

Connectivity and Misallocation in Multi-Unit Firms^{*}

Jasmina Chauvin¹, Juan Pablo Chauvin², and Giovanna Chaves³

¹*Georgetown University McDonough School of Business*

²*Inter-American Development Bank, Research Department*

³*New York University, Department of Economics*

December, 2025

Abstract

We study the causal effect of reducing distance-related frictions with headquarters on branch outcomes in multi-unit firms in Brazil. Exploiting the introduction of new airline routes, we compare branches that gain a direct flight to their headquarters with otherwise similar branches in the same location that do not. In contrast to prior findings from high-income country settings, improved connectivity with headquarters lowers 12-month survival and, conditional on survival, reduces service and production employment. To interpret these results, we develop a model with three distance-dependent frictions: (i) internal coordination costs of delivering headquarters inputs to branches; (ii) client-service costs when serving markets without a local branch; and (iii) a moral-hazard friction whereby distance amplifies local managers' incentives to over-hire. Consistent with the model's predictions, survival declines are larger at closer branches, while employment reductions are concentrated at more distant branches and in firms with lower headquarters bandwidth. The model clarifies how the impact of reduced frictions with headquarters depends on pre-existing distortions and reconciles heterogeneous connectivity effects observed across developing- and high-income contexts.

JEL Codes: D21, D22, D24

^{*}We are grateful to Camila Gomes, Juliana Pinillos, Rafael Macedo Rubiao, and Haydée Svab for outstanding research assistance. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent. All errors are our own.

1 Introduction

Research on the spatial organization of multi-unit firms documents a consistent finding: despite the widespread adoption of low-cost information and communication technologies, geographic distance from firm headquarters ("HQ") negatively affects the likelihood of branch entry and branch performance conditional on entry. Multiple empirical studies that exploit quasi-experimental reductions in travel time—due to the introduction of new airline or high-speed rail routes—find that greater proximity to headquarters increases investment and employment growth in the treated branches, including in the United States ([Giroud, 2013](#); [Giroud and Mueller, 2015](#)), France ([Charnoz et al., 2018](#)), and Germany ([Gumpert et al., 2022](#)). A leading explanation for these positive effects is that proximity to headquarters facilitates monitoring and increases managerial attention devoted to the focal branch. This, in turn, enables headquarters to identify high-performing branches and allocate capital and labor toward them.

Although improved presence of headquarters' operatives at the branch enables the firm to identify opportunities and increases productivity, it can also reveal resource misallocation and inefficiencies. Resource misallocation is significantly more pronounced in developing economies such as China and India than in developed economies, and is reflected in a larger tail of low-productivity firms ([Hsieh and Klenow, 2009](#)). Misallocation is evident in not only substantial performance differences across firms, but also in large disparities among different branches *within* firms (e.g., [Adhvaryu et al., 2023](#)). Developing economies have also been shown to exhibit worse management practices and to have a lower supply of professional managerial talent ([Bloom et al., 2012](#)). These empirical patterns raise the question: How do reductions in frictions between headquarters and local branches affect performance in contexts where inefficiency is prevalent? In this paper, we examine this question in the context of Brazil, an economy in which multi-unit firms play a similarly large role as they do in developed economies, but where geographic frictions are severe and inefficiency is pervasive (e.g., [Busso et al., 2013](#); [Morten and Oliveira, 2024](#)).¹

We begin by examining whether key empirical regularities documented in developed-country contexts relating distance to headquarters to branch outcomes also hold in the Brazilian context. The data for the analysis comes from the *Relação Anual de Informações Sociais* ("RAIS"), which is a Brazilian administrative dataset that allows us to observe the universe of formal-sector firms and all of their establishments. We document several

¹Table 1 shows that multi-unit firms account for more than 30 percent of employment in Brazil.

cross-sectional patterns. One, distance to headquarters negatively correlates with the probability that a firm opens a branch, which is consistent with prior findings in the U.S. and Germany (e.g., [Giroud, 2013](#); [Gumpert et al., 2022](#)). Second, conditional on operating a branch, the branch's distance to headquarters correlates negatively with branch survival and branch size, patterns also documented by [Kalnins and Lafontaine \(2013\)](#) in the U.S. lodging industry. Third, distance to headquarters correlates positively with the presence of professionals and managers at the local branch. This finding is consistent with [Charnoz et al. \(2018\)](#), who find that the number of skilled managers allocated to support activities decreases at remote affiliates as these become better connected to headquarters in French corporate groups. Overall, we find that the cross-sectional patterns exhibited by Brazilian multi-unit firms closely resemble those documented in high-income economies.

Next, we investigate the causal effects of proximity to headquarters by examining how branch outcomes change when connectivity to headquarters increases. Our identification strategy follows prior studies and leverages the introduction of new direct air connections between the city of a firm's headquarters and a focal branch. During our study period (2000–2017), Brazil experienced rapid growth in air travel, with the number of passenger trips made by air roughly tripling (e.g., [Oliveira et al., 2020](#)). This growth reflects not only increased utilization on existing routes but also the geographic expansion of the route network, largely driven by the entry of new low-cost carriers who frequently expanded into new, previously under-served areas.

The sample for the reduced-form analysis includes all formal sector multi-unit firms that were active at the start of our sample period (year 2000), whose headquarters were in a location served by an airport, which operated at least one other establishment in addition to the headquarters in a different commuting zone which did not already have a direct air connection to headquarters, and employed 10 or more employees in at least one year. These criteria yield 3,938 multi-unit firms and their 10,288 establishments. Because the introduction of routes ("treatment") is staggered, we estimate stacked event-study models (e.g., [Cengiz et al., 2019](#); [Butters et al., 2022](#)), where each stack consists of all multi-unit firm establishments located in the same city and year. Our empirical specification compares outcomes for establishments that received a direct connection to the city of their headquarters due to a new route introduction ("treated") to those of establishments in the same city for whom the new route did not link to their headquarter's location ("control"), before and after the introduction of the new route.

The causal analysis unveils two surprising findings. First, we find that a direct connec-

tion to headquarters *decreases* the 1-year establishment survival rate by 12.6 percentage points (over a mean baseline rate of 95 percent). Conditional on survival, gaining a direct connection to headquarters *decreases* employment of production workers by 7.5 percent, on average. We also detect a negative (but not statistically significant) effect on the earnings of production workers. Yearly event-study plots show that these effects are not temporary; rather, treated establishments remain smaller than control establishments for several years out. We observe no differences in trends between the two in the years preceding treatment, supporting the validity of our identification strategy.

These results hold across a number of robustness tests. First, we test that they do not reflect changes in sample composition or a particular weighting of heterogeneous treatment effects over different time periods by considering treatment effects only for a balanced set of branches that are observed for three years pre- and five years post treatment. Second, we assess robustness to alternative definitions of "treatment", including any reduction in air travel time with headquarters and travel time reductions above one and a half hours. We find production worker employment effects of a similar magnitude, and potentially even larger decreases in survival rates. Finally, we restrict our sample to multi-unit firms with only one branch other than their headquarters, to ensure that the results do not reflect adjustments stemming from changes at a firm's other branches. The results are also robust to this sample restriction, although estimates are less precisely estimated due to smaller sample sizes.

Motivated by these findings, we develop a model that rationalizes heterogeneous effects of increased proximity to headquarters on branch survival and employment. In the model, a firm chooses whether to enter a distant market, whether to serve it directly from headquarters or to operate a local branch, and, conditional on operating a branch, how many workers to employ. Production combines headquarter inputs (e.g., mentoring, troubleshooting, supervision) with local labor. The model embeds three distance-dependent frictions. First, when the firm serves the market from headquarters, client-facing activities become harder at longer distances, captured by a client-service cost that rises with distance (in the spirit of [Oldenski, 2012](#)). Second, when the firm operates a branch, delivering headquarter inputs on site entails internal coordination costs that also increase with distance and with the scale of those inputs. Finally, echoing the logic of ([Aghion and Tirole, 1997](#)), delegating hiring to a branch manager creates an agency friction: managers may over-hire because they enjoy private benefits from headcount, and misallocation is harder to discipline when headquarters is far away.

Taken together, these frictions imply cross-sectional patterns that map directly in the stylized facts in the data. Because both client–service and coordination costs rise with distance, overall profitability declines and there is a finite entry cutoff beyond which no positive profits can be obtained, making firms less likely to operate in a market the farther it is from headquarters. As a result, neither option is profitable beyond a certain distance. Furthermore, among the markets where the firm operates a branch, equilibrium branch employment is smaller in branches located farther from headquarters. These branches receive fewer headquarter inputs, which lowers the marginal productivity of labor and discourages hiring.

The model also predicts that improvements in connectivity between headquarters and a branch—such as those induced by new air connections—can lead to branch closures, particularly when such improvements allow the firm to serve the market profitably from headquarters (i.e., when the market is relatively close). In addition, the model provides a rationale for why improved connectivity with headquarters can lead to either growth or decline in branch employment. Specifically, the provision of headquarter inputs—facilitated by greater proximity—generates two opposing effects on employment. On the one hand, it boosts labor productivity, encouraging additional hiring. This mechanism is consistent with findings in high-income countries. On the other hand, it strengthens oversight and raises the likelihood of detecting unprofitable overhiring, leading to employment reductions. Which margin dominates depends both on the baseline distance and on the extent of ex-ante misallocation. This second mechanism is novel and helps reconcile our results with findings from high-income country contexts.

Informed by these insights, we return to the data and explore heterogeneity to test the model’s mechanisms more directly. We first split branches by their baseline distance to headquarters. Consistent with the model, declines in branch survival rates are larger for branches located closer to HQ, suggesting that improved connectivity makes it more attractive to serve those markets directly from headquarters. In contrast, reductions in employment are larger at more distant branches, particularly for service and production workers, in line with the idea that remote branches are more prone to overhiring due to weaker monitoring. We also examine whether effects vary with the degree to which headquarters are time-constrained, using the number of HQ managers per branch as a proxy for HQ bandwidth. Results show that both the negative survival and employment effects are larger among branches with lower HQ bandwidth, i.e., for which monitoring is more limited *ex ante* and allocative distortions are likely to be larger.

This paper contributes to the literature on how headquarter-branch distance frictions shape the organization and performance of multi-unit firms. Existing work exploits quasi-experimental travel-time shocks—notably new airline routes and high-speed rail—to show that connectivity and proximity shape plant outcomes in multi-unit firms. In the U.S., new airline routes raise investment and TFP and the effects extend to plant openings and closures (Giroud, 2013). In plant-level panels, headquarters access reallocates employment—rising at treated plants while falling at others, leaving firm-level employment roughly unchanged (Giroud and Mueller, 2015). In Germany, reductions in managerial travel frictions are associated with higher service/production employment at connected establishments (Gumpert et al., 2022), while French high-speed rail access re-centralizes support jobs to headquarters (Charnoz et al., 2018). Cross-sectional evidence similarly links greater headquarters distance to lower plant survival (Kalnins and Lafontaine, 2013). We extend this literature by providing causal evidence from a middle-income economy. Exploiting new direct headquarters–branch flights in Brazil, we find—contrary to most prior findings—that improved connectivity reduces survival and lowers non-managerial employment among survivors. We reconcile these results with earlier positive estimates by proposing a unifying framework in which the impact of increased connectivity depends on the ex-ante efficiency of internal resource allocation.

Our paper also contributes to the misallocation literature. Seminal work by Hsieh and Klenow (2009) showed that reallocation could deliver large TFP gains in developing economies. Several factors have been highlighted as contributing to misallocation, including policy distortions—such as size-dependent regulations (Bento and Restuccia, 2017) and inflexible labor legislation (Busso et al., 2013)—as well as financial frictions such as credit constraints and distorted capital allocation (Gopinath et al., 2017; Midrigan and Xu, 2014). A growing body of evidence also attributes part of the observed productivity dispersion to managerial quality and attention operating within firms (Adhvaryu et al., 2023; Bloom et al., 2012). More recently, spatial frictions—such as housing constraints (Hsieh and Moretti, 2019) and limited road connectivity (Morten and Oliveira, 2024)—have been linked to misallocation through their effect on higher internal migration costs, which hamper the efficient allocation of labor across locations (Bryan and Morten, 2019). We connect these managerial and spatial channels by showing that spatial frictions can also affect the delivery of productive managerial inputs and oversight to local branches, contributing to within-firm misallocation.

Finally, our paper adds to the literature on the economic effects of air connectivity. A

large body of work finds that air service catalyzes city growth—expanding employment, population, and incomes—especially in nontradable and service sectors (Brueckner, 2003; Sheard, 2014; Blonigen and Cristea, 2015; McGraw, 2020; Zhang and Graham, 2020), and that global air-link networks foster economic activity across distant places (Campante and Yanagizawa-Drott, 2018). Work in the U.S. context focusing on multi-unit firms (Giroud, 2013; Giroud and Mueller, 2015) provides micro-evidence of a mechanism behind these aggregate effects: air links enable the reallocation of resources from headquarters to distant affiliates. Our findings highlight that this mechanism can operate in the opposite direction in a middle-income economy, where resources are often misallocated. In such settings, greater connectivity can reduce over-hiring or allow firms to operate without inefficient branches. These adjustments may raise firm productivity but come at the cost of local employment and establishment survival. This contrast suggests that in developing economies the local effects of air connectivity may be more heterogeneous than previously recognized, opening avenues for future research.

The rest of this paper is organized as follows: Section 2 describes the data and institutional context, detailing our multi-unit firm sample and the changes in the institutional environment and air connectivity in Brazil during 2000–2017. Section 3 documents stylized facts relating distance to headquarters to location probability and branch outcomes. Section 4 introduces our empirical strategy, empirical model, and the main results, along with a series of robustness tests. Section 5 presents the model that examines how changes in proximity to headquarters affect branch outcomes. Section 6 presents additional empirical tests informed by the model. Section 7 concludes.

2 Data and Empirical Context

2.1 Data and Multi-Unit Firm Sample

Our main firm data comes from Brazil’s matched employer-employee dataset, *RAIS*.² Our RAIS sample is an unbalanced panel that covers all firms in the formal economy between 2000 and 2017. It contains information on the active establishments of each firm, including their location and industry. We also observe all employees in each establishment, as well as their occupational code, wages, and individual characteristics such as gender, age, and

²RAIS has been extensively used in prior research on Brazilian firm and worker outcomes. See, for example, Menezes-Filho et al. (2008), Helpman et al. (2010), Poole (2013), and Britto et al. (2022).

educational attainment.

We define as multi-unit those firms that have at least one active establishment other than their headquarters.³ Columns (1)–(2) of Table 1 show the contribution of multi-unit firms to the Brazilian labor market. From the universe of approximately 6.8 million firms operating between 2000 and 2017, less than 5 percent are multi-unit, but these firms account for 15 percent of all establishments and one third of employment in the country.

Table 1: Contribution of Multi-Unit Firms in Brazil (2000–2017)

	All Multi-Unit Firms		Restricted Sample	
	(1)	(2)	(3)	(4)
	Count	% of Total	Count	% of Total
Firms	322,082	4.72	29,762	0.44
Establishments	1,108,283	14.54	227,191	2.98
Avg. Annual Employment	12,278,641	33.08	5,728,943	15.43

Notes: Authors’ calculations using data from 2000–2017 (unbalanced panel). Columns 1 and 3 show the counts of unique firms and establishments, and average number of total employees in a year for all multi-unit firms and for our restricted sample, respectively. Columns 2 and 4 show each group’s share (in %) of the total number of firms, establishments, or employees in Brazil over the same period.

