if the minimum average cost is lower than for the traditional technology:

$$c^M(y^*) \leq \frac{w_p}{z_T} \iff e^\tau \left( \frac{w_m}{w_p} \right)^\alpha \leq \frac{z_M}{z_T}. \quad (8)$$

Intuitively, firms adopt the modern technology when the productivity advantage  $z_M/z_T$  dominates the additional costs coming from the managerial wage premium  $w_m/w_p$ , times the wedge  $e^\tau$ . A high cost of management deters adoption, with the strength of the effect depending on the managerial compensation share  $\alpha$ .

Free entry ensures that the industry produces a flexible amount of output at its minimum attainable average cost. Hence, aggregate industry behavior can be described using a representative firm that operates a linear production technology if the traditional technology offers lower costs or a Cobb-Douglas production function with managerial share  $\alpha$  if the modern technology does. This result greatly simplifies subsequent analysis by allowing us to focus on standard functional forms. Unlike most analyses with such production functions, we have a well-defined underlying notion of firm size and the number of firms, both of which jump discontinuously at the boundary between traditional and modern organization.<sup>11</sup>

<sup>11</sup>Firm size (measured as employment) is given by  $\bar{\ell}_p$  for traditional firms, although results do not vary much if we assume firm size is 1 (self-employment). The size of modern firms is given by:

$$\left( \frac{w_m}{w_p} \frac{\eta}{\gamma - \eta} \right)^{-\alpha} + \left( \frac{w_m}{w_p} \frac{\eta}{\gamma - \eta} \right)^{-(1-\alpha)}.$$

## 4.2 General Comparative Statics

Our parametric model is analytically convenient because it allows for a firm-level tradeoff between economies of scale and coordination costs but still aggregates to a standard industry-level Cobb-Douglas production function. Nonetheless, the underlying intuition about how the cost of management affects the incentive to adopt the modern business enterprise is quite general. Here, we show that this finding obtains under much weaker assumptions on technology. Starting from a general set of techniques with the shared features of economies of scale in the use of production workers and decreasing returns to scale in coordination, we show that a higher managerial wage premium causes a decrease in scale within every technique, as well as a switch towards techniques with a smaller degree of scale economies.

**Technology specification.** Formally, we assume that firms have access to a set of techniques  $T = \{1, \dots, T\}$ , which all use production workers and managers to produce output. Each technique features economies of scale in the use of production workers, summarized by a continuous function  $f_t$  mapping output  $y$  to the number of production workers  $\ell_p^t$  used per unit of output:

$$\frac{\ell_{p,t}(y)}{y} = f_t(y),$$

where scale economies are captured by assuming that  $f_t$  is weakly decreasing in  $y$ . Furthermore, we assume that  $t = 1, \dots, T$  are ordered in terms of an increasing degree of scale economies, where a technology  $t$  is said to feature a higher degree of scale economies than  $t'$  if its relative use of production workers  $f_t(y)/f_{t'}(y)$  is strictly decreasing in  $y$ .<sup>12</sup>

All technologies use managers to coordinate production workers, with the required number of managers governed by a common weakly increasing function  $g$ :

$$\frac{\ell_{m,t}}{y} = \frac{g[\ell_{p,t}(y)]}{y} = \tilde{g}_t(y),$$

where  $\tilde{g}_t(y) = \frac{g[yf_t(y)]}{y}$ . Apart from being weakly increasing, our only assumptions on  $g$  is continuity, and that it grows sufficiently fast so that  $\tilde{g}_t$  is weakly increasing in  $y$  and goes to  $\infty$  as  $y \rightarrow \infty$  for all  $t$  (this ensures that all technologies feature a finite production scale for every managerial wage premium).

---

<sup>12</sup>Since the relationship of featuring more scale economies is transitive, it induces a partial ordering on the space of continuous, non-decreasing functions. Assuming strict monotonicity simplifies the proofs.

**Average cost and comparative statics.** Normalizing the production worker wage to 1, the average cost function of technology  $t$  is

$$\bar{c}_t(y; w) \equiv \frac{\ell_{p,t}(y)}{y} + w_m \frac{\ell_{m,t}(y)}{y} = f_t(y) + w_m \tilde{g}_t(y).$$

We write  $y_t^*(w_m)$  for the output level that minimizes the average cost for each technology,  $t^*(w_m)$  for the technology that attains the lowest average cost, and  $y^*(w_m) \equiv y_{t^*(w_m)}^*$  for the cost-minimizing output level in the resulting technology.

In a competitive market, the industry operates on the bottom of its average cost curve. Our main finding is that a higher management cost reduces the optimal firm size within every technology, induces firms to switch to a technologies with less scale economies, and reduces the resulting firm size level. We prove the following proposition.

**Proposition 1.** *The functions  $y_t^*(w_m)$ ,  $t^*(w_m)$ , and  $y^*(w_m)$  are weakly decreasing in  $w_m$ .*

*Proof.* See the appendix. □

Proposition 1 proves that the effect of expensive management holds in a much wider class than the earlier Cobb-Douglas setup. The key is the tradeoff between economies of scale and coordination costs. As long as the average production worker requirement falls with output and the average manager requirement rises, then more expensive management causes a shrinking of production size. The specification of  $f_t$  covers a wide range of cases. Examples include production worker requirements exhibiting power decay,  $f_t(y) = y^{-\eta_t}$ , or exponential decay  $f_t(y) = \exp(-\eta_t y)$ , with a large  $\eta_t$  corresponding to a higher degree of scale economies.<sup>13</sup>

### 4.3 Aggregate Economy

The aggregate economy consists of a continuum of industries like the one described in Section 4.1. Firms in all industries face the same wages for both types of labor and the same wedge if they choose the modern firm structure. However, industries vary in their productivity when organized along modern and traditional lines. Chandler notes that modern business enterprises and middle management were developed as solutions to new problems posed by products where it was possible to use

<sup>13</sup>The specification in section 4.1 is covered by setting  $\eta_t = 0$  for the small technology, with a slight modification to accommodate that both technologies should have the same managerial requirements. Formally, we assume managerial requirements is  $\tilde{g}(\ell_p) = [(\ell_p - \bar{\ell})^{1+\eta}]_+$ . This specification ensures that the small-scale firm does not need to use any management (and never operate beyond  $\bar{\ell}$ ). Since  $\bar{\ell}$  is small, the resulting managerial requirement for the large firm is close to the original,  $g(\ell_p) = \ell_p^{1+\eta}$ .

“capital-intensive, energy-consuming, continuous or large-batch production technology to produce for mass markets” (Chandler, 1977, p. 347). For example, cement, steel, or flour are straightforward to organize using continuous or batch production technologies, were among the first industries to adopt modern production methods in the United States, and are organized in this manner essentially everywhere today. By contrast, products that lacked these characteristics – those that were labor-intensive, did not use complex machinery, produced at low volume, or who could sell their products easily through existing wholesalers – remained dominated by small firms. For example apparel production or plumbing services have largely resisted modern organization so far.

We capture this idea by assuming that there is a continuum of industries indexed by  $k$  with the production structure describe above, but with industry-dependent productivities  $z_T(k)$  and  $z_M(k)$ . Productivities are draws from independent Fréchet distributions with scale parameters  $Z_T$  and  $Z_M$  and a common dispersion parameter  $\theta$ . A final goods producer aggregates the industry output using a constant elasticity of substitution production function with elasticity  $\sigma$ . Note that the price  $p(k)$  is taken as given by firms and drops out of the cost-minimization/technology adoption problem. The probability (share) of industries organized through the modern business enterprise is given by

$$\mathbb{P} \left( e^\tau \frac{w_p^{1-\alpha} w_m^\alpha}{z_M(k)} < \frac{w_p}{z_T(k)} \right) = \frac{[Z_M(w_m/w_p)^{-\alpha} e^{-\tau}]^\theta}{Z_T^\theta + [Z_M(w_m/w_p)^{-\alpha} e^{-\tau}]^\theta}. \quad (9)$$

The development and intuition for our calibration and quantitative results is eased if we exploit the fact, well-known in the trade literature, that this setup is isomorphic to one with a simple two-sector CES aggregator:

$$Y = \left( Y_T^{\frac{\theta}{\theta+1}} + Y_M^{\frac{\theta}{\theta+1}} \right)^{\frac{\theta+1}{\theta}}, \quad (10)$$

where  $Y_T$  and  $Y_M$  are total output of traditional and modern firms. These outputs can in turn be represented using stand-in production technologies that resemble the industry production functions derived in Section 4.1:

$$Y_T = F_T(L_{T,p}, L_{T,m}) = \Gamma \left( \frac{\theta + 1 - \sigma}{\theta} \right)^{\frac{1}{\sigma-1}} Z_T L_{T,p}, \quad (11)$$

$$Y_M = F_M(L_{M,p}, L_{M,m}) = \Gamma \left( \frac{\theta + 1 - \sigma}{\theta} \right)^{\frac{1}{\sigma-1}} Z_M L_{M,p}^{1-\alpha} L_{M,m}^\alpha, \quad (12)$$

We use  $L$  to distinguish aggregate labor used by the entire set of traditional or modern industries;  $\Gamma$  is the gamma function.

The representative household has a total labor endowment  $L$  and supplies it to the two sectors. It uses earnings to finance consumption given the budget constraint

$$PC \leq w_p L_p + e^{-\zeta} w_m L_m + T. \quad (13)$$

where  $\exp(-\zeta)$  is a tax that drives a wedge between the marginal product of managerial labor and the wage. It captures, for example, the idea that workers may be paid an efficiency wage. Both the managerial wage tax  $\zeta$  and the tax on modern firms  $\tau$  are rebated to the household via  $T$ .

We allow the relative labor supply of the household to be determined by a general function which we write as

$$\frac{L_p}{L_m} = G\left(\frac{w_p}{w_m}, \zeta, Q\right). \quad (14)$$

This formulation makes clear that relative labor supply depends on relative wages. It is also a function of two relative labor supply shifters. The parameter  $\zeta$ , introduced above, captures shifters that alter the relative wage workers receive. By contrast, the parameter  $Q$  captures shifters that make workers more willing to work as managers independent of the relative wage. In Section 6 we associate these two shifters primarily with efficiency wages and the quantity and quality of education in a country. Given this interpretation, we assume  $G$  is increasing in the first two arguments and decreasing in the third.

We now have all the necessary ingredients to define an equilibrium, which is a set of prices  $\{w_p, w_m, P, P_T, P_M\}$  and quantities  $\{L_p, L_m, C, Y, Y_T, Y_M\}$  such that they solve the household problem,  $Y$  satisfies (10),  $Y_T, Y_M$  satisfy (11) and (12),  $P_T, P_M$  satisfy the usual CES pricing equations implied by (10), transfers satisfy

$$T = (1 - e^{-\zeta})w_m L_m + (1 - e^{-\tau})P_M Y_M,$$

and labor markets clear

$$\begin{aligned} L_p &= L_{T,p} + L_{M,p} \\ L_m &= L_{M,m}. \end{aligned}$$

## 5 Quantification

The previous section presents an appropriate technology adoption model in which the relevant technology is the modern business enterprise, which provides economies of scale but requires firms to hire managers to coordinate production. Our goal in this section is to calibrate the model and use it to quantify the importance of the relative price of management for explaining cross-country variation in the employment of middle management and adoption of modern business enterprises.

### 5.1 Calibration Strategy

Much of our calibration strategy and quantitative results can be understood through equation (9), which gives the share of industries that adopt the modern business enterprise. If we denote that share by  $s_M$ , then we can re-arrange to yield a more convenient formulation:

$$\log\left(\frac{s_M}{1-s_M}\right) = \theta \left[ \log\left(\frac{Z_M}{Z_T}\right) - \alpha \log\left(\frac{w_m}{w_p}\right) - \tau \right] \quad (15)$$

There are three main determinants of the size of the modern sector: the relative productivity of the modern technology; the relative wage of managers; and the wedge that captures other factors not modeled.

The parameter  $\theta$  here controls the dispersion of technology in each sector, with lower values of  $\theta$  corresponding to more dispersion. It also governs the elasticity of the modern sector with respect to each of the driving forces. Intuitively, as  $\theta$  becomes smaller and productivity becomes more dispersed, a smaller share of industries will find it optimal to change their firm structure in response to a given change in relative wages or relative mean productivities.

We use this equation to help understand the sources of cross-country variation in the share of modern firms. We think of all countries as having access to the same technology menu in the sense that  $Z_M/Z_T$  and  $\theta$  are common across countries, with any unbiased technology differences being captured by a TFP term  $A$ , and any bias in the choice across technologies captured in the wedge  $\tau$ . We observe from the Company database the cross-country variation in relative wages  $w_m/w_p$ . We want to assess how far this goes in helping understand cross-country variation in technology adoption and how much is left to the residual wedge  $\tau$ . The challenge is that we do not know the value for  $\theta$ . Following the previous discussion, if  $\theta$  is sufficiently large (so that relative productivities are sufficiently concentrated across industries), then observed differences in wages can explain all of

the cross-country variation in modern technology adoption. On the other hand, if  $\theta$  is small, then the observed changes in wages explain little and we are left to infer large wedges in developing countries.

We implement a two-part calibration strategy. In section 5.2, we calibrate  $\theta$  by turning to historical evidence from the expansion of middle management and modern enterprise in the United States in the early 20th century. In section 5.3 we calibrate the remaining, more standard parameters given  $\theta$ .

## 5.2 Calibrating $\theta$ : U.S. Historical Experience

Our calibration strategy for the key parameter  $\theta$  is to choose it such that the model matches the expansion of modern business enterprise in the early 20th century in the United States. We start by showing this expansion. We measure the spread of modern business enterprises through the expansion of middle managers, shown in Figure 3a. We prefer this indirect method because it is challenging to categorize firms or industries in the entire economy as modern or traditional, particularly prior to the expansion of the scope of economic censuses. As we see, the employment share of middle managers tripled over a sixty year period.

