Firms Unchained: Human Capital, the Emergence of Modern Services and 21st Century Development

[PRELIMINARY DRAFT]

Alvaro Cox [*] Matthew Schwartzman [†] Hannes Tepper [‡]

March 27, 2026

Abstract

Can the service sector drive productivity growth in developing economies? This paper argues that it can, but only when supported by a sufficiently skilled workforce. In contrast to manufacturing, productive non-tradable services scale by replicating complex operations across multiple locations, a process that requires high-skilled labor to provide managerial oversight. We argue that human capital shortages in developing countries constrain service-sector productivity growth by limiting the emergence of modern, multi-market firms. To test this hypothesis, we exploit a major higher education reform in Brazil that exogenously expanded access to college in the late 1990s. We provide causal evidence that increased local college attainment led to a disproportionate rise in employment at skill-intensive, multi-market service firms, accompanied by wage gains and broader labor reallocation toward the service sector. We interpret these findings through a theoretical framework in which skilled labor is a necessary input for firm expansion across space. Our results highlight the central role of human capital in enabling service-led growth and position multi-market firms as key engines of productivity growth in the 21st century.

[*]Universidad Carlos III Madrid: acox@eco.uc3m.es

[†]University of Michigan: schwam@umich.edu

[‡]Paris School of Economics: hannes.tepper@psemail.eu

We wish to thank the participants in the Applied Economics and Trade seminars at the University of Oslo for useful comments. We also thank Alessandro Tomarchio for outstanding research assistance. Hannes Tepper thanks the grant ANR-17-EURE-0001 for its support.

1 Introduction

From Britain in the late 18th century to China in the late 20th, economic development has historically been synonymous with industrialization: the reallocation of resources from agriculture to manufacturing. In the 21st century, the link between development and industrialization has frayed. Advances in automation and global competition have made manufacturing increasingly skill- and capital-intensive, limiting its capacity to absorb labor and deliver inclusive growth at scale (Rodrik, 2015; Acemoglu and Restrepo, 2022). This shift has prompted growing interest in the potential of the service sector to serve as an alternative engine of development (Chen et al., 2023; Fan et al., 2023; Hsieh and Rossi-Hansberg, 2023; Rodrik and Sandhu, 2024; Rodrik and Stiglitz, 2024). Yet while the mechanisms underpinning industrial growth, such as export orientation and scale economies in tradable goods, are well understood and grounded on a rich set of historical examples, it is less clear how services can generate sustained productivity growth.

Productivity growth in services arises from two complementary processes: innovation – the development of new ideas and business models – and implementation – the replication of these business models across multiple local markets. A key feature distinguishing most services from manufacturing is their limited tradability. Whereas manufacturing firms achieve scale by concentrating production in centralized facilities and relying on logistics networks to distribute output, firms in non-tradable services must expand by reproducing the same service technology in many locations simultaneously. In this context, managerial coordination plays a role analogous to logistics in manufacturing: it enables the efficient implementation of innovative business models across space. When service firms can scale across markets, the expected returns to innovation rise with the number of markets served. As a consequence, these two dimensions – innovation and implementation – reinforce each other. The firm's capacity to implement its business model across space determines both the extent of scale economies and the expected returns to innovation, linking spatial expansion to productivity growth in services.

The effective replication of non-tradable services across markets relies on organizational capabilities, and, in particular, on the availability of managerial talent. Implementing standardized processes across varied environments requires managers who can interpret and execute central directives, adapt them to local conditions, and maintain quality control. In developing economies, such managerial inputs are often scarce, limiting the expansion of productive service firms even when profitable business models exist.

This paper argues that low productivity in non-tradable services stems from a shortage of managerial human capital. We propose a mechanism in which higher education expands

the supply of workers capable of performing managerial roles. A limited managerial supply restricts firms' ability to scale across space, reducing the payoff to innovation and dampening aggregate productivity growth. Expanding access to college education thus increases the pool of potential managers, facilitates the spatial expansion of successful service firms, and reallocates employment toward more productive establishments. This process generates a complementarity between human capital accumulation and productivity growth in non-tradable services.

We empirically test this hypothesis in the context of Brazil, where a large-scale higher education reform in the 1990s substantially expanded access to college (Cox, 2025). The reform generated regional variation in the rise of college-educated workers, creating a natural setting to study how human capital affects the spatial scaling of service firms. Regions experiencing larger increases in college attainment saw the entry and expansion of productive multi-market service firms, which opened new establishments, reallocated labor toward these firms, and contributed to local economic growth. To quantify the aggregate implications of this mechanism, we develop a model of multi-location service firms in which the local supply of skilled labor constrains the ability of distantly headquartered firms to operate local establishments.

We begin by establishing a set of empirical facts that guide our analysis, before turning to causal identification and a theoretical model. These facts document Brazil's structural transformation and highlight the joint evolution of college attainment and multi-market service firms. Brazil's structural shift mirrors the experience of many developing economies outside East Asia. While agricultural employment has declined and self-employment has given way to wage employment, the share of employment in manufacturing has remained stagnant at around 15%. In contrast, the service sector has expanded steadily over the past three decades and now accounts for nearly 80% of total employment. These shifts coincided with two striking trends: a sharp increase in the yearly flow of college graduates beginning in the late 1990s, and a simultaneous rise in employment within multi-market firms. Both developments were concentrated in services, which experienced the largest gains in both skill intensity and the prevalence of multi-establishment firms across the economy.

We next examine the composition of employment within these firms to show that these trends are closely connected. Multi-market firms employ skilled workers at significantly higher rates than single-market firms, even comparing firms of the same employment size within narrowly defined industries and locations. This pattern indicates that skilled labor is a key input into multi-market production. Consequently, increases in the supply of college-educated workers are likely to disproportionately benefit multi-market firms.

We then explore the association between the rise of multi-market firms and income growth.

Using administrative matched employer-employee data, we document four novel facts about multi-market firms. First, multi-market firms pay substantially higher wages than their single-market counterparts, even for a given worker, indicating a sizable wage premium associated with spatial scaling. Second, we study the wage implications of expanding firm scope into new markets for *incumbent* workers, controlling for tenure and a rich set of fixed effects. We find that, on average, a firm that expands from one to two markets increases incumbent workers' wages by about 6%. Third, we estimate an event study around the first establishment opening experienced during a worker–firm job spell. Wage dynamics are flat in the five years prior to expansion and become positive and increasing thereafter, consistent with wage gains that materialize following the expansion event. Finally, we show that over long time horizons, firm employment increases more than proportionally with the number of markets served by the firm. Firms that increase their spatial scope of operations also increase their employment in every market where they operate.

Taken together, these facts suggest that expanding the spatial scope of the firm is accompanied by productivity upgrading, generating a feedback loop between innovation and implementation. Consistent with this mechanism, regions that experience larger increases in the employment share of multi-market firms also exhibit faster income growth, driven both by reallocation of workers toward these firms and by wage growth within them. Overall, the evidence indicates that the expansion of multi-market firms has been an important force for income growth in Brazil's service sector.

While these descriptive patterns align with our proposed mechanism, they do not establish causality. To identify the effect of skilled labor supply on the spatial expansion of service firms, we exploit a 1996 higher education reform in Brazil that lowered entry barriers for private colleges. Before the reform, the Federal Council of Education (FCE) exercised control over the creation of new private institutions, limiting capacity growth despite rising demand. These constraints, fueled by lobbying from public universities and professional associations, created a mismatch between the potential demand for higher education and the supply of colleges. The reform dismantled these barriers by deregulating entry and introducing a regulatory framework to ensure instructional quality. The resulting policy change generated a rapid and spatially heterogeneous expansion in college access. Following Cox (2025), who documents the reform and its implications for educational expansion, we construct an instrumental variable based on the pre-reform ratio of higher education capacity to potential demand—proxied by the local population aged 6 to 15 in 1991, who would reach college age after the reform. Intuitively, regions with tighter pre-reform capacity constraints were more likely to attract new private institutions once entry was deregulated.

Our baseline estimates show that a 1pp increase in the share of college-educated individ-

uals leads to a 3.5pp increase in the employment share of multi-market firms. This effect is driven primarily by the entry of new establishments belonging to multi-location firms in regions that experienced larger gains in college attainment. Consistent with our mechanism, these new establishments – particularly those intensive in skilled labor – emerged predominantly in the service sector. In summary, the natural experiment generated by Brazil's higher education reform shows that expanding the supply of skilled workers increases the prevalence of new multi-market firms in the service sector.

To distill our empirical results, we propose a novel theoretical framework that builds on that of Hsieh and Rossi-Hansberg (2023). We allow for skill heterogeneity among workers, and introduce a modification that underscores the importance of different types of labor to firm expansion: operating a new establishment requires on-site high-skilled workers to serve as managers. These managers, who oversee the implementation of the firm-level technology in new markets, are part of the fixed cost of operating establishments at distance.

Guided by our empirical facts, we assume that this fixed cost is greater at longer distances, so that multi-market firms whose operations cover large areas are more skill-intensive. Firms therefore select into operating establishments in distant markets. In the style of Hopenhayn (1992) and Melitz (2003), they operate only in markets where their productivity exceeds a cutoff governed by the fixed cost of local management. Firm productivity, however, is not given exogenously. Instead, firms make investments in a replicable "headquarters technology" that can be implemented in every establishment of the firm. As in Kleinman (2024), firms make greater headquarters investments when they expect to operate across a broad scope of markets. Finally, we incorporate a key feature of the Brazilian service landscape: formal firms compete not only against one another, but against a local "traditional" service sector that is less skill-intensive and that thrives on the demand of low-income consumers (Schwartzman, 2025).

In the model, regions with severe skill shortages are too prohibitively expensive for almost any multi-market firm to establish operations. Relieving the local skill shortage – for example, by increasing access to college – has transformative effects. Most directly, by reducing the relative price of managerial labor, expansion of the skill supply lowers the fixed cost to operate establishments, prompting more distantly-headquartered firms to select into operation in a given region. Owing to these fixed costs, the marginal distant firms who select into operation are more productive than local single-market firms, so that their entry represents reallocation of resources to more efficient service producers. These direct effects are augmented by two indirect effects that magnify the productivity-enhancing impact of skill expansion. First, increasing the skill of the labor force raises local incomes and reduces the price of modern services relative to less skilled traditional services. These changes grow the

market for modern services by shifting consumer demand. Second, the rising modern service share and falling establishment fixed costs create an opportunity to *implement* productive technologies at scale. Firms respond to this opportunity by increasing their *innovation*: they make greater investments in the replicable headquarters technology. Skill expansion therefore generates productivity upgrades in services that echo throughout the economy.

Related literature — We contribute to four main strands of the literature.

First, we contribute to the debate on whether the service sector can serve as an engine of economic development. Since the seminal work of Baumol (1967), the dominant view has been that services offer limited potential for productivity growth. Consequently, economic development was long believed to depend on a vibrant manufacturing sector. However, in the 21st century, manufacturing has increasingly lost its traditional role as a labor-absorbing engine of broad-based prosperity (Acemoglu and Restrepo, 2022; Rodrik and Sandhu, 2024; Rodrik and Stiglitz, 2024).

In response, a growing body of research has begun to reassess the growth potential of the service sector. Fan et al. (2023) document sustained productivity growth in India's service sector, while Hsieh and Rossi-Hansberg (2023) identify an "industrial revolution in services" in the United States. In China, Chen et al. (2023) find that service sector productivity growth may have even outpaced that of manufacturing. Yet, despite these developments, there is limited guidance on how to catalyze such a transformation in other contexts. Our paper contributes to this literature by documenting sustained productivity growth in Brazil's service sector, driven by the expansion of multi-market firms. We further show that this transformation was contingent on a substantial increase in the supply of college-educated workers. Together, our findings suggest that, when supported by an adequate supply of skilled labor, multi-market firms can act as engines of service-led growth.

Second, our paper also contributes to the literature on the modernization of the service sector and the mechanisms that underpin it. Buera and Kaboski (2012) develop a theory in which rising income levels drive a transition from home production to more specialized, skill-intensive service provision. A complementary line of work emphasizes the role of information and communication technologies (ICT) as a key driver of productivity growth in services (Hjort and Poulsen, 2019; Eckert et al., 2022; Hsieh and Rossi-Hansberg, 2023; Jiang, 2023; Hjort and Tian, 2025). More recently, Schwartzman (2025) proposes a theory in which the shift toward modern services is itself a source of aggregate growth. We contribute to this literature by providing empirical evidence that human capital plays a central role in enabling the modern provision of services. To the best of our knowledge, we are the first to establish this causal relationship empirically. Moreover, we extend the framework of Hsieh and Rossi-Hansberg (2023) by introducing worker heterogeneity to capture the pivotal role

of college-educated workers in managing multi-market firms. In doing so, we highlight the importance of managerial human capital as an input in the production of modern services.

Third, our work contributes to a broad literature on human capital and firm growth. Foundational contributions by Chandler (1977) and Lucas (1978) underscore the importance of organizational structures that support managerial delegation as a key enabler of firm expansion. More recent work has documented substantial international variation in management practices and their close association with human capital accumulation (Bloom and Van Reenen, 2007; 2010; Bloom, Lemos, et al., 2014). Cox (2025) provides direct evidence that increases in human capital can drive firm expansion and catalyze structural transformation, and Engbom et al. (2024) develop a theory where firms choose to adopt managerial capitalism based on the supply of skills in the economy. We build on this literature by highlighting the role of multi-market firms in the service sector as an important channel through which human capital affects aggregate outcomes. We document the productivity advantage of multi-market firms and show that rising college attainment enables firms to expand across markets and absorb larger shares of employment, thereby facilitating broad-based structural transformation.

Fourth and finally, we contribute to the literature on multi-establishment firms. While some studies model these firms as knowledge-based hierarchies (Garicano and Rossi-Hansberg, 2015; Gumpert et al., 2022), we follow Hsieh and Rossi-Hansberg (2023) in emphasizing the role of market entry and geographic expansion. Mariscal (2018), for instance, examines how access to information and communication technologies shapes the organizational structure of multi-market firms. In contrast, our work highlights the importance of human capital – specifically, the availability of skilled workers who can serve as managers – as a key input in enabling the emergence and expansion of multi-market firms. In this vein, Kleinman (2024) studies how increases in the returns to *headquarters* labor have driven rising wage inequality in the United States. We complement research on the importance of headquarters labor by showing that skilled managers at the *establishment* level are another key input for multi-location production. Our paper also represents some of the first research on multi-market firms in a developing country setting.

The paper is structured as follows. Section 2 describes the datasets and presents empirical facts that document the rise of multi-market service firms in Brazil, their reliance on skilled labor, and their productivity advantage. Section 3 introduces the education reform, outlines the empirical strategy, and presents the main empirical results. Section 4 develops a simple model of firm size and market entry with heterogeneous workers. Section 5 concludes and highlights key challenges for future research.

2 Empirical Analysis

2.1 Data

Our primary data set is the universe of formal employment contracts from the Relação Anual de Informações Sociais (RAIS) for the period 1995–2018. This administrative data set tracks the universe of formal contracts at the establishment level, providing information on worker demographics (education level), contract location (municipality), and industry classification. This granular, contract-level detail allows us to track employment composition, sectoral changes, and earnings at the establishment level.

The RAIS data also contain unique firm and establishment identifiers. The two identifiers allow us to track the number and spatial distribution of establishments within a firm. In Brazil, every firm must be registered in the national firm registry (CNPJ), which assigns a 14-digit ID to each firm. The first 8 digits of a firm's CNPJ identify the firm, while the last 4 digits identify the establishment. This structure enables us to identify a firm's headquarters.[1]

For results that rely on worker-level regressions, we take a random 3% sample of all workers who appear in RAIS over our period of study. For the firm-level regressions, we take a 10% sample of firms. In both cases, we do this to reduce the computational burden from large sample sizes.

Brazil has a significant share of its labor force working without formal contracts. To account for this, we supplement the RAIS data with two data sources that include the informal sector. First, we use the National Household Sample Survey (PNAD), which collects information on the labor conditions of individuals and enables us to identify the aggregate share of informal workers throughout our period of analysis. Second, we use the decennial Census for the years 1991, 2000, and 2010, which provides individual-level data on employment status, including informal contracts, industry of employment, income, and location of residence. Because it includes location, the Census allows us to account for informality at the microregion level.

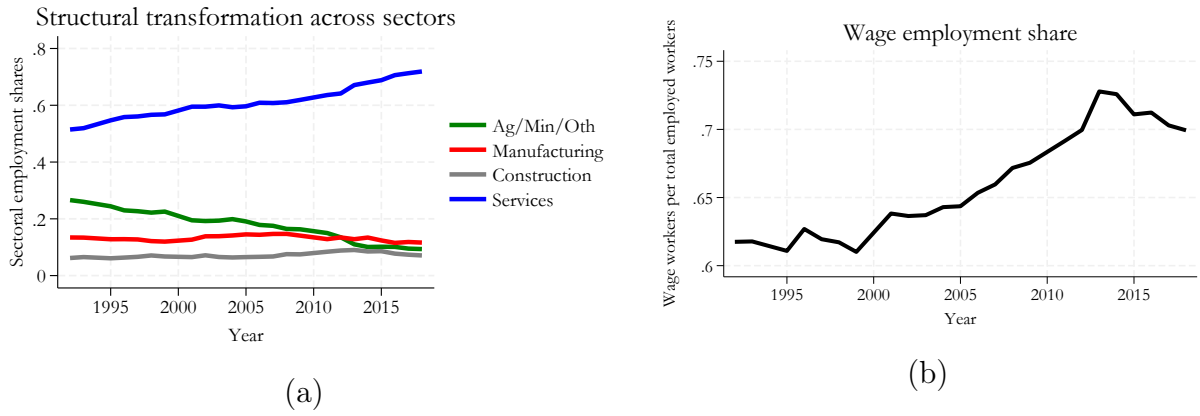
2.2 Facts For The Brazilian Economy

2.2.1 Structural Transformation Towards Services

Our first fact documents the process of structural transformation, which we define as the reallocation of labor from agriculture and self-employment to wage employment in modern firms. Figure (1) presents these aggregate trends for the period 1991–2018. Panel (a) shows

[1]The legal appendix n°82/1999 requires each firm to have one "matriz," or headquarter, which is designated by the number "0001."