Our empirical analyses using changes in air connectivity restrict the broader multi-unit sample to business entities that (i) were active in December 2000; (ii) operated at least one establishment in a different urban commuting zone (*arranjo populacional*, henceforth “*arranjo*”) from their headquarters; (iii) employed 10 or more workers in at least one year; and (iv) had their headquarters located in an *arranjo* served by an airport.⁴ These sample selection criteria ensure we focus on multi-location firms of a sufficient size, whose headquarters’ locations are plausibly unaffected by changes in air connectivity and whose evolution we can track over the subsequent years. The contribution of the final set of these

³We identify firms, establishments, and headquarters using the national registration number, the Cadastro Nacional de Pessoas Jurídicas (CNPJ). Each establishment is identified in RAIS with a unique 14-digit CNPJ; the first eight digits identify the firm and the subsequent six digits the establishment (e.g., Muendler et al., 2012). For headquarters, the last six digits start with “0001”. For more detail, see Data Appendix A.

⁴Headquarters location decisions may be influenced by the availability of non-stop flights (e.g., Bel and Fageda, 2008). We condition on firms that already existed in the year 2000 and do not include brand new entrants to reduce the possibility that the location of the headquarters in our sample is affected by the newly introduced routes that we study.

firms ("restricted sample") to establishments and employment in the country is displayed in columns (3)–(4) of Table 1.

We use each worker's occupational code (*Classificação Brasileira de Ocupações*, CBO 2002) to define their role within the firm. Drawing on definitions used in Gumpert et al. (2022), we categorize occupations into either professionals or service and production employees. Professionals are predominantly white-collar workers in management or specialized roles, such as company directors, engineers, doctors, technicians, middle managers and administrative supervisors. On the other hand, service and production workers are typically employed in occupations related to service and sales, the production of industrial goods and services, maintenance and repair, and agriculture.⁵ Following the literature, we restrict our employee sample to those active in December of each year.

2.2 Empirical Context: Brazil's Airline Route Expansions, 2000–2017

Our causal analysis exploits variation in connectivity to headquarters induced by the introduction of direct air routes between city pairs. This section describes the sources of air connectivity data and presents statistics on changes in connectivity across Brazilian cities between 2000 and 2017.

Airline Routes Data and Definitions

Our primary source of information on air connectivity is flight data provided by Brazil's National Civil Aviation Agency (ANAC).⁶ The agency compiles detailed monthly records of all flights operated within Brazilian territory. For each flight, the data include origin and destination airports, departure and arrival dates, airline, type of service (passenger or cargo), and the number of tickets sold, among other variables. We aggregate the number of regular passenger flights between each Brazilian origin-destination municipality pair by month for the period 2000–2017.⁷

⁵Professionals are classified based on having the first digit of their CBO 2002: 1, 2, 3 or 4. Service and production workers are classified based on having the first digit of their CBO 2002: 5, 6, 7, 8 or 9. For more detail, see Data Appendix A.

⁶Downloaded from: <https://www.gov.br/anac/pt-br/assuntos/regulados/empresas-aereas/envio-de-informacoes/microdados/microdados>.

⁷According to ANAC, "regular" flights meet the following criteria: (i) they transport passengers or cargo; (ii) they are available for public purchase; and (iii) they operate either on a published schedule or with a frequency that constitutes a systematic series. Regular flights exclude chartered flights, maintenance operations, employee transport, and extra flights added under exceptional circumstances. See

Key to our analysis is the definition of a "new route," i.e., a direct air connection between a pair of cities that did not previously exist. We define new routes in two steps. First, we classify a route as "active" in a given month if: (i) it has at least one flight during that month; and (ii) looking ahead over the subsequent 11 months, it also operates in at least 5 of those months. That is, an "active" route must provide passenger service in at least 6 months within a 12-month window. This ensures that we capture sustained (rather than one-off) connections. Second, we define a route as "new" if it is an active route that appears for the first time since the year 2000.

Changes in Air Connectivity Over the Period of Study

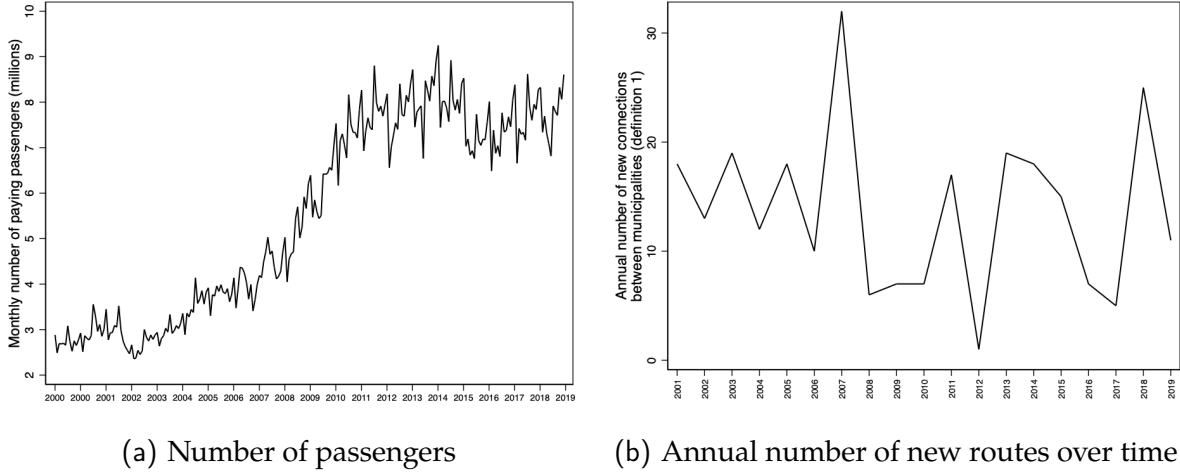
Figure 1 shows that Brazil experienced rapid growth in air travel beginning in the mid-2000s. The number of monthly air passenger trips rose from around 3 million in 2000 to a peak of approximately 9 million in 2014—a threefold increase—before stabilizing at a level between 7 and 9 million per month. According to [McKinsey&Company \(2010\)](#), 60% of domestic trips are for business purposes, and the majority of these (80%) are taken by frequent flyers—individuals who travel three or more times per year.

This rapid growth of air travel in Brazil during the 2000s occurred alongside a significant geographic expansion of the domestic air route network. A key driver of this expansion was the liberalization of what had previously been a highly regulated industry. This, coupled with favorable demand trends, spurred the entry and growth of new low-cost carriers, particularly GOL and Azul, which entered the market in 2001 and 2008, respectively. These airlines prioritized routes serving lower-cost, second- and third-tier cities, including mid-sized cities and regional capitals ([Oliveira et al., 2020](#)).

Panel B of Figure 1 displays the annual number of new routes introduced between Brazilian municipalities from 2000 to 2017. The year with the largest number of new route introductions was 2007.

<https://www.anac.pt/vpt/generico/regeconomica/autorizacoesdevoos/Paginas/AutorizacoesdeVoos.aspx>. For a small number of observations with missing information on service type, we classify flights with a positive number of purchased tickets as "passenger service," and those with zero passenger tickets but nonzero cargo volume as "cargo service."

Figure 1: Air passenger volumes and new route introductions over time



Notes: Authors' calculations using data from ANAC. Domestic passenger flights only. A new route is defined as the existence of an active air connection between a city pair (undirected) that appears for the first time after the year 2000.

3 Distance to Headquarters and Branch Outcomes

Table 2 presents cross-sectional correlations between distance to headquarters and establishment outcomes. The sample for the analyses are all multi-unit firms in Brazil at the beginning of our sample period (year 2000).

Columns (1)–(2) model the probability that a multi-unit firm operates an establishment in a location (*arranjo*) as a function of the location's distance to firm headquarters. The empirical model is estimated with industry and HQ city fixed effects, thus controlling for different patterns of geographic dispersion of firms across industries and the overall concentration of economic activity and of headquarters in Brazil (e.g., in the Center-East). The estimated coefficient in Column (1) implies that, controlling for industry and headquarter city, doubling a location's distance to headquarters is associated with an approximately 1.7 percentage points lower probability of a firm operating an establishment there. This implies that a firm with headquarters in the city of São Paulo, for instance, would be almost 4 percentage points more likely to operate a branch in São José dos Campos, 96 km away, than one in Brasília, 880 km away from headquarters. The estimated coefficient is little changed with the introduction of firm fixed effects in Column (2). Overall, this initial result is similar to that of [Gumpert et al. \(2022\)](#), who also find that the further a location is

to a firm’s headquarters, the less likely the firm is to operate an establishment there.

Columns (3)–(4) examine establishment survival conditional on operating. After controlling for firm fixed effects, the results indicate a negative correlation between distance and the establishment still operating at the end of our sample period (2017). The estimated coefficient in Column (4) implies that doubling the distance to headquarters is associated with a decrease of 0.3 percentage points in survival likelihood. Besides being less likely to survive, Column (6) shows that establishments operating farther away from headquarters are also smaller. Both patterns—a decline in within-firm survival and branch size with increasing distance to headquarters—mirror results documented by prior studies, which emphasize the benefits of more efficient investment and monitoring closer to headquarters (e.g., [Giroud, 2013](#); [Giroud and Mueller, 2015](#); [Kalnins and Lafontaine, 2013](#)).

The remaining columns of Table 2 examine measures of organizational structure, in particular, the share of employment in professional and managerial (vs. service and production) roles. More distant establishments have fewer total employees than closer establishments; however, they have a relatively higher share of professionals relative to service and production workers employed at the branch (columns 7-8). This result also echoes recent findings that improved connectivity to headquarters reduces the necessity for provision of managerial and support activities locally and thus increases functional specialization at the level of the branch. Patterns consistent with this notion are documented by [Charnoz et al. \(2018\)](#) in France, [Acosta and Lyngemark \(2021\)](#) in Denmark, and [Gumpert et al. \(2022\)](#) in Germany. To illustrate relationship sizes, our estimates imply that a branch of a firm with headquarters in São Paulo is likely to exhibit around 4.7% fewer employees, and roughly 0.9 percentage points more managers in the workforce if it is located in distant Brasília as opposed to the nearby city of São José dos Campos.

4 Causal Effects of Changes in Connectivity

4.1 Identification and Estimating Equation

Next, we move beyond the descriptive analysis to examine the causal effects of improvements in connectivity to headquarters. Following the well-established approaches in the literature that we build on, our empirical approach will estimate difference-in-differences models comparing establishments that benefited from a direct air connection to headquarters with establishments located in the same city that did not benefit from a direct air

Table 2: Distance to HQ and Establishment Outcomes, Cross-sectional Correlations

	Location probability		Survival to 2017		Log Employees		Mgmt & Professionals share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log dist. to HQ	-0.017*** (0.001)	-0.017*** (0.001)	0.007*** (0.001)	-0.003** (0.001)	0.044*** (0.003)	-0.021*** (0.006)	0.021*** (0.001)	0.004*** (0.001)
Observations	15,829,856	15,829,856	106,087	86,677	99,894	80,978	98,651	79,808
Firms	33,256	33,256	32,477	13,460	30,486	12,273	30,274	12,160
R-squared	0.057	0.096	0.151	0.670	0.067	0.567	0.192	0.772
HQ sector FE	Y	Y	Y	Y	Y	Y	Y	Y
HQ city FE	Y	Y	Y	Y	Y	Y	Y	Y
Estab. city FE	N	N	Y	Y	Y	Y	Y	Y
Firm FE	N	Y	N	Y	N	Y	N	Y

Notes: 2000 cross-section. The table presents coefficient estimates from an OLS model. Standard errors, in parentheses, are clustered by HQ arranjo. The sample includes all multi-unit business entities in our dataset in 2000, defined as having at least one establishment other than headquarters, regardless of their location, and at least 10 employees overall in that year. We exclude headquarters from the analysis. The dependent variable in columns (1)-(2) is an indicator for whether firm f operates at least one establishment in arranjo c ; in columns (3)-(4), *Survival* is an indicator variable for whether establishment j observed in year t (2000) is also observed at the end of the panel (2017); in columns (5)-(6), *Employees* is the log number of employees in establishment j at year-end; and in columns (7)-(8) is the share of professional workers over total employees. Professionals are defined as employees with a first digit of 1-4 of the CBO occupation code. Log distance to HQ is the log geodesic distance (in km) between the centroid of the most populous municipality of each arranjo c and the HQ arranjo for firm f . Sectors definitions use the Brazilian CNAE 1.0 codes at the 2-digit level.

connection (e.g., [Giroud, 2013](#); [Giroud and Mueller, 2015](#); [Catalini et al., 2020](#)).

We select our sample for this analysis starting with the almost 230,000 establishments that fit the sample selection criteria detailed in Section 2, i.e., which belonged to a multi-unit firm that was active in year 2000 and employed at least 10 employees in at least one year during the sample period.

From this sample, we initially identify establishments that received a new, direct flight connection to the *arranjo* of their headquarters. To identify “treated” establishments for the empirical specification, we restrict this sample to establishments that were located at a different *arranjo* than their headquarters at the time the new connection was introduced (and, by extension, the year prior to that). We further exclude 60,482 establishments that were “always treated” (i.e., which already enjoyed a direct air connection to headquarters from when they first appear in our panel), 9,189 establishments that were “partly treated” (which received a direct connection to headquarters which was not subsequently maintained for all future years in our sample), and 2,592 that did not receive a direct flight

connection but which observed a flight time reduction of 30 minutes or more to the *arranjo* of their headquarters at any point during the sample period (e.g., due to changes in the airline route network that affect the travel time on indirect flights).

Control establishments, on the other hand, are those that did not receive a direct flight to their headquarters and did not see a reduction in their flight time to headquarters of 30 minutes or more during the sample period.

Because treatment timing is staggered, since new direct routes are introduced throughout the sample period, traditional two-way fixed effects model estimates may suffer from bias in presence of treatment effect heterogeneity. To circumvent these biases, we create stacks of observations of treated branches and control establishments, similar to [Cengiz et al. \(2019\)](#). A stack is a function of location and treatment timing. Each stack consists of all treated establishments located in *arranjo* c and treated in calendar year t , and all control establishments that were also located in *arranjo* c in year t but were never treated (i.e., their headquarters is not located in the city that became connected via the new direct flight connecting the *arranjo*). Because our empirical specification will include period-stack fixed effects, establishments that are in a *arranjo*-treatment year (i.e., "stack") that do not contain at least one treated and control observation drop out of the sample.⁸ Our final estimation sample consists of 10,288 establishments across 3,938 different firms, operating between 2000 and 2017. 2,587 of these establishments are treated, and 7,701 serve as controls.