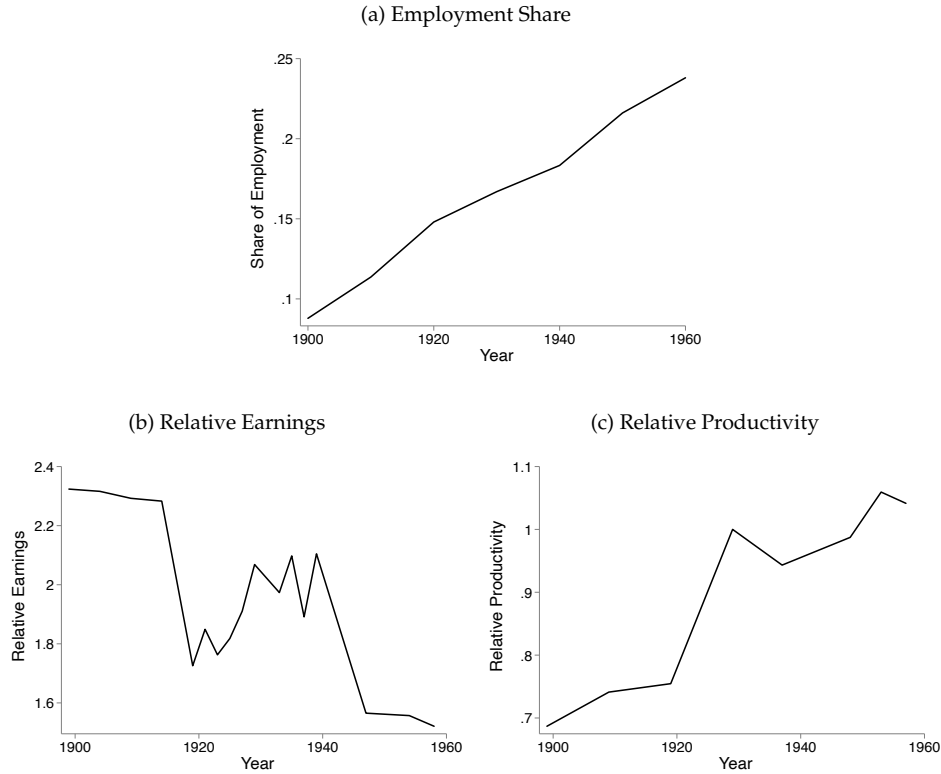
Returning to equation (15), there are in principle three changes over time that could help explain this expansion: faster relative growth of modern technologies; declining relative wages of managers; or reductions in wedges and barriers. There are two key advantages to using the United States experience here. First, the transition is well-documented by economic historians, who provide evidence on the extent of changes in wages and productivity growth during this period. We use this to measure the first two driving forces. Second, this expansion happened during an era where policy towards integrated modern firms was relaxed and barriers were few. Certainly, there is little evidence that the expansion can be attributed towards a large reduction in any such barriers.<sup>14</sup> Based on this, we abstract from changes in  $\tau$  when calibrating  $\theta$ .

Both of the remaining driving forces play an important role. We describe here briefly the source of these data; a more thorough description as well as discussion of trends from alternative sources is available in Appendix A.3. Figure 3b shows the earnings of salaried (white-collar) workers compared to wage (production, blue-collar) workers from the Census of Manufactures. Relative earnings of managers fell from 2.3 to 1.6 times that of production workers. It is worth noting that there is no indication that relative wages were ever as high as we find in developing countries today. A special survey of "industrial combinations" (large, leading conglomerates formed through mergers) in the

---

<sup>14</sup>The primary policy development was a stepping up of trust-busting, but this pushes in the opposite direction and as Chandler (1977) notes, trusts were not particularly modern or management-intensive.

FIGURE 3: MIDDLE MANAGEMENT IN U.S. HISTORY



1900 Census of Manufactures shows that salaried workers in industrial combinations earned 3.0 times as much as manufacturing production workers, and even the officers of such firms earned just 11.1 times more.

The relative productivity of modern technologies also increased during this period. To document this, we use the work of Chandler to identify two sets of industries that were either fully modernized early in this period or that never modernized (discarding those that switch at an intermediate point). We then measure average productivity growth of each set of industries from [Kendrick \(1961\)](#). Figure 3c shows the resulting series for relative labor productivity, normalized to be 1 in 1929. Relative productivity grows from around 0.7 in 1900 to 1.1 in 1960.

With measures of these driving forces in hand we can calibrate  $\theta$ . We convert the share of middle managers to the share of modern firms by noting that middle management accounts for around 30 percent of employment in developed countries and in large firms in countries around the world. Combining this fact with the rising employment share in Figure 3a implies that the left-hand side of



equation (15) rose by 3.7 log points per year. On the right hand side, relative productivity rose by 0.7 log points per year while relative earnings fell by 0.7 log points as well. All that remains is the weight  $\alpha$  on the price of middle management. Combining the 30 percent employment share with the observed earnings premium of 1.5 today, we calibrate  $\alpha = 0.4$ . This implies that in total the right hand side rose by a little more than 1 log point per year. The ratio of the left-hand side and the right-hand side gives us our calibrated value of  $\theta = 3.7$ .

This value of  $\theta$  is close to but slightly smaller than the value of  $\theta = 4$  that is often used in the international trade literature (Simonovska & Waugh, 2014). The implication is that the relative productivity of modern and traditional technology across industries is slightly more dispersed than the relative productivity of different countries in the same industry. As we see in equation (15), using a larger value of  $\theta$  will only amplify the importance of wages for modern firm adoption. Instead, we explore the sensitivity of our results to using a lower value  $\theta = 2$ . This would be a more appropriate if, for example, part of the expansion in modern adoption over time in the United States was driven by an unmeasured decline in wedges.

### 5.3 Calibration of Remaining Parameters

The rest of our parameters are chosen to fit a mixture of data from national accounts, representative labor force surveys, and Company data given the value for  $\theta$ . To make sure that our findings do not reflect any single country, we calibrate to fit the data from a stylized "rich" country that includes data for all countries with PPP GDP per worker above \$100,000 and a stylized "poor" country that includes data for all countries poorer than Bolivia (approximately \$18,500).

In addition to  $\theta$  and  $\alpha$ , the parameters  $Z_M$  and  $Z_T$  are common to all countries. We normalize  $Z_T = 1$  and calibrate  $Z_M$  to fit the employment share of middle managers in rich countries, taken from labor force surveys.

Countries differ in their relative price of management, their total factor productivity  $A$ , and their wedge  $\tau$ . We measure the relative price of management using the relative compensation of managers to non-managers that we observe within the Company's database (Table 4). This is a conservative choice. If we use instead average wages from outside sources or aggregate labor income per worker as the relevant price for non-managers, we find much more price variation and a larger role for the relative price of management.

We use bars over variables to denote the rich country and bars below variables to denote the poor

TABLE 5: CALIBRATION TABLE

Parameter	Description	Value	Target
$\alpha$	Manager share, modern technology	0.40	Middle manager compensation share
$Z_M / Z_T$	Relative productivity of technologies	1.5	Middle management employment share
$\theta$	Adoption elasticity w.r.t costs	3.7	See section 5.2
$\underline{\tau}, \bar{\tau}$	Modern technology wedge	0.18, 0	Equation (15)+ normalization
$\underline{A}, \bar{A}$	TFP level	5026, 26191	Real GDP per worker
$\frac{w_m}{w_p}, \frac{\bar{w}_m}{\bar{w}_p}$	Relative price of management	5.4, 1.5	Company database

country. We assume that the rich country is undistorted,  $\bar{\tau} = 0$ , and we calibrate  $\bar{A}$ ,  $\underline{A}$ , and  $\underline{\tau}$  to fit PPP GDP per worker from [World Bank \(2018\)](#) as well as the middle management employment share in developing countries.

The results are displayed in Table 5. The results from Section 3 are reflected in the high relative price of management in poor countries. The adoption wedge  $\underline{\tau}$  is 0.18 in poor countries, implying considerable barriers other than management prices to the adoption of modern business enterprises.

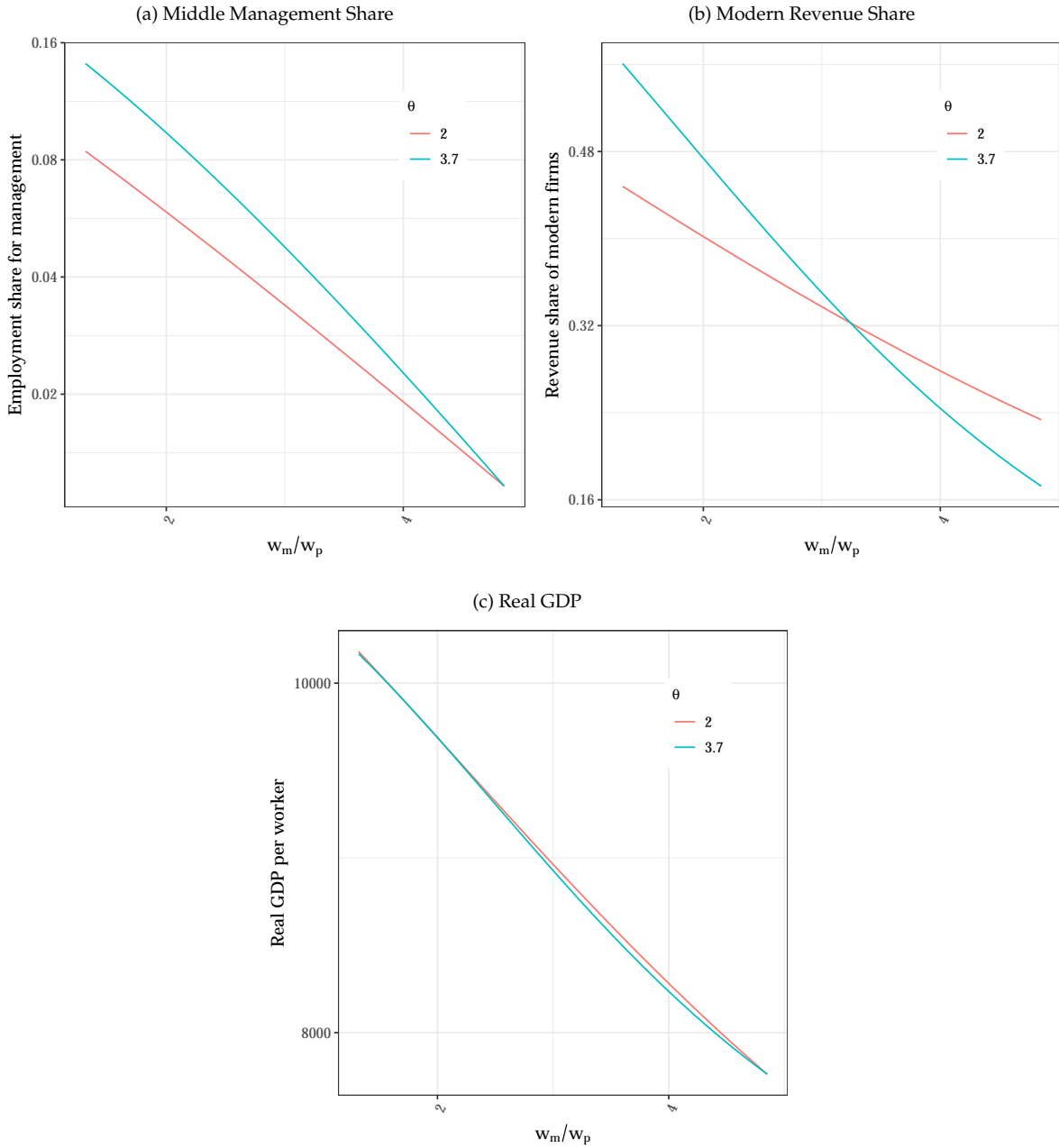
## 5.4 Quantitative Experiments

Our goal is to isolate the importance of the relative price of middle management in explaining cross-country variation in the adoption of modern firms. To do so, we lower the relative wage of managers in the poor country in the model until it is the same as the rich country. Following equation (14), we have in mind changes to  $\zeta$  or  $Q$  that shift relative labor supply and lower the equilibrium relative wage. We return in Section 6 to evidence on distortions and education systems.

Figure 4 shows the effect of changing relative wages on the employment share of middle managers, the revenue share of the modern sector, and real output in the developing country. Each outcome is plotted against relative wages, with the x-axis ranging from 1.5 (the measured value in the rich country) to 5.4 (the measured value in the poor country). Our benchmark results are for the economy calibrated with the value of  $\theta = 3.7$ . We also show the sensitivity of our results to allowing for a lower value  $\theta = 2$ . These results come from a model where the other parameters have been re-calibrated to hit the same targets in Table 5.

Figure 4a shows the results for the employment share of middle managers. Lowering the relative wage to rich country levels would increase the employment share from 1 to 16 percent. This effect follows directly from the fact that we measure large cross-country differences in relative prices, which we multiply by  $\theta\alpha \approx 1.5$ . Allowing for a lower value of  $\theta = 2$  decreases the importance of relative

FIGURE 4: COUNTERFACTUAL RESULTS



prices proportionally, consistent with equation (15).

Figure 4b shows the revenue share of the modern sector. A lower relative wage expands the size of the modern sector substantially, from 17 to 56 percent. Again, the effect is smaller when we consider a lower value of  $\theta$ . Finally, figure 4c shows the effects on real output. The result is an increase in

TABLE 6: DECOMPOSITION OF OUTPUT CHANGES

Term	Value of $\theta$	
	$\theta = 2$	$\theta = 3.7$
$\Delta_{\text{efficiency}}$	0.00	0.00
$\Delta_{\text{between}}$	0.23	0.21
$\Delta_{\text{within}}$	0.04	0.06
$\Delta \log Y$	<b>0.27</b>	<b>0.27</b>

output of 27 log points (31 percent), which is nearly independent of  $\theta$ . This change closes 12 percent of the output gap to rich countries. The remainder of the output gap is attributed to  $\tau$ , the distortion to modern technology adoption, and gaps in total factor productivity.

We gain additional insights if we decompose our output results into three underlying channels as in Basu & Fernald (2002).<sup>15</sup> The first channel is an *efficiency effect*, which captures increases in the available factor inputs. Since we have assumed that labor supply shifters are preference-based, this channel does not operate. The second is a *between-industry* reallocation effect, which captures the importance of changing wages for reallocating labor between the traditional and modern sectors. In our context, this term captures that the modern sector has higher average wages, as well as higher markups because of the distortion. Hence, reallocating labor towards this sector raises output. The third is a *within-industry* reallocation effect, which captures how changes in wages alter the managerial intensity in the modern sector. In our context, this term captures that managers earn higher wages and so reallocating workers towards management in the modern sector raises output.

Table 6 shows the results of the decomposition. As described above, there is no efficiency effect. We view this as being a conservative choice. A more general model would allow labor supply to change via more hours worked or human capital accumulation, yielding larger output effects. Of the two reallocation effects, the between-sector is the more important force in our model. This channel is strengthened by the management wage interacting with the distortion  $\tau$  to adopting the modern firm structure. As noted by Baqaee & Farhi (2020), the effect of reallocating factors is larger in the face of other wedges, which is captured here through the markup that modern firms charge to cover the distortion  $\tau$ .