Figure 1: Structural Transformation in Brazil



Note.— Panel (a) shows the share of workers whose main job was in one of four broad sectors, including Manufacturing, Construction, Services, and agriculture, mining, or other activities. Panel (b) displays the share of workers who are wage employed.

a substantial decline in agricultural employment, with the resulting labor largely absorbed by the service sector. In contrast, the employment share in manufacturing and construction remains largely unchanged over this period. Panel (b) depicts a parallel increase in the share of wage employment at the expense of self-employment.

2.2.2 College expansion and the rise of the multi-location firm

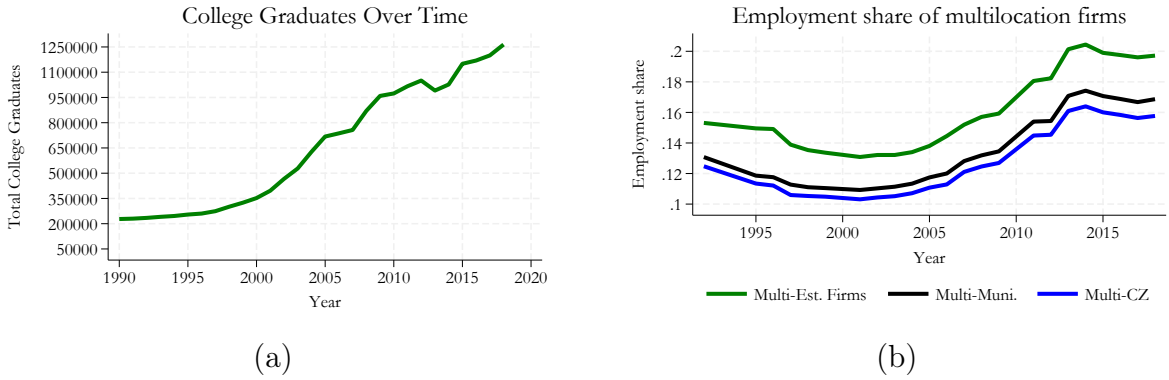
During the study period, the structural transformation of the economy was accompanied by two key trends: a substantial increase in human capital accumulation and a rise in the spatial scope of firms.

The first trend is the significant expansion of higher education. Facilitated by the 1996 education reform, college access expanded substantially, leading to a large increase in the number of college graduates (Cox, 2025). As shown in Panel (a) of Figure (2), the number of college graduates rose close to fivefold between 1995 and 2010, with a marked increase starting in 2002, five years after the reform's implementation. This surge in human capital suggests a significant shift in the economy's skill composition.

The second transformation involves the rise of multi-market firms, which we define as firms with a presence in more than one distinct geographic market. We measure this expansion using three nested definitions of "market": establishments within a municipality, municipalities, and Commuting Zones (CZs).[2] Panel (b) of Figure (2) shows a substantial increase in the employment share of multi-market firms across all three definitions. This

[2]We use the definition of microregion in Dix-Carneiro and Kovak (2017), which groups municipalities into 486 microregions analogous to U.S. commuting zones, with each representing a single labor market.

Figure 2: Structural Transformation in Brazil



(a) (b)

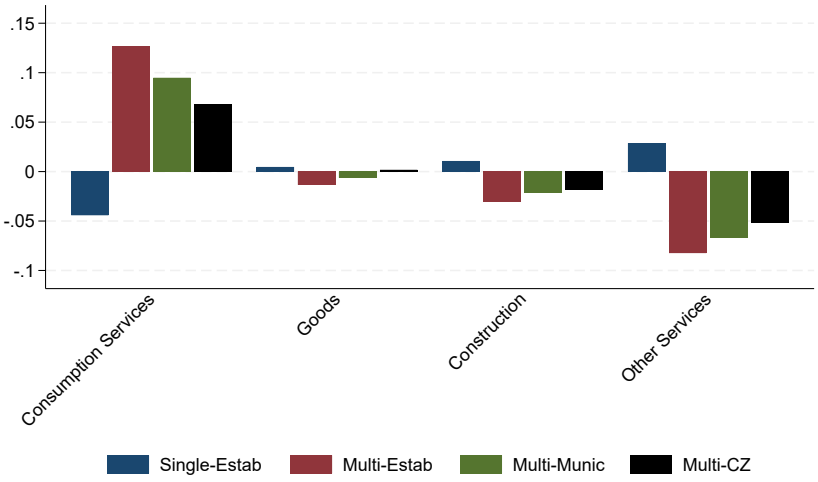
Note.— Panel (a) shows the evolution of the total number of college graduates for the period 1990-2018. Panel (b) displays the employment share in multi-market firms.

trend began shortly after the surge in college graduates and stabilized after 2014 at a significantly higher level. The co-movement of these two trends – human capital accumulation and the spatial expansion of firms – suggests a potential link between the availability of a skilled workforce and the organizational structure of firms.

2.2.3 The rise of multi-market firms was more pronounced in Services

The structural transformation of the Brazilian economy has been characterized by a substantial reallocation of labor toward the service sector. This shift is not uniform; instead, it is closely linked to the rise of multi-market firms. As we now document, multi-market firms are most prominent within the service sector, particularly among non-tradable services.

Figure 3: Sectoral Over/Under-Representation of Firms by Type (2018)

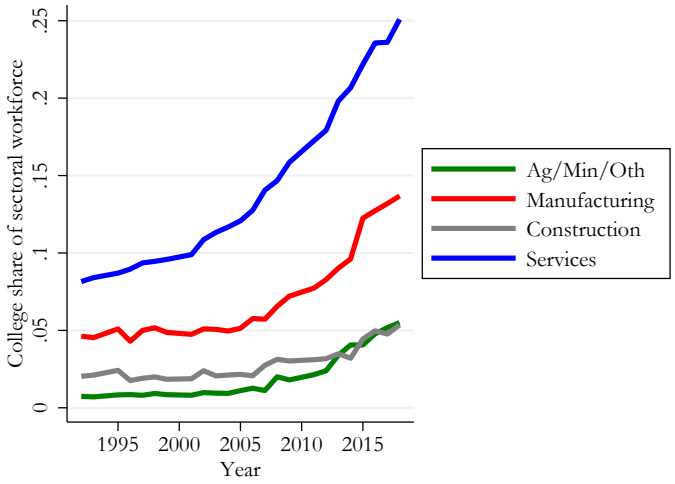


Note.— Figure displays the sectoral over-/under-representation of firms by organizational structure for the year 2018. Therefore, for a given firm type, it computes the difference between a sector's share within that firm type and the sector's share among all firms (all firm types combined). Consequently, bars of the same color sum to zero across all sectors. Firms are categorized as follows: single-establishment firms (blue); multi-establishment firms (red); multi-establishment firms operating in at least two municipalities (green); and multi-establishment firms operating in at least two commuting zones (black). 'Consumption Services' corresponds to retail, hospitality, and restaurants. 'Goods' denotes agriculture, manufacturing, mining, and other industries not elsewhere classified. 'Other Services' refers to the telecommunications and communication industries, education, health, and public service sectors.

To provide sectoral evidence, Figure 3 shows that, in 2018, multi-market firms were over-represented in the consumption services sector, which in our classification includes retail, hospitality, and restaurants—a pattern that holds across all three of our measures (establishments, municipalities, and commuting zones). More importantly, it also reveals significant heterogeneity within the service sector itself, as multi-market firms are overrepresented only in consumption services, in contrast to all other service categories. This finding suggests that a significant share of employment growth in services is driven by the geographic expansion of firms.

The rise of multi-market firms within the service sector was accompanied by a significant deepening of human capital, a trend that was especially pronounced within the service sector itself. As Figure 4 shows, the college-educated share of the sectoral workforce increased most rapidly in services over the 1990–2018 period. This pattern is consistent with the literature on skills and services, which posits a link between human capital accumulation and the growth of the service sector (Buera and Kaboski, 2012).

Figure 4: Evolution of the skill intensity within sectors



Note.— Figure displays the evolution of the skill intensity within four broad sectors of the Brazilian economy for the period 1991 to 2018. Sectoral skill intensity is measured as the share of the sector's workers who completed college.

2.2.4 Multi-Market Firms are More Skill-Intensive

We now present two additional stylized facts concerning the characteristics of multi-market firms: they are more skill-intensive and their skill intensity increases with the geographic distance between a firm's headquarters and its establishments. We define skill intensity as the share of college-educated workers at either the establishment or the firm level.

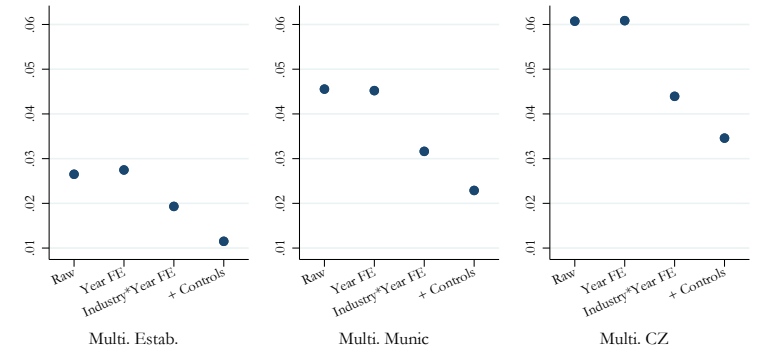
To document the first fact, we estimate the following regression:

$$\text{Skill Intensity}_{jit} = \alpha D_{jit}^{MM} + \theta_t + \theta_i + X_{jit}\Gamma + \varepsilon_{jit}, \tag{1}$$

where $\text{Skill Intensity}_{jit}$ denotes the share of college-educated workers in firm j, operating in industry i at time t. D_{jit}^{MM} is an indicator variable for multi-market firm status. The specification includes year fixed effects (θ_t), industry fixed effects (θ_i, and a vector of firm-level controls (X_{jit}), including firm size. We focus on the most saturated specification, which compares firms of similar size within the same industry and year.

Figure 5 plots the estimated coefficient α for our three distinct measures of multi-market status. The results consistently show that multi-market firms are significantly more skill-intensive. For instance, in our preferred specification (the last column of each plot), multi-market firms—defined by a presence in multiple commuting zones—have a skill intensity that is 3.5 percentage points higher than their single-market counterparts. This robust correlation suggests that a skilled workforce is a crucial input for firms that expand their geographic footprint.

Figure 5: Skill Intensity and multi-market firms



Note.— The figure plots estimates of α across alternative specifications of estimation equation (2). The unit of observation is an establishment-year pair. Data are sourced from RAIS.

This finding leads to a more granular analysis. We explore whether a firm's skill intensity is correlated with the geographic distance between its headquarters and its establishments. We focus on multi-establishment firms and estimate the following regression:

$$\log dist_{imjt}^{HQ,estab} = \alpha \, \text{Skill Intensity}_{imjt} + \theta_t + \theta_j + X_{jit}\Gamma + \varepsilon_{jit}, \tag{2}$$

where $\log dist_{imjt}^{HQ,estab}$ is the average distance between a firm's headquarters and its establishments.

Table 1 presents the estimates for α. As before, we include year and industry fixed effects, and controls including firm size. The results indicate that firms with higher skill intensity have a larger average distance between their headquarters and establishments. This relationship holds whether we measure skill intensity at the headquarters or the establishments. This suggests that skilled labor, both in the central headquarters and in local markets, is a critical input in the production technology of multi-market firms. Skill enables firms to manage and coordinate operations across a broader spatial scale. In this sense, we see skilled labor as a critical input in the production technology of multi-market firms.

Table 1: College decomposition: Relative firm size effect (MMCs)

Dep. Variable	ΔDistance HQ-Establishments		
Specification	(1)	(2)	(3)
Skill intensity, firm	1.27***	0.75***	
	(0.01)	(0.01)	
Skill intensity, HQ			0.58***
			(0.01)
Skill intensity, estab.			0.18***
			(0.01)
Year FE	Yes	Yes	Yes
Sector FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Obs.	848053	848053	778648
R^2	0.22	0.26	0.27

Standard errors in parentheses

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

Note.— The table presents regression results for the log distance between an establishment and its headquarters as a function of skill intensity measured at the firm level, headquarters level, and establishment level. The unit of observation is the establishment. Data are sourced from RAIS.

2.2.5 Multi-Location Firms Are More Productive

So far, we have documented that multi-market firms are concentrated in consumer services and are more skill-intensive. In the next set of facts, we show that these firms are also more productive than single-market firms. While we lack data on revenues or profits, we exploit our matched employer-employee data to compare indirect indicators of firm productivity between multi-market and single-market firms. In particular, we document that multi-market firms pay higher wages – even conditional on worker fixed effects – and that the more markets a firm operates in, the greater its employment is *within* each market. In any model where firms face an upward-sloping labor supply curve within each labor market, a firm's within-market wages and employment both increase in productivity, which acts as a labor demand shifter for the firm.

We begin with wages. We exploit the panel dimension of the match employer-employee database to isolate a productivity differential from worker sorting. Specifically, we test whether the *same worker* earns higher wages when employed by a multi-market firm. By absorbing time-invariant worker heterogeneity with worker fixed effects, we purge the estimate from permanent differences in worker ability, education, and other fixed attributes

that may correlate with sorting into multi-market employers. We further control flexibly for tenure and for time-varying shocks at the industry–region, industry–time, and region–time levels. Our baseline regression is:

$$\log w_{ijkrt} = \alpha\, \mathbf{1}\{n_j(t) > 1\} + \gamma\big(t - t_0(ij)\big) + \phi_i + \omega_{kr} + \omega_{kt} + \omega_{rt} + \varepsilon_{ijkrt}, \tag{3}$$

where $\log w_{ijkrt}$ is the logarithm of wages of worker i employed by firm j in sector k, region r, and year (or period) t. The indicator $\mathbf{1}\{n_j(t) > 1\}$ equals one if the worker's employer operates in more than one market at time t, where "markets" can be defined using alternative geographic partitions (e.g., establishments, municipalities, or commuting zones), and $n_j(t)$ denotes the resulting number of markets served by firm j. The term $t_0(ij)$ is the period in which worker i begins working for employer j, so $\big(t - t_0(ij)\big)$ captures worker–firm match tenure and γ is interpreted as a tenure (experience-with-employer) profile common across workers. The worker fixed effect ϕ_i absorbs all time-invariant worker characteristics. We include a full set of industry–region fixed effects ω_{kr} to absorb persistent differences in wage levels across sector–location cells, as well as industry–time and region–time fixed effects, ω_{kt} and ω_{rt}, to control for shocks that are common to all regions within an industry and common to all industries within a region, respectively. The residual ε_{ijkrt} captures idiosyncratic wage variation.

The coefficient α is the within-worker multi-market wage premium. A positive estimate, $\alpha > 0$, implies that the same worker is paid more when employed by a multi-market firm than when employed by a single-market firm, conditional on tenure and sectoral and regional time-varying controls. We interpret $\alpha > 0$ as evidence that multi-market firms are more productive (or generate higher surplus per worker), consistent with a multi-market production technology that raises marginal products through managerial scale, better organization, or knowledge diffusion across establishments. Because identification leverages within-worker comparisons, the estimate cannot driven by static sorting of higher-ability workers into multi-market firms.

Results in Table 2 show that multi-market firms confer a substantial wage premium when compared to single-market firms. These results hold stable across our three market definitions: establishment, municipality, and microregion. Specifically, when we control for worker fixed-effects, employment in a multi-market firm yields a 10% wage premium. Without worker-fixed effects, the multi-market firm has a wage differential of 19%. This difference tells us that multi-market firms do tend to employ higher-ability workers, but this sorting accounts for less than half of the total multi-market wage premium.

Table 2: Multi-market wage premium

Market definition:	Establishment		Municipality		Microregion	
	(1)	(2)	(3)	(4)	(5)	(6)
	Log(Wage)	Log(Wage)	Log(Wage)	Log(Wage)	Log(Wage)	Log(Wage)
Multi-market	0.185***	0.100***	0.187***	0.101***	0.193***	0.103***
	(0.000596)	(0.000375)	(0.000636)	(0.000389)	(0.000661)	(0.000402)
Tenure controls	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Worker FE	No	Yes	No	Yes	No	Yes
Obs.	30,643,269	30,229,942	30,643,269	30,229,942	30,643,269	30,229,942
Adj.R^2	0.39	0.76	0.39	0.76	0.39	0.76
Avg Wage	758.07	761.78	758.07	761.78	758.07	761.78
N. Workers	3,312,316	2,899,062	3,312,316	2,899,062	3,312,316	2,899,062

Standard errors in parentheses, clustered at the worker level.