We estimate the following empirical model on the sample of treated and control establishments:

$$Y_{ist} = \alpha + \beta \mathbf{1}\{W\}_{ist} + \alpha_{is} + \alpha_{ts} + \epsilon_{ist} \quad (1)$$

where the sub-indexes denote non-HQ establishments (i), calendar time periods (t), and stacks (s), respectively. W is an indicator denoting that establishment i already received "treatment" (i.e., improved its connection with its headquarter) in period t , α_{is} is an establishment fixed effect, estimated within each stack (since control branches can appear in more than one stack). α_{ts} is a period-stack fixed effect, and ϵ_{ist} is the error term. Key to isolating the effects of connectivity to HQ specifically as opposed to effects of connectivity more broadly is the inclusion of period-stack fixed effects, α_{ts} . These absorb city-level shocks that effect both treated and control establishments as well as city-wide

⁸To ensure we observe at least one pre-treatment year, we also require establishments and headquarters to appear in the panel and not move locations between years $t - 1$ and t ; that is, establishments will necessarily exist during these two consecutive years, but may enter or exit at different times, as our panel is unbalanced. Table [B10](#) reports the main effects excluding any movers.

benefits of improved air connectivity, which can affect population, employment growth, a city's sectoral composition, commuting patterns, and investment (e.g., [Brueckner, 2003](#); [Sheard, 2014](#); [Blonigen and Cristea, 2015](#); [Tveter, 2017](#); [Campante and Yanagizawa-Drott, 2018](#); [Coscia et al., 2020](#)).

Table 3 presents descriptive statistics for the estimation sample. Over the sample period, 95% of establishments survive until the following year-end. Establishments have, on average, 39 employees, of which almost 60 percent are professionals in managerial or technical roles earning around 40 percent more than their counterparts in service and production roles.

Table 3: Summary Statistics of Main Estimation Sample

	Mean	SD	p25	p50	p75	p95
Treated (direct connection)	0.30	0.46	0.00	0.00	1.00	1.00
Survive to t+1	0.95	0.22	1.00	1.00	1.00	1.00
Employees	38.73	178.48	4.00	11.00	26.00	132.00
Professionals	17.71	81.74	2.00	6.00	13.00	61.00
Production workers	23.59	139.04	0.00	5.00	13.00	78.00
Professional share	0.57	0.34	0.29	0.50	1.00	1.00
Average earnings (R\$ 2013)	2,285	2,083	1,124	1,672	2,644	5,881
Avg. professional earnings (R\$ 2013)	2,406	2,078	1,204	1,771	2,905	6,030
Avg. production worker earnings (R\$ 2013)	1,743	1,570	940	1,339	1,952	4,176
Distance to HQ (km)	1,077	835	351	848	1,703	2,658
No. of affiliates in same city	0.71	9.13	0.00	0.00	0.00	2.00
Total affiliates	21.49	76.93	3.00	5.00	13.00	76.00
HQ employees	237.57	698.56	15.00	50.00	187.00	980.00
HQ professionals	130.56	434.28	7.00	25.00	93.00	554.00

Notes: Observations are establishments of multi-unit firms that meet the sample criteria: the firms was active in December 2000, had at least one branch outside the HQ's *arranjo* from 2000–2017, employed more than 10 workers in at least one year, and had their HQ in an *arranjo* with an airport. The table only displays summary statistics for the establishments that either received a direct connection to their HQ (treated) that was maintained for the duration of the sample period or that never received such a connection (control), between 2001 and 2017. *Survival* is defined here as the annual survival probability; i.e., an indicator that equals 1 if the establishment survives until the following year-end. The number of total affiliates excludes the headquarters. Table shows statistics over the 2000-2017 period.

4.2 Causal effects of increased connectivity with headquarters

Table 4 presents the estimated average effects of receiving a direct air connection to headquarters on establishment-level outcomes. As it becomes easier for headquarters employees

to visit a firm’s establishment, and vice versa, the establishment becomes 12.6 percentage points more likely to shut down by the end of the following year, as seen in Column (1). We also observe clear shifts in branch workforce composition. Specifically, employment of professionals increases slightly (column 3), while the number of service and production workers declines by 7.5% (column 4). Further, while there appear to be no significant changes on average earnings in Columns (5)–(6), Column (7) suggests a negative, but not statistically significant, effects on the earnings of production workers. This indicates a reallocation away from occupations typically associated with operational and service functions.

Table 4: Causal Effects of a Direct Air Connection to HQ on Establishment Outcomes

	Survive t+1	Log employees	Log professionals	Log prod. workers	Log avg. earnings	Log avg. prof. earnings	Log avg. prod. earnings
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Treated	-0.126*** (0.006)	-0.004 (0.021)	0.044* (0.020)	-0.075** (0.025)	0.005 (0.008)	0.003 (0.009)	-0.011 (0.008)
Observations	109,791	102,556	97,021	77,271	101,998	97,021	77,271
N. control	7,700	7,249	6,681	5,051	7,217	6,681	5,051
N. treated	2,586	2,453	2,281	1,668	2,441	2,281	1,668
R-squared	0.431	0.887	0.880	0.872	0.913	0.901	0.902
Estab. FE	Y	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y	Y

Notes: This table shows the results of estimating Equation (1) over the 2000–2017 period. We consider three years prior to the treatment year as the pre-treatment period, and five years after the treatment year as the post-treatment period. *Treated* is an indicator that equals one in the year an establishment first receives a direct air connection to their headquarters and thereafter. *Survive* is an indicator for whether an establishment is still operating in the following year-end (excluding 2017 as we do not observe 2018). Employment and earnings are measured at year-end. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

We then estimate a stacked event study model to examine these treatment effects over time with the following specification:

$$Y_{ist} = \alpha + \sum_{k \neq 1} \beta_k \mathbf{1}\{t - T_{is} = k\} + \alpha_{is} + \alpha_{ts} + \epsilon_{ist}, \quad (2)$$

where the sub-indexes denote non-HQ establishments (i), time periods (t), years to/from treatment (k), and the stack (s), respectively; $\mathbf{1}\{t - T_{is} = k\}$ are event-time indicators relative to each establishment’s treatment year T_{is} , defined only for treated

observations; α_{is} is an establishment-stack fixed effect, α_{ts} is a period-stack fixed effect, and ϵ_{ist} is the error term.

Figure 2 illustrates the evolution of the effects of gaining a direct air connection on overall establishment-level outcomes over time. Subfigure 2a shows that already in the treatment year, the 1-year survival rate declines; that is, the likelihood of closure by the end of year $t + 1$ increases by 10 percentage points. In line with the results reported in Table 4, we find no significant effects on total employment (subfigure 2b) or average earnings (subfigure 2c).

Similarly, Figures 3 and 4 present the corresponding event study results for professionals and service and production workers, respectively. While average earnings remain largely unaffected in both groups (subfigures 3b and 4b), their employment trajectories seem markedly different after treatment. Improved connectivity leads to a significant and persistent decrease in service and production employment, reaching a 12.5 percent reduction five years post-treatment (subfigure 4a). In contrast, the number of professionals employed at these branches increases slightly and gradually, with gains of around 9–14 percent beginning in the third year (subfigure 3a).

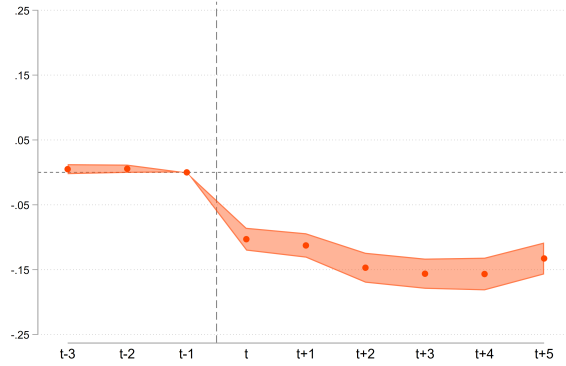
These findings are surprising in light of the existing literature. In particular, our results on service and production workers diverge sharply from those of Gumpert et al. (2022), who document an *increase*, rather than decrease, of almost the same magnitude in service and production employment. That these effects are not temporary but persist for several years after treatment also suggests that reducing geographic frictions between headquarters and local establishments can affect performance differently in contexts characterized by high inefficiency.

4.3 Robustness Tests of Main Effects

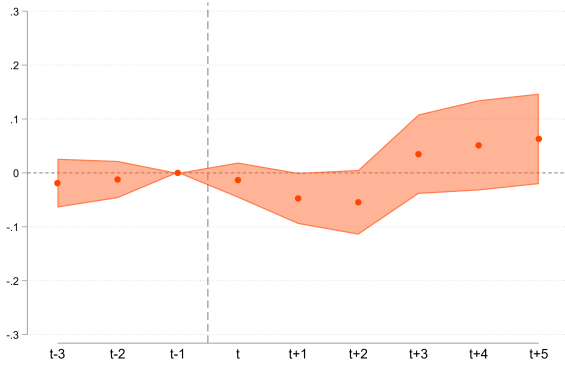
In Table 5, we show that our main results are robust to different empirical choices. Effects on survival rates are displayed in Panel A, while those on service and production employment are in Panel B.

Column (1) considers a panel of establishments treated at any time between 2003 and 2012, inclusive. This ensures that we could potentially observe establishments for the entire duration of the estimation panel; that is, three years prior to treatment (2000, for establishments treated in 2003) and five year after treatment (2017, for establishments treated in 2012). It addresses the concern that stacked DiD estimators may fail to recover

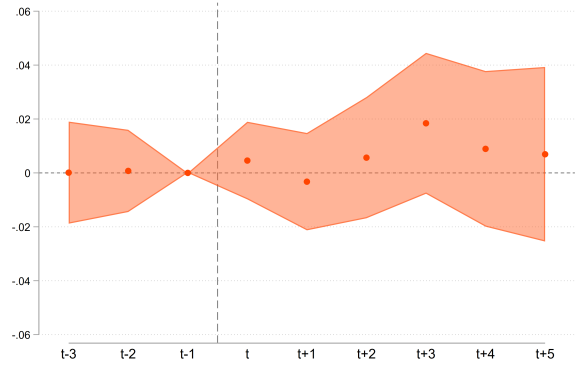
Figure 2: Causal effects of a direct air connection to HQ on establishment outcomes over time



(a) Likelihood of survival until the following year-end



(b) Log total employees

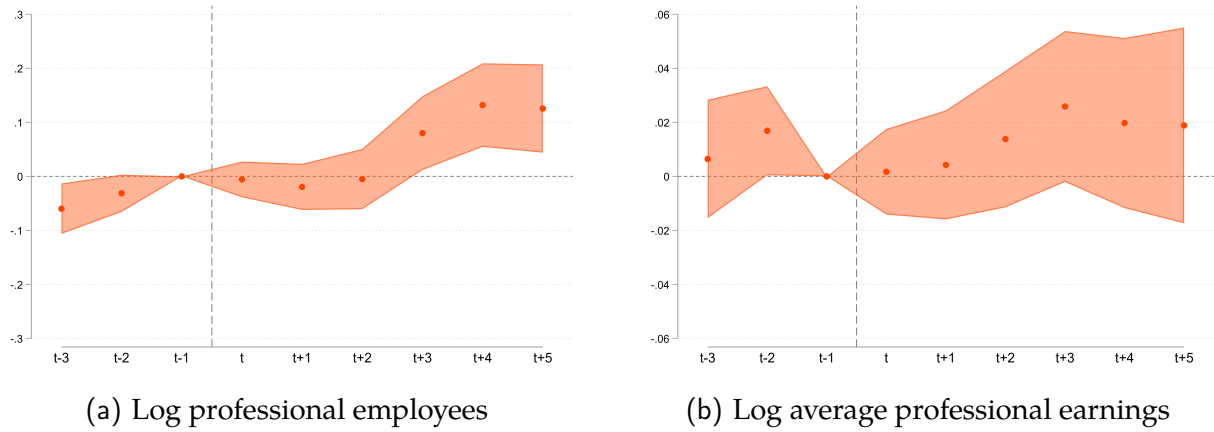


(c) Log average earnings

Note: This figure shows the results of estimating Equation (2) over the 2000–2017 period. We consider three years prior to the treatment year as the pre-treatment period, and five years after the treatment year as the post-treatment period. *Survive* is an indicator for whether an establishment is still operating in the following year-end (excluding 2017, as we do not observe 2018). Employment and earnings are measured at year-end. All variables are winsorized at the first and 99th percentiles. Standard errors clustered at the establishment level.

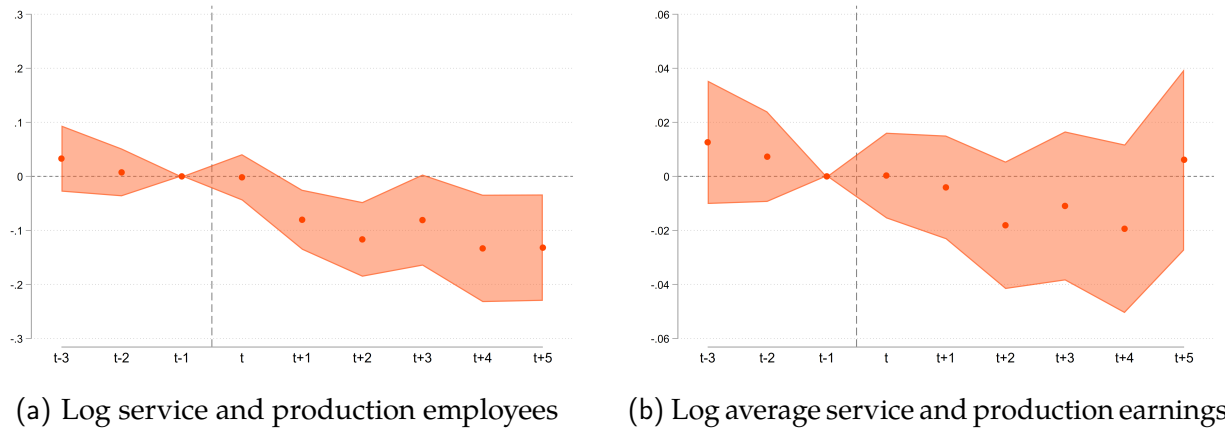
the average treatment effect (ATT) because it averages treatment effects observed over different time windows (Wing et al., 2024). The estimated effects in Column (1) are very similar to our main estimates and slightly larger. Columns (2)–(3) look exclusively at a subset of our main estimation sample of firms that only operate one branch other than HQ during the entire panel and those that have at least two other branches, respectively. The coefficients in column (2) are less precisely estimated due to the much smaller sample size but the overall patterns are similar. Finally, columns (4)–(5) use different treatment definitions. Rather than using direct flights, they consider establishments to be treated if

Figure 3: Causal effects of a direct air connection to HQ on outcomes for professionals over time



Note: This figure shows the results of estimating Equation (2) over the 2000–2017 period. We consider three years prior to the treatment year as the pre-treatment period, and five years after the treatment year as the post-treatment period. Employment and earnings are measured at year-end. All variables are winsorized at the first and 99th percentiles. Standard errors clustered at the establishment level.

Figure 4: Causal effects of a direct air connection to HQ on outcomes for service and production workers over time



Note: This figure shows the results of estimating Equation (2) over the 2000–2017 period. We consider three years prior to the treatment year as the pre-treatment period, and five years after the treatment year as the post-treatment period. Employment and earnings are measured at year-end. All variables are winsorized at the first and 99th percentiles. Standard errors clustered at the establishment level.

they had *any* reduction in air travel time greater than 30 minutes (including via indirect flights) in column (4) or *large* reductions in air travel time (greater than 1.5 hours) in column (5).

Our results are consistent across specifications, as they all point to significant decreases in survival rates of anywhere between 12.1 to 33.2 percentage points, and reductions in employment of production workers of around 7 to 9.4 percent. In fact, these findings indicate that our main specification in Table 4, which uses the full panel and a more strict treatment definition of exclusively direct flights, might be on a lower bound of the effect size.