This intuition helps explain why the output results in Figure 4c are relatively insensitive to the choice of  $\theta$ . Changing  $\theta$  in this model has two offsetting effects. On the one hand, a larger value of  $\theta$  implies a larger change in modern technology adoption in response to changing wages (as shown in

<sup>15</sup>See Appendix B.4 for a formal statement of the decomposition and a proof that it holds in our context.

Figures 4a and 4b). On the other hand, a larger value of  $\theta$  leads us to calibrate a smaller distortion  $\tau$  to be consistent with the observed share of modern firms. This smaller  $\tau$  reduces the output effect of reallocating labor. The overall output effect is thus relatively insensitive.

## 6 Understanding Middle Manager Compensation

So far we have established that the quality-adjusted price of middle management varies little or not at all with development. This fact implies large variation in the relative price of middle management, which through the lens of our quantitative model is a significant deterrent to the adoption and expansion of modern business enterprises. In this section we discuss several candidate explanations for these empirical patterns.

### 6.1 Global Labor Market

One important explanatory factor is that the labor market for skilled workers is increasingly global. Migration can explain why real prices vary little across countries. This factor provides a natural explanation for our finding that the elasticity is lower for more skilled workers and falls to roughly zero for medium and high-skilled managers: their labor market is more global. It would require a striking coincidence to generate the same result through offsetting supply and demand shifts for countries across a wide range of development. An important role for migration can also help explain why today's developing countries appear to have much higher relative prices for middle managers than the United States did in the early 20th century.

The broader business and management literature supports an important role for migration in both directions. Brain drain of skilled workers from developing countries is a well-documented phenomenon (Docquier & Rapoport, 2012). Educated workers are particularly likely to exit and flow into a small subset of OECD countries (Kerr *et al.*, 2016). Hiring local workers remains an option for top firms, but in many cases they find placements abroad to be an important recruitment and training tool (Hsieh *et al.*, 1999). Workers also flow in the other direction. Some firms target local-born, foreign-educated workers who might be willing to return to their birth country (Hsieh *et al.*, 1999). Finally, expatriate workers continue to fill a significant share of management roles in developing and emerging markets (Hsieh *et al.*, 1999; Cho, 2018).

Migration policy is a natural policy tool to address global labor flows. However, blunt policies

that seek to limit migration may not be desirable. Limiting immigration into developing countries (including expatriates) would lower the supply of management in developing countries. More subtly, the literature on brain drain has long acknowledged an offsetting brain gain from return migrants who bring back valuable skills (Docquier & Rapoport, 2012). Also, evaluations of migration limitations have to account for the costs borne by prospective migrants.

## 6.2 Education and Scarcity of Potential Managers

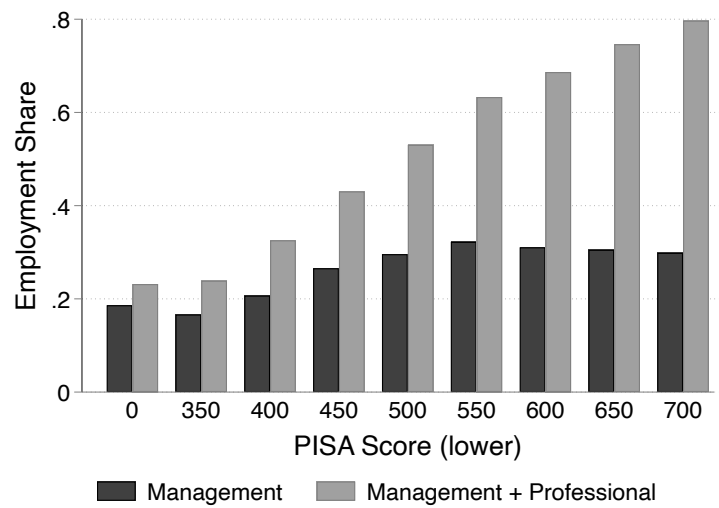
Although global labor markets are likely an essential part of the story for high wages, they are unlikely to be the entire story. A second explanatory factor is the shortage of locally-trained workers capable of assuming middle manager roles. It is well-known that developing countries are particularly scarce in secondary- and tertiary-educated workers (Barro & Lee, 2013). Results from the occasional participation of developing countries in internationally standardized achievement tests such as the OECD PISA also reveal much lower average test scores than rich countries (Hanushek & Woessmann, 2012; Cubas *et al.*, 2016).

Cross-country test score differences are large but also somewhat abstract. To put them into context, we note that the average secondary school student in many developing countries scores at reading level 1b on PISA assessments. PISA characterizes this reading level as "Tasks at this level require the reader to locate a single piece of explicitly stated information in a prominent position in a short, syntactically simple text ..." (OECD, 2014, p. 191). They also provide a sample assessment question for students who read at this level. The question asks students to read Aesop's fable "The Miser and his Gold", which is a one-paragraph story that opens with the sentence, "A miser sold all that he had and bought a lump of gold, which he buried in a hole in the ground by the side of an old wall." Students are asked, "How did the miser get a lump of gold?" (OECD, 2014, p. 212).

We hypothesize that students reading at or below this level are not capable of storing, retrieving, and processing information at the level necessary to act as middle managers in modern business enterprises. To formalize this idea, we develop novel empirical results utilizing the Longitudinal Surveys of Australian Youth (LSAY). The important feature of this dataset is that it tracks students who take the PISA exams in Australia as late as age 25, allowing us to measure how PISA test scores map into subsequent occupational choices in a fixed country with fixed wages. Details are available in Appendix A.4.

Figure 5 shows the main result from the LSAY, which is the share of workers making various

FIGURE 5: PISA TEST SCORES AND OCCUPATIONAL CHOICES



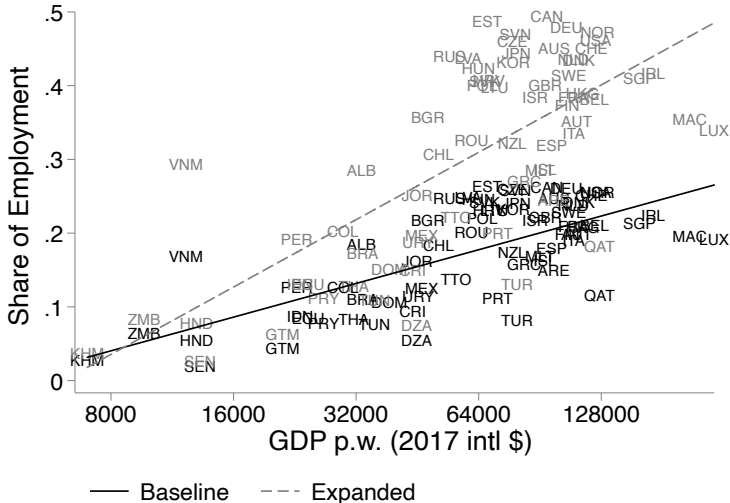
occupational choices by test score bin. The black bars show the share of workers in each bin who join middle manager occupations, which rises from 18 to 30 percent. While there is a notable trend, this probably understates the importance of test scores for the capacity to be a manager because many high-scoring Australians choose other education-intensive occupations. To make this point, the gray bars show the share choosing manager or professional occupations, which rises from 20 to over 80 percent as a function of test scores.

Essentially all Australians attend school long enough to be eligible for the PISA exams (PISA is administered to 15-year olds). Further, the average reading score is sufficiently high (503 in the 2018 round) to generate a substantial number of potential and actual managers. The situation in many developing country is very different: most workers do not attend school long enough to even be eligible for PISA and the test scores among those who do so are much lower.

We start with a simple calculation to make this point precise. First, we use the data from [Barro & Lee \(2013\)](#) to compute the share of each country's working age population that has some secondary or more schooling, while assuming that the rest lack the literacy skills necessary to become effective middle managers. We then use the country's distribution of PISA reading scores (from any round in which they participated) multiplied by the fraction of Australians in each test score bin who become middle managers (Figure 5) to get a sense of the importance of age-15 reading skill differences for potential occupational choices. One way to think about this counterfactual is that it is the share of

each country’s workforce that would become managers if faced with Australian relative wages.

**FIGURE 6: SHARE OF POTENTIAL MANAGERS IMPLIED BY EDUCATION DATA**



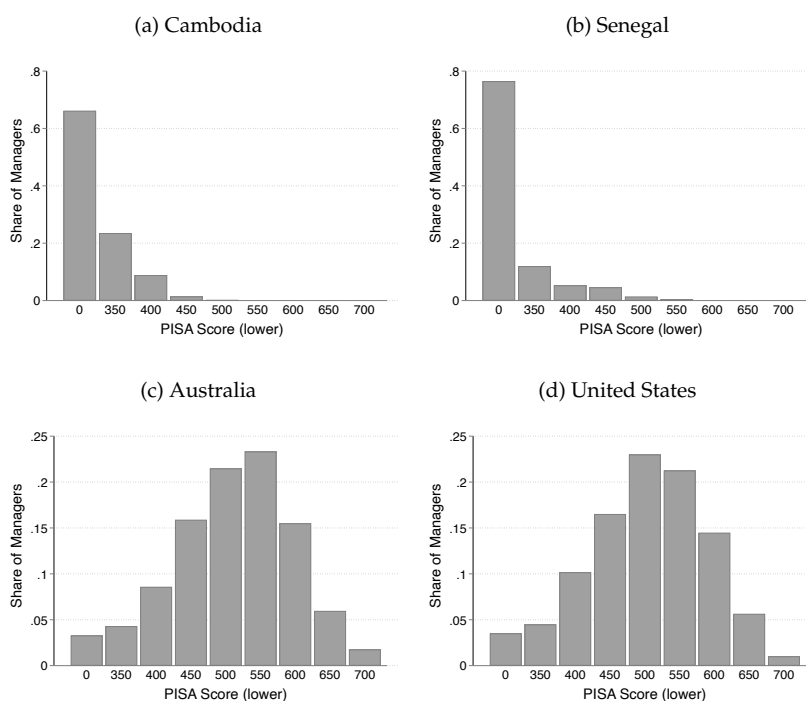
The results of this calculation are shown in black and labeled Baseline in Figure 6. The share of managers is strongly increasing in development. Most rich countries would be expected to have 20–30 percent of their workforce as managers, in line with the actual data. By contrast, many developing countries would expect less than a ten percent management share. Cambodia’s share of 4 percent aligns closely with the actual figure of 5 percent in their most recent census.

Given that developing countries have higher relative wages for managers, we should expect some substitution of workers into management occupations. Two additional calculations suggest that there is limited scope for such substitution. First, we re-compute the set of potential middle managers by taking each country’s distribution of PISA reading scores and multiplying it by the fraction of Australians in each test score bin who become middle managers or professionals (the gray bars in Figure 5). The results are shown in gray and labeled Expanded in Figure 6. This calculation greatly expands the pool of potential managers for rich countries but barely changes it at all among most developing countries.

Second, we compute the test score distribution among the implied managers from this Expanded calculation. The results are shown for select countries in Figure 7. For the two sample rich countries, we see that the implied managers have high literacy levels, with median and mode test scores above 500 (the OECD average). By contrast, for the two sample poor countries only a tiny share of the



FIGURE 7: COUNTERFACTUAL DISTRIBUTION OF TEST SCORES AMONG MANAGERS



implied managers score above 500. Roughly 70 percent of the implied managers have reading scores below 350, suggesting low literacy levels. This calculation captures the underlying challenge, which is that there are simply too few students who enter high school and score at a reasonably high level on reading exams to allow for much expansion of the pool of managers.

These findings also help shed light on the findings of Bloom *et al.* (2014), who find that average management quality is strongly correlated with development. As Figure 7 shows, this finding could simply reflect the failure of education systems to provide a pool of graduates with the necessary skills, which is in turn consistent with the broader evidence of large cross-country differences in human capital (Hendricks & Schoellman, 2018). An important role for skill is also growing evidence that management training interventions improve the quality of management and firm profitability (Bloom *et al.*, 2013; Giorelli, 2019; Bianchi & Giorelli, 2021).

### 6.3 Efficiency Wages

A third hypothesis is that efficiency wage considerations contribute to high relative wages for managers in developing countries. There are several reasons to think that efficiency wages may play

a more important role in these labor markets. First, contracting is generally more difficult in such economies given the poorly functioning legal systems and courts (Acemoglu *et al.*, 2005; Boehm & Oberfield, 2020). Sadka *et al.* (2018) provide details on the specific information frictions and institutional features of the legal system that lower the efficiency of labor courts in the context of Mexico. Second, modern business enterprises rely on advantages conveyed by superior technologies or stocks of intangible capital. The role of middle managers as local coordinators gives them access to sensitive business information.<sup>16</sup> Providing insufficient incentives would thus be particularly costly. Third, middle managers who are not able to emigrate face a thin labor market. Given this, employers might find it optimal to increase pay to replace the motivation usually supplied by outside career options.

Existing work shows that firms respond to these incentives by limiting how much decision-making they decentralize in developing countries (Bloom *et al.*, 2012; Akcigit *et al.*, 2021). Small firms also compensate by relying more on family members, albeit at the cost of poorer management quality (Bloom & Van Reenen, 2007; Bloom *et al.*, 2013). To the extent that modern business enterprises cannot fully centralize sensitive information and decision-making (e.g., cost accounting), it is natural to suspect that at least some of the wage premium reflects a classic efficiency wage designed to align middle managers' incentives with the central office's. However, we are not aware of direct estimates of the quantitative impact of efficiency wages in this market.

#### 6.4 Firm-wide Wage-setting procedures

Finally, in related work, Hjort *et al.* (2020) use the same database we use in this paper to show that wages in a firm's headquarters have a direct, causal effect on wages for the same jobs in the firm's foreign affiliates.<sup>17</sup> They show evidence that this is because many employers use firm-wide wage-setting procedures, which helps rationalize in particular the high wages for workers in low-skill occupations in foreign establishments (see also Goldschmidt & Schmeider, 2017; Deroncourt *et al.*, 2021). Alfaro-Urena *et al.* (2021) also show that multinational firms pay a premium in Costa Rica; the premium is larger there for less skilled workers. We also find a particularly low elasticity of compensation within firms (Table 2, Column 5). However, we note that our results do not appear to be driven particularly by multinational firms (Table 3).

---

<sup>16</sup>For example, general managers may be privy to the firm's strategic initiatives, while accountants monitor cash flows and report figures to the central office.

<sup>17</sup>The sample analyzed in Hjort *et al.* (2020) includes public sector employers, but only multinational employers.

## 7 Conclusion

This paper consists of three main exercises. First, we use the proprietary database of a compensation consulting company to document that the quality-adjusted price of middle management varies little or not at all with development. Second, we quantify the importance of the high relative price of management for the adoption of modern business enterprises in a model of technology adoption. We find that giving poor countries the same relative price of management as rich ones would lead them to expand the revenue share of modern firms from 17 to 56 percent, raising output by 31 percent. Third, we provide preliminary evidence on why relative wages vary systematically with development, including new evidence on the supply of workers with the requisite literacy skills to attain those positions.