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

Note.— The table presents regressions of log wages on an indicator for the multi-market status of the firm and a set of fixed effects. The unit of observation is a worker-year. In columns 1 and 2, multi-market firms are defined as multi-establishment firms; in columns 3 and 4, they are multi-municipality firms; in columns 5 and 6, they are multi-microregion firms. Even-numbered columns control for worker fixed effects, while odd-numbered columns do not.

Wage gains increase with the spatial scope of the firm Our theory of service productivity growth emphasizes the complementarity between innovation and implementation: the scope of the firm determines its incentives to improve its own production technology, which we assume to be non-rival across markets within the firm. This complementarity between firm scope and productivity-enhancing innovation implies that wage premia should not only reflect cross-sectional differences between multi- and single-market employers, but should also *grow endogenously* when a given firm expands its geographic scope. We therefore test whether expansions in firm scope – measured by the opening of establishments in additional markets – are associated with wage gains for *incumbent* workers who were already employed at the firm. Because these workers are in "old" markets of the firm, this exercise allows us to test whether expansion is associated with firm-wide productivity improvements, rather than productivity growth localized to the firm's new establishments.

We exploit the panel structure of RAIS to construct long differences in wages over each worker–firm job spell. For a worker i employed at firm j in sector k and region r, we compare the log wage in a later period t to the log wage at the beginning of the job spell, $t_0(ij)$. The key regressors are (i) a change in the firm's multi-market status over the spell, and (ii) the change in the log number of markets served by the firm, both measured at the firm level between t_0 and t. Intuitively, this design asks: among workers who stay with their employer,

do those employed at firms that expand geographically experience faster wage growth than those employed at firms whose scope is stable? Our baseline long-difference specification is:

$$\log w_{ijkrt} - \log w_{ijkrt_0} = \alpha\, \Delta\text{Multi-market status}_j^{t_0 \to t} + \beta\, \Delta \log(\#\text{ markets})_j^{t_0 \to t} + \Big[X_{ijkr}(t_0)'\gamma_0\Big] \cdot (t - t_0(ij)) + \omega_{kr} + \omega_{kt} + \omega_{rt} + \varepsilon_{ijkrt}. \quad (4)$$

The dependent variable is wage growth from t_0 to t. The term $\Delta\text{Multi-market status}_j^{t_0 \to t}$ captures whether the firm transitions into multi-market operation between the start and current period, while $\Delta \log(\#\text{ markets})_j^{t_0 \to t}$ measures the intensity of scope expansion on the extensive margin (entering additional markets). The vector $X_{ijkr}(t_0)$ contains worker and firm characteristics measured at job-spell start (e.g., initial wage, initial multi-market status, and initial log number of markets), interacted with tenure $(t - t_0(ij))$ to allow for heterogeneous wage growth profiles across different initial conditions. We include industry–region fixed effects ω_{kr} and flexible industry–time and region–time fixed effects, ω_{kt} and ω_{rt}, to absorb common shocks operating at the sectoral and local levels.

The coefficient β is the main object of interest: $\beta > 0$ implies that an increase in firm scope is associated with greater wage growth for incumbent workers. This pattern is consistent with the view that expanding across space is accompanied by improvements in firm productivity that raise the wage schedule for continuing employees. The coefficient α captures discrete wage-growth effects from crossing the multi-market threshold per se (e.g., moving from one to multiple markets), distinct from the continuous expansion margin captured by $\Delta \log(\#\text{ markets})$.

Results in Table 3, which defines a market as a microregion, suggest that the increment from 1 to 2 markets confers an increase of 5.5% wage growth.[3] More generally, the elasticity of worker wage growth with respect to growth of the firm's number of markets is 0.11-0.12; a 1% increase in the number of markets is associated with a 0.11-0.12% increase in wage growth. This result is driven both by managers and by production workers, as observed in Table 10 in the Appendix. Greater wage growth among incumbents is suggestive of stronger incentives (and realized outcomes) to raise productivity when firms expand their operational footprint.

[3]The total effect of going from 1 to 2 markets is $\alpha + \beta \log(2)$. In column (3), $\hat{\alpha} = -0.0225$ and $\hat{\beta} = 0.112$, yielding a total effect of 0.055.

Table 3: Within-job-spell: Testing for wage gains led by firm expansion

	(1)	(2)	(3)
	Δ Log(Wage)	Δ Log(Wage)	Δ Log(Wage)
Δ Multi-market	-0.0373***	-0.0443***	-0.0225***
	(0.00195)	(0.00186)	(0.00187)
Δ Log(N of Markets)	0.116***	0.122***	0.112***
	(0.00159)	(0.00151)	(0.00151)
Tenure	Yes	Yes	Yes
Industry-Region FE	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes
Region-Time FE	Yes	Yes	Yes
Initial log(Wage) × Tenure	No	Yes	Yes
Initial log(NumMarkets) × Tenure	No	No	Yes
Initial Multi-market × Tenure	No	No	Yes
Obs.	30,581,126	30,581,126	30,581,126
Adj.-R^2	0.13	0.19	0.19
Avg Wage	757.90	757.90	757.90
N. Workers	3,309,019	3,309,019	3,309,019

Standard errors in parentheses, clustered at the worker level.

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

Note.— The table presents long-difference regressions of log wages on changes in firm scope over the course of a worker's job spell. In all specifications, the unit of observation is a worker-year, and differences are taken with respect to the start of the worker's job spell. We define the firm's number of markets as the number of microregions where the firm operates. We progressively control for heterogeneous tenure effects by initial wage and initial markets of the firm.

To further discipline the interpretation of the within-job-spell wage-growth results, we estimate an event study that traces the dynamics of incumbent workers' wages around the *first* multi-market expansion experienced during a given worker–firm spell. This design serves two purposes. First, it visualizes whether wage gains materialize at the time of expansion or accumulate over time. Second, by allowing us to test for pre-trends, it sharpens our interpretation of the long-difference results: if wages were already rising faster prior to expansion, then the long-difference estimates could reflect anticipatory growth, rather than a response of innovation to firm scope.

For each worker–firm job spell (i, j), we define the "event" as the first period in which firm

j increases its number of markets while employing worker i. Denote the period of the event by $t_e(ij)$. We then define event time as $\tau(ij,t) = t - t_e(ij)$, so that $\tau = 0$ is the period of first expansion. Separately, we measure job-spell tenure by $t - t_0(ij)$, where $t_0(ij)$ is the start of the spell; controlling flexibly for tenure ensures that the estimated event-time coefficients do not mechanically pick up standard within-job wage growth profiles. We estimate the following event-study specification:

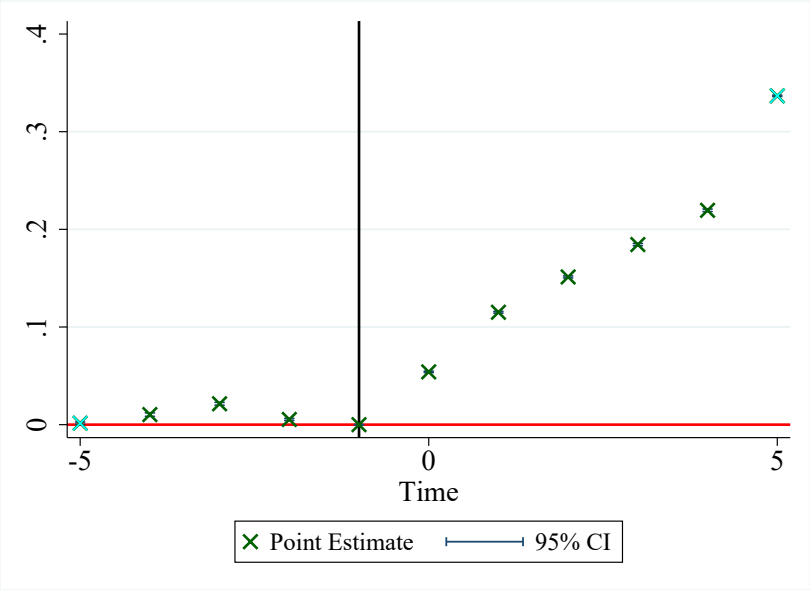
$$\log w_{ijkrt} = \sum_{\tau=-5}^{5} \beta_\tau \, \mathbf{1}\{\tau(ij,t) = \tau\} + \gamma\big(t - t_0(ij)\big) + X_{ijkr}(t_0)'\delta_0 + \phi_i + \omega_{kr} + \omega_{kt} + \omega_{rt} + \varepsilon_{ijkrt}. \tag{5}$$

The outcome is the log wage of worker i employed by firm j in sector k, region r, and period t. The coefficients $\{\beta_\tau\}$ trace wage differences relative to a reference period (e.g., $\tau = -1$ omitted) around the first expansion event. We pool all observations with $\tau(ij,t) \leq -5$ into the $\tau = -5$ bin and all observations with $\tau(ij,t) \geq 5$ into the $\tau = 5$ bin to avoid noise from sparsely populated tails. As before, the worker fixed effect ϕ_i absorbs time-invariant worker heterogeneity. The fixed effects ω_{kr}, ω_{kt}, and ω_{rt} flexibly net out persistent industry–region wage differences and time-varying sectoral and local shocks. The vector $X_{ijkr}(t_0)$ includes job-spell initial conditions—such as the starting wage and the firm's initial number of markets—so that comparisons around the event are made among spells with similar baseline levels. We retain spells in which the firm never expands during the worker–firm match as part of the control group.

We plot the estimated event study coefficients and confidence intervals (which are barely visible) in Figure 6. Prior to firm expansion, the event study coefficients are mostly flat, showing there is no supplemental wage growth in the lead-up to expansion. After expansion, wages rise upon impact ($\hat{\beta}_0 = 0.06$) and continue to grow in subsequent periods, cumulating to $\hat{\beta}_4 = 0.22$ and $\hat{\beta}_5 = 0.33$. These results are consistent with the hypothesis that expanding the scope of the firm generates a substantial increase in earnings by workers. They suggest that firm-level productivity growth is shared across the firm's workforce and, by extension, across the firm's set of markets. We interpret this substantial increase in earnings as evidence of the feedback effect that the spatial scope of the firm has on the productivity upgrading of the firm. In other words, this fact provides evidence that at-scale implementation of service production incentivizes technological improvements.

Firm employment and scope We have comprehensively documented a positive association between worker wages and firm scope, as measured by the number of markets. Now we test whether firm scope increases the firm's employment more than one-for-one over the

Figure 6: Firm Expansion into New Microregion. Dynamics of Incumbent Workers' Log Wages



Note.— The figure presents an event study with a worker's log wages as the dependent variable. Event-time $\tau = 0$ is defined as the first year during the worker's job spell that their employing firm expanded into a new microregion (commuting zone). Our event study specification includes controls for linear job tenure trends, the worker's log wage at the start of the job spell, the firm's log number of markets at the start of the job spell, and fixed effects for sector-region, sector-year, and region-year. Point estimates are represented by X's, and 95% confidence intervals (barely visible) are represented by capped spikes.

long run. If firm employment scales more than proportionally with the number of markets, then employment-per-market also increases in firm scope. This result would provide further evidence that growth of firm scope is associated with growth of productivity, and hence labor demand, in every market of the firm.

We use firm-level long-difference regressions to estimate the relationship between a firm's employment growth and growth in the number of markets it serves. In particular, we estimate the following regression:

$$\Delta_h \log(Emp_{ft}) = \alpha_h \Delta_h \log N_{ft} + \alpha_{0,h} \log N_{ft-h} + \beta_{0,h} \log Emp_{ft-h} + \omega_{krh} + \omega_{kth} + \omega_{rth} + \varepsilon_{fth}, \tag{6}$$

where Δ_h represents the change in a variable Y from $t-h$ to t: $\Delta_h Y_{ft} = Y_{ft} - Y_{ft-h}$. Our parameter of interest in Equation (6) is α_h, which represents the elasticity of the firm's employment growth with respect to growth in the firm's number of markets N_{ft} over time horizon h. We control for the firm's initial log employment and log number of establishments at time $t-h$, to account for heterogeneous employment trajectories by initial conditions. At every time horizon h, we also control for sector-region, sector-time, and region-time fixed effects. We estimate Equation (6) separately for four time horizons: $h = 1, 2, 5, 10$.

The results are in Table 4, which uses the microregion as its definition of a market. The leftmost four columns show results for the full sample of firms, while the rightmost four columns show results specifically for firms in non-tradable services. In all cases, the estimated elasticity α_h is positive and significant. However, its magnitude changes with respect to h. In the short run, for $h = 1$, firm employment increases less than proportionally with the number of markets: $\alpha_1 < 1$. This is natural given that the firm's employment in new markets may not reach its full potential over a short horizon. But in the long run, for $h = 5$ or $h = 10$, α_h is significantly greater than 1. The long-run elasticity suggests that as the firm expands its scope, it also deepens its presence and its market share across all of its markets. Together with the response of the firm's wages, this result is consistent with long-run productivity improvements that respond to the scope of the firm, reinforcing the feedback loop between innovation and implementation.

Table 4: Number of Workers

	Dependent Variable: $\Delta_h \log(Emp_{ft})$							
	Full Sample				Non-Tradable			
	(h=1)	(h=2)	(h=5)	(h=10)	(h=1)	(h=2)	(h=5)	(h=10)
$\Delta_h \log N_{ft}$	0.811***	1.008***	1.261***	1.449***	0.855***	1.069***	1.329***	1.547***
	(0.00959)	(0.00954)	(0.0121)	(0.0180)	(0.0142)	(0.0139)	(0.0166)	(0.0238)
$\log Emp_{ft-h}$	-0.0845***	-0.140***	-0.230***	-0.300***	-0.0915***	-0.151***	-0.247***	-0.326***
	(0.000306)	(0.000541)	(0.00112)	(0.00212)	(0.000405)	(0.000711)	(0.00149)	(0.00294)
$\log N_{ft-h}$	0.190***	0.320***	0.541***	0.724***	0.215***	0.353***	0.589***	0.773***
	(0.00247)	(0.00444)	(0.00954)	(0.0187)	(0.00355)	(0.00630)	(0.0137)	(0.0256)
Industry-Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Region-Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Obs.	4,377,229	3,719,313	2,420,155	1,201,373	2,310,139	1,944,830	1,236,553	584,717
Adj.-R^2	0.05	0.09	0.14	0.18	0.06	0.09	0.15	0.18
Avg Y	0.02	0.02	0.07	0.18	0.02	0.03	0.10	0.25
N. Firms	655,975	539,931	346,800	184,268	364,261	296,742	187,097	96,459

Standard errors in parentheses, clustered at the firm level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Note.— The table presents long-difference regressions of firm employment growth $\Delta_h \log Emp_{ft}$ on growth in the firm's number of markets $\Delta_h \log N_{ft}$ over different time horizons h. In all specifications, the number of markets of the firm is the number of microregions where the firm operates. All specifications control for initial log employment and initial log markets of the firm. The rightmost four columns limit the sample to firms in non-tradable service industries.

2.2.6 Reallocation to multi-market firms drove local wage gains

As our last motivating fact, we show that local economies tend to benefit from the expansion of employment in multi-location firms. In particular, for each Brazilian commuting zone r, we compute total growth of average wages from 1995 to 2018 as $\frac{\overline{w}_{r,2018} - \overline{w}_{r,1995}}{\overline{w}_{r,1995}}$ using RAIS, then use OLS to regress this wage growth on the change in the employment share of multi-location firms, $\Delta \left(\frac{L_r^M}{L_r} \right)$, where L^M is employment in multi-location firms and L is total local

employment in RAIS, during the same period. The result of this regression is in Column 1 of Table 5. The expansion of multi-establishment firms was positively correlated with wage growth; a one percentage point increase in the multi-establishment employment share is associated with an additional 0.11 percentage points of wage growth.

The association between multi-market firm expansion and wage growth could operate primarily through wage growth *within* the set of multi-market and single-market firms, or by *reallocation* towards multi-market firms, which pay greater wages. We refer to this reallocation towards multi-market firms as the "modernization" margin. To understand the mechanics of the association between multi-market expansion and wage growth, we perform a decomposition exercise in the remaining columns of Table 5. Define $\bar{w}_{rt}$ as the average wages in region r in year t. For each region, the growth in wages can be decomposed as:

$$\frac{\bar{w}_{rt} - \bar{w}_{rt-1}}{\bar{w}_{rt-1}} = l^M_{rt-1} \underbrace{\frac{\bar{w}^M_{rt} - \bar{w}^M_{rt-1}}{\bar{w}^M_{rt-1}}}_{\substack{\textit{Wage growth within} \\ \textit{multi-market firms}}} + l^S_{rt-1} \underbrace{\frac{\bar{w}^S_{rt} - \bar{w}^S_{rt-1}}{\bar{w}^S_{rt-1}}}_{\substack{\textit{Wage growth within} \\ \textit{single-market firms}}} + \frac{1}{\bar{w}_{rt-1}} \underbrace{\left(\bar{w}^M_{rt} - \bar{w}^S_{rt}\right)}_{\substack{\textit{Multi-market} \\ \textit{wage gap}}} \underbrace{\left(l^M_t - l^M_{t-1}\right)}_{\substack{\textit{"Modernization"} \\ \textit{margin}}} \tag{7}$$

where l^j_{rt} is the employment share of firm type $j = \{M, S\}$ in region r on year t, while $\bar{w}^j_{rt}$ is the average wage of firm type j in region r on year t. The types of firms are multi-market firm (M) and single establishment firm (S). The extent to which different margins are associated with wage growth helps us understand whether multi-market firms directly drove wage gains by pulling in more workers, or if they simply tended to enter in places that saw strong income growth, without paying a significant premium.