5 A Model of Spatial Frictions in Multi-Unit Firms

In this section, we develop a model to analyze how distance to headquarters shapes a firm’s decision to enter a remote market, whether to operate a local branch, and how many workers the branch employs. The framework aims to account for the patterns we document in Brazil while remaining consistent with contrasting findings from high-income economies (e.g. [Giroud 2013](#); [Charnoz et al. 2018](#); [Gumpert et al. 2022](#)).

The model nests three distance-related frictions that are all increasing in distance d from headquarters to the target market. The first is a *client–service friction* that makes it harder to serve customers from afar when the firm does not operate a branch. The second is an *internal coordination friction* that raises the cost of bringing headquarter inputs on site when the firm does operate a branch. The third is a *moral–hazard friction*: delegating hiring to a local manager creates scope for over-hiring that is harder to discipline from afar.

5.1 Model Setup

The firm can serve a remote market either by producing locally with a branch or by producing at headquarters and serving the market directly from there. Production technology is Cobb–Douglas:

$$Q = L^\alpha H Q^{1-\alpha}, \quad \alpha \in (0, 1), \quad (3)$$

where L is local labor and HQ is a bundle of headquarter inputs (training, mentoring, problem-solving, crisis management, supervision) that must be delivered in person.⁹ There is a maximum feasible level of headquarter input $\overline{HQ} > 0$, which can be interpreted as a technological or logistical constraint, or a CEO’s time constraint (as in [Gumpert et al. 2022](#)).

⁹The model can be modified to assume, instead, that HQ inputs can be partially provided remotely, to study the effects of increased connectivity and the rise of remote work in recent years. We abstract from this dimension of analysis, since it is beyond the scope of this paper.

Table 5: Robustness Tests of Main Effect

Panel A: Survive t+1					
	Balanced panel (five years pre & post)	MU firms with one unit	MU firms with 2+ units	Any air travel time reduction	Air travel time reduction > 1.5 hour
	(1)	(2)	(3)	(4)	(5)
Treated	-0.136*** (0.008)	-0.150*** (0.028)	-0.121*** (0.007)	-0.305*** (0.009)	-0.332*** (0.011)
Observations	73,543	7,461	99,538	1,256,680	780,752
N. control	4,777	565	7,044	51,886	46,235
N. treated	1,749	225	2,325	2,686	1,719
R-squared	0.390	0.548	0.434	0.332	0.387
Estab. FE	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y

Panel B: Log prod. workers					
	Balanced panel (five years pre & post)	MU firms with one unit	MU firms with 2+ units	Any air travel time reduction	Air travel time reduction > 1.5 hour
	(1)	(2)	(3)	(4)	(5)
Treated	-0.088** (0.032)	-0.226 (0.134)	-0.070** (0.026)	-0.094*** (0.026)	-0.079* (0.034)
Observations	50,692	4,650	69,092	786,294	486,216
N. control	3,004	297	4,566	29,628	26,875
N. treated	1,115	121	1,490	1,627	1,077
R-squared	0.865	0.873	0.875	0.875	0.877
Estab. FE	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y

Notes: This table shows the results of estimating our empirical specification over alternative samples. In column (1), we consider only establishments treated between 2003 and 2012, inclusive. In column (2) we consider the subset of firms with only one branch other than HQ during the whole panel, while in column (3) we estimate the effects over the subset of firms with at least two branches other than HQ at any point in time. In columns (4) and (5), we consider treatment to be receiving any travel time reduction of at least 30 minutes or more than 1.5 hours, respectively. *Survive* is an indicator for whether an establishment is still operating in the following year-end (excluding 2017 as we do not observe 2018). *Log prod. workers* is measured at year-end. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

The output price is normalized to one, and the local wage w is exogenous.¹⁰

If the firm *does not* operate a branch, it produces at headquarters and ships the product.

¹⁰Our empirical specification includes location-year fixed effects, which accounts for systematic wage differences over time and across locations.

Shipping is assumed to be costless,¹¹ but conducting client-facing tasks from a distance entails additional costs. For example, geographic separation can hinder engagement with key customers or partners, limit the ability to gather information and adapt to local market conditions, complicate negotiations with local regulators, and impede other context-sensitive activities that benefit from physical presence. We model this *client–service friction* as an iceberg cost $\tau(d)$, normalizing $\tau(0) = 1$, and with $\tau'(d) > 0$ and $\tau''(d) \geq 0$, such that distance makes customer-facing tasks disproportionately harder.

If the firm *operates* a branch, it provides HQ inputs on site incurring *internal coordination costs*, $\kappa(d, HQ)$ —for example, travel, oversight, information transmission, and on-site problem-solving become costlier at distant branches.¹² We assume $\kappa(d, HQ)$ is strictly increasing and convex in each argument, with nonnegative complementarity: deploying more HQ inputs is disproportionately costly both at higher volumes and at greater distances (e.g., fatigue and scheduling frictions), and distance and HQ scale reinforce each other in raising coordination costs.¹³

In the spirit of standard agency models (e.g., [Aghion and Tirole, 1997](#); [Aghion et al., 2021](#)), we assume that branch decisions are delegated to a local manager whose interests may diverge from those of headquarters. The manager receives as compensation a share $\beta \in (0, 1)$ of revenues, net of labor costs. Beyond formal pay, the manager derives a private benefit from over-hiring that rises with distance, captured by $\gamma(d)$. For instance, they might enjoy per-employee perks or personal utility from discretionary labor usage. We assume these private benefits rise with distance to headquarters ($\gamma'(d) \geq 0$), as smaller, more socially interconnected remote markets make discretionary hiring more visible and can grant the manager reputational or social-status gains from providing jobs. Over-hiring is deterred by expected penalties that scale with headquarters presence, at rate $\delta > 0$. The threshold

$$HQ^{\text{crit}}(d) \equiv \frac{\gamma(d)}{\delta}$$

marks the level of headquarter inputs above which incentives are aligned and the manager does not over-hire.

¹¹We abstract from shipping costs—which are typically emphasized in the FDI and international trade literature (e.g., [Markusen, 1984](#); [Brainard, 1997](#); [Helpman et al., 2004](#))—to simplify the exposition and because such costs are unlikely to be affected by improvements in passenger air connectivity. Most domestic cargo in Brazil moves by water, truck, or train.

¹²The distinction between client-facing costs and intra-firm coordination costs was first introduced in the international trade literature by [Oldenski \(2012\)](#).

¹³Formally, $\frac{\partial \kappa(d, HQ)}{\partial d} > 0$; $\frac{\partial^2 \kappa(d, HQ)}{\partial d^2} \geq 0$; $\frac{\partial \kappa(d, HQ)}{\partial HQ} > 0$; $\frac{\partial^2 \kappa(d, HQ)}{\partial d \partial HQ} \geq 0$; and $\frac{\partial^2 \kappa(d, HQ)}{\partial HQ^2} \geq 0$.

The sequence of decisions starts with the firm deciding whether entry can be profitable at all. Conditional on entry, the firm chooses an operating mode: either it produces locally at a branch or it produces at headquarters and serves the market from there. Given the mode, headquarters chooses HQ , the bundle of on-site headquarter inputs. Finally, if a branch is operated, the local manager hires L taking HQ as given. We solve by backward induction (see details in Appendix C).

5.2 Profits and Entry

When the firm operates a branch at distance d and chooses (L, HQ) , it generates a profit of

$$\pi_b(d) = (1 - \beta)[L^\alpha HQ^{1-\alpha} - wL] - \kappa(d, HQ). \quad (4)$$

When the firm serves the market from headquarters, it produces on site, sets $HQ = \overline{HQ}$,¹⁴ and chooses L to maximize

$$\pi_{HQ}(d) = \frac{L^\alpha \overline{HQ}^{1-\alpha}}{\tau(d)} - wL. \quad (5)$$

Let $\pi_b^*(d)$ and $\pi_{HQ}^*(d)$, denote maximized profits under each mode. The firm's overall maximized profit is thus given by

$$\Pi(d) \equiv \max\{\pi_b^*(d), \pi_{HQ}^*(d), 0\}. \quad (6)$$

The firm enters the market if and only if $\Pi(d) > 0$. Both mode-specific profit schedules decline in d : with a branch, coordination becomes costlier; without a branch, client service becomes harder. As the firm's profit decreases monotonically with distance, there comes a point d^{entry} where no mode can generate positive profits, and the firm does not enter the market. Formally:

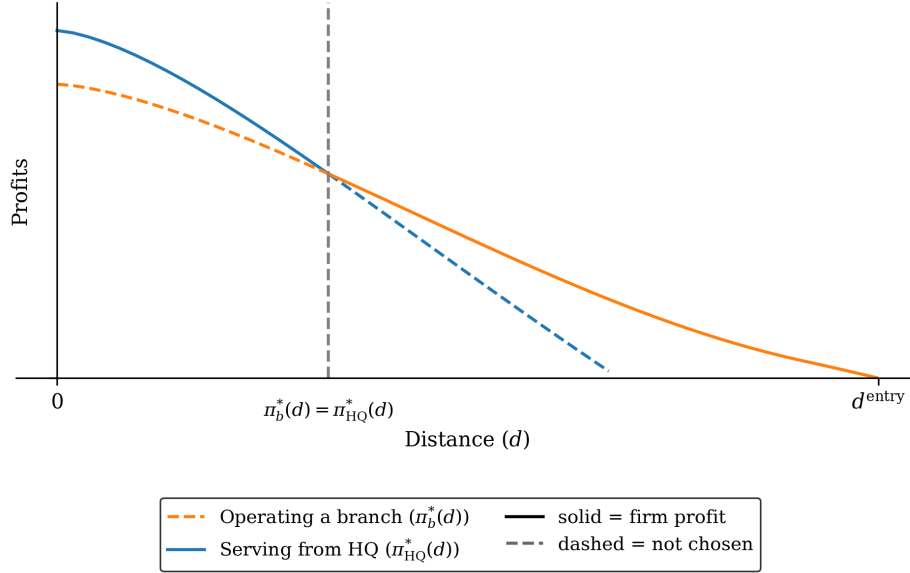
Proposition 1. Entry cutoff and monotonicity. *There exists a unique d^{entry} such that $\Pi(d) > 0$ for $d < d^{\text{entry}}$ and $\Pi(d) = 0$ for $d \geq d^{\text{entry}}$. Moreover, $\Pi(d)$ is continuous and strictly decreasing on any interval where a single operating mode is profit-maximizing; at a mode switch, the left and right derivatives of $\Pi(d)$ with respect to d are strictly negative.*

Proof: see Appendix C.

This is illustrated in Figure 5, which plots the maximized profits under each operating mode, $\pi_b^*(d)$ and $\pi_{HQ}^*(d)$, against distance d . The solid portions of each schedule indicate

¹⁴Since coordination costs $\kappa(\cdot)$ do not apply in this mode and the production term is increasing in HQ , the firm optimally sets HQ at the capacity constraint.

Figure 5: Firm Profits, Entry, and Mode Choice



Notes: The figure shows maximized profits by mode as functions of distance d , plotted where profits are positive. We impose $\tau(0) = 1$ using $\tau(d) = 1 + \eta_\tau d^{1.5}$ and specify $\kappa(d, HQ) = \kappa_0 + \eta_d d^{1.5} + \eta_{HQ} HQ^{1+\xi}$ with $\eta_d, \eta_{HQ}, \xi > 0$. The vertical dashed line denotes d^{entry} ; the solid segments indicate the chosen mode. Parameter values: $\alpha = 0.30$, $w = 1.00$, $\beta = 0.05$, $\delta = 1.50$, $\overline{HQ} = 2.0$, $\eta_\tau = 0.045$, $\kappa_0 = 0.090$, $\eta_d = 0.012$, $\eta_{HQ} = 0.06$, $\xi = 0.10$, $\gamma(d) = 0.25 + 0.005 d$.

the region where that mode yields the highest profit, and is thus chosen by the firm. The slope of $\Pi(d)$ to the left of d^{entry} coincides with the slope of whichever operating mode is currently profit-maximizing (branch or HQ service), and entry occurs when at least one schedule remains above zero.

5.3 Distance Frictions and Mode Choice

Write the branch-profit premium as

$$\Delta\pi(d) \equiv \pi_b^*(d) - \pi_{HQ}^*(d),$$

such that, conditional on entry, the firm will choose to operate a branch if $\Delta\pi(d) > 0$, and to serve the market directly from headquarters otherwise.

To see how distance affects this decision, we differentiate with respect to d and apply the envelope theorem, yielding the three-friction decomposition

$$\frac{\partial \Delta \pi(d)}{\partial d} = \underbrace{-\frac{\partial \kappa(d, HQ^*(d))}{\partial d}}_{\text{internal coordination}} + \underbrace{(1 - \beta) \Psi(d) \gamma'(d)}_{\text{moral hazard } (\leq 0)} + \underbrace{[L_{HQ}^*(d)]^\alpha \overline{HQ}^{1-\alpha} \frac{\tau'(d)}{\tau(d)^2}}_{\text{client-service } (>0)}. \quad (7)$$

where $\Psi(d) \leq 0$ is a moral-hazard envelope term that equals zero when incentives are aligned and is defined precisely in Appendix C. The first two terms push the firm away from operating a branch as distance rises; the third pushes it toward operating a branch. The sign is therefore *a priori* ambiguous.

When the client-service term dominates, better connectivity (lower d) tilts the firm away from branches. This is the situation illustrated in Figure 5: On the right of the mode switch threshold ($\Delta \pi(d) = 0$), the branch-profit premium is positive, and the firm chooses to operate a branch up to distance d^{entry} . The choice to operate a branch is less attractive at shorter distances. On the left of the threshold the branch-profit premium is negative, and the firm prefers to operate directly from headquarters. The opposite is true if the coordination and moral-hazard terms dominate. In this case, better connectivity tilts the firm towards operating branches.

5.4 Headquarter Inputs and Distance

When the firm operates a branch, headquarters sets HQ anticipating the manager's choice of L . Because the marginal cost of HQ rises with distance and the manager's private benefit from over-hiring does not fall with distance, the marginal return to HQ weakens as d grows.

Proposition 2. *Headquarter inputs fall with distance. As long as the marginal benefit of headquarter inputs is weakly decreasing in distance d , the equilibrium headquarter inputs $HQ^*(d)$ are weakly decreasing in d .*

Proof: See Appendix C.

5.5 Branch Employment: Productivity vs. Discipline

Conditional on operating a branch, the manager's labor choice reflects two distinct forces. To see this, write $L^*(d) = L^\dagger(HQ^*(d); d)$ and differentiate, to obtain the identity

$$\frac{dL^*(d)}{dd} = \frac{\partial L^\dagger}{\partial HQ} \frac{\partial HQ^*}{\partial d} + \frac{\partial L^\dagger}{\partial \gamma} \gamma'(d). \quad (8)$$

The first term is the *HQ channel*: better connectivity raises on-site HQ^* ; if marginal HQ mainly acts as a productive input, this pushes up the manager's desired L , whereas if it mainly acts as monitoring, it pushes L down. The second term is the moral-hazard channel: a larger d strengthens the over-hiring incentive and pushes up L directly. Which one dominates depends on whether HQ is working mostly as a productive input or mostly as a disciplinary tool at the margin.

Under a mild single-crossing regularity (stated in Proposition 3), there is at most one distance $\bar{d}$ at which the marginal role of HQ flips from production to discipline. For small d , headquarters' time mainly boosts production and the productivity channel dominates. For large d , headquarters' time mainly reins in over-hiring and the discipline channel dominates. We summarize this as follows:

Proposition 3. *Employment response by regime.* Let $L^*(d) = L^\dagger(HQ^*(d); d)$ be equilibrium branch employment.