We view this evidence on the source of relative wage patterns as preliminary rather than definitive. More evidence especially on the quantitative importance of the various channels would be welcome. Further work is also needed to translate these ideas into policy proposals that could be implemented to reduce the wages of managers and encourage modern business enterprise. For example, Vietnam stands out as a clear outlier in terms of the supply of skilled workers and potential managers given its income level. Does this imply that Vietnamese managerial wages will fall, and the modern business sector grow? If so, what policies in Vietnam encouraged this development? These are promising avenues for future research.

## References

- ACEMOGLU, DARON, & ZILIBOTTI, FABRIZIO. 2001. Productivity Differences. *Quarterly Journal of Economics*, **116**(2), 563–606.
- ACEMOGLU, DARON, JOHNSON, SIMON, & ROBINSON, JAMES A. 2005. Institutions as a Fundamental Cause of Long-Run Growth. *Chap. 6, pages 385–472 of: AGHION, PHILIPPE, & DURLAUF, STEVEN N. (eds), Handbook of Economic Growth*, vol. 1A. Elsevier.
- AITKEN, BRIAN, HARRISON, ANN, & LIPSEY, ROBERT. 1996. Wages and Foreign Ownership: A Comparative Study of Mexico, Venezuela, and the United States. *Journal of International Economics*, **40**(3-4), 345–371.
- AKCIGIT, UFAK, ALP, HARUN, & PETERS, MICHAEL. 2021. Lack of Selection and Limits to Delegation: Firm Dynamics in Developing Countries. *American Economic Review*, **111**(1), 231–275.
- ALFARO-URENA, ALONSO, MANELICI, ISABELA, & VASQUEZ, JOSE P. 2021. *The Effects of Multinationals on Workers: Evidence from Costa Rica*. Mimeo Universidad de Costa Rica.
- AUSTRALIAN GOVERNMENT DEPARTMENT OF EDUCATION, SKILLS, & EMPLOYMENT. 2020a. *Longitudinal Surveys of Australian Youth, 2003 cohort (Version 7.1)*. <http://dx.doi.org/10.4225/87/5I0BPG>.
- AUSTRALIAN GOVERNMENT DEPARTMENT OF EDUCATION, SKILLS, & EMPLOYMENT. 2020b. *Longitudinal Surveys of Australian Youth, 2006 cohort (Version 10.1)*. <http://dx.doi.org/10.4225/87/WFEYRU>.
- AUSTRALIAN GOVERNMENT DEPARTMENT OF EDUCATION, SKILLS, & EMPLOYMENT. 2020c. *Longitudinal Surveys of Australian Youth, 2009 cohort (Version 9.0)*. <http://dx.doi.org/10.4225/87/6BW27V>.
- BANERJEE, ABHIJIT V., & DUFLO, ESTHER. 2005. Growth Theory through the Lens of Development Economics. *Chap. 7, pages 473–554 of: AGHION, PHILIPPE, & DURLAUF, STEVEN N. (eds), Handbook of Economic Growth*, vol. 1A. Elsevier.
- BAQAEE, DAVID REZZA, & FARHI, EMMANUEL. 2020. Productivity and Misallocation in General Equilibrium. *Quarterly Journal of Economics*, **135**(1), 105–163.

- BARRO, ROBERT J., & LEE, JONG WHA. 2013. A New Data Set of Educational Attainment in the World, 1950–2010. *Journal of Development Economics*, **104**, 184–198.
- BASU, SUSANTO, & FERNALD, JOHN G. 2002. Aggregate Productivity and Aggregate Technology. *European Economic Review*, **46**(6), 963–991.
- BASU, SUSANTO, & WEIL, DAVID N. 1998. Appropriate Technology and Growth. *Quarterly Journal of Economics*, **113**(4), 1025–1054.
- BAYER, CHRISTIAN, & KUHN, MORITZ. 2019. *Which Ladder to Climb? Decomposing Life Cycle Wage Dynamics*. Mimeo, University of Bonn.
- BECKER, GARY S., & MURPHY, KEVIN M. 1992. The Division of Labor, Coordination Costs, and Knowledge. *Quarterly Journal of Economics*, **107**(4), 1137–1160.
- BENTO, PEDRO, & RESTUCCIA, DIEGO. 2017. Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics*, **9**(3), 267–303.
- BENTO, PEDRO, & RESTUCCIA, DIEGO. 2021. On Average Establishment Size across Sectors and Countries. *Journal of Monetary Economics*, **117**, 220–242.
- BIANCHI, NICOLA, & GIORCELLI. 2021. *The Dynamics and Spillovers of Management Interventions: Evidence from the Training Within Industry Program*. mimeo, Kellogg School of Management, Northwestern University.
- BLOOM, NICHOLAS, & VAN REENEN, JOHN. 2007. Measuring and Explaining Management Practices Across Firms and Countries. *Quarterly Journal of Economics*, **122**(4), 1351–1408.
- BLOOM, NICHOLAS, SADUN, RAFFAELLA, & VAN REENEN, JOHN. 2012. The Organization of Firms Across Countries. *Quarterly Journal of Economics*, **127**(4), 1663–1705.
- BLOOM, NICHOLAS, EIFERT, BENN, MAHAJAN, APRAJIT, MCKENZIE, DAVID, & ROBERTS, JOHN. 2013. Does Management Matter? Evidence from India. *Quarterly Journal of Economics*, **128**(1), 1–51.
- BLOOM, NICHOLAS, LEMOS, RENATA, SADUN, RAFFAELLA, SCUR, DANIELA, & VAN REENEN, JOHN. 2014. The New Empirical Economics of Management. *Journal of the European Economic Association*, **12**(4), 835–876.

- BOEHM, JOHANNES, & OBERFIELD, EZRA. 2020. Misallocation in the Market for Inputs: Enforcement and the Organization of Production. *The Quarterly Journal of Economics*, **135**(4), 2007–2058.
- BUERA, FRANCISCO J., & KABOSKI, JOSEPH P. 2012. Scale and the Origins of Structural Change. *Journal of Economic Theory*, **147**(2), 684–712.
- BUERA, FRANCISCO J, KABOSKI, JOSEPH, & SHIN, YONGSEOK. 2011. Finance and Development: A Tale of Two Sectors. *American Economic Review*, **101**(5), 1964–2002.
- CASELLI, FRANCESCO. 2005. Accounting for Cross-Country Income Differences. *Chap. 9, pages 679–741 of: AGHION, PHILIPPE, & DURLAUF, STEVEN N. (eds), Handbook of Economic Growth*, vol. 1A. Elsevier.
- CASELLI, FRANCESCO. 2016. *Technology Differences Over Time and Space*. Princeton, NJ: Princeton University Press.
- CASELLI, FRANCESCO, & COLEMAN, II, WILBUR JOHN. 2006. The World Technology Frontier. *American Economic Review*, **96**(3), 499–522.
- CAVALLO, ALBERTO, CRAVINO, JAVIER, & DRENİK, ANDRES. 2019. *Using Online Data for International Wage Comparisons*. mimeo, University of Michigan.
- CHANDLER, JR., ALFRED D. 1977. *The Visible Hand: The Managerial Revolution in American Business*. Harvard University Press.
- CHO, JAEHAN. 2018. Knowledge Transfer to Foreign Affiliates of Multinationals through Expatriation. *Journal of International Economics*, **113**, 106–117.
- COASE, RONALD H. 1937. The Nature of the Firm. *Economica*, **4**(16), 386–405.
- COMIN, DIEGO, & HOBIJN, BART. 2010. An Exploration of Technology Diffusion. *American Economic Review*, **100**(5), 2031–2059.
- CUBAS, GERMAN, RAVIKUMAR, B., & VENTURA, GUSTAVO. 2016. Talent, Labor Quality, and Economic Development. *Review of Economic Dynamics*, **21**, 160–181.
- DERENONCOURT, ELLORA, NOELKE, CLEMENS, & WEIL, DAVID. 2021. *Spillover Effects from Voluntary Employer Minimum Wages*. mimeo, University of California, Berkeley.

- DOCQUIER, FRÉDÉRIC, & RAPOPORT, HILLEL. 2012. Globalization, Brain Drain, and Development. *Journal of Economic Literature*, **50**(3), 681–730.
- EATON, JONATHAN, & KORTUM, SAMUEL. 2002. Technology, Geography, and Trade. *Econometrica*, **70**(5), 1741–1779.
- FENG, ANDY, & VALERO, ANNA. 2020. Skill-Biased Management: Evidence from Manufacturing Firms. *Economic Journal*, **130**(628), 1057–1080.
- FRIED, STEPHIE, & LAGAKOS, DAVID. 2020. *Electrification and Firm Productivity: A General Equilibrium Approach*. mimeo, Boston University.
- GARICANO, LUIS, & ROSSI-HANSBERG, ESTEBAN. 2006a. The Knowledge Economy at the turn of the Twentieth Century: The Emergency of Hierarchies. *Journal of the European Economic Association Papers & Proceedings*, **4**(2-3), 396–403.
- GARICANO, LUIS, & ROSSI-HANSBERG, ESTEBAN. 2006b. Organization and Inequality in a Knowledge Economy. *Quarterly Journal of Economics*, **121**(4), 1383–1435.
- GIORCELLI, MICHELA. 2019. The Long-Term Effects of Management and Technology Transfers. *American Economic Review*, **109**(1), 121–152.
- GOLDSCHMIDT, DEBORAH, & SCHMEIDER, JOHANNES F. 2017. The Rise of Domestic Outsourcing and the Evolution of the German Wage Structure. *Quarterly Journal of Economics*, **132**(3), 1165–1217.
- GOLLIN, DOUGLAS. 2008. Nobody’s business but my own: Self-employment and small enterprise in economic development. *Journal of Monetary Economics*, **55**(2), 219–233.
- GUMPERT, ANNA, STEIMER, HENRIKE, & ANTONI, MANFRED. 2019. *Firm Organization with Multiple Establishments*. CESifo Working Paper No. 7435.
- GUNER, NEZIH, PARKHOMENKO, ANDRII, & VENTURA, GUSTAVO. 2018. Managers and Productivity Differences. *Review of Economic Dynamics*, **29**, 256–282.
- HANUSHEK, ERIC A., & WOESSMANN, LUDGER. 2012. Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. *Journal of Economic Growth*, **17**(4), 267–321.
- HENDRICKS, LUTZ, & SCHOELLMAN, TODD. 2018. Human Capital and Development Accounting: New Evidence from Wage Gains at Migration. *Quarterly Journal of Economics*, **133**(2), 665–700.

- HERRENDORF, BERTHOLD, ROGERSON, RICHARD, & VALENTINYI, AKOS. 2014. Growth and Structural Transformation. In: AGHION, PHILIPPE, & DURLAUF, STEVEN N. (eds), *Handbook of Economic Growth*, vol. 2. Elsevier.
- HJORT, JONAS, LI, XUAN, & SARSONS, HEATHER. 2020. *Across-Country Wage Compression in Multinationals*. NBER Working Paper No. 26788.
- HSIEH, CHANG-TAI, & KLENOW, PETER J. 2014. The Life Cycle of Plants in India and Mexico. *Quarterly Journal of Economics*, **129**(3), 1035–1084.
- HSIEH, TSUN-YAN, LAVOIE, JOHANNE, & SAMEK, ROBERT A. P. 1999. Think global, hire local. *McKinsey Quarterly*, 92–101.
- KENDRICK, JOHN W. ASSISTED BY PECH, MAUDE R. 1961. *Productivity Trends in the United States*. Princeton, NJ: Princeton University Press.
- KERR, SARI PEKKALA, KERR, WILLIAM, ÖZDEN, ÇAĞLAR, & PARSONS, CHRISTOPHER. 2016. Global Talent Flows. *Journal of Economic Perspectives*, **30**(4), 83–106.
- KUZNETS, SIMON. 1973. Modern Economic Growth: Findings and Reflections. *American Economic Review*, **63**(3), 247–258.
- LAGAKOS, DAVID. 2020. Urban-Rural Gaps in the Developing World: Does Internal Migration Offer Opportunities? *Journal of Economic Perspectives*, **34**(3), 174–192.
- LIPSEY, ROBERT E., & SJOHOLM, FREDRIK. 2004. Foreign Direct Investment, Education and Wages in Indonesian Manufacturing. *Journal of Development Economics*, **73**(1), 415–422.
- MINNESOTA POPULATION CENTER. 2019. *Integrated Public Use Microdata Series, International: Version 7.2 [dataset]*. Minneapolis, MN: IPUMS. Available online at <https://doi.org/10.18128/D020.V7.2>.
- NIEMI, ALBERT W. 1972. Structural and Labor Productivity Patterns in United States Manufacturing, 1849-1899. *Business History Review*, **46**(1), 67–84.
- OECD. 2014. *PISA 2012 Results: What Students Know and Can Do – Student Performance in Mathematics, Reading and Science (Volume 1, Revised edition, February 2014)*. <http://dx.doi.org/10.1787/9789264208780-en>.



- OI, WALTER Y, & IDSON, TODD L. 1999. Firm Size and Wages. *Chap. 33, pages 2165–2214 of: ASHENFELTER, ORLEY C., & CARD, DAVID (eds), Handbook of Labor Economics, vol. 3b.* Elsevier.
- POSCHKE, MARKUS. 2018. The Firm Size Distribution across Countries and Skill-Biased Change in Entrepreneurial Technology. *American Economic Journal: Macroeconomics*, **10**(3), 1–41.
- ROBERT WALTERS. 2017. *Salary Survey*.
- RUGGLES, STEVEN, FLOOD, SARAH, GOEKEN, RONALD, GROVER, JOSIAH, MEYER, ERIN, PACAS, JOSE, & SOBEK, MATTHEW. 2019. *IPUMS USA: Version 9.0 [dataset]*. Minneapolis, MN: IPUMS. Available online at <https://doi.org/10.18128/D010.V9.0>.
- SADKA, JOYCE, SEIRA, ENRIQUE, & WOODRUFF, CHRISTOPHER. 2018. *Information and Bargaining through Agents: Experimental Evidence from Mexico's Labor Courts*. Tech. rept. National Bureau of Economic Research.
- SIMONOVSKA, INA, & WAUGH, MICHAEL E. 2014. The Elasticity of Trade: Estimates and Evidence. *Journal of International Economics*, **92**(1), 34–50.
- STEWART, FRANCES. 1977. *Technology and Underdevelopment*. London, UK: The Macmillan Press, Ltd.
- WAUGH, MICHAEL. 2010. International Trade and Income Differences. *American Economic Review*, **100**(5), 2093–2124.
- WORLD BANK. 2018. *World Development Indicators*. Washington DC: World Bank.