Table 5: Decomposed Wage Growth and Multi-Establishment Employment Growth

	(1)	(2)	(3)	(4)
	Total wage growth	Within multi-mkt.	Within single-mkt.	Modernization
Δ Multi-est. emp. share	0.109*	0.226***	-0.336***	0.220***
	(0.0509)	(0.0501)	(0.0401)	(0.0103)
Observations	486	486	485	486

Standard errors in parentheses

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

Note.— The table presents and decomposes the association between microregion-level wage growth and the change in the employment share of multi-market firms from 1995 to 2018. Column (1) presents the raw association as estimated from an OLS regression. The remaining columns decompose wage growth into three components, as per Equation (7), and regresses each of these components on the employment share of multi-market firms.

Columns 2-4 of Table 5 perform the decomposition in Equation (7). They show that the modernization margin did drive significant wage gains, and that the association with within-firm-type wage growth was heterogeneous. Modernization – the reallocation of the workforce towards multi-establishment firms – contributes 0.22 percentage points of wage growth for every percentage point increase in the multi-establishment employment share, explaining more than the entire association between multi-establishment employment growth and wage growth. Within firm types, multi-establishment employment growth is associated with significant wage gains in multi-establishment firms (Column 2) and wage *losses* in single-establishment firms (Column 3).

Since the negative coefficient for single-establishment firms is larger than the positive coefficient for multi-establishment firms, multi-establishment employment growth is associated with overall negative wage growth within firm type, suggesting that multi-establishment firms did not simply enter in places that would otherwise have seen strong wage growth. The stark difference between the trajectory of multi- and single-establishment wages suggests that multi-establishment entry may have driven productivity upgrading in multi-establishment firms, while negatively impacting single-establishment firms through competition.

2.3 Natural Experiment

2.3.1 Institutional Context

A detailed account of Brazil's 1996 higher education reform is provided in Cox (2025), on which this section builds. The reform fundamentally reshaped the regulation of higher education, triggering a rapid and spatially heterogeneous expansion in college access.

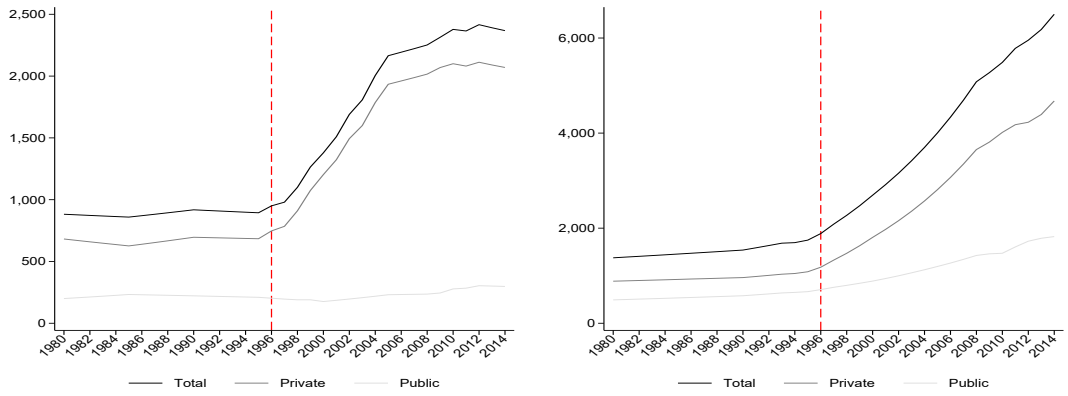
Prior to the reform, the Federal Council of Education (FCE), established in 1961, tightly

controlled the creation of private higher-education institutions. The Council approved new colleges and programs, determined curricula, and set enrollment limits. This centralized oversight made the system highly responsive to political and professional pressures, which over time constrained the expansion of higher education. As a result, the number of institutions stagnated through the 1980s and early 1990s.

Two structural constraints limited college access before the reform. First, geographic disparities in college availability forced many prospective students to travel long distances to take entrance exams and attend classes. Second, existing institutions operated at or near full capacity, implying binding supply constraints. As demand for higher education rose, these geographic and capacity-based frictions generated substantial regional variation in unmet demand, setting the stage for the heterogeneous response to the 1996 reform.

The 1996 reform dismantled these constraints by overhauling the regulatory framework. It simplified the approval process for new colleges and introduced standardized quality-assurance mechanisms to monitor instruction. Deregulation allowed private, for-profit providers to enter the market, leading to a large-scale, market-driven expansion in capacity. As shown in Figure 7, both the number of colleges and total enrollment rose sharply after 1996, driven entirely by private institutions. This expansion was geographically broad, with new campuses opening in previously underserved regions, lowering the cost of access and facilitating local enrollment.

Figure 7: Higher Education Reform

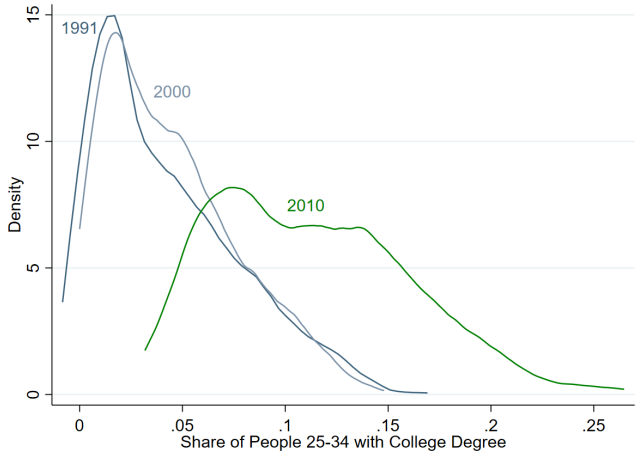


(a) Number of Colleges

(b) Enrollment (1000s)

Note.— Panel (a) plots the time series of the total number of colleges and the number of colleges by type of institution, public or private. Panel (b) reproduces the same information using college enrollment measured in thousands. In both panels, the red line highlights 1996, when the education law was enacted.

Figure 8: Expansion in College Attainment



Note.— The figure shows the distribution across CZs (labor markets) of the share of people in the 25 to 34 age group with a college degree. The figure shows the distribution of this share for 1991, 2000, and 2010.

The rise in enrollment translated into a substantial increase in the supply of college graduates. Figure 8 documents this evolution: in 1991, five years before the reform, the median commuting zone (CZ) had 3.3% of its youth population (ages 25–34) with a college degree. By 2000, three years after the reform, this figure remained largely unchanged, as most students had not yet graduated. However, by 2010, the median had risen to 10.5%, reflecting a large rightward shift in the distribution of college attainment across CZs.

Although the rapid growth of private institutions raised concerns about instructional quality, new oversight mechanisms – including a national exit exam (ENADE) and periodic accreditation reviews – were implemented to monitor educational outcomes. Evidence in Cordeiro and Cox (2025) indicates that for-profit colleges established after the reform delivered instruction of comparable quality to pre-existing non-profit institutions, and that the occupational composition and fields of study of graduates remained stable across cohorts.

Overall, the 1996 reform dramatically expanded higher-education capacity, increased enrollment, and raised the share of college-educated individuals in the workforce, particularly in previously underserved regions. This spatially heterogeneous expansion in skill supply provides the foundation for our identification strategy.

2.3.2 Empirical Strategy

This section outlines the empirical strategy used to assess the consequences of rising college attainment on the expansion of multi-market firms. The approach builds on the identification framework developed in Cox (2025), to which we refer readers for additional institutional and methodological details.

We measure the rise in college attainment as the change in the share of individuals aged 25 to 34 with a college degree between 2000 and 2010. This cohort enrolled in higher education roughly between 1995 and 2005—a period during which the effects of the 1996 higher education reform, particularly the expansion of private colleges, became most salient.

$$\Delta Y_r = \delta_0 + \beta \Delta h_r + X_r \gamma + \theta_r^s + \epsilon_r \tag{8}$$

Equation (8) presents the baseline specification, where ΔY_r denotes the change in an economic outcome of interest in region r, such as the growth in the number of multi-market firms or the change in their employment share. The explanatory variable of interest, Δh_r, measures the change in college attainment among young adults (ages 25–34) in the region between 2000 and 2010. The coefficient β captures the effect of rising educational attainment on economic development. Control variables X_r capture pre-reform development patterns in human capital and economic development, primarily from the 1991 and 2000 censuses. Finally, ϵ_r denotes unobserved determinants of economic outcomes across regions.

Estimating β by Ordinary Least Squares is unlikely to yield a causal estimate. Human capital accumulation and multi-market firm expansion both respond to unobserved regional fundamentals. Economically dynamic areas may attract new establishments and induce greater investment in education, generating simultaneity through omitted-variable bias and reverse causality.

To address these concerns, we adopt a two-part identification strategy. First, we flexibly control for pre-reform levels and trends in economic development using rich covariates in X_r, which proxy for observable underlying regional fundamentals. Second, and most importantly, we instrument for Δh_r using a measure of pre-reform higher education scarcity relative to *future* potential demand. The instrument is constructed as the ratio of higher-education capacity in 1995 to the size of the local population aged 6–15 in 1991—those who would reach college age after the reform and thus benefit from the newly deregulated expansion of higher-education supply:

$$Z_r = \frac{HEC_{r,1995}}{\text{Population aged 6–15}_{r,1991}}$$

Here, $HEC_{r,1995} = \sum_{m \in \Omega_r} \omega_m Colleges_{m,1995}^{100km}$, where Ω_r denotes the set of municipalities in region r, ω_m is a population weight from the 1991 Census, and $Colleges_{m,1995}^{100km}$ is the number of colleges within 100 kilometers of municipality m in 1995. The 100km radius is supported by administrative evidence showing that over 95% of private-sector students enroll within this distance from their place of residence.[4]

[4]See the details in Cox (2025).

The first-stage relationship captures the idea that, following the deregulation of the higher education sector, private colleges were more likely to expand in areas that had been under-supplied relative to demographic demand. This exogenous supply shock drives regional variation in the increase in college attainment across Brazil. This identification strategy is grounded in the historical regulatory structure that preceded the reform. Until 1996, the Federal Council of Education (FCE) maintained centralized control over college authorization, often failing to allocate new capacity where future demand was greatest. This regulatory mismatch—alongside the post-reform liberalization that removed barriers to private entry—generates quasi-experimental variation in the expansion of higher education. An alternative instrument based on observed post-reform entry of colleges would risk endogeneity, as investors may have targeted economically dynamic regions. Our measure, by contrast, captures pre-reform constraints and demographic structure, thus offering a more credible source of exogenous variation.

The main empirical results are presented in Section 2.3.3. We report OLS and IV estimates and assess the validity of the identification assumption through a series of robustness checks and falsification exercises discussed in the latter part of that section.

2.3.3 Results

This section evaluates the impact of rising college attainment on the prevalence and expansion of multi-market firms by estimating equation (8). Appendix A.1 reports the results for the first stage analysis. The analysis proceeds in four parts.

First, Table 6 presents the estimated effects of rising college attainment on two key outcomes: the change in the share of firms that operate across multiple markets, and the change in the employment share of these firms. We consider three definitions of "multi-market" status: (i) firms with multiple establishments, (ii) firms operating in multiple municipalities, and (iii) firms spanning multiple commuting zones (CZs). These complementary definitions capture different degrees of geographic scope.

Second, Table 7 explores the extensive margin by analyzing the change in the number of multi-market firms. This allows us to distinguish between two potential mechanisms underlying the rise in the employment share of multi-market firms: the entry or transformation of new firms into multi-market organizations versus the expansion of existing multi-market firms.

Third, the same table disaggregates the results by sector to assess heterogeneity in the responsiveness of firm structure to human capital accumulation. This analysis provides insight into the types of activities more likely to benefit from a more educated workforce through organizational scaling and geographic expansion.

Finally, Figure 9 presents the dynamic effects. The evolution of the estimated coefficients over time illustrates how the impact of rising college attainment unfolds throughout the period and provides supporting evidence for the parallel trends assumption underlying the identification strategy.

Table 6: Implication for the prevalence of multi-market firms, 1996-2018

Dep.Variable	Δ share of multi-market firms						Δ employment share in multi-market firms					
Market Def.	Establishment		Municipality		CZ		Establishment		Municipality		CZ	
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δh_{25-34}	0.06	0.29	0.03	0.17	0.02	0.15	0.75**	3.57***	0.59*	3.73***	0.56*	3.64***
	(0.05)	(0.15)	(0.05)	(0.18)	(0.05)	(0.16)	(0.27)	(0.98)	(0.26)	(1.02)	(0.26)	(1.01)
R^2	0.10	0.04	0.09	0.06	0.08	0.06	0.06	.	0.04	.	0.04	.
F-test		46.93		46.93		46.93		46.93		46.93		46.93
CV 5%		37.42		37.42		37.42		37.42		37.42		37.42
CV 10%		23.11		23.11		23.11		23.11		23.11		23.11
Obs.	486	486	486	486	486	486	486	486	486	486	486	486

Standard errors in parentheses

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

Note.— The table presents regression results where the dependent variables are: (1) the change in the share of multi-market firms among all firms, and (2) the change in the employment share of multi-market firms among total employment. Both outcomes are regressed on measures of college attainment.

Table 6 reports the estimates for the effects of rising college attainment on the prevalence of multi-market firms. Columns (1)–(6) examine the change in the share of firms classified as multi-market between 1996 and 2018, while columns (7)–(12) focus on the corresponding change in the employment share of these firms. Odd-numbered columns present OLS estimates; even-numbered columns report Two-Stage Least Squares (2SLS) estimates using the instrumental variable strategy outlined in Section 2.3.2.

The results indicate that rising college attainment had no discernible effect on the share of firms classified as multi-market. However, they did lead to a substantial reallocation of employment toward these firms. Specifically, a 1 percentage point increase in the share of college-educated individuals is associated with a 3.5 percent increase in the employment share of multi-market firms over the 22-year period. This effect is robust across all three definitions of market scope—multiple establishments, multiple municipalities, and multiple commuting zones—suggesting that the expansion of these firms spans large geographic areas, not merely within narrowly defined markets.

A comparison of the OLS and IV estimates reveals important distinctions in both interpretation and magnitude. The OLS estimates approximate the average treatment effect on the treated (ATT). However, because changes in college attainment may respond endogenously to broader economic trends, these estimates are susceptible to bias. One plausible source is skill-biased technical change, which disproportionately raises the productivity of college-educated workers and lowers the fixed or managerial costs of operating across multiple markets. While such technological shifts facilitate firm expansion, they may simultaneously displace lower-skilled workers, attenuating the net employment response. As a result, the OLS estimates may understate the true effect of rising human capital on firm-level outcomes.

The IV estimates, by contrast, capture the local average treatment effect (LATE) for regions with relatively high pre-reform scarcity of higher education. These are the areas most affected by the relaxation of entry restrictions following the 1996 reform. That the IV estimates for the change in the employment share of multi-market firms are consistently larger than the OLS estimates suggests substantial heterogeneity in the marginal return to college attainment across regions. In particular, if the supply of skilled labor was a binding constraint on firm growth in these initially underserved areas, then relaxing this constraint through higher college attainment would yield outsized organizational and employment responses. This finding supports the hypothesis that the transition from traditional, single-market production to modern, multi-market structures is skill-intensive, and that expanding access to college education enables firms to scale operations and reorganize along more complex dimensions.

We next examine whether the observed increase in the employment share of multi-market firms is driven by the entry of new firms into the multi-market category or by the expansion of employment within already-established multi-market firms. Table 7 reports the estimated effects of rising college attainment on the growth in the number of firms. Columns (1), (2), (5), (6), (9), and (10) consider all firms, regardless of whether they employ college-educated workers. These estimates capture the overall expansion at the extensive margin. Columns (3), (4), (7), (8), (11), and (12) restrict attention to firms that report employing at least one college graduate, thereby allowing us to assess whether the expansion is mediated through the hiring of skilled labor—a key channel in our proposed mechanism. As in prior tables, odd-numbered columns report OLS estimates, and even-numbered columns report Two-Stage Least Squares (2SLS) estimates based on the instrumental variable strategy.

Table 7: Growth in the number of multi-market firms, 1996-2018

Dep.Variable	Growth number of multi-market firms											
Market def.	Establishments				Municipalities				CZ			
Sample	All		w/Skill		All		w/Skill		All		w/Skill	
Specification	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Δh_{25-34}	-0.96	8.21**	1.80	19.90***	-1.00	7.22**	1.54	17.75**	-0.94	6.74*	1.45	17.56**
	(0.77)	(2.90)	(1.05)	(5.31)	(0.73)	(2.65)	(1.00)	(5.63)	(0.73)	(2.66)	(0.99)	(5.43)
R^2	0.46	0.26	0.33	.	0.41	0.23	0.32	.	0.39	0.23	0.31	.
F-test		46.93		46.93		46.93		46.93		46.93		46.93
CV 5%		37.42		37.42		37.42		37.42		37.42		37.42
CV 10%		23.11		23.11		23.11		23.11		23.11		23.11
Obs.	486	486	486	486	486	486	486	486	486	486	486	486

Standard errors in parentheses

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

Note.— The table reports regression results where the dependent variable is the growth in the number of multi-market firms over the period 1996–2018, regressed on measures of college attainment.