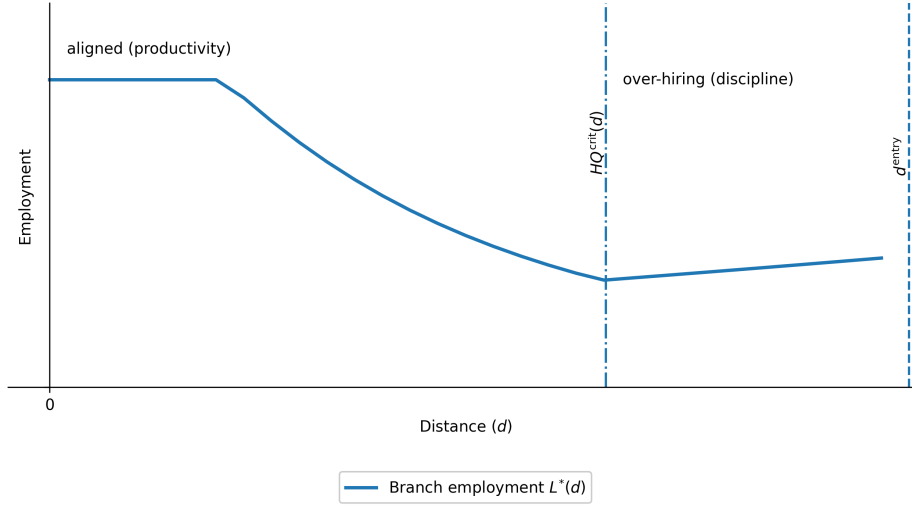
On the aligned-incentives region, $HQ^*(d) \geq HQ^{\text{crit}}(d)$, employment is weakly decreasing in distance d .

On the over-hiring region, $HQ^*(d) < HQ^{\text{crit}}(d)$, assume the marginal effect of HQ on the manager's desired employment, evaluated at $HQ^*(d)$, changes sign at most once as d increases (single-crossing). Let $\bar{d}$ denote the unique distance (if it exists) at which an extra unit of HQ leaves the manager's desired employment unchanged. Then for $d > \bar{d}$ (discipline-dominated), $L^*(d)$ is weakly increasing in d , and strictly increasing when $HQ^*(d)$ is interior (so that $dHQ^*/dd < 0$). For $d < \bar{d}$ (productivity-dominated), the sign of $\frac{dL^*(d)}{dd}$ depends on the relative strength of the HQ channel versus the over-hiring channel; in particular, whenever the HQ channel weakly dominates on $[0, \bar{d})$, $L^*(d)$ is weakly decreasing there.

Proof: See Appendix C.

Figure 6 illustrates these regimes and the single-crossing logic. At very short distances, there is high headquarters presence and manager incentives are aligned with those of the firm's. Employment scales proportionally with HQ , producing the nearly flat left segment. As d increases but remains below $\bar{d}$, headquarters cut back on HQ (Proposition 2). When the HQ channel dominates the over-hiring channel, the manager's desired L falls and $L^*(d)$ slopes downward (as in Figure 6). Beyond the dashed line at $HQ^{\text{crit}}(d) = \gamma(d)/\delta$, the value of HQ time is mainly disciplinary, and the curve tilts upward (i.e., better connectivity reduces over-hiring and therefore lowers L).

Figure 6: Branch Employment and Distance



Notes: The figure plots equilibrium branch employment $L^*(d)$ against distance d . The vertical dot–dash line marks $HQ^{\text{crit}}(d) = \gamma(d)/\delta$, the boundary between aligned incentives (no over-hiring) and over-hiring. The vertical dashed line marks the entry cutoff d^{entry} . For $d < \bar{d}$ (productivity-dominated), better connectivity (lower d) raises HQ^* and—when the HQ channel dominates the over-hiring channel—raises employment (as in this parameterization). For $d > \bar{d}$ (discipline-dominated), better connectivity raises monitoring intensity and reduces over-hiring, lowering employment. Parameter values match Figure 5.

The distinction between these two regimes help explain why the effect of improved connectivity on branch employment can differ across settings. In high-income economies, there is typically higher labor formality, reducing the scope for branch managers to engage in patronage (i.e. lower γ), and making over-hiring less likely (lower $HQ^{\text{crit}} = \gamma/\delta$). In such environments, improved connectivity (lower d) typically raises HQ^* and *increases* employment—as found in prior high-income country studies (e.g., Giroud, 2013; Gumpert et al., 2022). In low and middle-income economies, such as Brazil, higher γ makes it more likely that branches are situated in the discipline-dominated region ($d > \bar{d}$), where better connectivity strengthens monitoring and *reduces* over-hiring, so employment falls. The single-crossing condition in Proposition 3 rules out multiple flips between regimes and guarantees a unique $\bar{d}$.

6 Additional Empirical Tests

We now return to the data to conduct additional empirical tests of the model’s predictions. First, we examine how the main effects discussed in Section 4 vary with the baseline distance

between headquarters and the branch. The results are presented in Table 6.

Table 6: Causal effects of a direct air connection on establishment outcomes by distance to HQ

Dist. to HQ	Survive t+1		Log professionals		Log prod. workers	
	Below	Above	Below	Above	Below	Above
	Median (1)	Median (2)	Median (3)	Median (4)	Median (5)	Median (6)
Treated	-0.181*** (0.012)	-0.091*** (0.009)	0.007 (0.032)	0.063* (0.030)	-0.007 (0.040)	-0.122** (0.038)
Observations	51,704	55,153	45,081	49,374	38,467	36,368
N. control	3,647	3,657	3,096	3,224	2,509	2,204
N. treated	1,231	1,295	1,026	1,196	772	846
R-squared	0.497	0.410	0.897	0.878	0.889	0.875
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 4, except that models are estimated in sub-samples of establishments based on their geodesic distance to their headquarters in the year prior to treatment. Columns labeled *below* [(1), (3), (5)] are the results for those establishments below or at the median distance, while columns labeled *above* [(2), (4), (6)] show the effects on establishments above the median distance. The median distance to headquarters corresponds to approximately 812km. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Columns (1) and (2) estimate the effects of new air connections to headquarters on branch survival through the following year-end, separately for branches located below and above the median headquarters-branch distance. The results show that the negative effect on survival is more pronounced for branches located closer to headquarters: treated establishments in closer markets are 18.1 percentage points less likely to survive, compared to a 9.1 percentage point decline in survival among more distant branches. This is fully consistent with the model's prediction that improved connectivity makes branch closure more likely when distance falls below a threshold at which it becomes more profitable to serve the market directly from headquarters, without maintaining a branch.

Columns (3) through (6) report the effects of new air connections on branch employment, again estimated separately for closer and more distant branches. In contrast to the survival results, these results show that employment reductions are concentrated among more distant branches, with a particularly strong and statistically significant decline among

service and production workers. These patterns are consistent with the model’s assumption that local managers in more remote locations face stronger incentives to overhire, resulting in inefficient baseline allocations of labor. In such cases, increased visits from headquarters strengthen monitoring and lead to employment reductions.

We next investigate whether the impact of improved air connectivity depends on how thinly headquarters’ personnel are spread across the firm’s network of establishments. Following [Giroud \(2013\)](#), we construct a proxy for headquarters time constraints. Specifically, we define *headquarters-bandwidth*, equal to the number of managers or supervisors employed at the headquarters divided by the number of non-HQ branches operated by the firm, at the year prior to treatment.¹⁵ A lower manager-to-branch ratio indicates that the headquarters is more time-constrained and, in terms of the model, corresponds to a smaller firm-level endowment of headquarter inputs $\bar{H}$. We split the sample at the median of this ratio and estimate the baseline specifications separately for establishments whose headquarters are more and less time-constrained. The results are reported in Table 7.

These estimates also align with the model’s predictions. Columns (1) and (2) show that positive shocks to connectivity with headquarters decrease the probability of survival for both subsamples, but the effect is stronger when headquarters bandwidth is low: survival falls by 13.9 percentage points for branches with *ex ante* more time-constrained HQ versus 10.4 percentage points for branches with less time-constrained HQ. In the model, because branch profit is increasing in the baseline stock of headquarter inputs, $\partial \pi_b^* / \partial \bar{H} > 0$, constrained firms lie closer to the break-even frontier. A given fall in distance d therefore pushes more of them below the exit cutoff.

In turn, the results in columns (3)–(6) show that the negative effect on the employment of service and production workers is concentrated in the subsample of branches whose HQ is highly time-constrained, who experience a 7 percent employment decline. In the model, when headquarters face higher time constraints, more distant branches receive less monitoring and tend to be overstaffed. Therefore, even if the increase in headquarters time on site following improved connectivity is smaller among bandwidth-constrained firms, its disciplinary effect is amplified because the ex-ante distortion is larger.

¹⁵Unlike [Giroud \(2013\)](#), the Brazilian employer–employee records allow us to identify managers and supervisors directly (see Data Appendix A)

Table 7: Causal effects of a direct air connection on establishment outcomes by HQ constraints

Num. HQ managers per estab.	Survive t+1		Log professionals		Log prod. workers	
	Below Median (1)	Above Median (2)	Below Median (3)	Above Median (4)	Below Median (5)	Above Median (6)
Treated	-0.139*** (0.008)	-0.104*** (0.011)	0.037 (0.025)	0.019 (0.036)	-0.070* (0.028)	-0.055 (0.052)
Observations	71,628	34,917	62,903	31,370	52,465	22,696
N. control	4,939	2,492	4,265	2,217	3,408	1,496
N. treated	1,511	997	1,313	905	989	632
R-squared	0.447	0.459	0.880	0.890	0.866	0.893
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 4, except that models are estimated in sub-samples of establishments based on the headquarters' constraints, defined as the number of managers or supervisors at headquarters divided by the number of branch establishments the firm operates, similar to Giroud (2013), at the year prior to treatment. Columns labeled *below* [(1), (3), (5)] are the results for those establishments below or at the median number of managers per branch (those with more time-constrained HQs), while columns labeled *above* [(2), (4), (6)] show the effects on establishments above the median number of managers per branch (those with less time-constrained HQs). The median corresponds to approximately 1 manager or supervisor at headquarters per branch establishment. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

7 Conclusion

This paper shows that reducing geographic frictions between headquarters and branches can have markedly different consequences when allocative inefficiency is prevalent. Using Brazil's matched employer–employee data and plausibly exogenous changes in headquarter–branch travel time that result from the staggered roll-out of new airline routes during 2000–2017, we document two main findings. First, the familiar cross-sectional patterns observed in high-income settings are also present in Brazil: markets farther from headquarters are less likely to host a branch, surviving branches are smaller, and employ a larger share of managers. Second, causal estimates reveal a striking difference relative to prior findings: gaining a direct air connection to headquarters *lowers* one-year survival probabilities by about 13 percentage points and, among survivors, *reduces* service and production employment by about 8 percent—effects that persist for at least five post-treatment years.

To reconcile these findings with earlier evidence of positive employment effects in developed economies, we develop a simple model with three distance-dependent frictions—client–service costs when serving the market from HQ, internal coordination costs when operating a branch, and a moral-hazard friction that raises over-hiring by local managers at distance—that rationalizes both the cross-sectional facts and the causal effects. Improved connectivity raises productivity but also strengthens monitoring, making over-staffing more likely to be detected and corrected (a novel “disciplinary” mechanism).

Additional heterogeneity analyses reveal that the model’s predictions match the heterogeneous patterns in the data. Specifically, the decline in survival is larger for branches closer to headquarters (consistent with serving those market directly, without a branch), whereas employment reductions are larger at distant branches and when headquarters are more time-constrained (lower manager-to-branch ratios), where ex-ante over-hiring is more likely. These results suggest that, in the presence of agency problems and inefficiency, improvements in connectivity can lead to a “pruning” effect, as headquarters is better able to correct such inefficiencies.

Overall, our results underscore that the net impact of connectivity depends not only on distance, but also on baseline misallocation: in contexts with weak monitoring and agency problems, improved headquarter access can discipline branches and reduce non-managerial employment, even as it shifts activity closer to headquarters.

References

- Acosta, C. and Lyngemark, D. H. (2021). The internal spatial organization of firms: Evidence from Denmark. *Journal of Urban Economics*, 124:103366.
- Adhvaryu, A., Nyshadham, A., and Tamayo, J. (2023). Managerial quality and productivity dynamics. *The Review of Economic Studies*, 90(4):1569–1607.
- Aghion, P., Bloom, N., Lucking, B., Sadun, R., and Van Reenen, J. (2021). Turbulence, firm decentralization, and growth in bad times. *American Economic Journal: Applied Economics*, 13(1):133–169.
- Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations. *Journal of Political Economy*, 105(1):1–29.
- Bel, G. and Fageda, X. (2008). Getting there fast: globalization, intercontinental flights and location of headquarters. *Journal of Economic Geography*, 8(4):471–495.
- Bento, P. and Restuccia, D. (2017). Misallocation, Establishment Size, and Productivity. *American Economic Journal: Macroeconomics*, 9(3):267–303.
- Blonigen, B. A. and Cristea, A. D. (2015). Air service and urban growth: Evidence from a quasi-natural policy experiment. *Journal of Urban Economics*, 86:128–146.
- Bloom, N., Genakos, C., Sadun, R., and Van Reenen, J. (2012). Management practices across firms and countries. *Academy of Management Perspectives*, 26(1):12–33.
- Brainard, L. (1997). An empirical assessment of the proximity-concentration trade-off between multinational sales and trade. *American Economic Review*, 87:520.
- Britto, D. G., Pinotti, P., and Sampaio, B. (2022). The effect of job loss and unemployment insurance on crime in Brazil. *Econometrica*, 90(4):1393–1423.
- Brueckner, J. K. (2003). Airline traffic and urban economic development. *Urban Studies*, 40(8):1455–1469.
- Bryan, G. and Morten, M. (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy*, 127(5):000–000.
- Busso, M., Madrigal, L., and Pagés, C. (2013). Productivity and resource misallocation in Latin America. *The BE Journal of Macroeconomics*, 13(1):903–932.
- Butters, R. A., Sacks, D. W., and Seo, B. (2022). How do national firms respond to local cost shocks? *American Economic Review*, 112(5):1737–1772.
- Campante, F. and Yanagizawa-Drott, D. (2018). Long-range growth: economic development in the global network of air links. *The Quarterly Journal of Economics*, 133(3):1395–1458.