# Online Only Appendix

## Table of Contents

<b>A Data Appendix</b>	<b>43</b>
A.1 Representative Data Sources . . . . .	43
A.2 Wage Trends in Representative Data Sources . . . . .	44
A.3 Details on U.S. Historical Data . . . . .	45
A.4 Details on Longitudinal Surveys of Australian Youth Analysis . . . . .	47
A.5 Occupational Codes for Middle Managers . . . . .	49
<b>B Theory Appendix</b>	<b>53</b>
B.1 Aggregate industry production function . . . . .	53
B.2 Isomorphism to economy with continuum of sectors and Frechét shocks . . . . .	54
B.3 Proof of Proposition 1 . . . . .	59
B.4 Formal Statement and Proof of Output Decomposition . . . . .	64

## A Data Appendix

### A.1 Representative Data Sources

The Company's database covers a very particular population of jobs and firms - middle managers at modern business enterprises. It is not well-suited for studying representative workers or firms in developing countries because those workers and firms do not engage the Company's services and so do not appear in the Company's database. We assemble nationally representative datasets to provide this information and the necessary context for our results.

We draw on two main data sources. The bulk of our data comes from censuses from [Minnesota Population Center \(2019\)](#). We focus on countries that include an employment status variable and record occupation using the ISCO-08 occupation scheme at the 2-, 3-, or 4-digit level. We augment this data with labor force surveys from Bangladesh, Bolivia, Canada, France, Guyana, and the United States. These surveys are available to us from past work, nationally representative, and include occupation recorded using the ISCO-08 scheme with the necessary detail. The lone exception is the United States Current Population Survey (CPS) data. The United States generally uses its own occupation coding schemes that resemble but are not the same as the ISCO schemes. We include them because the United States is a useful benchmark for many counterfactuals.

We combine data from both types of sources and treat them symmetrically. We use them to construct facts about the distribution of employment and about wages for middle managers. When studying the distribution of employment we impose minimal sample selection criteria so that our results align as closely as possible with aggregate facts. We focus on people who are employed and age 16–70 with valid responses to a few key questions of interest (employment status, occupation, and weight).

We use the occupation codes to define which workers are middle managers. Tables giving the full list of included codes are available in [Appendix A.5](#). Focusing on the ISCO-08 scheme, we include all managers; business professionals; business associate professionals; and all clerks as middle managers. We include as similar as possible a set of workers in the United States. All remaining workers are non-managers.

In [Figure 1](#) we compare the distribution of employment across ISCO-08 1-digit occupations between the Company's database and various benchmarks. For the representative data sources from poor countries this is simply the weighted employment share by 1-digit occupation code. The com-

parison to the U.S. business service sector is more involved. We start with the U.S. CPS data. We restrict attention to workers in the business service sector, which includes the industries advertising; business management and consulting; computer programming services; accounting, auditing, and bookkeeping services; personnel services; and commercial research, development, and testing. To permit a comparison of the occupational distribution for this subsample we construct a crosswalk from U.S. occupation codes to ISCO-08 1-digit occupations.

We also use the representative data sources to provide results on the compensation of typical middle managers. Here we impose stricter sample selection criteria. We focus on wage workers aged 16–70 with valid reports of earnings. Earnings are not collected from the self-employed in most countries. We have experimented with using Mincer earnings regressions to impute earnings for the self-employed and found very similar results, available upon request. Note that many countries do not provide earnings data, especially in [Minnesota Population Center \(2019\)](#), so our sample here is smaller. For each country we convert the provided earnings figure to annual earnings (to be comparable with the Company’s database). For the most part this entails, say, multiplying monthly earnings by months worked or twelve if months worked is not provided. We then convert earnings to year 2017 international dollars exactly as we did for the Company database. Finally, we merge on 2017 GDP per worker from [World Bank \(2018\)](#).

## **A.2 Wage Trends in Representative Data Sources**

Section 3.1 describes our findings on the compensation of middle managers from representative data sources; here we provide the details. All findings are computed using the sample as described in the previous subsection. We parallel our treatment of the Company database as closely as possible. We regress log annual earnings on log GDP per worker, treating each worker as an observation. We normalize the provided weights for each country to sum to one and then use the weights and robust standard errors in the regression. The resulting compensation elasticity shown in Table 2 is 0.63.

The representative database includes information on the worker’s occupation, but not on the firm. In column (2) we explore controlling for occupation fixed effects, which is the exact ISCO-08 2-digit occupation for each worker. This somewhat reduces the elasticity to 0.57. Still, the results are notably different from those in Table 2.

**TABLE A.1: COMPENSATION ELASTICITY: REPRESENTATIVE DATA**

	(1)	(2)
Log GDP p.w.	0.630*** (0.00521)	0.566*** (0.00548)
Fixed Effects	None	Job
R-squared	0.258	0.334
N	1537701	1533595

Standard errors in parentheses

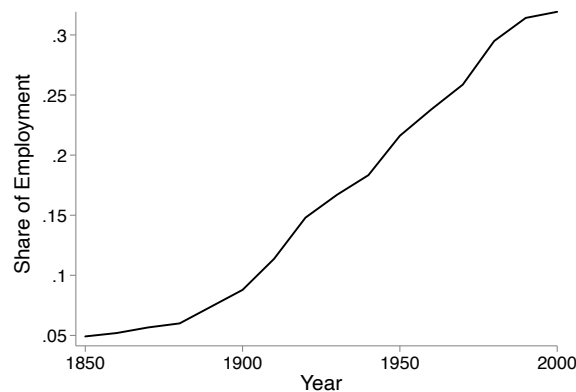
\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

### A.3 Details on U.S. Historical Data

Section 5.2 explains how we use the U.S. historical experience to calibrate the key parameter  $\theta$ . Here, we explain the underlying data sources and data construction. We also consider alternative sources and a longer time period for additional context.

Data on the employment share of middle managers over time (Figure 3a) is built using population census microdata from [Ruggles et al. \(2019\)](#). We limit the sample to employed 16–70 year olds with valid non-zero weights and valid responses to occupation and gender. We use the harmonized occupation variable `occ1950`, which translates original occupational responses into the 1950 coding scheme. Table A.4 shows the codes that are included in middle management. The employment share of middle managers is the year-specific weighted average of the share of workers with the corresponding occupations. The data are available from 1850. Figure A.1 shows that there was an acceleration in the employment of middle management from around 1880, with the employment share leveling off around 1980.

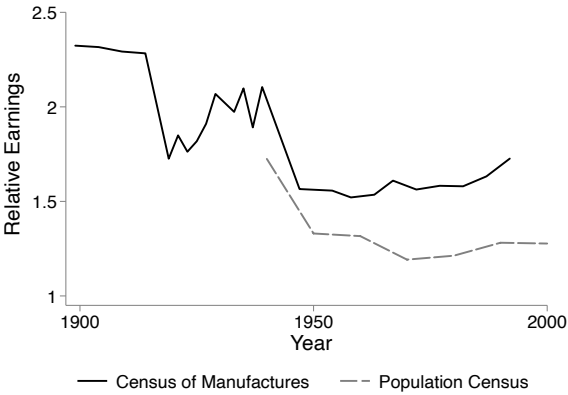
**FIGURE A.1: MIDDLE MANAGEMENT EMPLOYMENT SHARE: LONG RUN**



Data on relative earnings of middle managers (Figure 3b) come from two sources. In the text we focus on earnings data from the Census of Manufactures, which are taken from various years' general reports. They typically list the average number of salaried and wage workers over the course of the year (later, non-production and production workers) and the total bill for each type of worker. We compute the average salary by dividing the total bill by the total employment for each type, then take the ratio.

From 1940 onward we can also construct the relative earnings from population census microdata. We use the same sample as for employment shares, but further restrict our attention to workers who worked for wages and salary (excluding the self-employed) and who report positive wage and salary income. We construct mean log annual income for workers with middle manager and non-middle manager occupations to help limit the effect of outliers and the skewness of the wage distribution. We exponentiate and take the ratio for each year. Figure A.2 compares the two series for the longest period possible, 1900–2000. We see that the population census tracks the Census of Manufactures fairly closely for the period of overlap and the level is if anything slightly lower.

**FIGURE A.2: RELATIVE EARNINGS OF MANAGERS: LONG-RUN, ALTERNATIVE SOURCES**

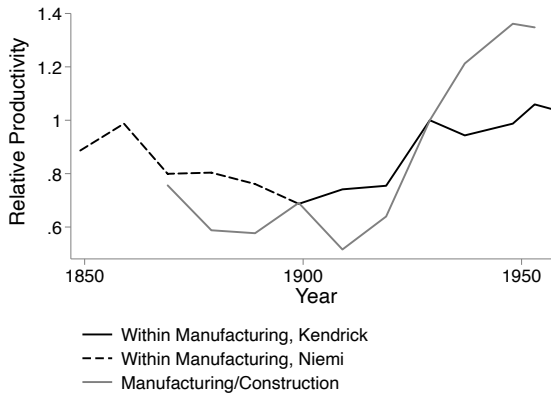


The relative productivity growth data in Figure 3c draw on the work of Kendrick (1961), who provides an index of real labor productivity (output per hour worked) for twenty manufacturing industries from 1899–1957, drawing again on the Census of Manufactures. We take the modern sector to consist of food, chemicals, petroleum, primary metals, non-electrical machinery, and transportation equipment. The traditional sector is defined as textiles, apparel, furniture, printing, leather, and instruments. Some remaining industries are omitted from this comparison because they were only partly modernized by 1919 (tobacco, lumber, paper, rubber, and stone/clay/glass; these exhibit even

faster subsequent productivity growth) or because they were not mentioned at all (beverages, fabricated metals, and electrical machinery).

We consider two extensions of this analysis. First, we consider using instead a comparison of overall manufacturing to construction productivity. Manufacturing benefited greatly from modern technologies while construction remains dominated by smaller firms even today. Here we use again with [Kendrick \(1961\)](#), who provides an index of real labor productivity for manufacturing and construction from 1869–1953. As we see in Figure A.3, the trends are similar from 1900 onward.

**FIGURE A.3: RELATIVE MODERN PRODUCTIVITY: LONG-RUN, ALTERNATIVE SOURCES**



Second, we consider sources that enable us to study relative productivity in the longer run. Kendricks’ figures already extend back to 1869 for construction and manufacturing and show no evidence of a trend in productivity prior to 1910. We add to this by collecting estimates of real productivity growth by manufacturing industry from [Niemi \(1972\)](#). Niemi follows Kendricks by using reported volumes of the Census of Manufactures to estimate productivity back as far as 1839, although he makes fewer adjustments and corrections than Kendricks does for later years. We chain his real labor productivity growth series on to those of [Kendrick \(1961\)](#) to provide preliminary evidence on labor productivity changes prior to 1899. As shown in Figure A.3, there is again little evidence of any prior productivity trend, consistent with new, major developments from around 1900.

**A.4 Details on Longitudinal Surveys of Australian Youth Analysis**

The Longitudinal Surveys of Australian Youth is a long-running research project that tracks the progress of students through school and into the early workforce. It is managed and funded by

the Australian Government Department of Education, Skills and Employment, with support from various level of the Australian government. Since 2003, the initial wave of the survey has been integrated with the Organization for Economic Cooperation and Development (OECD) Programme for International Student Assessment (PISA). Thus, the initial wave contains PISA scores for about 14,000 15-year old students per wave. Respondents are tracked for up to ten years, to age 25, with information on progress through schooling and then entry into the labor market collected over time.

Given the ten-year time horizon for the data, three waves of the survey are completed: the 2003, 2006, and 2009 cohorts ([Australian Government Department of Education & Employment, 2020a,b,c](#)). We collect data from all three waves and pool them for our analysis. Each contains similar information in terms of PISA test scores and employment and occupation outcomes at later waves. Pooling helps especially with increasing our sample size for students with low PISA test scores, which is important given low average test scores in developing countries.

We focus on reading test scores since literacy is important for management roles. PISA does not assign each worker a unique score. Instead, it assigns five “plausible values” per subject, which is designed to account for sampling variation in test scores. We implement the preferred approach of repeating the analysis for each potential score and then averaging the outcomes.

Our primary outcome of interest is adult occupation. We use the occupation at age 25 whenever possible. Some young adults lack an occupation because they are not working, do not provide enough occupational detail to permit coding, or have attrited from the survey. To combat this, we iterate backwards from age 25 for those who lack a valid occupation and explore whether they provide one at an earlier age. If they do, we use the latest possible occupation, although we disregard occupations provided before age 21.

We translate occupations into middle manager and professional roles. The LSAY uses the ANZSCO first edition occupation coding scheme, which is a modified but recognizable version of ISCO coding schemes. [Table A.5](#) gives the mapping from this scheme into management occupations. We define professionals as anything in the 1-digit category 2: Professionals.

Our analysis simply computes the share of workers in various test score ranges who make the occupational choices. All analyses are weighting using the provided longitudinal weights that adjust for attrition.



**TABLE A.2: OCCUPATIONAL CODES FOR MIDDLE MANAGERS: ISCO-08**

Codes	Title
11	Chief Executives, Senior Officials and Legislators
12	Administrative and Commercial Managers
13	Production and Specialized Services Managers
14	Hospitality, Retail and Other Services Managers
24	Business and Administration Professionals
33	Business and Administration Associate Professionals
41*	General and Keyboard Clerks
42*	Customer Services Clerks
43*	Numerical and Material Recording Clerks
44*	Other Clerical Support Workers

Codes reported at the 2-digit level. \* indicates occupations only included in the broad definition of middle management.

### **A.5 Occupational Codes for Middle Managers**

This appendix provides the occupational codes that are included in middle management in various data sources.

**TABLE A.3: OCCUPATIONAL CODES FOR MIDDLE MANAGERS: US CPS**

Codes	Title
003–019	Executive, Administrative, and Managerial Occupations
023–037	Management Related Occupations
303–307*	Supervisors, Administrative Support Occupations
308–309*	Computer Equipment Operators
313–315*	Secretaries, Stenographers and Typists
316–323*	Information Clerks
325–336*	Records Processing Occupations, Except Financial
337–344*	Financial Records Processing Occupations
345–347*	Duplicating, Mail, and Other Office Machine Operators
348–353*	Communications Equipment Operators
354–357*	Mail and Message Distributing Occupations
359–374*	Material Recording, Scheduling, and Distributing Clerks
375–378*	Adjusters and Investigators
379–389*	Miscellaneous Administrative Support Occupations

Codes refer to the 1988 and 1993 U.S. Current Population Surveys, which in turn used 1980 and 1990 U.S. Census occupation codes. \* indicates occupations excluded from the narrower definition of middle management used in robustness checks.