The results indicate that the growth in the employment share of multi-market firms reflects, at least in part, an expansion along the extensive margin: regions with greater increases in college attainment also experienced a rise in the number of firms operating across multiple markets. This finding suggests that the reform facilitated a structural transition from traditional, localized firms to more complex, geographically dispersed organizations. Importantly, the results are stronger for multi-market firms that employ college-educated workers, consistent with the idea that skilled labor plays a central role in enabling such organizational change. This pattern reinforces the plausibility of our mechanism: the expansion of higher education increased the availability of skilled workers, which in turn allowed more firms to adopt multi-market structures.

Table 8: Growth in the number of multi-market firms by sector, 1996-2018

Dep.Variable	Growth number of multi-Market firms w/skill					
Sector	Manufacture		Infrastructure		Services	
Specification	OLS	IV	OLS	IV	OLS	IV
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Establishments						
Δh_{25-34}	2.06	6.34	0.82	5.12	1.05	19.80***
	(1.59)	(6.94)	(1.12)	(4.01)	(1.14)	(4.90)
R^2	0.19	0.17	0.32	0.29	0.26	.
Panel B: Municipalities						
Δh_{25-34}	2.15	4.70	0.83	-0.28	0.48	15.49***
	(1.53)	(5.96)	(1.10)	(5.28)	(1.08)	(4.24)
R^2	0.16	0.18	0.32	0.34	0.24	.
Panel C: Commuting Zones						
Δh_{25-34}	2.45	4.32	0.92	-0.07	0.15	14.54***
	(1.52)	(5.91)	(1.12)	(5.39)	(1.07)	(4.24)
R^2	0.15	0.18	0.31	0.33	0.23	.
F-test		46.93		46.93		46.93
CV 5%		37.42		37.42		37.42
CV 10%		23.11		23.11		23.11
Obs.	486	486	486	486	486	486

Standard errors in parentheses

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

Note.— The table reports regression results by sector where the dependent variable is the growth in the number of multi-market firms over the period 1996–2018, regressed on measures of college attainment.

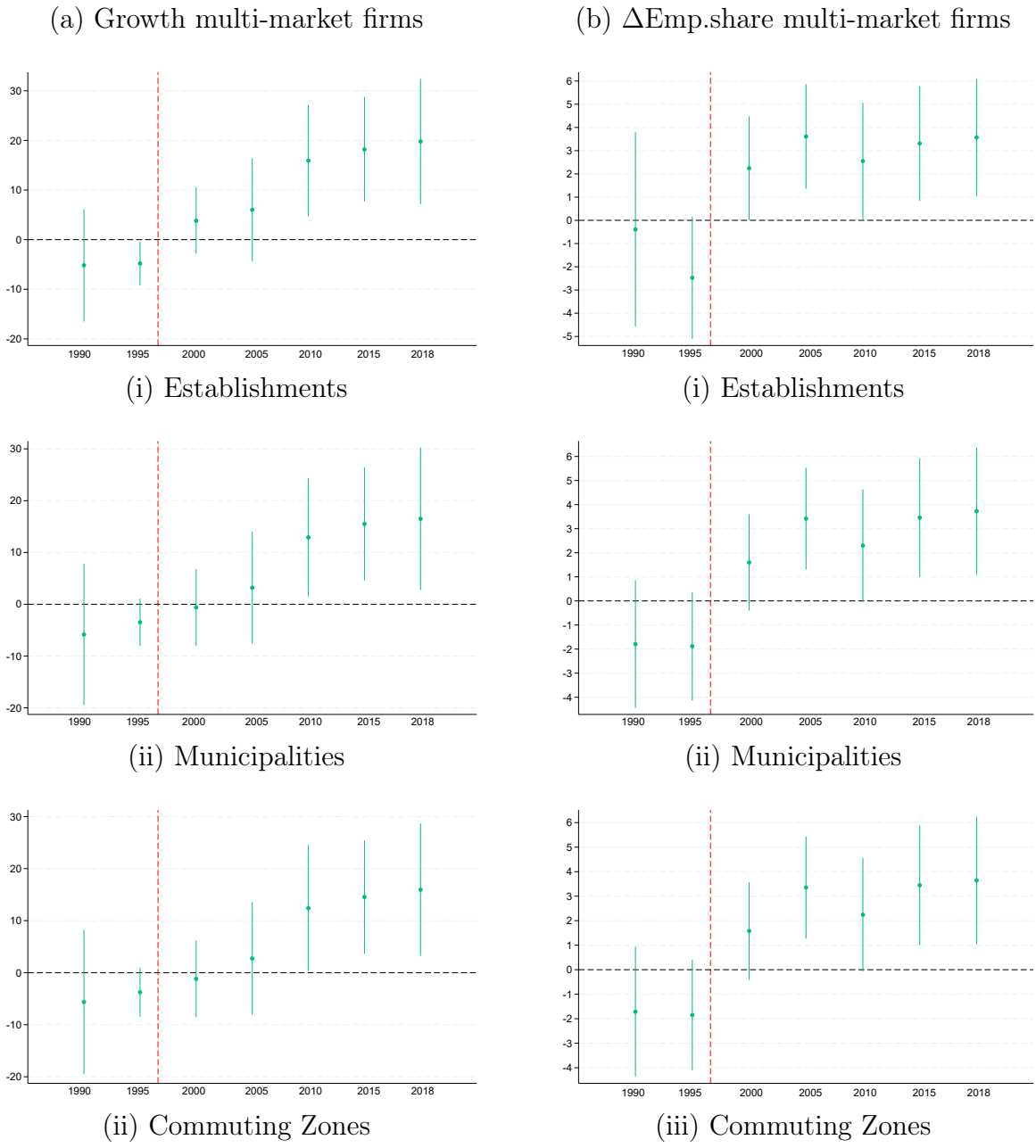
We further investigate sectoral heterogeneity in the effects of rising college attainment, motivated by the hypothesis that the transformation toward multi-market firm structures is more pronounced in service-oriented, non-tradable industries. Table 8 presents estimates disaggregated by sector. Columns (1) and (2) report results for manufacturing; columns (3) and (4) focus on infrastructure and construction; and columns (5) and (6) report estimates for service industries. Within each sector, we distinguish between OLS (odd-numbered columns) and IV (even-numbered columns) specifications. The table is organized into three panels, each corresponding to a different definition of market scope. Panel A uses the number of

multi-establishment firms as the outcome variable, capturing firm expansion through the creation of multiple physical units. Panel B shifts the focus to firms operating in multiple municipalities, while Panel C measures multi-market presence at the level of commuting zones (CZs), reflecting the broadest geographic reach.

The results confirm that the expansion of multi-market firms is concentrated in the service sector. Across all definitions of market scope—establishments, municipalities, and commuting zones—the number of multi-market firms grew most rapidly in services. This pattern aligns with the hypothesis that service industries, typically non-tradable and more reliant on skilled labor to scale, are especially responsive to improvements in local human capital. By relaxing constraints on the availability of college-educated workers, rising college attainment enables firms in these sectors to expand geographically and adopt more complex organizational structures.

Finally, Figure 9 illustrates the dynamic relationship between rising college attainment and firm-level outcomes over time. Panel (a) plots the IV-estimated effect of increased college attainment on the growth in the number of multi-market firms, while Panel (b) shows the corresponding effect on the change in their employment share. Each panel reports results separately for the three definitions of market scope—multi-establishment, multi-municipality, and multi-CZ firms. The horizontal axis indicates the final year used to compute the change in outcomes relative to the baseline year of 1996. For example, the estimate at 2010 reflects the effect of college attainment on the change in the outcome between 1996 and 2010, with 95% confidence intervals. The estimate at 1990 captures the pre-reform relationship and serves as a test of the parallel trends assumption. The timing of the 1996 higher education reform is indicated by the vertical red dashed line.

Figure 9: Tracking the effect over time



Note.— The figure reports IV-estimated effects of increased college attainment on (a) the growth in the number of multi-market firms and (b) the growth in the employment share of multi-market firms. Regressions are conducted separately for three definitions of market scope: (i) establishments, (ii) municipalities, and (iii) commuting zones.

2.3.4 Additional evidence from the college wage premium

Our main results show that an increase in local college attainment led to a disproportionate expansion of multi-market firms, particularly in services. This finding is consistent with our central mechanism: multi-market production relies disproportionately on skilled labor

because operating standardized activities across space requires managerial and organizational capabilities that are intensive in college-educated workers. The previous section established that multi-market firms are more skill-intensive than single-market firms, that this skill intensity rises with geographic scope, and that these firms pay a substantial wage premium. Taken together, these results suggest that the expansion of multi-market firms changes the local relative demand for skilled labor.

A natural implication is that the relationship between multi-market firm expansion and the college wage premium should depend on whether skilled labor is locally scarce. In labor markets where college-educated workers are scarce, the growth of multi-market firms should raise the relative demand for skill and therefore increase the college premium. By contrast, once the higher education reform relaxed local skill constraints and expanded the supply of college-educated workers, the same increase in the employment share of multi-market firms should generate less upward pressure on skilled wages. If the supply response is sufficiently strong, the relationship may even turn negative. This pattern would provide additional support for our interpretation that the reform operated in part by relaxing a managerial-skill bottleneck that had previously constrained the expansion of modern service firms. This interpretation builds on the paper's broader argument that college-educated workers are a key input in the multi-market production technology and that their pre-reform scarcity constrained service-led structural transformation.

Accordingly, we test whether the association between the change in the college premium and the change in the employment share of multi-market firms differs sharply across the pre-reform and post-reform periods. The hypothesis is not simply that multi-market firms are correlated with higher skilled wages, but more specifically that the sign of this relationship changes as the economy moves from a regime of skill scarcity to one where the skill scarcity constraint relaxes. In the period before the reform-induced rise in college attainment had materialized, commuting zones with larger increases in the employment share of multi-market firms should also exhibit larger increases in the college premium, reflecting stronger relative demand for scarce managerial skill. In the later period, when the reform had substantially increased the local supply of college-educated workers, commuting zones with larger increases in the employment share of multi-market firms should exhibit a weaker relationship, or potentially a negative relationship with the college premium, when the skill supply force dominates.

To test this hypothesis, we estimate the following specification:

$$\Delta \log \left(\frac{w_r^h}{w_r^\ell} \right)_t = \alpha_t + \beta_t \Delta MMShare_{rt} + X_{rt}\gamma_t + u_r$$

where $\Delta \log\left(\frac{w_r^h}{w_r^\ell}\right)$ denotes the change in the college premium in region r and $\Delta MMShare_r$ the change in the employment share of multi-market firms. We use two periods, $t = \{1991-2000, 2000-2010\}$, and test the hypothesis that $\beta_{2000} > 0$, while $\beta_{2010} << \beta_{2000}$, or the stronger prediction that $\beta_{2010} \approx 0$ or $\beta_{2010} < 0$. Table 9 shows the results.

Table 9: College premium and multi-market firm expansion

Dep.Variable	Δ College premium					
Period	1991-2000			2000-2010		
Market	Establishment	Municipality	MMC	Establishment	Municipality	MMC
Specification	(1)	(2)	(3)	(1)	(2)	(3)
$\Delta MMShare_r$	0.78**	0.75*	0.75*	-0.01	-0.07	-0.06
	(0.26)	(0.29)	(0.29)	(0.16)	(0.18)	(0.18)
Obs.	482	482	482	482	482	482
R^2	0.61	0.60	0.60	0.36	0.36	0.36

Note.— The table reports the relationship between the rise in the employment share of multi-market firms and the change in the college-premium for two periods, 1991-2000 and 2000-2010. We use three definitions of multi-market firm: multi-establishment, multi-municipality, and multi-MMC. Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

The estimates reveal a clear change in the relationship between multi-market firm expansion and the college premium across periods. In 1991–2000, the coefficient on the change in the employment share of multi-market firms is positive and statistically significant, ranging from 0.75 to 0.78, indicating that locations with faster multi-market expansion also experienced larger increases in the college premium. This pattern is consistent with the idea that, when skilled labor was scarce, the expansion of multi-market firms increased the relative demand for scarce skilled labor. In contrast, in 2000–2010 the estimated coefficients become slightly negative, between -0.01 and -0.07, but are not statistically distinguishable from zero. This suggests that once the reform expanded the supply of college-educated workers, the positive association between multi-market expansion and the college premium disappeared, consistent with a substantial relaxation of local skill constraints.

3 Model

To interpret the empirical patterns documented above and to quantify the economic effects of expanding college access, we develop a simple model of regional production and skill allocation. The framework builds on Hsieh and Rossi-Hansberg (2023), where firms in non-tradable industries can operate across multiple local markets by replicating their technology across establishments. We introduce a key modification: each establishment requires high-skilled workers to serve as managers. This feature creates a direct link between the local supply of college-educated workers and the spatial distribution of productive activity.

In the present draft, we set up a two-region version of the model with perfectly competitive labor markets to build intuition. Future versions of the paper will feature an arbitrary set of N regions and employer monopsony in the labor market for quantitative analysis.

3.1 Environment

Consider two locations, denoted $n \in C, R$, where C represents a central city (e.g., São Paulo) and R a rural hinterland. From the perspective of a given location n, let n' denote the other location. In each location, there are two non-tradable service industries – traditional and modern services – and a single tradable sector whose output can be exchanged across regions. Each location n is endowed with a measure L_n of low-skilled workers and H_n of high-skilled workers, which are the only factors of production. The model is static, so establishments' entry and operation decisions coincide.

3.2 Production

We consider three production sectors: one tradable good G and two non-tradable services, traditional services T and modern services M.

3.2.1 Tradable good

The tradable good G is produced in each location by a representative firm using a constant-returns-to-scale technology that employs only low-skilled labor with exogenous productivity $A_{G,n}$:

$$G = A_{Gn}\ell_n$$

Because G is perfectly tradable, its price equalizes across locations. We normalize this common price to one and use it as the numeraire. The tradable sector captures the range of exportable commodities – such as soybeans, oil, and manufactures – that generate variation in nominal income (that is, in the quantity of currency available for expenditure) across Brazilian regions.

3.2.2 Non-tradable modern services

We begin by specifying the production structure of modern services, which constitute the core of our analysis. Traditional services share many of the structural features of modern services but differ in key dimensions related to their scalability across markets and their reliance on high-skilled labor. These differences, introduced in Section 3.2.6, are designed to

capture the empirically observed contrast between high-productivity, skill-intensive, multi-market service firms and more localized, low-productivity service activities.

Modern service firms: Each firm i in industry j is endowed with a firm-wide productivity level A_{ij}. Firms produce output according to a simple linear production function:

$$Y_{ijn} = A_{ij}\ell^P_{ijn}, \tag{9}$$

where ℓ^P_{ijn} is the measure of low-skilled workers hired for production by the firm-establishment ijn. The cost to hire a low-skilled worker in market n is $w_{n,\ell}$.

The aggregate output of an industry j in market n combines output of every establishment in (j,n) according to a CES aggregator with elasticity of substitution σ. Denoting the endogenous set of operating establishments in (j,n) as M_{jn}, industry-market output is:

$$Y_{jn} = \left(\int_{i\in M_{jn}} Y_{ijn}^{\frac{\sigma-1}{\sigma}} di\right)^{\frac{\sigma}{\sigma-1}} \tag{10}$$

Firms in (j,n) therefore engage in monopolistic competition and charge prices p_{ijn} at a constant markup over marginal costs, $p_{ijn} = \frac{\sigma}{\sigma-1}\frac{w_{n,\ell}}{A_{ij}}$. This implies that establishment ijn's production labor demand and profits gross of operating costs, π^G_{ijn}, are isoelastic to productivity A_{ij}:

$$\begin{aligned}
\pi^G_{ijn} &= \frac{1}{\sigma}\frac{A_{ij}^{\sigma-1}}{\int_{i'\in M_{jn}} A_{i'j}^{\sigma-1} di'} E_{jn} \\
\ell^P_{ijn} &= \frac{\sigma-1}{\sigma}\frac{A_{ij}^{\sigma-1}}{\int_{i'\in M_{jn}} A_{i'j}^{\sigma-1} di'}\frac{E_{jn}}{w_{n,\ell}},
\end{aligned} \tag{11}$$

where $E_{jn} = P_{jn}Y_{jn}$ is total expenditure in industry-market (j,n), and $P_{jn}^{1-\sigma} = \int p_{ijn}^{1-\sigma} di$ is the price index (j,n).

3.2.3 Operating establishments

While M_{jn} denotes the set of j-industry firms operating an establishment in location n, let μ^n_j denote the set of firms in industry j headquartered in location n. We define $M^n_{jn} = M_{jn} \cap \mu^n_j$ and $M^{n'}_{jn} = M_{jn} \cap \mu^{n'}_j$ as the subsets of M_{jn} consisting of firms headquartered in n and n', respectively.