- Catalini, C., Fons-Rosen, C., and Gaulé, P. (2020). How do travel costs shape collaboration? *Management Science*, 66(8):3340–3360.
- Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.
- Charnoz, P., Lelarge, C., and Trevien, C. (2018). Communication costs and the internal organisation of multi-plant businesses: evidence from the impact of the French high-speed rail. *The Economic Journal*, 128(610):949–994.
- Coscia, M., Neffke, F. M., and Hausmann, R. (2020). Knowledge diffusion in the network of international business travel. *Nature Human Behaviour*, 4(10):1011–1020.
- Giroud, X. (2013). Proximity and investment: Evidence from plant-level data. *Quarterly Journal of Economics*, 128(2):861–915.
- Giroud, X. and Mueller, H. M. (2015). Capital and labor reallocation within firms. *The Journal of Finance*, 70(4):1767–1804.
- Gopinath, G., Kalemli-Özcan, , Karabarbounis, L., and Villegas-Sanchez, C. (2017). Capital allocation and productivity in south europe. *Quarterly Journal of Economics*, 132(4):1915–1967.
- Gumpert, A., Steimer, H., and Antoni, M. (2022). Firm Organization with Multiple Establishments. *Quarterly Journal of Economics*, 137(2):1091–1138.
- Helpman, E., Itskhoki, O., and Redding, S. (2010). Inequality and unemployment in a global economy. *Econometrica*, 78(4):1239–1283.
- Helpman, E., Melitz, M. J., and Yeaple, S. R. (2004). Export versus FDI with heterogeneous firms. *American economic review*, 94(1):300–316.
- Hsieh, C.-T. and Klenow, P. J. (2009). Misallocation and manufacturing TFP in China and India. *The Quarterly Journal of Economics*, 124(4):1403–1448.
- Hsieh, C. T. and Moretti, E. (2019). Housing constraints and spatial misallocation. *American Economic Journal: Macroeconomics*, 11(2):1–39.
- Kalnins, A. and Lafontaine, F. (2013). Too far away? The effect of distance to headquarters on business establishment performance. *American Economic Journal: Microeconomics*, 5(3):157–179.
- Markusen, J. R. (1984). Multinationals, multi-plant economies, and the gains from trade. *Journal of international economics*, 16(3-4):205–226.
- McGraw, M. J. (2020). The role of airports in city employment growth, 1950–2010. *Journal of Urban Economics*, 116(March 2017):103240.

- McKinsey&Company (2010). Estudo do setor de transporte aéreo do Brasil: Relatório consolidado.
- Menezes-Filho, N. A., Muendler, M.-A., and Ramey, G. (2008). The structure of worker compensation in Brazil, with a comparison to France and the United States. *The Review of Economics and Statistics*, 90(2):324–346.
- Midrigan, V. and Xu, D. Y. (2014). Finance and misallocation: Evidence from plant-level data. *American Economic Review*, 104(2):422–458.
- Morten, M. and Oliveira, J. (2024). The effects of roads on trade and migration: Evidence from a planned capital city. *American Economic Journal: Applied Economics*, 16(2):389–421.
- Muendler, M.-A., Rauch, J. E., and Tocoian, O. (2012). Employee spinoffs and other entrants: Stylized facts from Brazil. *International Journal of Industrial Organization*, 30(5):447–458.
- Oldenski, L. (2012). Export Versus FDI and the Communication of Complex Information. *Journal of International Economics*, 87(2):312–322.
- Oliveira, R. P., Oliveira, A. V., Lohmann, G., and Bettini, H. F. (2020). The geographic concentrations of air traffic and economic development: A spatiotemporal analysis of their association and decoupling in Brazil. *Journal of Transport Geography*, 87:102792.
- Poole, J. P. (2013). Knowledge transfers from multinational to domestic firms: Evidence from worker mobility. *Review of Economics and Statistics*, 95(2):393–406.
- Sheard, N. (2014). Airports and urban sectoral employment. *Journal of Urban Economics*, 80:133–152.
- Tveter, E. (2017). The effect of airports on regional development: Evidence from the construction of regional airports in Norway. *Research in Transportation Economics*, 63:50–58.
- Wing, C., Freedman, S. M., and Hollingsworth, A. (2024). Stacked difference-in-differences. Technical report, National Bureau of Economic Research.
- Zhang, F. and Graham, D. J. (2020). Air transport and economic growth: a review of the impact mechanism and causal relationships. *Transport Reviews*, 40(4):506–528.

Appendix

Table of Contents

A	Data & Sample Construction	1
A1	Raw data sources	1
A2	Data cleaning and sample selection procedure	4
A3	Headquarters identification	4
A4	Sector and occupation classification	5
B	Additional Figures & Tables	7
B1	Contribution of multi-unit firms in Brazil by year	7
B2	Descriptive statistics of main estimation sample	8
B3	Additional heterogeneity analyses of the main effects	9
B	Robustness Tests	15
B1	Robustness to different establishment samples	15
B2	Robustness to different worker type definitions	16
B3	Robustness to alternative stacking procedures	17
B4	Robustness to different estimation methods	22
C	Model Appendix	24
C1	Manager's Employment Choice and the Over-Hiring Cutoff	24
C2	Headquarters' choice of HQ	26
C3	Employment Comparative Statics	27
C4	Serving the Market from Headquarters	29
C5	Mode Choice: Three Channels and a Sufficient Condition	29
C6	Entry Cutoff	30
C7	Proofs of Propositions	31

A Data & Sample Construction

A1 Raw data sources

A1.1 Firm data

RAIS (Relação Anual de Informações Sociais) is the annual census of all formal employment contracts in Brazil, in which every establishment electronically reports the universe of active and terminated contracts held during the reference year. Each record includes worker identifiers, establishment and firm CNPJs, occupation, wages, hours contracted, admission and separation dates, and a comprehensive set of establishment characteristics.

To create establishment-level variables, we aggregate worker-level RAIS data for every year between 1995 and 2017. First, we retain a single observation per worker–establishment (worker ID–CNPJ) contract spell within each year. A “spell” here refers to a unique contract entry (a row in RAIS); a worker may have more than one contract in the same establishment within a year. When that occurs, we keep the last; if two contracts end in the same month, we retain the one with highest wage. We drop any contract with zero wages, in line with the existing literature. Then, we aggregate these contracts to the establishment-year level. For this, we consider only contracts that are active in December of that year. Total employment and average wages over the month of December are calculated, dropping duplicate worker-CNPJ pairs.

A1.2 Airline routes

Our main source of information on air connectivity consists of flight data provided by the Brazilian National Civil Aviation Service (*Agência Nacional de Aviação Civil*, ANAC).¹⁶ The regulatory agency compiles detailed information on flights operated inside Brazilian borders on a monthly basis. For each flight, they report the airport and date of departure and arrival, airline, type of service offered (cargo or passenger), type of airplane, number of seats offered and purchased, cargo weight offered and purchased, distance covered, fuel consumption, and hours flown.

We use data from 2000 to 2017 to construct a panel dataset of monthly flights between an origin-destination municipality pair by airline and type of service offered. We focus only on regular flights whose origin and destination are inside Brazil. According to the ANAC,

¹⁶The datasets can be obtained [here](#) and a description of variables is available [here](#).

regular flights are those with the following characteristics: (i) carried out to transport passengers or cargo; (ii) with services available for purchase by the public; (iii) operated to ensure traffic between two or more airports, either according to a published timetable or through flights that, due to their regularity or frequency, constitute a systematic series.¹⁷ These exclude chartered flights, flights carried out for maintenance or transportation of employees, and any extra flights established by airlines due to particular situations.

Using only regular, passenger flights, we first collapse the flight logs to the origin-destination-month-year level, aggregating flights operated between the same municipality pair ("route") in a month and counting the number of months in which at least one flight is operated. We then define a route as "active" in a calendar year if it operates at least one flight every month for at least 6 months that year. These need not be consecutive months.

Next, we compute the shortest feasible itinerary between every arranjo pair that is serviced by an airport, as follows:

1. Keep only our active routes. That is, we treat routes that do not operate flights at least half the year as not having direct flights in that calendar year;
2. Generate the full matrix of municipality origin–destination combinations;
3. For combinations without a direct leg, search for one-stop and two-stop connections, adding 1 hour per layover to the published block time; any path requiring 3+ stops is coded as no air connection.

Because airports in adjacent municipalities may service the same commuting zone (i.e., "arranjo populacional", see section A1.3), we aggregate these results to the arranjo origin–arranjo destination level. Within an arranjo pair, we keep the direct flight when it exists; otherwise we select the connection with the shortest total travel time (including layovers). Pairs inside the same arranjo or outside arranjos are discarded.

The resulting panel (arranjo-pair $\times$ year) contains:

- origin arranjo
- destination arranjo
- calendar year
- connection indicator (1 = a route of at most two layovers exists, 0 = none)

¹⁷Description provided [here](#).

- shortest total flight time
- number of layovers
- municipality of layover

A1.3 *Arranjos populacionais*

Arranjos populacionais are harmonized urban commuting zones defined by IBGE (Brazilian Institute of Geography and Statistics), which cluster neighboring municipalities based on strong demographic integration.

We use IBGE's 2015 definition, which identifies 478 unique arranjos. We define the central coordinates for each arranjo as the centroid of its most populous municipality.

A2 Data cleaning and sample selection procedure

Procedure	Num. Firms	Num. Estab.
	8,474,967	11,460,350
<i>1. Data cleaning</i>		
Drop invalid identifiers	8,474,966	11,460,348
Keep only establishments that existed in 2000 or later	6,837,081	8,997,734
Keep only firms that are always CNPJ-registered	6,823,323	7,620,987
Drop if establishment or HQ are ever in munis created after 2012	6,822,777	7,620,396
<i>2. Sample selection</i>		
Drop if firms are always single-establishment	322,082	1,109,554
Keep only firms with employees in Dec. 2000	132,005	586,751
Drop firms that are not always businesses entities	125,191	503,036
Drop firms with always less than 10 employees	91,907	429,844
Drop if firms are never multi-establishment across arranjos	43,055	297,318
Keep only firms with HQs in arranjo with airport	29,762	227,191
<i>3. Estimation sample</i>		
Drop firms where all establishments are out of sample	19,758	192,073
Drop establishments that are always out of sample	19,758	96,458
Drop establishments that were always treated	10,531	35,976
Drop establishments that were partly treated	8,441	26,787
Drop control establishments with a time reduction > 30 mins	8,080	24,195
Drop treated establishments that we do not observe at t-1	8,054	24,081
Drop treated estabs that switched arranjo from t-1 to t0	8,022	23,999
Drop treated estabs with HQ that switched arranjo from t-1 to t0	7,996	23,901
Assign control establishments to each treatment stack	3,938	10,291
Drop estabs in stacks missing treated or control	3,938	10,288
	3,938	10,288

Notes: Number of unique firms and establishments after each step of the sample selection procedure. Out of sample means that either: (i) the establishment is in the same arranjo as HQ; (ii) the establishment is HQ; (iii) the establishment is located in an arranjo that does not have an airport. Always treated means the establishment has a direct flight to headquarters since the beginning of the panel. Partly treated means the establishment had an inconsistent direct flight to headquarters, which was not retained throughout the panel.

A3 Headquarters identification

In line with the literature, we use each establishment's CNPJ (*Cadastro Nacional de Pessoas Jurídicas*, or National Registry of Legal Entities) number to identify the firm's headquarters and non-headquarter branches. Headquarter CNPJs follow the pattern XX.XXX.XXX/0001-XX, where the first 8 digits constitute the firm identifier and the sequence "0001" denotes the head office. A non-headquarter branch, on the other hand, replaces "0001" with a sequential code ("0002", "0003", etc), usually indicating their order among the firm's units.

However, roughly 3% of the establishments did not report a headquarters in at least

one year between 2000 and 2017 (and 0.30% did not have one in any year). The majority of these firms (roughly 85%) are single-establishment in the years in which they are missing a reported headquarters; thus, we assign these establishments to be headquarters in these years. For the remaining, we sequentially assign headquarters status to the establishment with the largest number of managers, the highest total payroll, the greatest number of total employees, or the oldest establishment. This definition overlaps with the CNPJ definition 98% of the times. Appendix Table [B11](#) reports the main results excluding these firms.

A4 Sector and occupation classification

A4.1 Sector

We assign each establishment to one of three broad sectors using the first two digits of the CNAE 1.0 class reported for its main economic activity. Divisions 01–14 (agriculture, forestry, fishing and mining) and 40–41 (electricity, gas and water) are grouped as *Agriculture*. Divisions 15–37 (all manufacturing industries) and 45 (construction) as classified as *Manufacturing/Industry*. Every remaining division (50–99), covering wholesale/retail, transport, finance, real estate and other services is classified as *Services*.

A4.2 Occupation

Worker occupations are harmonized to the six-digit CBO-02. For years 2000–2002, the original CBO-94 codes are cross-walked to CBO-02. Each code is then mapped into four mutually exclusive categories:

1. Managers and directors (CBO-02 first digit = 1)
2. Professionals (CBO-02 first digit = 2)
3. Technicians and administrators (CBO-02 first digit = 3 or 4)
4. Service and production workers (CBO-02 first digit = 5–9)

In our main analyses, we combine categories 1, 2 and 3 into a general “professionals” group.

Additionally, we build a complementary hierarchy variable that splits workers into managers or supervisors (CBO-02 first digit = 1 *or* third-digit = 0), other white-collar workers (CBO-02 first digit = 2–4, excluding third-digit = 0), and again blue-collar service

and production workers (CBO-02 first digit = 5–9, excluding third-digit = 0). Robustness results using this split are presented in Appendix [B](#).

B Additional Figures & Tables

B1 Contribution of multi-unit firms in Brazil by year

Table B1: Contribution of Multi-Unit Firms in Brazil, year by year

Year	Metric	All Multi-Unit Firms		Estimation Sample	
		Count	% of Total	Count	% of Total
2000	Firms	49,805	3.07	976	0.06
	Establishments	181,056	9.98	1,999	0.11
	Employment	8,263,682	33.75	97,195	0.40
2001	Firms	21,609	7.54	475	0.17
	Establishments	78,693	24.80	1,002	0.32
	Employment	8,471,444	32.65	112,851	0.43
2002	Firms	16,707	6.25	337	0.13
	Establishments	57,561	19.52	686	0.23
	Employment	8,844,284	32.33	126,316	0.46
2003	Firms	13,421	5.39	232	0.09
	Establishments	44,323	16.12	508	0.18
	Employment	8,905,959	31.64	125,985	0.45
2004	Firms	12,565	4.84	210	0.08
	Establishments	41,551	14.44	464	0.16
	Employment	9,517,139	31.88	134,381	0.45
2005	Firms	12,143	4.54	213	0.08
	Establishments	40,722	13.75	557	0.19
	Employment	10,028,099	31.75	147,865	0.47
2006	Firms	12,442	4.73	209	0.08
	Establishments	41,948	14.25	533	0.18
	Employment	10,882,822	32.33	166,845	0.50
2007	Firms	13,039	4.83	159	0.06
	Establishments	44,220	14.62	527	0.17
	Employment	11,605,706	32.61	180,133	0.51
2008	Firms	15,820	5.22	214	0.07
	Establishments	52,330	15.31	582	0.17
	Employment	12,332,051	33.01	194,866	0.52
2009	Firms	15,990	5.07	153	0.05
	Establishments	53,683	15.14	516	0.15
	Employment	12,617,975	32.31	192,088	0.49
2010	Firms	17,730	4.98	173	0.05
	Establishments	57,929	14.56	597	0.15
	Employment	13,773,582	32.92	219,405	0.52
2011	Firms	18,040	4.82	164	0.04
	Establishments	60,676	14.52	506	0.12
	Employment	14,668,940	33.39	237,286	0.54
2012	Firms	17,819	5.00	106	0.03
	Establishments	61,996	15.43	474	0.12
	Employment	15,094,242	33.43	259,011	0.57
2013	Firms	17,935	4.80	92	0.02
	Establishments	61,459	14.72	403	0.10
	Employment	15,666,150	33.63	260,228	0.56
2014	Firms	16,805	4.68	115	0.03
	Establishments	56,719	14.24	391	0.10
	Employment	15,795,813	34.01	263,362	0.57
2015	Firms	15,633	5.00	53	0.02
	Establishments	53,259	15.21	294	0.08
	Employment	15,211,389	33.89	242,676	0.54
2016	Firms	15,545	5.37	38	0.01
	Establishments	52,572	16.18	189	0.06
	Employment	14,639,161	34.01	211,197	0.49
2017	Firms	19,034	6.40	19	0.01
	Establishments	67,586	20.22	60	0.02
	Employment	14,697,097	34.01	184,845	0.43

Notes: 2000 cross-section. Columns 2 & 4 show the counts of unique firms and establishments, and number of total employees in the year for all multi-unit firms and for our estimation sample, respectively. Columns 3 & 5 show each group's share (in %) of the total number of firms, establishments, or employees in Brazil during the same period.