**TABLE A.4: OCCUPATIONAL CODES FOR MIDDLE MANAGERS:  
HISTORICAL US CENSUS**

Codes	Title
000	Accountants and auditors
072	Personnel and labor relations workers
200	Buyers and department heads, store
201	Buyers and shippers, farm products
204	Credit men
205	Floormen and floor managers, store
210	Inspectors, public administration
250	Officials and administrators (n.e.c.), public administration
260	Officials, lodge, society, union, etc.
270	Postmasters
280	Purchasing agents and buyers (n.e.c.)
290	Managers, officials, and proprietors (n.e.c.)
310	Bookkeepers
321*	Collectors, bill and account
322*	Dispatchers and starters, vehicle
325*	Express messengers and railway mail clerks
335*	Mail carriers
341*	Office machine operators
342*	Shipping and receiving clerks
350*	Stenographers, typists, and secretaries
365*	Telegraph operators
370*	Telephone operators
390*	Clerical and kindred workers (n.e.c.)
400	Advertising agents and salesmen
450	Insurance agents and brokers
470	Real estate agents and brokers
480	Stock and bond salesmen

Codes refer to occ1950, the standardized 1950 U.S. Census codes used by [Rugles et al. \(2019\)](#) to harmonize occupations over time. \* indicates occupations excluded from the narrower definition of middle management used in robustness checks.

**TABLE A.5: OCCUPATIONAL CODES FOR MIDDLE MANAGERS:  
ANZSCO FIRST EDITION**

Codes	Title
1111–1113	Chief Executives, General Managers and Legislators
1311–1399	Specialist Managers
1411–1499	Hospitality, Retail and Service Managers
2211–2212	Accountants, Auditors and Company Secretaries
2221–2223	Financial Brokers and Dealers, and Investment Advisers
2231–2233	Human Resource and Training Professionals
2244	Intelligence and Policy Analysts
2245	Land Economist and Valuers
2247	Management and Organization Analysts
2249	Other Information and Organization Professionals
2251–2254	Sales, Marketing and Public Relations Professionals
5111–5999	Clerical and Administrative Workers
5111	Contract, Program and Project Administrators <sup>†</sup>
5122	Practice Managers <sup>†</sup>
5211	Personal Assistants <sup>†</sup>
5512	Bookkeepers <sup>†</sup>
5522	Credit and Loans Officers <sup>†</sup>
5991	Conveyancers and Legal Executives <sup>†</sup>
5992	Court and Legal Clerks <sup>†</sup>
5995	Inspectors and Regulatory Officials <sup>†</sup>
5996	Insurance Investigators, Loss Adjusters and Risk Surveyors <sup>†</sup>

Codes refer to ANZSCO first edition, used to code occupations of young adults in the LSAY. <sup>†</sup> indicates the subset of clerical and administrative workers included in the narrow definition of middle management.

## B Theory Appendix

### B.1 Aggregate industry production function

We show that there exists an aggregate industry production function of the form (B.1) by showing that (B.1) generates the same aggregate supply correspondence as the full industry model with free entry, where the supply correspondence is defined as a mapping from a price vector to a set of profit maximizing output-input combinations.

$$F^{ind}(h_p, h_m; z_T, z_M) = \begin{cases} z_T h_p & \text{if } e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} > \frac{w_p}{z_T} \\ \kappa z_M h_p^{1-\alpha} h_m^\alpha & \text{if } e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \leq \frac{w_p}{z_T} \end{cases} \quad (\text{B.1})$$

We start with the supply correspondence of the proposed aggregate production function (B.1), using the notation  $(\ell_p, \ell_m, y)$  for an input-output combination with production services  $\ell_p$ , management services  $\ell_m$ , and output  $y$ .

First, we note the trivial fact that given constant returns to scale, there is no profit maximizing input-output combination when the price is below unit cost, and the profit maximizing combination is unbounded when the price is above unit costs:

$$S^{agg}(P, w_p, w_m) = \begin{cases} \emptyset & \text{if } P < \min \left\{ \frac{w_p}{z_T}, e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \right\}, \\ \tilde{S}_{P^*}^{agg}(w_p, w_m) & \text{if } P = \min \left\{ \frac{w_p}{z_T}, e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \right\}, \\ \{(\infty, \infty, \infty)\} & \text{if } P > \min \left\{ \frac{w_p}{z_T}, e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \right\}. \end{cases} \quad (\text{B.2})$$

where  $\tilde{S}_{P^*}^{agg}$  is the supply correspondence when the price equals the unit cost, and is given by

$$\tilde{S}_{P^*}^{agg}(w_p, w_m) = \begin{cases} \left\{ \left( \frac{y}{z_T}, 0, y \right) : y \geq 0 \right\} & \text{if } \frac{w_p}{z_T} < e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \\ \left\{ \left( \frac{y}{z_M} (1-\alpha) \left( \frac{w_m}{w_p} \right)^\alpha, \frac{y}{z_M} \alpha \left( \frac{w_m}{w_p} \right)^{-(1-\alpha)}, y \right) : y \geq 0 \right\} & \text{if } \frac{w_p}{z_T} \geq e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \end{cases}, \quad (\text{B.3})$$

where the two cases represent the optimal input-output vectors if the traditional versus modern technology is selected. Due to constant returns to scale, all vectors yield the same (zero) profit, and the multiplicative constant on  $y$  in the inputs reflects the factor requirement of a unit of output given the technology choice and the factor price vector.

To calculate the supply correspondence with multiple firms and free entry, we note that the supply

is only non-zero and well defined if the the price equals the minimum point on the average cost curve. If the price is lower, there will be no entry, and if the price is higher, there will be unlimited entry. The minimum of the average cost curve is  $\min \left\{ \frac{w_p}{z_T}, e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M} \right\}$ , that is, the same as the unit cost for the proposed industry representative firm. Hence (B.2) also holds for the supply correspondence in the multi-firm case.

For the case when  $P$  equals the minimum of the average cost curve, it is clear that (B.3) holds for the case when the minimum average cost is attained by the traditional technology. Indeed, in this case, free entry ensures that total output is  $y = M z_T \bar{h}$  and total production service input is  $M \bar{h}$  where  $M$  is the number of entering firms. Neglecting integer constraints on  $M$  it is clear that  $\left\{ \left( \frac{y}{z_T}, 0, y \right) : y \geq 0 \right\}$  is the set of profit-maximizing input-output combinations, just as in (B.3).

Finally, when the modern technology is selected, we use that the average cost minimizing output level is

$$y^* = \kappa z_M \left( \frac{w_m \gamma}{w_u \eta} \right)^{-\frac{1}{\eta+\gamma}} = \kappa z_M \left( \frac{w_m / \alpha}{w_u / (1-\alpha)} \right)^{-\frac{1}{\eta+\gamma}},$$

which implies

$$\begin{aligned} \ell_p^* &= \left( \frac{y^*}{\kappa z_M} \right)^{\frac{1}{1+\tilde{\eta}}} = \left( \frac{y^*}{\kappa z_M} \right)^{1-\eta} = \frac{y^*}{\kappa z_M} \left( \frac{1-\alpha}{\alpha} \right)^\alpha \left( \frac{w_m}{w_u} \right)^\alpha = \frac{y^*}{z_M} (1-\alpha) \left( \frac{w_m}{w_u} \right)^\alpha \\ \ell_m^* &= (\ell_p^*)^{1+\tilde{\gamma}} = \left( \frac{y^*}{\kappa z_M} \right)^{\frac{1+\tilde{\gamma}}{1+\tilde{\eta}}} = \left( \frac{y^*}{\kappa z_M} \right)^{1+\gamma} = \frac{y^*}{z_M} \alpha \left( \frac{w_m}{w_p} \right)^{-(1-\alpha)}, \end{aligned}$$

and we recover the same supply correspondence as in (B.2).

## B.2 Isomorphism to economy with continuum of sectors and Fréchet shocks

**Proposition 2.** *Assume that the economy has the same household sector as the baseline model, but a production sector with  $Y = \left( \int_0^1 y(k)^{\frac{\sigma-1}{\sigma}} dk \right)^{\frac{\sigma}{\sigma-1}}$ , where  $y(k)$  is given by (B.1), and  $\frac{z_T(k)}{Z_T}, \frac{z_M(k)}{Z_M} \stackrel{iid}{\sim} \text{Fréchet}(\theta)$ . The equilibrium in that economy has the same wages, prices, aggregate outputs, sectoral employments and sectoral revenues as the baseline economy.*

*Proof.* As the case for the two sector model, we reparameterize the model so that it is in terms of labor

instead of efficiency units. This means that the production functions are

$$F^{ind}(h_p, h_m; z_T, z_M) = \begin{cases} \tilde{z}_T \ell_p & \text{if } e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M} > \frac{\hat{w}_p}{\tilde{z}_T} \\ \kappa \tilde{z}_M \ell_p^{1-\alpha} h_m^\alpha & \text{if } e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M} \leq \frac{\hat{w}_p}{\tilde{z}_T} \end{cases} \quad (\text{B.4})$$

with  $\tilde{z}_T = z_T Q_p$  and  $\tilde{z}_M = z_M Q_p^{1-\alpha} Q_m^\alpha$ , which implies that

$$\frac{\tilde{z}_T}{Z_T Q_p}, \frac{\tilde{z}_M}{Z_M Q_p^{1-\alpha} Q_m^\alpha} \stackrel{iid}{\sim} \text{Frechét}(\theta).$$

As in the two-sector model, consumer optimization implies that we have  $\frac{\hat{w}_m}{\hat{w}_p} = e^\tau$ , where  $\hat{w}_p = Q_p w_p$  and  $\hat{w}_m = Q_m w_m$ .

Given the cost functions for the individual industries and cost minimization by the final output producer

$$p(k) = \min \left\{ \frac{\hat{w}_p}{\tilde{z}_T(k)}, e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M(k)} \right\}$$

$$y(k) = Y \left( \frac{p(k)}{P} \right)^{-\sigma}.$$

We write  $I_M = \left\{ k : e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M} < \frac{\hat{w}_p}{\tilde{z}_T} \right\}$  for the set of industries which choose the modern industry.

This set has measure

$$\mu(I_M) \equiv s_M = \mathbb{P} \left( e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M} < \frac{\hat{w}_p}{\tilde{z}_T} \right) = \frac{Z_M^\theta (Q_p^{1-\alpha} Q_m^\alpha)^\theta e^{-\tilde{\tau}\theta} e^{-\tau\alpha\theta}}{Z_T^\theta Q_p^\theta + Z_M^\theta (Q_p^{1-\alpha} Q_m^\alpha)^\theta e^{-\tilde{\tau}\theta} e^{-\tau\alpha\theta}}.$$

Furthermore, we can write labor use per unit of output as

$$l(k) = \begin{cases} \frac{1}{\tilde{z}_T(k)} & \text{if } \frac{\hat{w}_p}{\tilde{z}_T(k)} \leq e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M(k)} \\ \frac{(1-\alpha)e^{\tau\alpha} + \alpha e^{-(1-\alpha)\tau}}{\tilde{z}_M(k)} & \text{if } \frac{\hat{w}_p}{\tilde{z}_T(k)} > e^{\tilde{\tau}} \frac{\hat{w}_p^{1-\alpha} \hat{w}_m^\alpha}{\tilde{z}_M(k)}. \end{cases}$$

The labor market clearing condition then is

$$L = \int_0^1 l(k) y(k) dk = Y \left( \int_{[0,1]-I_M} l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} dk + \int_{I_M} l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} dk \right). \quad (\text{B.5})$$

By reparameterizing in terms of distributions of  $\tilde{z}_m$  and  $\tilde{z}_t$ , we obtain

$$\begin{aligned} \int_{I_m^c} l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} dk &= (1 - \pi) \mathbb{E} \left[ l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} \mid k \notin I_m \right] \\ &= (1 - \pi) \frac{1}{\hat{w}_p P^{-\sigma}} \mathbb{E} [\tilde{p}(k)^{1-\sigma} \mid k \notin I_m] \end{aligned} \quad (\text{B.6})$$

$$\begin{aligned} \int_{I_m} l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} dk &= \pi \mathbb{E} \left[ l(k) \left( \frac{p(k)}{P} \right)^{-\sigma} \mid k \in I_m \right] \\ &= e^{-\tilde{\tau}} \pi \frac{[(1 - \alpha) + \alpha e^{-\tau}]}{\hat{w}_p P^{-\sigma}} \mathbb{E} [p(k)^{1-\sigma} \mid k \in I_m]. \end{aligned} \quad (\text{B.7})$$

The properties of the Frechét function implies that

$$\begin{aligned} \mathbb{E}[p(k)^{1-\sigma} \mid k \in I_m] &= \mathbb{E}[\min \{p_m(k)^{1-\sigma}, p_t(k)^{1-\sigma}\} \mid p_m(k) < p_t(k)] \\ &= \mathbb{E}[\min \{p_m(k)^{1-\sigma}, p_t(k)^{1-\sigma}\}] \\ &= \mathbb{E}[p(k)^{1-\sigma}] \\ &= P^{1-\sigma}, \end{aligned} \quad (\text{B.8})$$

where the last step uses the standard price index formula for CES aggregators. Putting (B.6)-(B.8) back into (B.5) implies

$$\hat{w}_p L = PY (1 - \pi + \pi e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]).$$