Each establishment in industry j incurs a fixed overhead cost that is independent of its

scale of production, but that depends on the distance between the establishment and its headquarters. Specifically, for an establishment in region n, the fixed cost is lower when the firm is headquartered locally (n) than when it is headquartered elsewhere (n'). The establishment is set up using high- and low-skilled labor according to a constant-returns-to-scale Cobb–Douglas technology, with high-skilled labor used at intensity q_j. Denoting the wage of high-skilled labor in market n by $w_{n,h}$, the fixed cost of operating an establishment in location n is given by:

$$f_{ijn}(w_{n,h}, w_{n,\ell}) = \begin{cases} \overline{f}_j w_{n,h}^{q_j} w_{n,\ell}^{1-q_j} & \text{if } i \in \mu_j^n \\ \tau_j \overline{f}_j w_{n,h}^{q_j+\kappa} w_{n,\ell}^{1-q_j-\kappa} & \text{if } i \in \mu_j^{n'} \end{cases} \tag{12}$$

Hence, operating establishments away from the firm headquarters is more skill-intensive than operating locally: implementation of the firm's production technology at a distance requires a greater skill input. The parameter κ captures this additional skill intensity, while τ_j represents an overall cost shifter associated with remote operation. In the limit as $\tau_j \to \infty$, the economy converges to one without multi-location firms.

Firms choose to operate in a location only if profits net of operating costs are non-negative. Accordingly, there exist two productivity thresholds, $\underline{A}_{jn}^n$ and $\underline{A}_{jn}^{n'}$, for firms headquartered in n and n', respectively, to open an establishment in n. These thresholds satisfy

$$\begin{aligned} \pi_{jn}(\underline{A}_{jn}^n) &= \overline{f}_j w_{n,h}^{q_j} w_{n,\ell}^{1-q_j} \\ \pi_{jn}(\underline{A}_{jn}^{n'}) &= \tau_j \overline{f}_j w_{n,h}^{q_j+\kappa} w_{n,\ell}^{1-q_j-\kappa} \end{aligned} \tag{13}$$

This immediately implies that

$$\left(\frac{\underline{A}_{jn}^{n'}}{\underline{A}_{jn}^{n}} \right)^{\sigma-1} = \tau_j \hat{w}_n^{\kappa} \tag{14}$$

where $\hat{w}_n = \frac{w_{n,h}}{w_{n,\ell}}$ denotes the local skill premium. Thus, a higher skill premium restricts operations by distant firms more than by local firms, as scarce high-skilled labor makes remote management more costly.

3.2.4 Establishment entry

We now impose structure on the distribution of firm productivity, A_{ij}, which is drawn at the time of firm entry. For firms headquartered in n, we assume that $A_{ij}^{\sigma-1}$ follows a Pareto distribution with shape parameter γ and an endogenously determined scale parameter x_{ij}^n;

equivalently, the cumulative distribution function (CDF) is $G_{ij}^n(z) = 1 - \left(\frac{z}{x_{ij}^n}\right)^{-\gamma}$ for $z \geq x_{ij}^n$. As described in the next section, firms in (j,n) choose x_{ij}^n before observing A_{ij}; by symmetry of the problem, the optimal choice is common across firms, so $x_{ij}^n = x_j^n$ for all i. This parametrization yields closed form expressions for aggregate industry-market productivity in (j,n), which we present in the appendix. In particular, aggregate productivity increases in the measure of firms headquartered in each location $\left(|\mu_j^n|, |\mu_j^{n'}|\right)$ and in the technology choice in each location $\left(x_j^n, x_j^{n'}\right)$, and decreases in the establishment operation cutoffs $\left(\underline{A}_{jn}^n, \underline{A}_{jn}^{n'}\right)$.

Furthermore, the measure of distantly headquartered firms relative to locally headquartered firms depends on the local skill premium, the cost shifter τ, and the measure of headquarters and technology in n' relative to n.

$$\frac{|M_{jn}^{n'}|}{|M_{jn}^n|} = (\hat{w}_n^\kappa \tau_j)^{-\gamma} \left(\frac{x_j^{n'}}{x_j^n}\right)^\gamma \frac{|\mu_j^{n'}|}{|\mu_j^n|} \tag{15}$$

The measure of establishments headquartered elsewhere is more sensitive to the local skill premium, $\hat{w}_n$, than the measure of locally headquartered establishments. All else equal, increases in the local skill premium therefore shift the composition of businesses operating in n towards locally headquartered firms. We show in the appendix that for some parameter configurations, an increase in $\hat{w}_n$ can even raise the overall number of local establishments, as a higher skill premium provides protection for local firms by deterring entry of distant firms.

Thus far, we have taken the sets of firms headquartered in each location, μ_j^n and $\mu_j^{n'}$, as given. However, these headquarters, as well as the productivity shifters x_j^n and $x_j^{n'}$, are endogenous equilibrium objects. We now turn to characterizing their determination.

3.2.5 Headquarters entry

Firm entry into n – which occurs by establishing a headquarters there – takes place *ex ante* of firm i learning its productivity A_{ij}. Firm-level productivity $A_{ij}^{\sigma-1}$ is drawn from a Pareto distribution whose shape is γ and whose scale (floor) is the endogenous choice x_{ij}^n. By choosing x_{ij}^n prior to observing the draw, the firm pins down the lower bound of its *ex ante* productivity distribution. Equivalently, the firm chooses its CDF $G_{ij}^n(z) = 1 - \left(\frac{z}{x_{ij}^n}\right)^{-\gamma}$.

The choice of x_{ij}^n Firms choose x_{ij}^n to maximize expected profits over the induced distribution of productivity draws. Opening a firm with technology level x requires a fixed "headquarters" cost $F_j^n(x)$, which has an intercept – the fixed cost of running a headquarters, regardless of technology – and an isoelastic convex term in x – the cost of investment

in technology. This fixed cost is denominated in goods and is subject to congestion in the mass of firms headquartered in n. In contrast, we assume that technology choice is subject to *local knowledge spillovers*: the cost of technological upgrading is lower when the local average technology x_j^n is greater. In particular:

$$F_j^n(x|\mu_j^n) = \Phi_j|\mu_j^n|^\chi + \frac{\overline{F}_j}{\eta}\frac{x^\eta}{(x_j^n)^\varphi}, \tag{16}$$

where $\eta > \gamma$ ensures an interior solution for x, and χ parameterizes congestion as the mass of headquarters $|\mu_j^n|$ rises. Local knowledge spillovers are parameterized by $\varphi \geq 0$.

Having established the costs of installing technology x, we now consider the expected returns to investments in technology. First note that post-entry profits, $\tilde{\Pi}_j^n(A_{ij})$, are a piecewise function of the realized productivity draw A_{ij}:

$$\tilde{\Pi}_j^n(A_{ij}) = \begin{cases} 0 \text{ if } A_{ij} \leq \underline{A}_{jn}^n \\ w_{n,\ell}\left(\frac{1}{\sigma}\frac{E_{jn}}{w_{n,\ell}}\frac{A_{ij}^{\sigma-1}}{\int_{i'\in M_{jn}} A_{i'j}^{\sigma-1}di'} - \overline{f}_j\hat{w}_n^{q_j}\right) \text{ if } \underline{A}_{jn}^n < A_{ij} \leq \underline{A}_{jn'}^n \\ w_{n,\ell}\left(\frac{1}{\sigma}\frac{E_{jn}}{w_{n,\ell}}\frac{A_{ij}^{\sigma-1}}{\int_{i'\in M_{jn}} A_{i'j}^{\sigma-1}di'} - \overline{f}_j\hat{w}_n^{q_j}\right) + w_{n',\ell}\left(\frac{1}{\sigma}\frac{E_{jn'}}{w_{n',\ell}}\frac{A_{ij}^{\sigma-1}}{\int_{i'\in M_{jn'}} A_{i'j}^{\sigma-1}di'} - \tau_j\overline{f}_j\hat{w}_{n'}^{q_j+\kappa}\right) \text{ if } \underline{A}_{jn'}^n < A_{ij} \end{cases} \tag{17}$$

Notably, the option to serve distant markets increases the value of high draws $A_{ij} > \underline{A}_{jn'}^{,n}$, strengthening incentives to invest in x. Taking the expectation over all possible draws of A_{ij}, expected profits are isoelastic in x, with an endogenous coefficient Ω_{jn} that represents the returns to technology.

$$\mathbb{E}\left[\tilde{\Pi}_j^n(A_{ij})|x\right] = \Omega_{jn}x^\gamma, \tag{18}$$

In the appendix, we present the full expression for Ω_{jn}. The returns to technology decrease in the market-level cutoffs $\underline{A}_{jn}^n$ and $\underline{A}_{jn'}^n$: technology is more valuable when establishments are less costly and the firm can profitably operate in more markets.

Firms choose their technology shifter x_{ij}^n to maximize expected profits net of the fixed entry cost:

$$x_{ij}^n = \arg\max_x \left\{\mathbb{E}\left[\tilde{\Pi}_j^n(A_{ij})|x\right] - \frac{\overline{F}_j}{\eta}\frac{x^\eta}{(x_j^n)^\varphi} - \Phi_j|\mu_j^n|^\chi\right\} \tag{19}$$

Since firms choose x ex ante, before observing their idiosyncratic productivity, symmetry

implies that all entrants in (j,n) select the same technology shifter x_j^n; that is,

$$x_j^n = \left(\frac{\gamma}{\overline{F}_j}\Omega_{jn}\right)^{\frac{1}{\eta-\varphi-\gamma}} \tag{20}$$

Evaluated at x_j^n, expected profits net of entry costs are

$$\Pi_{jn}^e = \frac{\eta-\gamma}{\eta}\left(\frac{\gamma}{\overline{F}_j}\right)^{\frac{\gamma}{\eta-\varphi-\gamma}}\Omega_{jn}^{\frac{\eta-\varphi}{\eta-\varphi-\gamma}} - \Phi_j|\mu_j^n|^\chi \tag{21}$$

Π_{jn}^e is increasing and isoelastic in Ω_{jn} with elasticity $\eta/(\eta-\gamma)$, and decreasing in the unit technology cost $\overline{F}_j$ with elasticity $\gamma/(\eta-\gamma)$. The congestion term $\Phi_j|\mu_j^n|^\chi$ shifts profits downward one-for-one.

Free entry The measure of local headquarters $|\mu_j^n|$ therefore adjusts to satisfy the free entry condition, $\Pi_{jn}^e \leq 0$. Under any equilibrium with positive entry, this implies a positive relationship between $|\mu_j^n|$, the technology choice x_j^n, and the technological returns Ω_{jn}:

$$|\mu_j^n| = \left(\frac{\frac{\eta-\gamma}{\eta}\left(\frac{\gamma}{\overline{F}_j}\right)^{\frac{\gamma}{\eta-\varphi-\gamma}}\Omega_{jn}^{\frac{\eta-\varphi}{\eta-\varphi-\gamma}}}{\Phi_{jn}}\right)^{\frac{1}{\chi}} \quad ; \quad x_j^n = \left(\frac{\eta\gamma}{\eta-\gamma}\frac{\Phi_j}{\overline{F}_j}|\mu_j^n|^\chi\right)^{\frac{1}{\eta-\varphi}}, \tag{22}$$

so that the effect of Ω_{jn} on x_j^n is summarized by $|\mu_j^n|$.

Using the expression for Ω_{jn} in Equation (33) of the appendix, together with the expressions for expected profits conditional on operating an establishment, the free-entry condition can be written as

$$\begin{aligned}\Omega_{jn} &= \frac{\overline{f}_j}{\gamma-1}\left[\left(\underline{A}_{jn}^n\right)^{-\gamma(\sigma-1)} w_{n,\ell}\hat{w}_n^{q_j} + \left(\underline{A}_{jn'}^n\right)^{-\gamma(\sigma-1)}\tau_j w_{n',\ell}\hat{w}_{n'}^{q_j+\kappa}\right] \\ &= \left(\frac{\eta}{\eta-\gamma}\Phi_j|\mu_j^n|^\chi\right)^{\frac{\eta-\varphi-\gamma}{\eta-\varphi}}\left(\frac{\overline{F}_j}{\gamma}\right)^{\frac{\gamma}{\eta-\varphi}},\end{aligned} \tag{23}$$

where the first line of (23) summarizes how the cutoffs $\underline{A}_{jn}^n, \underline{A}_{jn}^{n'}$ to operate establishments diminish the incentives for innovative activities, i.e. headquarters entry and investment. In particular, the possibility to operate in multiple locations implies regional interdependence: innovative activity in one region n depends on economic conditions in the other region n'. When these conditions render implementation of technology in distant regions difficult, innovation declines for the entire economy.

3.2.6 Traditional Services

Traditional services share many structural features with the production framework described above for modern services, but differ along two sector-specific dimensions. First, traditional services are less skill-intensive – captured by a lower q_j – when opening an establishment for operation. Second, we assume that multi-location production is impossible in traditional services: the industry-specific cost shifter for operating distant establishments is infinite ($\tau_T \to \infty$). Traditional services can therefore be thought of as a special case of the industry-level structure laid out above.

3.3 Demand

We specify demand across sectors as a nested, non-homothetic CES function of per-capita real income y_n, with an outer nest allocating expenditure between goods and services, and an inner nest allocating expenditure between traditional and modern services. In this specification, increases in income disproportionately raise the demand for modern services relative to traditional services.

$$y_n = \left(\beta^{\frac{1}{\zeta}} \left(\omega^{\frac{1}{\xi}} \left(\frac{y_{Mn}}{y_n^{\varepsilon_M - 1}} \right)^{\frac{\xi-1}{\xi}} + (1-\omega)^{\frac{1}{\xi}} \left(\frac{y_{Tn}}{y_n^{\varepsilon_T - 1}} \right)^{\frac{\xi-1}{\xi}} \right)^{\frac{\xi}{\xi-1}\frac{\zeta-1}{\zeta}} + (1-\beta)^{\frac{1}{\zeta}} y_{Gn}^{\frac{\zeta-1}{\zeta}} \right)^{\frac{\zeta}{\zeta-1}} \tag{24}$$

where $y_{jn} = \frac{Y_{jn}}{L_n + H_n}$ is per-capita production of sector j in location n.

We further assume that $\zeta \to 1$, so that real income is Cobb–Douglas between goods G and a non-homothetic CES aggregate of services. Under this demand system, the relative demand for modern versus traditional services is given by:

$$\frac{E_{Mn}}{E_{Tn}} = \frac{\omega}{1-\omega} \left(\frac{P_{Mn}}{P_{Tn}} \right)^{1-\xi} y_n^{(\varepsilon_M - \varepsilon_T)(1-\xi)}, \tag{25}$$

with $(\varepsilon_M - \varepsilon_T)(1-\xi) \geq 0$ so that modern services are a luxury relative to traditional services.

This demand system implies the following aggregate expenditure shares:

$$\begin{aligned}
\vartheta_X(e_n, \vec{P}_n) &= 1 - \beta \\
\vartheta_M(e_n, \vec{P}_n) &= \beta \frac{\omega \left(P_{Mn} y_n^{\varepsilon_M - 1}\right)^{1-\xi}}{\omega \left(P_{Mn} y_n^{\varepsilon_M - 1}\right)^{1-\xi} + (1-\omega)\left(P_{Tn} y_n^{\varepsilon_T - 1}\right)^{1-\xi}} \\
\vartheta_T(e_n, \vec{P}_n) &= \beta \frac{(1-\omega) \left(P_{Tn} y_n^{\varepsilon_T - 1}\right)^{1-\xi}}{\omega \left(P_{Mn} y_n^{\varepsilon_M - 1}\right)^{1-\xi} + (1-\omega)\left(P_{Tn} y_n^{\varepsilon_T - 1}\right)^{1-\xi}},
\end{aligned} \tag{26}$$

where $y_n = y(e_n, \vec{P}_n)$ is defined implicitly by:

$$y(e_n, \vec{P}_n) = \frac{e_n}{P_{Sn}^{\beta} P_{Gn}^{1-\beta}} \quad \text{where } P_{Sn} = \left(\omega \left(P_{Mn} y_n^{\varepsilon_M - 1}\right)^{1-\xi} + (1-\omega)\left(P_{Tn} y_n^{\varepsilon_T - 1}\right)^{1-\xi}\right)^{\frac{1}{1-\xi}} \tag{27}$$

Because firm profits average out to zero due to free entry, all income and hence all expenditures come from wages. Aggregate nominal expenditure is $E_n = e_n(L_n + H_n) = w_{\ell,n}(L_n + \hat{w}_n H_n)$. Therefore (j,n)-level expenditures satisfy:

$$E_{jn} = \vartheta_j(e_n, \vec{P}_n)\left(w_{n,\ell} L_n + w_{n,h} H_n\right) \tag{28}$$

3.4 Market Equilibrium

The equilibrium determines, for each location $n \in C, R$: the nominal wage of low-skilled labor $w_{\ell,n}$, the skill premium $\hat{w}_n \equiv w_{n,h}/w_{n,\ell}$, the allocation of labor across sectors, and the masses of headquartered firms $|\mu_j^n|, |\mu_j^{n'}|$ for $j \in T, M$.

We normalize the price of the tradable good to one. With G produced under constant returns using only low-skilled labor, the zero-profit condition implies $w_{\ell,n} = A_{G,n}$. Hence, nominal low-skill wages are pinned down by local tradable-sector productivity.