B2 Descriptive statistics of main estimation sample

Table B2: Balance table of main estimation sample

Variable	(1) Total		(2) Control		(3) Treated		(2)-(3) Pairwise t-test	
	N	Mean/(SE)	N	Mean/(SE)	N	Mean/(SE)	N	Mean difference
Survive to t+1	10,233	1.00 (0.00)	7,647	1.00 (0.00)	2,586	1.00 (0.00)	.n	.n
Employees	10,233	35.20 (1.45)	7,647	33.45 (1.64)	2,586	40.39 (3.11)	10,233	-6.94**
Professionals	10,151	15.63 (0.69)	7,586	14.50 (0.74)	2,565	18.95 (1.61)	10,151	-4.45***
Production workers	10,151	19.82 (1.12)	7,586	19.17 (1.27)	2,565	21.74 (2.34)	10,151	-2.57
Professional share	10,151	0.58 (0.00)	7,586	0.57 (0.00)	2,565	0.61 (0.01)	10,151	-0.04***
Average earnings	10,161	2,225.37 (21.43)	7,593	2,195.00 (23.94)	2,568	2,315.17 (46.66)	10,161	-120.17**
Avg. manager earnings	9,496	2,371.79 (22.66)	7,079	2,347.25 (25.56)	2,417	2,443.67 (48.21)	9,496	-96.42*
Avg. nonmanager earnings	7,310	1,666.10 (17.29)	5,548	1,637.78 (18.67)	1,762	1,755.26 (41.08)	7,310	-117.48***
Distance to HQ (km)	10,233	1,046.15 (8.31)	7,647	1,033.15 (9.63)	2,586	1,084.60 (16.44)	10,233	-51.45***
No. of affiliates in same city	10,233	3.35 (0.10)	7,647	3.87 (0.13)	2,586	1.81 (0.12)	10,233	2.06***
Total affiliates	10,233	144.01 (2.85)	7,647	157.97 (3.48)	2,586	102.75 (4.55)	10,233	55.22***
HQ employees	10,180	495.88 (13.11)	7,615	481.16 (15.06)	2,565	539.56 (26.57)	10,180	-58.39*
HQ professionals	9,963	381.01 (12.02)	7,443	375.66 (14.24)	2,520	396.83 (22.13)	9,963	-21.18

Notes: Balance test for treated and control establishments used in the main estimation sample.

B3 Additional heterogeneity analyses of the main effects

Table B3: Causal Effect of Direct Air Connection to HQ by Sector

Sectors	Survive t+1			Log professionals			Log prod. workers		
	Agr. (1)	Manuf. (2)	Serv. (3)	Agr. (4)	Manuf. (5)	Serv. (6)	Agr. (7)	Manuf. (8)	Serv. (9)
Treated	-0.150* (0.068)	-0.135*** (0.021)	-0.124*** (0.007)	0.301 (0.256)	0.011 (0.074)	0.044* (0.021)	0.278 (0.431)	-0.314** (0.099)	-0.060* (0.025)
Observations	1,263	11,423	96,141	973	9,778	85,276	904	8,438	66,882
N. control	90	785	6,751	64	657	5,889	55	531	4,387
N. treated	38	325	2,204	26	289	1,945	30	238	1,378
R-squared	0.801	0.515	0.432	0.963	0.911	0.877	0.912	0.912	0.866
Estab. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the establishment sector at the year prior to treatment. Columns labeled *Agr.* [(1), (4), (7)] are the results for those establishments in agriculture, columns labeled *Manuf.* [(2), (5), (8)] show the effects on establishments in manufacturing, and columns *Serv.* [(3), (6), (9)] are for the service sector. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B4: Causal Effect of Direct Air Connection to HQ by Num. of Affiliates

Num. affiliates	Survive t+1		Log professionals		Log prod. workers	
	Below Median (1)	Above Median (2)	Below Median (3)	Above Median (4)	Below Median (5)	Above Median (6)
Treated	-0.143*** (0.013)	-0.118*** (0.007)	0.042 (0.041)	0.049* (0.024)	-0.109* (0.055)	-0.052 (0.028)
Observations	33,447	76,229	28,055	68,796	23,121	53,953
N. control	2,367	5,319	1,891	4,773	1,465	3,566
N. treated	826	1,756	696	1,579	515	1,148
R-squared	0.448	0.453	0.869	0.887	0.864	0.887
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the total number of affiliates at the firm-level in the year prior to treatment. Columns labeled below [(1), (3), (5)] are the results for those establishments below or at the median number, while columns labeled above [(2), (4), (6)] show the effects on establishments above the median number. The median number of affiliates corresponds to approximately 4. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B5: Causal Effect of Direct Air Connection to HQ by Num. of Affiliates in the Same Arranjo

Num. oth. affil. in same arranjo	Survive t+1		Log professionals		Log prod. workers	
	Below Median (1)	Above Median (2)	Below Median (3)	Above Median (4)	Below Median (5)	Above Median (6)
Treated	-0.129*** (0.008)	-0.100*** (0.010)	-0.005 (0.026)	0.098** (0.032)	-0.090** (0.033)	0.003 (0.042)
Observations	65,944	43,772	57,600	39,321	42,882	34,222
N. control	4,792	2,905	4,114	2,563	2,825	2,212
N. treated	1,763	820	1,552	726	1,080	584
R-squared	0.438	0.481	0.874	0.900	0.878	0.885
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the total number of other affiliates at the firm-level in the same arranjo the year prior to treatment. Columns labeled below [(1), (3), (5)] are the results for those establishments below or at the median number, while columns labeled above [(2), (4), (6)] show the effects on establishments above the median number. The median number of other affiliates in the same arranjo corresponds to approximately 0. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B6: Causal Effect of Direct Air Connection to HQ by Num. of Locations

Num. locations other than HQ	Survive t+1		Log professionals		Log prod. workers	
	Below Median (1)	Above Median (2)	Below Median (3)	Above Median (4)	Below Median (5)	Above Median (6)
Treated	-0.161*** (0.022)	-0.121*** (0.007)	0.023 (0.069)	0.052* (0.021)	-0.209* (0.102)	-0.063* (0.026)
Observations	15,009	94,450	12,027	84,591	10,265	66,467
N. control	1,050	6,621	809	5,849	653	4,352
N. treated	392	2,185	299	1,966	232	1,420
R-squared	0.513	0.436	0.872	0.883	0.869	0.878
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the total number of arranjos where the firm has establishments in the year prior to treatment. Columns labeled below [(1), (3), (5)] are the results for those establishments below or at the median number, while columns labeled above [(2), (4), (6)] show the effects on establishments above the median number. The median number of arranjos corresponds to approximately X. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B7: Causal Effect of Direct Air Connection to HQ by Num. of HQ Managers per Kilometer

Num. HQ managers per km	Survive t+1		Log professionals		Log prod. workers	
	Below Median (1)	Above Median (2)	Below Median (3)	Above Median (4)	Below Median (5)	Above Median (6)
Treated	-0.120*** (0.008)	-0.129*** (0.011)	0.046 (0.025)	0.014 (0.035)	-0.067* (0.031)	-0.079 (0.047)
Observations	67,520	39,045	59,643	34,650	48,356	26,818
N. control	4,582	2,843	3,991	2,489	3,107	1,796
N. treated	1,400	1,105	1,237	978	910	707
R-squared	0.431	0.480	0.878	0.895	0.869	0.890
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the number of managers at HQ divided by the total distance between HQ and all its affiliates (in km), in the year prior to treatment. Columns labeled below [(1), (3), (5)] are the results for those establishments below or at the median number, while columns labeled above [(2), (4), (6)] show the effects on establishments above the median number. The median number of managers in HQ per km corresponds to approximately X. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B8: Causal Effect of Direct Air Connection to HQ by Estab. Region

Estab. Region.	Survive t+1			Log professionals			Log prod. workers		
	N (1)	NE (2)	SE (3)	N (4)	NE (5)	SE (6)	N (7)	NE (8)	SE (9)
Treated	-0.060** (0.019)	-0.084*** (0.010)	-0.161*** (0.012)	0.118* (0.050)	0.104** (0.037)	0.008 (0.038)	-0.000 (0.068)	-0.110* (0.044)	-0.079 (0.056)
Observations	8,665	28,372	35,409	7,680	24,875	31,393	5,588	20,628	24,391
N. control	716	1,838	2,401	650	1,613	2,059	486	1,308	1,500
N. treated	256	820	799	240	759	672	188	549	488
R-squared	0.362	0.383	0.462	0.835	0.882	0.893	0.861	0.861	0.874
Estab. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Estab. Region.	Survive t+1		Log professionals		Log prod. workers	
	S (10)	CO (11)	S (12)	CO (13)	S (14)	CO (15)
Treated	-0.163*** (0.016)	-0.151*** (0.018)	-0.023 (0.051)	-0.009 (0.052)	-0.078 (0.050)	-0.033 (0.054)
Observations	24,825	12,520	22,007	11,066	17,654	9,010
N. control	1,844	901	1,573	786	1,162	594
N. treated	417	294	355	255	243	200
R-squared	0.464	0.395	0.872	0.865	0.891	0.861
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the region of location of the establishment the year prior to treatment. Columns labeled N [(1), (4), (7)] refer to establishments located in the North, columns labeled NE [(2), (5), (8)] refer to establishments located in the Northeast, columns labeled SE [(3), (6), (9)] refer to establishments located in the Southeast, columns labeled S [(10), (12), (14)] refer to establishments located in the South, and columns labeled CO [(11), (13), (15)] refer to establishments located in the Midwest. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B9: Causal Effect of Direct Air Connection to HQ by HQ Region

HQ Region.	Survive t+1			Log professionals			Log prod. workers		
	N (1)	NE (2)	SE (3)	N (4)	NE (5)	SE (6)	N (7)	NE (8)	SE (9)
Treated	-0.172*** (0.051)	-0.156*** (0.022)	-0.139*** (0.011)	0.127 (0.119)	0.190* (0.087)	0.030 (0.032)	0.034 (0.130)	-0.113 (0.091)	-0.123** (0.042)
Observations	4,415	18,992	48,750	3,804	16,624	43,867	3,184	15,321	32,977
N. control	313	1,272	3,344	259	1,126	2,954	222	937	2,032
N. treated	113	267	1,562	108	230	1,382	100	207	928
R-squared	0.507	0.471	0.472	0.894	0.888	0.895	0.865	0.882	0.899
Estab. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

HQ Region.	Survive t+1		Log professionals		Log prod. workers	
	S (10)	CO (11)	S (12)	CO (13)	S (14)	CO (15)
Treated	-0.245*** (0.020)	-0.099*** (0.028)	-0.124 (0.066)	0.033 (0.083)	-0.111 (0.073)	0.149 (0.134)
Observations	30,182	6,004	25,887	5,325	19,374	4,816
N. control	2,173	445	1,805	387	1,369	339
N. treated	434	186	372	166	267	144
R-squared	0.463	0.537	0.888	0.890	0.869	0.876
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 6, except that models are estimated in sub-samples of establishments based on the region of location of HQ the year prior to treatment. Columns labeled N [(1), (4), (7)] refer to establishments with HQ located in the North, columns labeled NE [(2), (5), (8)] refer to establishments with HQ located in the Northeast, columns labeled SE [(3), (6), (9)] refer to establishments with HQ located in the Southeast, columns labeled S [(10), (12), (14)] refer to establishments with HQ located in the South, and columns labeled CO [(11), (13), (15)] refer to establishments with HQ located in the Midwest. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

B Robustness Tests

B1 Robustness to different establishment samples

Table B10: Causal Effect of Direct Air Connection to HQ, Excluding Movers

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.136*** (0.007)	0.049* (0.021)	-0.068* (0.027)
Observations	94,857	82,032	64,698
N. control	6,685	5,648	4,228
N. treated	2,253	1,942	1,412
R-squared	0.439	0.885	0.876
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, except that it excludes establishments that switched locations at least once during the panel or establishments whose HQs switched locations at least once. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B11: Causal Effect of Direct Air Connection to HQ, Excluding Establishments Without CNPJ-identified HQs

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.122*** (0.007)	0.018 (0.021)	-0.075** (0.027)
Observations	92,908	82,296	64,591
N. control	6,512	5,633	4,237
N. treated	2,200	1,947	1,439
R-squared	0.440	0.887	0.879
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, except that it excludes establishments whose HQs are not identified via their CNPJ number, as detailed in Appendix A. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

B2 Robustness to different worker type definitions

Table B12: Causal Effect of Direct Air Connection to HQ, With Alternative Worker Categories

	Log mng./super.	Log white other	Log prod. workers
Treated	-0.013 (0.020)	0.044* (0.021)	-0.070** (0.025)
Observations	76,249	91,940	74,784
N. control	4,800	6,169	4,862
N. treated	1,611	2,121	1,595
R-squared	0.838	0.873	0.872
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table uses an alternative hierarchical categorization of worker types, as described in Appendix A. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B13: Causal Effect of Direct Air Connection to HQ, by CBO 1-digit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	-0.004 (0.019)	0.031 (0.038)	0.016 (0.032)	0.071** (0.022)	-0.072** (0.025)	0.348 (0.323)	-0.026 (0.039)	-0.053 (0.115)	-0.091 (0.051)
Observations	63,804	24,171	43,419	84,019	63,704	829	39,127	6,773	13,490
N. control	3,945	1,668	2,596	5,383	4,023	58	2,187	428	779
N. treated	1,255	650	993	1,877	1,305	20	747	159	320
R-squared	0.823	0.865	0.854	0.859	0.871	0.927	0.890	0.900	0.903
Estab. FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Notes: This table divides employees by the first-digit CBO-2002 class. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

B3 Robustness to alternative stacking procedures

Table B14: Causal Effect of Direct Air Connection to HQ, Controls Assigned to First Stack

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.098*** (0.008)	-0.035 (0.022)	-0.104*** (0.028)
Observations	56,917	49,272	38,335
N. control	7,691	6,671	5,034
N. treated	2,555	2,249	1,633
R-squared	0.491	0.883	0.879
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in their first stack. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B15: Causal Effect of Direct Air Connection to HQ, Controls Assigned to First Stack

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.062*** (0.008)	0.023 (0.022)	-0.082** (0.027)
Observations	56,587	48,804	38,144
N. control	7,690	6,673	5,068
N. treated	2,569	2,259	1,646
R-squared	0.523	0.885	0.881
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in the stack with the fewest number of controls. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B16: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.116*** (0.008)	0.010 (0.023)	-0.100** (0.031)
Observations	23,696	21,117	16,222
N. control	2,092	1,830	1,348
N. treated	2,099	1,825	1,335
R-squared	0.446	0.879	0.882
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers distance to HQ, size of the establishment and age of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B17: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Distance to HQ