Moreover, we have

$$\begin{aligned} P &= [\mathbb{E}p(k)^{1-\sigma}]^{\frac{1}{1-\sigma}} = \gamma \hat{w}_p \left[ (Q_p Z_t)^\theta + [Z_m (Q_p^{1-\alpha} Q_m^\alpha) e^{-\tilde{\tau}} e^{-\alpha\tau}]^\theta \right]^{-\frac{1}{\theta}} \\ &= \hat{w}_p \left( \tilde{Z}_t^\theta + [\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^\theta \right)^{-\frac{1}{\theta}} \end{aligned}$$



and we obtain

$$\begin{aligned}
Y &= L \frac{\left(\tilde{Z}_t^\theta + [\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^\theta\right)^{\frac{1}{\theta}}}{1 - \pi + \pi e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \\
&= L \frac{\left(\tilde{Z}_t^\theta + [\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^\theta\right)^{\frac{\theta+1}{\theta}}}{\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \\
&= L \left( \frac{\tilde{Z}_t^\theta}{\left(\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]\right)^{\frac{\theta}{\theta+1}}} + \frac{[\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^\theta}{\left(\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]\right)^{\frac{\theta}{\theta+1}}} \right)^{\frac{\theta+1}{\theta}} \\
&= L \left( \left[ \frac{\tilde{Z}_t^{\theta+1}}{\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \right]^{\frac{\theta}{\theta+1}} + \left[ \frac{[\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^{\theta+1}}{\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \right]^{\frac{\theta}{\theta+1}} \right)^{\frac{\theta+1}{\theta}} \\
&= L \left( \left[ \frac{\tilde{Z}_t^{\theta+1}}{\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \right]^{\frac{\theta}{\theta+1}} + \left[ \frac{[\tilde{Z}_m e^{-\tilde{\tau}} e^{-\alpha\tau}]^{\theta+1}}{\tilde{Z}_t^\theta + \tilde{Z}_m^\theta e^{-\tilde{\tau}} [1 - \alpha + \alpha e^{-\tau}]} \right]^{\frac{\theta}{\theta+1}} \right)^{\frac{\theta+1}{\theta}} \\
&= \left( (\tilde{Z}_t L \times e_t)^{\frac{\theta}{\theta+1}} + \left[ \frac{\tilde{Z}_m L}{e^{\tau\alpha} [(1 - \alpha) + \alpha e^{-\tau}]} \times e_m \right]^{\frac{\theta}{\theta+1}} \right)^{\frac{\theta+1}{\theta}} \\
&= \left( Y_T^{\frac{\theta}{\theta+1}} + Y_M^{\frac{\theta}{\theta+1}} \right)^{\frac{\theta+1}{\theta}}.
\end{aligned}$$

where  $Y_T = \tilde{Z}_t L e_t$  and  $Y_M = \frac{\tilde{Z}_m L}{(1 - \alpha)e^{\tau\alpha} + \alpha e^{-\tau(1 - \alpha)}} e_m$ , where  $e_m$  and  $e_t$  are given by

$$e_M = 1 - e_T \equiv \frac{\ell_{m,p} + \ell_{m,m}}{L} = \frac{\tilde{Z}_M^\theta e^{-\tilde{\tau}\theta} e^{-\tau\alpha\theta} [e^{-\tilde{\tau}} (1 - \alpha + \alpha e^{-\tau})]}{\tilde{Z}_T^\theta + \tilde{Z}_M^\theta e^{-\tilde{\tau}\theta} e^{-\tau\alpha\theta} [e^{-\tilde{\tau}} (1 - \alpha + \alpha e^{-\tau})]}.$$

Noting that the labor market clearing condition can be written

$$\begin{aligned}
L &= Y_T \frac{\partial \tilde{C}_T}{\partial \hat{w}_p} + Y_M \left( \frac{\partial \tilde{C}_T}{\partial \hat{w}_p} + \frac{\partial \tilde{C}_T}{\partial \hat{w}_m} \right) \\
&= \frac{Y_T}{Z_T} + \frac{Y_M}{Z_M} \left( \frac{\hat{w}_m}{\hat{w}_p} \right)^\alpha \left[ 1 - \alpha + \alpha \left( \frac{\hat{w}_m}{\hat{w}_p} \right)^{-1} \right] \\
&= \frac{P_T Y_T}{Q_p w_p} + \frac{Y_M P_M}{Q_p w_p} [1 - \alpha + \alpha e^{-\tau}], \tag{B.9}
\end{aligned}$$

and that

$$\frac{P_M Y_M}{P_T Y_T} = \left( \frac{P_M}{P_T} \right)^{-\theta} = e^{-\tilde{\tau}\theta} e^{-\alpha\tau\theta} \left( \frac{\tilde{Z}_M}{\tilde{Z}_T} \right)^\theta, \tag{B.10}$$

we obtain

$$Y_T = L \times e_T \times \tilde{Z}_T \quad (\text{B.11})$$

$$Y_M = L \times e_M \times \frac{\tilde{Z}_M}{(1-\alpha)e^{\tau\alpha} + \alpha e^{-(1-\alpha)\tau}}. \quad (\text{B.12})$$

Here, the terms  $e_T$  and  $e_M$  denote the share of labor employed in the modern and traditional sector respectively, and the term  $(1-\alpha)e^{\tau\alpha} + \alpha e^{-(1-\alpha)\tau}$  regulates the labor requirement in the modern sector, and it is minimized when wages are equalized ( $\tau = 0$ ).

□

**Economy with stochastic productivities.** The equilibrium is a set of prices  $w_p, w_m, \{p(k)\}, P$  and quantities  $H_p, H_m, C, Y, \{y(k)\}$  such that households optimize given (14) and (13), industry prices  $p(k)$  are given by the minimizing average cost, the aggregate price level is given by  $P = [\mathbb{E}p(k)^{1-\sigma}]^{\frac{1}{1-\sigma}}$ ,  $Y$  is given by (10)-(12),  $\{y(k)\}$  is a cost-minimizing choice given  $\{p(k)\}$  and  $Y$ , the good market clears,  $C = Y$ , transfers equal the total value going to markups:

$$T = [1 - \exp(-\tau)]w_m H_m + (e^{\tilde{\tau}} - 1) \int_{k \in I_M} [w_p h_p(k) + w_m h_m(k)] dk,$$

and labor markets clear  $H_x = \int_0^1 h_x(k) dk$  for  $x = p, m$ .

**Equilibrium.** Free entry implies that prices equal unit costs (including wedges)

$$p(k) = \min \left\{ \frac{w_p}{z_T(k)}, e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M(k)} \right\}.$$

Given these prices, the CES price index satisfies (Eaton & Kortum, 2002):

$$P = \left( \int_0^1 p(k)^{1-\sigma} dk \right)^{\frac{1}{1-\sigma}} = \gamma \left[ \left( \frac{w_p}{\bar{z}_T} \right)^{-\theta} + \left( e^{\tilde{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{\bar{z}_M} \right)^{-\theta} \right]^{-1/\theta}, \quad (\text{B.13})$$

where  $\gamma = \Gamma\left(\frac{\theta+1-\sigma}{\theta}\right)^{\frac{1}{\sigma-1}}$  and  $\Gamma$  is the gamma function (Eaton & Kortum, 2002). Normalizing the aggregate price level to  $P \equiv \gamma$  implies

$$1 = \left(\frac{w_p}{\bar{z}_T}\right)^{-\theta} + \left(e^{\bar{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{\bar{z}_M}\right)^{-\theta}. \quad (\text{B.14})$$

The share of industries with the large-scale technology is

$$\pi = \mathbb{P}\left(e^{\bar{\tau}} \frac{w_p^{1-\alpha} w_m^\alpha}{z_M(k)} \leq \frac{w_p}{z_T(k)}\right) = \frac{\bar{z}_M^\theta (w_p^{1-\alpha} w_m^\alpha)^{-\theta} e^{-\bar{\tau}\theta}}{\bar{z}_T^\theta w_p^{-\theta} + \bar{z}_M^\theta (w_p^{1-\alpha} w_m^\alpha)^{-\theta} e^{-\bar{\tau}\theta}}. \quad (\text{B.15})$$

The share  $\pi$  coincides with the expenditure share on large-scale industries, since the Fréchet-distribution of the  $z$ 's implies that the price distribution of the two technologies is the same, conditional on them being the lowest-cost industry.

### B.3 Proof of Proposition 1

We prove the result in three steps, starting with that  $y_t^*(w_m)$  is weakly decreasing in  $w_m$  for every  $t$ , then showing that the optimal technology  $t_m^*(w_m)$  is weakly decreasing in  $w_m$ , and lastly showing that  $y^*(w_m) \equiv y_{t_m^*(w_m)}^*$  is weakly decreasing in  $w_m$ .

**Result 1:**  $y_t^*(w_m)$  is weakly decreasing for every  $t$ . We prove the result by contradiction, so we assume that there exists a technology  $t$  and  $w'_m > w_m$  such that  $y_t^*(w'_m) > y_t^*(w_m)$ . For simplicity, we drop the subscript  $t$  and the explicit dependence on  $w_m$ , and we write  $y^*$  and  $y^{*'}$  for  $y_t^*(w_m)$ , and  $y_t^*(w'_m)$ . To obtain a contradiction, we show that if a firm produces output  $y^{*'}$  under the wage  $w_m$ , it obtains a lower average cost than when it produces  $y^*$ , contradicting that  $y^*$  minimizes unit costs.

In particular, the optimality of  $y^{*'}$  for  $w'_m$  implies

$$\bar{c}(y^{*'}; w'_m) \leq \bar{c}(y^*; w'_m) \iff f(y^{*'}) + w'_m \tilde{g}(y^{*'}) \leq f(y^*) + w'_m \tilde{g}(y^*), \quad (\text{B.16})$$

from which we can derive

$$\begin{aligned}
f(y^{*'}) + w_m \tilde{g}(y^{*'}) &= f(y^{*'}) + w'_m \tilde{g}(y^{*'}) + (w_m - w'_m) \tilde{g}(y_t^{*'}) \\
&\leq f_t(y_t^*) + w'_m \tilde{g}(y_t^*) + (w_m - w'_m) g_t(y_t^{*'}) \\
&= f_t(y^*) + w_m \tilde{g}(y^*) + (w'_m - w_m) \tilde{g}(y_t^*) + (w_m - w'_m) \tilde{g}(y^{*'}) \\
&= f_t(y^*) + w_m \tilde{g}(y^*) + (w'_m - w_m) [g_t(y^*) - g_t(y^{*'})] \\
&< f_t(y^*) + w_m \tilde{g}(y^*),
\end{aligned}$$

where the second line uses (B.16), and the last line uses  $w'_m > w_m$ ,  $y_t^{*'} > y_t^*$ , and that  $\tilde{g}$  is increasing. This expression implies that  $y^{*'}$  attains a lower average cost than  $y^*$ , contradicting the optimality of for  $w_m$ . Hence,  $y_t^*(w_m)$  is weakly increasing in  $w_m$  for all  $t$ .

**Result 2:**  $t^*(w_m)$  is weakly decreasing in  $w_m$ . This is the most complicated result, and we will present it using a series of definitions lemmas building up to the final result (the complexity derives from the weak assumptions we made about differentiability and convexity, which precludes standard reasoning based on first-order conditions).

Our first result is that the attained minimum average cost is continuous in the managerial wage level.

**Lemma 3.** For each technology  $t$ , the attained unit cost  $\bar{c}_t[y_t^*(w_m); w_m]$  is continuous in  $w_m$ .

*Proof of lemma.* Consider a change in the wage  $w'_m = w_m + \delta'$  for some small number  $\delta'$ . Again writing  $y^*$  and  $y^{*'}$  for  $y_t^*(w_m)$ , and  $y_t^*(w'_m)$ , we can write the optimality conditions

$$\begin{aligned}
c[y^{*'}, w'_m] &\leq c[y^*, w'_m] = c[y^*, w_m] + \delta \tilde{g}(y^*) \\
c[y^*, w_m] &\leq c[y^{*'}, w_m] = c[y^{*'}, w'_m] - \delta \tilde{g}(y^{*'}).
\end{aligned}$$

Here, the first line captures that you are better off with  $y^{*'}$  than  $y^*$  if the wage is  $w'_m$ , and the second line captures that you are better off with  $y^*$  if the wage is  $w'_m$ . The equal signs just substitute  $w'_m = w_m + \delta$  into the standard equation for the average cost.

Together, the two equations imply

$$c[y^*, w_m] + \delta \tilde{g}(y^{*'}) \leq c[y^{*'}, w'_m] \leq c[y^*, w_m] + \delta \tilde{g}(y^*),$$

capturing that the minimum average cost change from moving  $w_m$  is located between the extra wage cost coming from the pre-change average management use and the post-change average management use. Letting  $\delta \rightarrow 0$ , implies that  $\lim_{w'_m \rightarrow w_m} \bar{c}[y^*(w'_m), w'_m] = \bar{c}[y^*(w_m), w_m]$ . That is,  $\bar{c}$  is continuous in  $w_m$ .  $\square$

Using the result, we then establish exactly in what sense small-scale technologies choose smaller optimal output levels  $y_t^*$ . To this end, we define a collection of sets

$$Y_t = \left\{ y \in \mathbb{R}_{0\leq} : f_t(y) \leq \min_{t'} f_{t'}(y) \right\}, \quad t = 1, \dots, T$$

which, for each  $t$ , gives the set of output levels for which technology  $t$  has the lowest average use of production workers. While the average cost-minimizing output  $y_t^*$  for technology  $t$  might fall outside  $Y_t$ , it is possible to show that *whenever*  $t$  has the lowest global average cost,  $y_t^*$  does fall inside  $Y_t$ . We prove the following lemma.

**Lemma 4.** *If  $t^* = t$ , then  $y_t^* \in Y_t$ .*

*Proof.* Suppose that  $y_t^* \notin Y_t$ . Then there exists  $t'$  with  $f_{t'}(y_t^*) < f_t(y_t^*)$ . But then  $t'$  can obtain a lower average cost than  $t$ , so  $t^* \neq t$ .  $\square$

Next, we solve for the structure of the sets  $Y_t$  where  $t$  might have a globally optimal output level. We show that each set  $Y_t$  is an interval, and that they are arranged in increasing order, with potential overlap.

**Lemma 5.** *Each  $Y_t$  is a (possibly empty) closed interval in  $\mathbb{R}_+$ . The intervals jointly cover  $\mathbb{R}_+$  and are given in increasing order: if all intervals are non-empty, there exist  $0 = y_0 \leq y_1 \leq \dots \leq y_T$  and  $\bar{y}_1 \leq \dots \leq \bar{y}_T = \infty$  such that  $Y_1 = (y_0, y_1]$ ,  $Y_2 = [y_1, y_2]$ ,  $\dots$ ,  $Y_T = [y_{T-1}, y_T]$ . If  $Y_t$  is empty for any  $t$ , the same condition applies for the set of non-empty intervals.*

*Proof of lemma.*  $Y_t$  is an interval if  $y_1, y_2 \in Y_t$  with  $y_1 < y_2$  implies that  $z \in Y_t$  for all  $z \in (y_1, y_2)$ . To show that this holds, we proceed by contradiction, and assume that we can find  $y_1 < y_2, z$  and a technology  $t' \neq t$  such that

$$f_{t'}(z) < f_t(z) \iff \frac{f_{t'}(z)}{f_t(z)} < 1$$

However, now either  $t < t'$  and  $t > t'$ . If  $t < t'$ , the technology  $t'$  has a higher degree of scale economies than  $t$ ,  $\frac{f_{t'}(z)}{f_t(z)}$  is falling with  $z$ , and  $\frac{f_{t'}(y_2)}{f_t(y_2)} < 1$ , implying that  $y_2 \notin Y_t$ . If  $t > t'$ , the

technology  $t'$  has a lower degree of scale economies than  $t$ , and a symmetric reasoning can be used to establish that  $y_1 \notin Y_t$ . Hence, we obtain a contradiction, confirming that  $Y_t$  is an interval. The interval is closed in  $\mathbb{R}_+$ , since  $f_t(y) - \min_{t'} f_{t'}(y) \leq 0$  is the pre-image of a closed set under a continuous mapping.