The equilibrium of the economy is characterized by the free-entry condition in each location (equation (23)), the market-clearing condition for high-skilled labor, and market-clearing in modern and traditional service expenditures.[5]

In particular, the equilibrium condition for the market for high-skilled workers, used as overhead labor in establishments, satisfies:

$$H_n = \sum_{j \in T, M} \overline{f}_j \left(q_j |M_{jn}^n| \hat{w}_n^{q_j - 1} + \tau_j (q_j + \kappa) |M_{jn}^{n'}| \hat{w}_n^{q_j + \kappa - 1}\right) \tag{29}$$

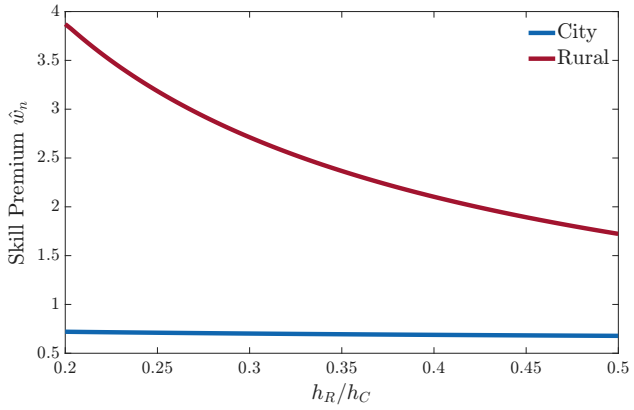
[5]The low-skilled labor market must also clear, but always does so because unit-elastic demand for goods production absorbs any low-skilled labor not used for service production.

3.5 Key Comparative Static: Skill Catch-Up in the Hinterland

Having specified the model, we now demonstrate its capacity to explain the effects of a skill expansion akin to the one enacted by Brazil's college reform. We model this skill expansion as an increase in the rural hinterland's high-skilled share, $h_R = \frac{H_R}{H_R+L_R}$, holding all other model primitives fixed. Specifically, we show how equilibrium model outcomes vary from progressively increasing the rural region's skilled share from 20% to 50% of the skilled share in the city. As proof of concept, we do this for a benchmark set of parameter values that we present in Table 12 in the Appendix. Future versions of the paper will use our empirical results above to discipline the parameters of the model.

The Local Wage Premium The key change underpinning all other comparative statics of the model is that skill expansion reduces the relative price of skills in the rural area $\hat{w}_R$. We demonstrate this change, which is consistent with the evidence from Cox (2025), in Figure 10. In our parameterization, the skill premium is modest in the city but starts from a dramatically high level in the rural hinterland. As skill becomes more abundant, the rural skill premium falls and converges towards that of the city.

Figure 10: Regional Skill Premia

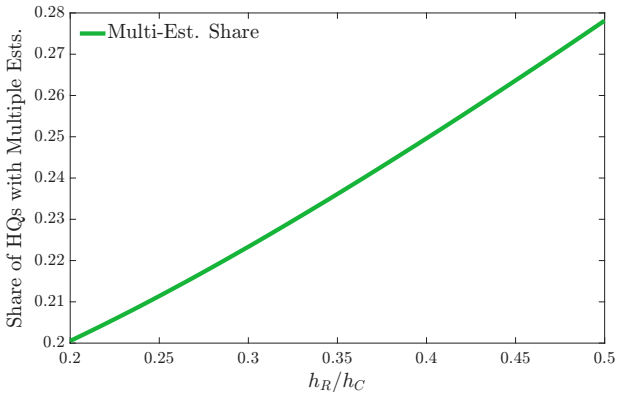


The Geographic Scope of Firms In response to the falling rural skill premium, firms increase their geographic scope. We display several relevant equilibrium objects in Figure 11, all of which show the growing importance of multi-establishment activity as the skill supply expands in rural areas. First, Figure 11a shows that a rising share of the aggregate economy's headquarters operate multiple production establishments, i.e. they operate in both the city and the rural hinterland. Second, 11b shows how the composition of city and rural establishments responds. A rising share of rural establishments are headquartered in the city, while the share of city establishments headquartered in rural areas experiences a

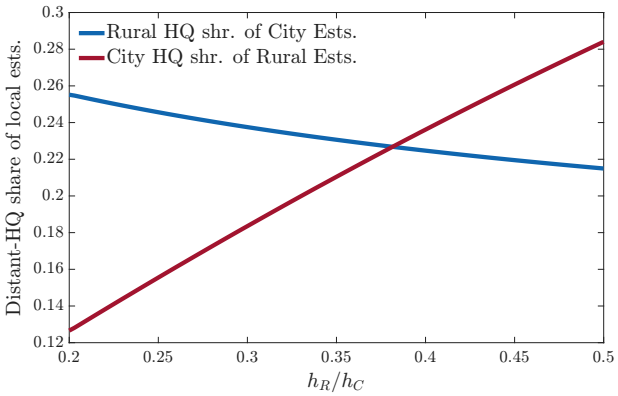
modest fall. A similar pattern holds for the share of sales: rural sales increasingly belong to firms headquartered in the city. Overall, the expansion of skill supply in rural areas allows city-headquartered firms to access new markets and claim a greater share of economic activity.

Figure 11: Multi-Establishment Activity

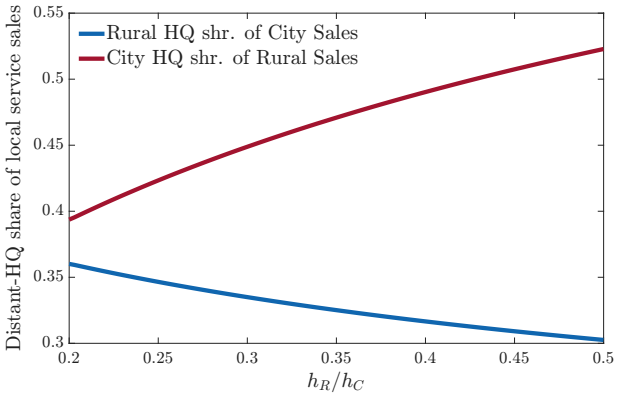
(a) Multi-Est. Share of Firms



(b) Distant HQ Share of Ests.



(c) Distant HQ Share of Service Sales

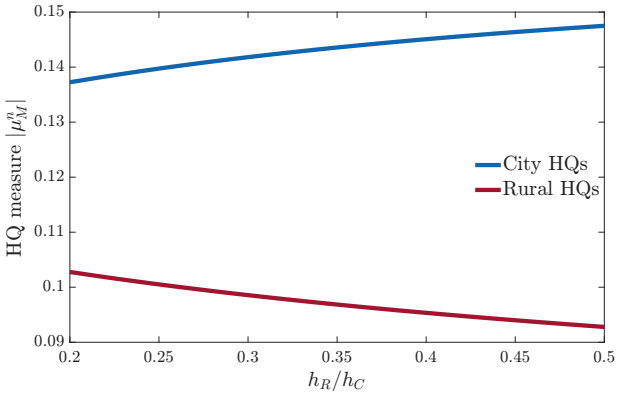


Innovative Activity As predicted by the model, the opportunity for city HQs to implement technology in distant markets increases their incentive for innovative activity. Figure 12 displays the relevant comparative statics. In response to the greater abundance of skill in rural areas, the city sees an increase in the flow of entering HQs. Simultaneously, firms that enter in the city make greater investments in their technology level x_M^C, because the probability that they can implement the technology in rural areas is greater when rural skill is abundant. However, the competitive pressure from innovative activity in the city actually reduces innovation in rural areas. The cutoff productivity for a rural-headquartered firm to operate locally increases in response to the inflow of city-headquartered firms, causing a decline in entry and investment in rural headquarters. This is a notable feature of the

model: convergence in the supply of skills across regions can lead to *divergence* in innovative activity, as more innovation concentrates in the city.

Figure 12: Headquarters Entry and Technology Investment

(a) HQ entry



(b) Technology Choice

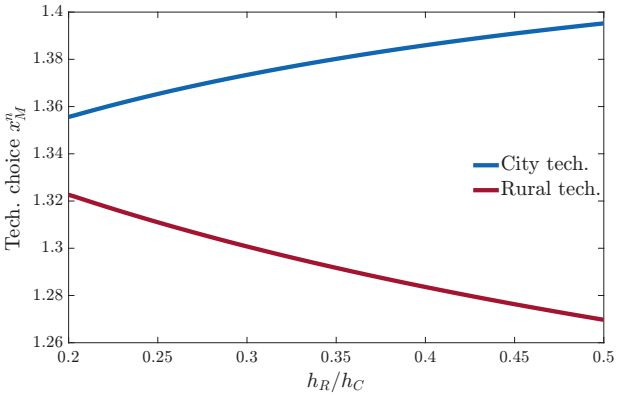
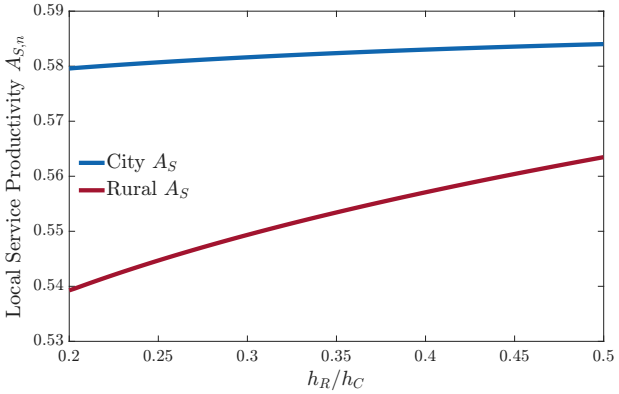
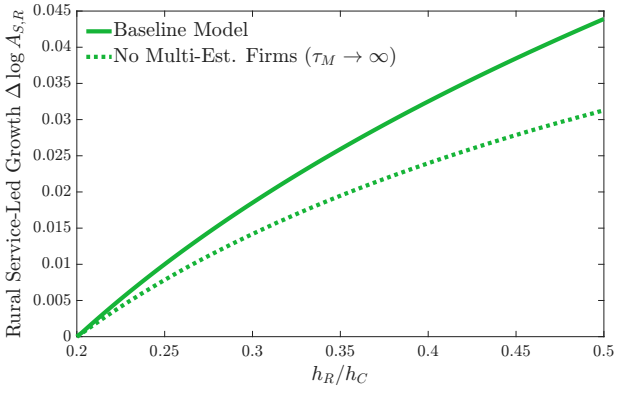


Figure 13: Service Productivity A_S

(a) By Region



(b) The Role of Multi-Est. Firms



Service-Led Growth Finally, in Figure 13 we display the impact of the rural skill expansion on service-led growth, as represented by service productivity $A_{S,n}$. Despite the divergence of innovative activity in the service sector, rural areas see convergence of service productivity (Figure 13a). The opportunity to open new service establishments implementing highly-productive technology makes the rural service sector more productive, even though no technological parameters of the model (such as Φ_j or $\overline{F}_j$) changed. To assess the contribution of multi-establishment firms to service-led growth, in Figure 13b we compare rural service productivity growth in our baseline model to an alternative model in which we impose $\tau_M \to \infty$, so that there are no multi-establishment firms. While services still grow in the model without multi-establishment firms – since high-skilled workers are still an

input to single-establishment production – growth is higher when multi-location production is possible.

4 Conclusion

This paper examines how the expansion of college-educated labor in Brazil enabled a new pathway of structural transformation, driven by the growth of multi-market firms, particularly in the service sector. Departing from the traditional industrialization-led narrative, we document a significant reallocation of employment toward higher-productivity, wage-based service jobs, coinciding with a rapid rise in the number and scale of multi-market firms following the 1996 higher education reform.

Our empirical findings highlight human capital formation through expanded access to tertiary education as a key driver of firm expansion and productivity gains. The reform removed longstanding regulatory barriers to private entry, disproportionately increasing college access in previously underserved regions. These same regions subsequently experienced stronger growth in multi-market firm activity, particularly among firms that hired skilled workers and offered wage premia relative to single-market counterparts. The associated reallocation of employment contributed meaningfully to local wage growth, underscoring the developmental role of firm-level organizational change.

We complement our empirical analysis with a theoretical framework that extends Hsieh and Rossi-Hansberg (2023) by introducing heterogeneous labor and distance-sensitive fixed costs of market entry. In this model, college-educated workers act as a critical managerial input, enabling firms to overcome the spatial and organizational frictions associated with operating across multiple markets. As the relative supply of skilled labor rises, firms are better positioned to expand geographically and disseminate modern production structures.

Together, these results challenge the view that services are inherently stagnant or low in productivity. When supported by investments in higher education, modern services – especially those delivered by scalable, multi-market firms – can serve as an engine of growth and structural transformation. This has important implications for development strategies in the 21st century, particularly for countries facing premature deindustrialization.

These findings open several avenues for future research. One promising direction is to explore how human capital interacts with emerging technologies—particularly Artificial Intelligence (AI)—in shaping the organization and productivity of multi-market firms in service-oriented economies. As AI increasingly automates tasks within firms, a critical question is whether it will complement or substitute for white-collar managerial occupations typically held by college-educated workers. Understanding the direction and magnitude of this inter-

action is essential for assessing how AI may reshape the demand for skilled labor and, by extension, the viability of human capital–driven service-led growth. Relatedly, future work should examine how the employment multipliers associated with human capital expansion evolve in the presence of AI and automation. More broadly, as firms adopt new technologies, how will this affect their ability to expand across markets and replicate complex service operations? Addressing these questions will be key to understanding the next phase of structural transformation in developing and advanced economies alike.

References

Acemoglu, D. and Restrepo, P., 2022. Tasks, Automation, and the Rise in U.S. Wage Inequality. *Econometrica* [Online], 90(5). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA19815, pp.1973–2016. Available from: https://doi.org/10.3982/ECTA19815 [Accessed November 7, 2024].

Baumol, W.J., 1967. Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *The american economic review* [Online], 57(3). Publisher: American Economic Association, pp.415–426. Available from: https://www.jstor.org/stable/1812111 [Accessed April 22, 2025].

Bloom, N., Lemos, R., Sadun, R., Scur, D., and Van Reenen, J., 2014. JEEA-FBBVA Lecture 2013: The New Empirical Economics of Management. *Journal of the european economic association* [Online], 12(4), pp.835–876. Available from: https://doi.org/10.1111/jeea.12094 [Accessed April 4, 2025].

Bloom, N. and Van Reenen, J., 2007. Measuring and Explaining Management Practices Across Firms and Countries*. *The quarterly journal of economics* [Online], 122(4), pp.1351–1408. Available from: https://doi.org/10.1162/qjec.2007.122.4.1351 [Accessed April 4, 2025].

Bloom, N. and Van Reenen, J., 2010. Why Do Management Practices Differ across Firms and Countries? *Journal of economic perspectives* [Online], 24(1), pp.203–224. Available from: https://doi.org/10.1257/jep.24.1.203 [Accessed January 21, 2025].

Buera, F.J. and Kaboski, J.P., 2012. The Rise of the Service Economy. *American economic review* [Online], 102(6), pp.2540–2569. Available from: https://doi.org/10.1257/aer.102.6.2540 [Accessed April 4, 2025].

Chandler, A., 1977. The visible hand: the managerial revolution in American business. *Cambridge, ma, harvard university press*, 51, pp.107–108.

Chen, X., Pei, G., Song, Z., and Zilibotti, F., 2023. Tertiarization Like China. *Annual review of economics* [Online], 15(1), pp.485–512. Available from: https://doi.org/10.1146/annurev-economics-071122-030026 [Accessed April 16, 2025].

Cordeiro, F. and Cox, A., 2025. *College quality and tuition subsidies in equilibrium.*

Cox, A., 2025. *Fostering Development Through Higher Education: College Attainment, Firms and Economic Growth* [Online]. Available from: https://www.dropbox.com/scl/fi/s6dj6m6pkcm31pukiqjh1/Alvaro-2024-Fostering-development-through-higher-education-submission.pdf?rlkey=32d5tsbbbeqyby5ffpluellvr&e=1&st=qtjtx45q&dl=0 [Accessed March 25, 2025].

Dix-Carneiro, R. and Kovak, B.K., 2017. Trade Liberalization and Regional Dynamics. *American economic review* [Online], 107(10), pp.2908–2946. Available from: https://doi.org/10.1257/aer.20161214 [Accessed March 20, 2025].

Eckert, F., Ganapati, S., and Walsh, C., 2022. *Urban-Biased Growth: A Macroeconomic Analysis* [Online]. Working Paper. Working Paper Series. National Bureau of Economic Research. Available from: https://doi.org/10.3386/w30515 [Accessed April 4, 2025].

Engbom, N., Malmberg, H., Porzio, T., Rossi, F., and Schoellman, T., 2024. Economic Development According to Chandler.

Fan, T., Peters, M., and Zilibotti, F., 2023. Growing Like India—the Unequal Effects of Service-Led Growth. *Econometrica* [Online], 91(4). _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.398 pp.1457–1494. Available from: https://doi.org/10.3982/ECTA20964 [Accessed March 31, 2025].

Garicano, L. and Rossi-Hansberg, E., 2015. Knowledge-Based Hierarchies: Using Organizations to Understand the Economy. *Annual review of economics* [Online], 7(1), pp.1–30. Available from: https://doi.org/10.1146/annurev-economics-080614-115748 [Accessed March 21, 2025].

Gumpert, A., Steimer, H., and Antoni, M., 2022. Firm Organization with Multiple Establishments*. *The quarterly journal of economics* [Online], 137(2), pp.1091–1138. Available from: https://doi.org/10.1093/qje/qjab049 [Accessed March 21, 2025].

Hjort, J. and Poulsen, J., 2019. The Arrival of Fast Internet and Employment in Africa. *American economic review* [Online], 109(3), pp.1032–1079. Available from: https://doi.org/10.1257/aer.20161385 [Accessed May 27, 2025].

Hjort, J. and Tian, L., 2025. The Economic Impact of Internet Connectivity in Developing Countries. *Annual review of economics* [Online]. Available from: https://doi.org/10.1146/annurev-economics-081224-102352 [Accessed May 27, 2025].

Hopenhayn, H.A., 1992. Entry, Exit, and firm Dynamics in Long Run Equilibrium. *Econometrica* [Online], 60(5). Publisher: [Wiley, Econometric Society], pp.1127–1150. Available from: https://doi.org/10.2307/2951541 [Accessed October 20, 2023].

Hsieh, C.-T. and Rossi-Hansberg, E., 2023. The Industrial Revolution in Services. *Journal of political economy macroeconomics* [Online], 1(1). Publisher: The University of Chicago Press, pp.3–42. Available from: https://doi.org/10.1086/723009 [Accessed November 8, 2024].

Jiang, X., 2023. *Information and Communication Technology and Firm Geographic Expansion* [Online]. SSRN Scholarly Paper. Rochester, NY: Social Science Research Network. Available from: https://doi.org/10.2139/ssrn.4466678 [Accessed March 21, 2025].

Kleinman, B., 2024. Wage inequality and the spatial expansion of firms.

Lucas, R.E., 1978. On the Size Distribution of Business Firms. *The bell journal of economics* [Online], 9(2). Publisher: [RAND Corporation, Wiley], pp.508–523. Available from: https://doi.org/10.2307/3003596 [Accessed April 4, 2025].

Mariscal, A., 2018. Firm Organization and Information Technology: Micro and Macro Implications. *2018 meeting papers* [Online]. Number: 1076 Publisher: Society for Economic Dynamics. Available from: https://ideas.repec.org//p/red/sed018/1076.html [Accessed March 21, 2025].

Melitz, M.J., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* [Online], 71(6). Publisher: Econometric Society, pp.1695–1725. Available from: https://ideas.repec.org/a/ecm/emetrp/v71y2003i6p1695-1725.html [Accessed November 3, 2022].

Olea, J.L.M. and Pflueger, C., 2013. A Robust Test for Weak Instruments. *Journal of business & economic statistics* [Online], 31(3). Publisher: [American Statistical Association, Taylor & Francis, Ltd.], pp.358–369. Available from: https://www.jstor.org/stable/43702731 [Accessed May 22, 2025].

Rodrik, D., 2015. *Premature Deindustrialization* [Online]. Working Paper. Working Paper Series. National Bureau of Economic Research. Available from: https://doi.org/10.3386/w20935 [Accessed April 25, 2023].

Rodrik, D. and Sandhu, R., 2024. *Servicing development: productive upgrading of labor-absorbing services in developing economies* [Online]. (technical report). National Bureau of Economic Research. Available from: https://www.nber.org/papers/w32738 [Accessed November 7, 2024].

Rodrik, D. and Stiglitz, J.E., 2024. A New Growth Strategy for Developing Nations.

Schwartzman, M., 2025. *From Street Markets to Shopping Malls: The Transformation of the Service Sector* [Online]. Available from: https://m-schwartzman.github.io/MatthewWebsite/Schwartzman_JMP_post.pdf [Accessed January 7, 2025].

A Appendix

A.0.1 Heterogeneity of wages gains from firm expansion across space

Table 10: Within-job-spell: Testing for wage gains led by firm expansion by type of worker

	Manager			Non-Manager		
	(1)	(2)	(3)	(4)	(5)	(6)
	Δ Log(Wage)	Δ Log(Wage)	Δ Log(Wage)	Δ Log(Wage)	Δ Log(Wage)	Δ Log(Wage)
Δ Multi-market	-0.137***	-0.137***	-0.0748***	-0.0472***	-0.0498***	-0.00950**
	(0.00878)	(0.00825)	(0.00848)	(0.00350)	(0.00327)	(0.00341)
Δ Log(N of Markets)	0.278***	0.276***	0.242***	0.123***	0.126***	0.104***
	(0.00691)	(0.00667)	(0.00670)	(0.00347)	(0.00330)	(0.00327)
Tenure	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Region FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Region-Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Initial log(Wage) × Tenure	No	Yes	Yes	No	Yes	Yes
Initial log(NumMarkets) × Tenure	No	No	Yes	No	No	Yes
Initial Multi-market × Tenure	No	No	Yes	No	No	Yes
Obs.	1,430,708	1,430,708	1,430,708	8,210,215	8,210,215	8,210,215
Adj.-R^2	0.19	0.25	0.25	0.16	0.21	0.21
Avg Wage	886.11	886.11	886.11	616.56	616.56	616.56
N. Workers	168,785	168,785	168,785	1,636,272	1,636,272	1,636,272

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

A.1 First Stage Analysis

The instrumental variable strategy exploits cross-regional variation in the scarcity of higher education capacity prior to the 1996 reform. The reform removed barriers to entry and authorized for-profit colleges to operate legally, thereby enabling a market-driven expansion of the higher education sector. Private institutions responded to these regulatory changes by expanding disproportionately into commuting zones (CZs) where centralized planning had previously limited the supply of higher education relative to potential demand. This expansion reduced geographic and capacity constraints in underserved areas, facilitating a subsequent rise in college attainment.

To evaluate this mechanism empirically, we estimate a first-stage specification in which the dependent variable, Δh_r, is either the change in college attainment among individuals aged 25–34 between 2000 and 2010 or the growth in higher education capacity between 1995 and 2005. The analysis restricts attention to capacity built before 2006, ensuring that changes in attainment by 2010 can plausibly reflect enrollment during the relevant post-reform window. The key explanatory variable is the pre-reform scarcity measure Z_r, defined as the ratio of higher education capacity in 1995 to the size of the 6–15-year-old population in

1991. Control variables X_r include pre-reform development trends in education and economic structure, measured between 1991 and 2000.

The identification strategy compares CZs with similar pre-reform characteristics but different values of Z_r. A lower value of Z_r implies greater relative scarcity of college access prior to the reform. Under the hypothesis that scarcity induced post-reform expansion, we expect a negative relationship between Z_r and both the change in higher education capacity and the rise in attainment—that is, $\hat{\alpha} < 0$.

Table 11: First-stage Analysis

Dep.Variable	Δ Colleges	Δ College attainment
Specification	(1)	(2)
Z_r	-260.08***	-7.57***
	(50.77)	(1.10)
R^2	486	486
F-test	46.93	46.93
Obs.	0.22	0.35

Standard errors in parentheses

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

Note.— The table reports first-stage regression coefficients where the dependent variable is either (1) the change in college attainment among individuals aged 25–34 between 2000 and 2010, or (2) the growth in higher education capacity between 1995 and 2005. Both outcomes are regressed on a pre-reform scarcity measure, defined as the ratio of higher education capacity in 1995 to the size of the 6–15 year-old population in 1991.

Table 11 reports the first-stage estimates. Column 1 shows that regions with lower Z_r experienced significantly greater increases in higher education capacity between 1995 and 2005. Column 2 presents results for the change in college attainment, which also exhibit a negative and statistically significant relationship with Z_r. All specifications include the full set of controls.

These findings support the relevance of the instrument: college attainment rose faster in CZs that were underserved prior to the reform, consistent with the idea that the 1996 policy change induced market entry in regions where capacity had lagged behind potential demand. The next section turns to the second-stage estimates and reports the robust weak-instrument F-statistic following Olea and Pflueger (2013).

B Theoretical Appendix

B.1 Local Aggregate Productivity and Cutoffs

Aggregating across the productivity draws of all operating firms in sector–location pair (j,n) yields

$$\int_{i'\in M_{jn}} A_{i'j}^{\sigma-1}di' = \frac{\gamma}{\gamma-1}\left(|\mu_j^n|\left(\underline{A}_{jn}^n\right)^{\sigma-1}\left(\frac{\left(\underline{A}_{jn}^n\right)^{\sigma-1}}{x_j^n}\right)^{-\gamma} + |\mu_j^{n'}|\left(\underline{A}_{jn}^{n'}\right)^{\sigma-1}\left(\frac{\left(\underline{A}_{jn}^{n'}\right)^{\sigma-1}}{x_j^{n'}}\right)^{-\gamma}\right) \tag{30}$$

Combining this aggregation with the selection conditions in (13) and the relationship $\left(\frac{\underline{A}_{jn}^{n'}}{\underline{A}_{jn}^{n}}\right)^{\sigma-1} = \tau_j\hat{w}_n^{\kappa}$ yields closed-form expressions for the entry thresholds $\underline{A}_{jn}^n, \underline{A}_{jn}^{n'}$:

$$\begin{aligned}\left(\underline{A}_{jn}^n\right)^{\gamma(\sigma-1)} &= \frac{\gamma}{\gamma-1}\sigma\overline{f}_j\frac{w_{n,\ell}}{E_{n,j}}\hat{w}_n^{q_j}\left(|\mu_j^n|\left(x_j^n\right)^{\gamma} + |\mu_j^{n'}|\left(x_j^{n'}\right)^{\gamma}\hat{w}_n^{\kappa(1-\gamma)}\tau_j^{1-\gamma}\right)\\ \left(\underline{A}_{jn}^{n'}\right)^{\gamma(\sigma-1)} &= \frac{\gamma}{\gamma-1}\sigma\overline{f}_j\frac{w_{n,\ell}}{E_{n,j}}\hat{w}_n^{q_j}\left(|\mu_j^n|\left(x_j^n\right)^{\gamma}\hat{w}_n^{\kappa\gamma}\tau_j^{\gamma} + |\mu_j^{n'}|\left(x_j^{n'}\right)^{\gamma}\hat{w}_n^{\kappa}\tau_j\right)\end{aligned} \tag{31}$$

The measure of firms headquartered in each location that operate establishments in n can then be written as $|M_{jn}^n| = |\mu_j^n| \times F\left(\left(\underline{A}_{jn}^n\right)^{\sigma-1}\right)$ and $|M_{jn}^{n'}| = |\mu_j^{n'}| \times F\left(\left(\underline{A}_{jn}^{n'}\right)^{\sigma-1}\right)$. Using the expression for the Pareto CDF and equation (31) yields:

$$|M_{jn}^n| = \frac{\frac{\gamma-1}{\gamma}\frac{1}{\sigma\overline{f}_j}\frac{E_{n,j}}{w_{n,\ell}}}{\hat{w}_n^{q_j} + \left(\frac{x_j^{n'}}{x_j^n}\right)^{\gamma}\frac{|\mu_j^{n'}|}{|\mu_j^n|}\hat{w}_n^{q_j+\kappa(1-\gamma)}\tau_j^{1-\gamma}} \qquad |M_{jn}^{n'}| = \frac{\frac{\gamma-1}{\gamma}\frac{1}{\sigma\overline{f}_j}\frac{E_{n,j}}{w_{n,\ell}}}{\hat{w}_n^{q_j+\kappa}\tau_j + \left(\frac{x_j^{n}}{x_j^{n'}}\right)^{\gamma}\frac{|\mu_j^{n}|}{|\mu_j^{n'}|}\hat{w}_n^{q_j+\kappa\gamma}\tau_j^{\gamma}} \tag{32}$$

B.2 Expected Profits as a Function of Cutoffs

$$\begin{aligned}\mathbb{E}\left[\tilde{\Pi}_j^n(A_{ij})|x\right] =& x^{\gamma}\Bigg[\left(\underline{A}_{jn}^n\right)^{-\gamma(\sigma-1)}w_{n,\ell}\left(\frac{1}{\sigma}\frac{E_{jn}}{w_{n,\ell}}\frac{\gamma}{\gamma-1}\frac{\left(\underline{A}_{jn}^n\right)^{\sigma-1}}{\int_{i'\in M_{jn}}A_{i'j}^{\sigma-1}di'} - \overline{f}_j\hat{w}_n^{q_j}\right)\\ &+ \left(\underline{A}_{jn'}^n\right)^{-\gamma(\sigma-1)}w_{n',\ell}\left(\frac{1}{\sigma}\frac{E_{jn'}}{w_{n',\ell}}\frac{\gamma}{\gamma-1}\frac{\left(\underline{A}_{jn'}^n\right)^{\sigma-1}}{\int_{i'\in M_{jn'}}A_{i'j}^{\sigma-1}di'} - \tau_j\overline{f}_j\hat{w}_{n'}^{q_j+\kappa}\right)\Bigg]\\ =& \Omega_{jn}x^{\gamma},\end{aligned} \tag{33}$$

so that it is immediate that expected profits decline in the cutoffs $\underline{A}_{jn}^n$ and $\underline{A}_{jn'}^n$.

B.3 Expanded Free Entry Condition

Substituting the formulas for $\left(\underline{A}_{jn}^n\right)^{-\gamma(\sigma-1)}$ and $\left(\underline{A}_{jn'}^n\right)^{-\gamma(\sigma-1)}$, yields equivalent free-entry expressions in terms of headquarters masses, expenditures, and wages:

$$\begin{aligned}\Omega_{jn} &= \frac{1}{\gamma\hat{\sigma}}\left[\frac{E_{jn}}{|\mu_j^n|\left(x_j^n\right)^{\gamma}+|\mu_j^{n'}|\left(x_j^{n'}\right)^{\gamma}\tau_j^{1-\gamma}\hat{w}_n^{\kappa(1-\gamma)}}+\frac{E_{jn'}}{|\mu_j^n|\left(x_j^n\right)^{\gamma}+|\mu_j^{n'}|\left(x_j^{n'}\right)^{\gamma}\tau_j^{\gamma-1}\hat{w}_{n'}^{\kappa(\gamma-1)}}\right] \\ &= \frac{1}{\gamma\hat{\sigma}}\left(\frac{\eta-\gamma}{\gamma\eta}\frac{\overline{F}_j}{\Phi_j}\right)^{\frac{\gamma}{\eta-\varphi}}\left[\frac{E_{jn}}{|\mu_j^n|^{1+\frac{\chi\gamma}{\eta-\varphi}}+|\mu_j^{n'}|^{1+\frac{\chi\gamma}{\eta-\varphi}}\tau_j^{1-\gamma}\hat{w}_n^{\kappa(1-\gamma)}}+\frac{E_{jn'}}{|\mu_j^n|^{1+\frac{\chi\gamma}{\eta-\varphi}}+|\mu_j^{n'}|^{1+\frac{\chi\gamma}{\eta-\varphi}}\tau_j^{\gamma-1}\hat{w}_{n'}^{\kappa(\gamma-1)}}\right] \\ &= \left(\frac{\eta}{\eta-\gamma}\Phi_j|\mu_j^n|^{\chi}\right)^{\frac{\eta-\varphi-\gamma}{\eta-\varphi}}\left(\frac{\overline{F}_j}{\gamma}\right)^{\frac{\gamma}{\eta-\varphi}}\end{aligned} \tag{34}$$

Hence,

$$\begin{aligned}\implies \quad & \frac{1}{\sigma}\left[\frac{E_{jn}}{|\mu_j^n|^{1+\frac{\chi\gamma}{\eta-\varphi}}+|\mu_j^{n'}|^{1+\frac{\chi\gamma}{\eta-\varphi}}\tau_j^{1-\gamma}\hat{w}_n^{\kappa(1-\gamma)}}+\frac{E_{jn'}}{|\mu_j^n|^{1+\frac{\chi\gamma}{\eta-\varphi}}+|\mu_j^{n'}|^{1+\frac{\chi\gamma}{\eta-\varphi}}\tau_j^{\gamma-1}\hat{w}_{n'}^{\kappa(\gamma-1)}}\right] \\ =& \frac{\eta\gamma}{\eta-\gamma}\Phi_j|\mu_j^n|^{\chi\frac{\eta-\varphi-\gamma}{\eta-\varphi}}\end{aligned} \tag{35}$$

We have two such conditions, one for each $n \in C, R$. Given location-level expenditures (E_{jC}, E_{jR}) and wages $\vec{w}$, the system (35) for $n = C$, (35) for $n = R$ determines $\left(|\mu_j^C|, |\mu_j^R|\right)$. Closing the model then requires expenditures consistent with the demand system and wages that clear factor markets.

B.4 Benchmark Parameter Values

Table 12: Parameter Values for Model Simulation

Parameter	Description	Value
$\overline{f}_M$	Modern est. fixed cost	0.4
$\overline{f}_T$	Trad. est. fixed cost	0.1
τ_M	Additional "distance cost" in modern	2.0
q_M	High-skilled intensity of modern overhead	0.8
q_T	High-skilled intensity of traditional overhead	0.2
κ	Additional high-skilled intensity at distance	0.2
σ	EoS among varieties	4.5
γ	Pareto shape parameter	1.5
χ	HQ congestion	1.0
$\Phi_{M,C}$	Modern entry cost in city	1.0
$\Phi_{M,R}$	Modern entry cost in rural	1.15
Φ_T	Traditional entry cost	0.5
$\overline{F}_M$	Modern tech. cost shifter	1.0
$\overline{F}_T$	Traditional tech. cost shifter	10.0
η	Tech. cost convexity	1.8
β	Service Cobb-Douglas weight	0.8
ω	Modern weight in service composite	0.5
ξ	EoS between modern/trad. services	4.0
$A_{G,C}$	City goods productivity	1.0
$A_{G,R}$	Rural goods productivity	1.0
$H_C + L_C$	City population	1.0
$H_R + L_R$	Rural population	1.0
h_C	City high-skilled share	0.05