To show that  $Y_t$  are in increasing order, consider two arbitrary technologies  $t < t'$ . To show that  $Y_{t'}$  is located weakly to the right of  $Y_t$ , assume for the purpose of contradiction that  $y' < y$  and  $y' \in Y_{t'}$  and  $y \in Y_t$ . In this case, we have

$$f_t(y') \geq f_{t'}(y') \implies f_t(y) > f_{t'}(y),$$

contradicting that  $y \in Y_t$ . Hence, no element in  $Y_{t'}$  is strictly to the left of any element in  $Y_t$ . Combined with  $\{Y_t\}$  being closed intervals covering the positive real line, the result follows.  $\square$

Using the lemmas, we can show the main result. Again, we proceed by contradiction and assume that there exists some interval where the optimal  $t$  is increasing in the wage. Formally, assume there exist two wages  $w_m < w'_m$  and  $t < t'$  such that  $\min t^*(w_m) = t$  and  $\min t^*(w'_m) = t'$ .

The proof strategy is to show that there has to exist a "switch point": a wage in the middle of  $w_m$  and  $w'_m$  where the low and high technology have exactly the same costs. Then we can use a perturbation method to show that a higher managerial wage is worse for the high technology, contradicting that a higher wage causes a switch to the high technology.

To find a candidate switch point, define

$$\hat{w}_m = \sup\{w''_m > w_m : \min t^*(w''_m) \leq t\}.$$

For  $\hat{w}_m$ , we can show that there exists  $\underline{\tilde{t}} \leq t < \bar{\tilde{t}}$  with  $\underline{\tilde{t}}, \bar{\tilde{t}} \in t^*(\hat{w}_m)$ . To see that it contains technologies larger than  $t$ , pick a sequence going down to  $\hat{w}_m$ :  $\tilde{w}_m \downarrow \hat{w}_m$  with  $\min t^*(\tilde{w}_m) > t$ . It is possible to find a technology  $\tilde{t}$  and an infinite subsequence  $\hat{w}_m^j$  so that  $c_{\tilde{t}}^*(\hat{w}_m^j) = \min_{t''} c_{t''}^*(\tilde{w}_m^j)$  for all elements of the subsequence. Since  $c_{\tilde{t}}^*$  and  $\min_{t''} c_{t''}^*$  are both continuous, taking the limit implies

$$c_{\tilde{t}}^*(\hat{w}_m) = \min_{t''} c_{t''}^*(\hat{w}_m).$$

To find  $\underline{\tilde{t}}$ , we first note that if  $\hat{w}_m = w_m$ , we trivially have  $c_{\tilde{t}}^*(\hat{w}_m) = \min_{t''} c_{t''}^*(w_m)$ . Otherwise, we can pick an increasing subsequence  $\tilde{w}_m \uparrow \hat{w}_m$  and since some  $\underline{\tilde{t}} \leq t$  has the minimum cost for an infinite

number of wages, we can pass to a sub-sequence to find  $\underline{t} \leq t$  with

$$c_{\underline{t}}^*(\hat{w}_m) = \min_{t'} c_{t'}^*(\hat{w}_m).$$

We now want to obtain a contradiction by showing that  $c_{\underline{t}}^*(\hat{w}_m + \delta) \leq c_{\tilde{t}}^*(\hat{w}_m + \delta)$  for any sufficiently small  $\delta$ . Indeed, this contradicts that  $\hat{w}_m$  is an infimum, since it shows that  $t^*(\hat{w}_m + \delta)$  contains  $\underline{t} \leq t$ .<sup>18</sup>

To show this, we write  $\underline{y}^*$  and  $\tilde{y}^*$  for the output level associated with  $\underline{t}$  and  $\tilde{t}$  respectively. From the lemma, we know that  $\underline{y}^* \in Y_t$  and  $\tilde{y}^* \in Y_{t'}$ , and that  $Y_t$  and  $Y_{t'}$  coincide at most in a singleton. Moreover, since  $\underline{t} < \tilde{t}$ , we have  $\underline{y}^* \leq \tilde{y}^*$ .

We start with the case  $\underline{y}^* < \tilde{y}^*$ , where neither is in the intersection of  $Y_t$  and  $Y_{t'}$ . In this case,  $\tilde{y}^* \notin Y_t$ , and we note that

$$f_{\tilde{t}}(\tilde{y}^*) < f_t(\tilde{y}^*) \leq f_t(\underline{y}^*),$$

where the first inequality is due to  $\tilde{y}^* \notin Y_t$ , and the second inequality reflects that  $f_t$  is weakly decreasing. Since the attained average cost is the same, we need that the high technology has higher costs of management:  $\tilde{g}_{\tilde{t}}(\tilde{y}^*) > \tilde{g}_{\underline{t}}(\underline{y}^*)$ . From the proof of the continuity lemma, we also know that if we increase the managerial wage by a small amount  $\delta$ , the increase in cost is bound above by the initial managerial level, and bounded below by the final managerial level. However, for sufficiently small increases in  $\delta$ , if  $\tilde{t}$  is still optimal, we know that the manager use for the high technology will still be lower than the initial managerial use for the low technology.<sup>19</sup> But then, the cost for  $\tilde{t}$  will be higher than for  $\underline{t}$ . Since this is true for all technologies with the same cost as  $\underline{t}$  at  $\hat{w}_m$ , this means that  $\tilde{t}$  beats all those technologies for  $\hat{w}_m$ , contradicting that  $\hat{w}_m$  is a supremum of wages where a technology less than  $t$  can be optimal.

Next, we consider three corner case. First, suppose that both technologies have their optimal output level at the intersection of  $Y_t$  and  $Y_{t'}$ , and consider  $\hat{w}_m + \delta$  for a small  $\delta$ . Since the optimal output level is weakly decreasing in  $w_m$ , the high technology will either have exactly output  $Y_t \cap Y_{t'}$  after the change, or a lower output. In the first case, the low technology can attain the same cost level, and in the second case,  $\tilde{t}$  cannot be the lowest average cost technology, since the cost-minimizing

<sup>18</sup>I assume here that we only use  $\underline{t} < \tilde{t}$  in the proof, so that the reasoning works out even though multiple  $t$ 's both below and above  $t$  might attain the minimum at  $\hat{w}_m$ .

<sup>19</sup>It is not completely obvious, since management use is not continuous in wages. However, if  $\tilde{t}$  is still optimal, the corresponding output level is in  $Y_{t'}$ , and so  $f_{\tilde{t}}$  at the optimal output level is bounded away from  $f_t(\tilde{y}^*)$ . Hence, since total cost is continuous, we need to have  $\tilde{g}_{\tilde{t}}(y_t^*(w_m + \delta))$  below  $\tilde{g}_{\underline{t}}(\underline{y}^*)$  for sufficiently small  $\delta$  to ensure that the total cost converge to  $c_t(y_t^*, w_m)$ .

output falls outside  $Y_{t'}$ .

Second, suppose that only the low technology has its optimal output level at the intersection. In that case, the high technology also attains one of its minima at that output level, and when we increase  $w_m$ , this output level will have the smallest increase in cost, and thus still be a global minima. But then the low technology also still attains a global minimum there, contradicting that supremum hypothesis.

Last, we note that we cannot have that only the high technology has an optimal output at the intersection. In that case, the low technology would also attain a global minimum there, and given the convention that firms choose the highest output that minimizes average costs, we would be in the case that both technologies have their optimal output at the intersection.

Having gone through the cases when none, one, or both technologies have their technology at the intercept, we can conclude that we never are able to remove  $\tilde{t}$  as a cost-minimizing technology when we increase  $w_m$ , contradicting that  $\hat{w}_m$  is the supremum of wages where there is an optimizing technology less than  $t$ .

**Result 3:**  $y^*$  is weakly decreasing in  $w_m$ . The result follows quite easily from result 1 and 2. Result 1 shows that the optimal output level is weakly decreasing in  $w_m$ , and result 2 shows that switches in technology goes towards the lower technology. To finalize the proof, we only need to establish that technology switches goes from a technology with a higher output level to a lower output level.

To do this, we note that whenever a technology shift happens between two technologies  $t < t'$ , the output level before the shift is in  $Y_{t'}$  and the output level after the shift is in  $Y_t$ . Since  $Y_t$  is weakly to the left of  $Y_{t'}$ , there is a weak decrease in output level after the shift.

## B.4 Formal Statement and Proof of Output Decomposition

To state the formal proposition, it is helpful to introduce notation that distinguishes between revenue shares, cost shares, and employment shares. For the modern sector, we write  $s_M \equiv \frac{P_M Y_M}{PY}$  for its revenue share,  $s_M^c \equiv \frac{w_p H_{M,p} + w_m H_{M,m}}{w_p H_p + w_m H_m}$  for its cost share, and  $e_M \equiv \frac{L_m}{L_p + L_m}$  for its employment share. Within the modern sector, we write  $e_{M,p} \equiv \frac{L_{M,p}}{L_M}$  and  $e_{M,m} \equiv \frac{L_{M,m}}{L_m}$  for the employment share of production workers and managers respectively (and the compensation shares being  $1 - \alpha$  and  $\alpha$  respectively due to the Cobb-Douglas production structure). Last, we write  $e_p \equiv \frac{L_p}{L}$  and  $e_m = \frac{L_m}{L}$  for the employment shares of production workers and managers.



We can now state the following proposition.

**Proposition 6.** *Given changes in  $w_m/w_p$ , the change in real output satisfies*

$$d \log \left( \frac{Y}{P} \right) = \Delta_Q + \Delta_{between} + \Delta_{within} \quad (\text{B.17})$$

where

$$\begin{aligned} \Delta_Q &= 0 \\ \Delta_{between} &= [(s_M - s_M^c) + (s_M^c - e_M)] \times (\theta + 1) \left[ -\alpha d \log \left( \frac{w_m}{w_p} \right) \right] \\ \Delta_{within} &= s_M \alpha (1 - \alpha) \left[ \frac{e_{M,p}}{1 - \alpha} - \frac{e_{M,m}}{\alpha} \right] \times \left[ -d \log \left( \frac{w_m}{w_p} \right) \right] \end{aligned}$$

*Proof.* Totally differentiating the labor market clearing condition implies

$$0 = e_{T,p} (d \log H_{T,p} - d \log Q_p) + e_{M,p} (d \log H_{T,m} - d \log Q_p) + e_{M,m} (d \log H_{M,m} - d \log Q_m)$$

which gives us

$$e_p d \log Q_p + e_m d \log Q_m = e_T d \log H_{T,p} + e_{M,p} d \log H_{T,m} + e_{M,m} d \log H_{M,m}.$$

In our derivation, we also use  $e_T = 1 - e_{M,p} - e_{M,m}$  and  $s_T = 1 - s_M$ . From these observations, we derive

$$\begin{aligned} d \log \left( \frac{Y}{P} \right) &= s_T d \log Y_T + s_M d \log Y_m \\ &= s_T d \log H_{T,p} + s_M [(1 - \alpha) d \log H_{M,p} + \alpha d \log H_{M,m}] \\ &= e_p d \log Q_p + e_m d \log Q_m + \\ &\quad (s_T - e_T) d \log H_{T,p} + [s_M(1 - \alpha) - e_{M,p}] d \log H_{M,p} + [s_M \alpha - e_{M,m}] d \log H_{M,m} \\ &= e_p d \log Q_p + e_m d \log Q_m + \\ &\quad [s_M(1 - \alpha) - e_{M,p}] d \log \left( \frac{H_{M,p}}{H_{T,p}} \right) + [s_M \alpha - e_{M,m}] d \log \left( \frac{H_{M,m}}{H_{T,p}} \right). \end{aligned}$$

To further simplify the expression, we note that  $d \log \left( \frac{H_{M,p}}{H_{T,p}} \right) = d \log \left( \frac{H_{M,p}}{Y_M} \right) + d \log \left( \frac{Y_M}{Y_T} \right)$ , since

$d \log H_{T,p} = d \log Y_T$ . Furthermore,

$$\begin{aligned} d \log \frac{H_{M,p}}{Y_m} &= (-\alpha) d \log \frac{H_{M,m}}{H_{M,p}}, \\ d \log \frac{H_{M,m}}{Y_M} &= (1 - \alpha) d \log \frac{H_{M,m}}{H_{M,p}}, \end{aligned}$$

From this, we derive

$$\begin{aligned} & [s_M(1 - \alpha) - e_{M,p}] d \log \left( \frac{H_{M,p}}{H_{T,p}} \right) + [s_M \alpha - e_{M,m}] d \log \left( \frac{H_{M,m}}{H_{T,p}} \right) \\ &= [s_M(1 - \alpha) - e_{M,p}] d \log \left( \frac{H_{M,p}}{Y_M} \right) + [s_M \alpha - e_{M,m}] d \log \left( \frac{H_{M,m}}{Y_m} \right) + (s_M - e_M) d \log \frac{Y_M}{Y_T} \\ &= \alpha(1 - \alpha) s_M \left[ \frac{e_{M,p}}{(1 - \alpha) s_M} - \frac{e_{M,m}}{\alpha s_M} \right] d \log \frac{H_{M,m}}{H_{M,p}} + [(s_M - s_M^c) + (s_M^c - e_M)] d \log \frac{Y_M}{Y_T}. \end{aligned}$$

Furthermore, (??) and the Cobb-Douglas production function in the modern sector implies

$$d \log \frac{Y_M}{Y_T} = -(\theta + 1) \alpha d \log \frac{w_m}{w_p}, \quad d \log \frac{H_{M,m}}{H_{M,p}} = -d \log \left( \frac{w_m}{w_p} \right),$$

and we obtain the final result. □

For  $\Delta_{between}$ , the term  $(s_M - s_M^c) + (s_M^c - e_M)$  shows that redistributing labor from the traditional to the modern sector improves output through two channels: by redistributing output to a high markup sector (captured by  $s_M - s_M^c$ ), and by redistributing workers to a high wage sector (captured by  $s_M^c - e_M$ ). The modern sector has a higher markup and a cost share that is higher than its employment share, so reallocation towards the modern sector raises real output.

On the other hand,  $\Delta_{within}$  captures the gain from increasing the management intensity in the high wage sector. Since both types of workers are subject to the same markup, the only gain comes from moving workers towards an occupation with higher wages. This effect is captured by managers having a lower employment share  $e_{M,m}$  than compensation share  $\alpha$ , so that  $\frac{e_{M,p}}{1 - \alpha} - \frac{e_{M,m}}{\alpha}$  is positive.