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.106*** (0.008)	0.023 (0.023)	-0.098** (0.031)
Observations	23,535	21,022	16,081
N. control	2,074	1,825	1,345
N. treated	2,092	1,824	1,341
R-squared	0.446	0.883	0.885
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers solely distance to HQ at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B18: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Establishment Age

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.116*** (0.008)	0.004 (0.023)	-0.086** (0.030)
Observations	23,763	21,184	16,453
N. control	2,111	1,851	1,366
N. treated	2,116	1,847	1,343
R-squared	0.445	0.882	0.885
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers solely age of the at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B19: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Establishment Size

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.114*** (0.008)	0.032 (0.023)	-0.118*** (0.030)
Observations	23,781	21,201	16,209
N. control	2,082	1,825	1,335
N. treated	2,085	1,820	1,335
R-squared	0.447	0.878	0.886
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers solely size of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B20: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Distance to HQ and Establishment Age

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.116*** (0.008)	0.020 (0.023)	-0.105*** (0.031)
Observations	23,640	21,077	16,289
N. control	2,091	1,833	1,368
N. treated	2,093	1,823	1,331
R-squared	0.447	0.884	0.886
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers distance to HQ and age of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B21: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Distance to HQ and Establishment Size

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.110*** (0.008)	0.020 (0.023)	-0.102** (0.031)
Observations	23,587	21,050	16,088
N. control	2,074	1,820	1,342
N. treated	2,095	1,825	1,337
R-squared	0.442	0.880	0.881
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers distance to HQ and size of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B22: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 on Establishment Age and Size

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.113*** (0.008)	0.009 (0.023)	-0.082** (0.030)
Observations	23,713	21,091	16,169
N. control	2,107	1,850	1,345
N. treated	2,110	1,838	1,344
R-squared	0.442	0.881	0.885
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, but only keeps controls in one stack. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers size and age of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B23: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 by Distance to HQ

Dist. to HQ	Survive t+1		Log professionals		Log prod. workers	
	(Below)	(Above)	(Below)	(Above)	(Below)	(Above)
Treated	-0.154*** (0.013)	-0.088*** (0.011)	-0.038 (0.039)	0.031 (0.033)	-0.046 (0.052)	-0.134** (0.049)
Observations	10,992	11,770	9,390	10,883	7,490	7,956
N. control	926	1,063	785	955	598	669
N. treated	1,062	971	866	895	662	616
R-squared	0.524	0.393	0.888	0.880	0.884	0.891
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 4, except that models are estimated in sub-samples of establishments based on their geodesic distance to their headquarters in the year prior to treatment. Columns labeled below [(1), (3), (5)] are the results for those establishments below or at the median distance, while columns labeled above [(2), (4), (6)] show the effects on establishments above the median distance. The median distance to headquarters corresponds to approximately 812km. It only keeps controls in one stack, and stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers distance to HQ, size of the establishment and age of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B24: Causal Effect of Direct Air Connection to HQ, Controls Matched 1:1 (2003-2012)

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.121*** (0.009)	0.030 (0.029)	-0.131** (0.041)
Observations	18,103	15,547	11,790
N. control	1,399	1,227	879
N. treated	1,407	1,226	863
R-squared	0.418	0.871	0.876
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, except it only keeps controls in one stack and restricts the sample to stacks that were treated between 2003 and 2012. Stacking is done through an algorithm that matches controls 1:1 to individual treated establishments, and assigns them to that stack. The match is based on a similarity score that considers distance to HQ, size of the establishment and age of the establishment at the year prior to treatment, using normalized Euclidean distance. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

B4 Robustness to different estimation methods

Table B25: Causal Effect of Direct Air Connection to HQ, Interacted with Number of Affiliates in the Same Arranjo

	Survive t+1	Log professionals	Log prod. workers
Treated	-0.128*** (0.007)	0.043* (0.020)	-0.087*** (0.026)
Treated X Affiliates	0.001** (0.000)	0.001 (0.001)	0.004** (0.002)
Observations	109,791	97,021	77,271
N. control	7,700	6,681	5,050
N. treated	2,586	2,281	1,668
R-squared	0.431	0.880	0.872
Estab. FE	Y	Y	Y
City-year FE	Y	Y	Y

Notes: This table is parallel to Table 4, except that models are estimated with an interaction term that captures the marginal effect of having an additional establishment in the same arranjo. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level.

Table B26: Causal Effect of Direct Air Connection to HQ, Interacted with Number of Affiliates in the Same Arranjo by Distance to HQ

Dist. to HQ	Survive t+1		Log professionals		Log prod. workers	
	(1)	(2)	(1)	(2)	(1)	(2)
Treated	-0.185*** (0.012)	-0.085*** (0.009)	-0.004 (0.033)	0.071* (0.031)	-0.018 (0.042)	-0.106* (0.041)
Treated X Affiliates	0.001** (0.000)	-0.002 (0.001)	0.005** (0.002)	-0.007* (0.003)	0.000 (0.002)	-0.004 (0.004)
Observations	51,704	55,153	45,081	49,374	38,467	36,368
N. control	3,647	3,657	3,097	3,224	2,509	2,204
N. treated	1,231	1,295	1,026	1,196	772	846
R-squared	0.498	0.411	0.897	0.878	0.889	0.875
Estab. FE	Y	Y	Y	Y	Y	Y
City-year FE	Y	Y	Y	Y	Y	Y

Notes: This table is parallel to Table 4, except that models are estimated with an interaction term that captures the marginal effect of having an additional establishment in the same arranjo, and in sub-samples of establishments based on their geodesic distance to headquarters in the year prior to treatment. Columns (1) are the results for those establishments below or at the median distance, while columns (2) show the effects on establishments above the median distance. The median distance to headquarters corresponds to approximately 825km. All variables are winsorized at the first and 99th percentiles. Establishment and year fixed effects are estimated by stack. Standard errors clustered at the establishment level

C Model Appendix

This appendix presents derivations and proofs of the results presented in the main part of the text. We start by recalling the key assumptions:

- Distance raises client–service costs, formally: $\tau'(d) > 0$ and $\tau''(d) \geq 0$. We normalize $\tau(0) = 1$.
- Internal coordination costs also increase with distance and are strictly convex in HQ : $\partial\kappa/\partial d > 0$, $\partial^2\kappa/\partial d^2 \geq 0$, $\partial\kappa/\partial HQ > 0$, $\partial^2\kappa/\partial HQ^2 > 0$, and $\partial^2\kappa/\partial d \partial HQ \geq 0$.
- Distance also strengthens the manager’s over-hiring incentive: $\gamma(d) \geq 0$ and $\gamma'(d) \geq 0$.
- Wages $w > 0$, the manager’s share $\beta \in (0, 1)$, and the detection slope $\delta > 0$ are fixed.

We solve the model by backward induction, starting with the manager’s choice of L given (HQ, d) ; then headquarters’ choice of HQ ; then equilibrium branch employment; and finally the decisions of mode choice and entry.

C1 Manager’s Employment Choice and the Over-Hiring Cutoff

The manager’s utility is given by:

$$U_m(L; HQ) = \begin{cases} \beta(L^\alpha HQ^{1-\alpha} - wL) + \gamma(d)\tilde{L} - \delta HQ\tilde{L} & \text{if } L > \hat{L}, \\ \beta(L^\alpha HQ^{1-\alpha} - wL) & \text{otherwise,} \end{cases} \quad (9)$$

where $\hat{L} = (\alpha/w)^{1/(1-\alpha)}HQ$ is the optimal labor choice under no moral hazard (i.e., the choice that maximizes the second case of equation 9), and $\tilde{L} \equiv L - \hat{L}$ denotes over-hiring.

For $L > \hat{L}$, the utility-maximizing choice of employment is:

$$\frac{\partial U_m}{\partial L} = \beta(\alpha L^{\alpha-1} HQ^{1-\alpha} - w) + \gamma(d) - \delta HQ.$$

Evaluated at $L = \hat{L}$, the bracket equals zero, so over-hiring occurs if and only if $\gamma(d) - \delta HQ > 0$. Define

$$HQ^{\text{crit}}(d) \equiv \frac{\gamma(d)}{\delta}.$$

When $HQ \geq HQ^{\text{crit}}(d)$, the manager chooses to hire $\hat{L}$. When $HQ < HQ^{\text{crit}}(d)$, the first-order condition $\partial U_m/\partial L = 0$ implies the best response function

$$L^\dagger(HQ; d) = HQ \left(\frac{\alpha\beta}{\beta w - \gamma(d) + \delta HQ} \right)^{\frac{1}{1-\alpha}}, \quad (10)$$

Note that the over-hiring best response in (10) is well-defined as long as

$$\beta w - \gamma(d) + \delta HQ > 0.$$

If this inequality fails, the manager's payoff is unbounded above in L , and such a branch would never be chosen in equilibrium. In what follows we restrict attention to the region where the inequality holds.

Lemma 1 (Continuity and monotonicity of $L^\dagger$). *Fix d . With $L^\dagger(HQ; d)$ given by equation (10) for $HQ < HQ^{\text{crit}}(d)$ and $L^\dagger(HQ; d) = \widehat{L}$ otherwise. Then:*

(i) **Continuity at $HQ^{\text{crit}}(d)$.** $L^\dagger(\cdot; d)$ is continuous on $[0, \infty)$ and $\lim_{HQ \uparrow HQ^{\text{crit}}(d)} L^\dagger(HQ; d) = \widehat{L}(HQ^{\text{crit}}(d))$.

(ii) **Monotonicity in HQ .** On $HQ \geq HQ^{\text{crit}}(d)$, $L^\dagger = \widehat{L}$ is linear and strictly increasing in HQ . On $HQ < HQ^{\text{crit}}(d)$,

$$\frac{\partial L^\dagger}{\partial HQ}(HQ; d) = \left(\frac{\alpha\beta}{\beta w - \gamma(d) + \delta HQ} \right)^{\frac{1}{1-\alpha}} \left[1 - \frac{1}{1-\alpha} \frac{\delta HQ}{\beta w - \gamma(d) + \delta HQ} \right],$$

which is strictly decreasing in HQ and crosses zero once at $\bar{H}(d) = \frac{1-\alpha}{\alpha} \frac{\beta w - \gamma(d)}{\delta}$ (if $\bar{H}(d) < HQ^{\text{crit}}(d)$). Thus, $L^\dagger$ is strictly increasing on $[0, \min\{\bar{H}(d), HQ^{\text{crit}}(d)\})$.

Proof. (i) **Continuity.** From equation (10), as HQ increases to $HQ^{\text{crit}}(d)$ the denominator $\beta w - \gamma(d) + \delta HQ$ converges to βw . Hence

$$\lim_{HQ \uparrow HQ^{\text{crit}}(d)} L^\dagger(HQ; d) = HQ^{\text{crit}}(d) \left(\frac{\alpha\beta}{\beta w} \right)^{\frac{1}{1-\alpha}} = \widehat{L}(HQ^{\text{crit}}(d)),$$

and continuity elsewhere follows.

(ii) **Monotonicity.** On $HQ \geq HQ^{\text{crit}}(d)$, $L^\dagger = \widehat{L} = (\alpha/w)^{1/(1-\alpha)} HQ$ is linear and strictly increasing. For $HQ < HQ^{\text{crit}}(d)$, differentiating equation (10) yields

$$\frac{\partial L^\dagger}{\partial HQ}(HQ; d) = \left(\frac{\alpha\beta}{\beta w - \gamma(d) + \delta HQ} \right)^{\frac{1}{1-\alpha}} \left[1 - \frac{1}{1-\alpha} \frac{\delta HQ}{\beta w - \gamma(d) + \delta HQ} \right].$$

The bracketed term is strictly decreasing in HQ , equals 1 at $HQ = 0$, and disappears once at $\bar{H}(d) = \frac{1-\alpha}{\alpha} \frac{\beta w - \gamma(d)}{\delta}$. Therefore the derivative is positive for $HQ < \bar{H}(d)$ and negative for $HQ > \bar{H}(d)$, implying the stated monotonicity on the over-hiring region and, in particular, that $L^\dagger$ is strictly increasing on $[0, \min\{\bar{H}(d), HQ^{\text{crit}}(d)\})$. □

Lemma 1 ensures the manager's reaction is continuous at $HQ^{\text{crit}}(d)$ and well behaved on each side. Together with the strict concavity of branch profits in HQ (next subsection),

this implies a unique interior maximizer $\widetilde{HQ}(d)$ whenever $\overline{HQ}$ does not bind and justifies the envelope derivatives used later for mode choice and entry.

C2 Headquarters' choice of HQ

Given the manager's reaction $L^\dagger(HQ; d)$ from equation (10), branch profits as a function of (d, HQ) are

$$\pi_b(d, HQ) = (1 - \beta) \left[(L^\dagger(HQ; d))^\alpha HQ^{1-\alpha} - w L^\dagger(HQ; d) \right] - \kappa(d, HQ).$$

On the region where the firm and the manager's incentives are aligned, $HQ \geq HQ^{\text{crit}}(d)$, $L^\dagger = \widehat{L}$ and the production term is linear in HQ . Given that κ is convex, the branch profit above is strictly concave.

On the over-hiring region $HQ < HQ^{\text{crit}}(d)$, it is convenient to define

$$\Omega_1(d, HQ) \equiv \frac{\alpha\beta}{\beta w - \gamma(d) + \delta HQ}, \quad \Omega_2(d, HQ) \equiv \frac{\delta HQ}{(1 - \alpha) [\beta w - \gamma(d) + \delta HQ]},$$

so that $L^\dagger = HQ \Omega_1^{1/(1-\alpha)}$. Differentiating yields the first-order condition

$$\frac{\partial \kappa(d, HQ)}{\partial HQ} = (1 - \beta) \left[\Omega_1^{\frac{\alpha}{1-\alpha}} (1 - \alpha \Omega_2) - w \Omega_1^{\frac{1}{1-\alpha}} (1 - \Omega_2) \right]. \quad (11)$$

To keep the comparative statics transparent and rule out non-monotone cases, we impose the following regularity condition on how the marginal benefit of headquarters varies with distance:

Assumption 1 (Monotone marginal benefit of HQ). *Along the relevant region, the marginal benefit of HQ does not rise with distance:*

$$\frac{\partial B(d, HQ)}{\partial d} \leq 0,$$

where $B(d, HQ)$ denotes the bracketed term on the right-hand side of equation (11).

We can then formalize the concavity and uniqueness result with the following lemma:

Lemma 2 (Concavity and uniqueness). *Assume $\frac{\partial^2 \kappa}{\partial HQ^2} > 0$. For any fixed d , $\pi_b(d, \cdot)$ is strictly concave in HQ on each region $[0, HQ^{\text{crit}}(d))$ and $[HQ^{\text{crit}}(d), \infty)$. Hence the interior optimum, when it exists, is unique; denote it by $\widetilde{HQ}(d)$, and define $HQ^*(d) \equiv \min\{\widetilde{HQ}(d), \overline{HQ}\}$.*

Proof. On $HQ \geq HQ^{\text{crit}}(d)$, $L^\dagger = \widehat{L}$ so the production term in π_b is linear in HQ ; strict concavity follows from the convexity of $\kappa(d, \cdot)$. On $HQ < HQ^{\text{crit}}(d)$, the marginal benefit

of HQ is the right-hand side of equation (11). As HQ rises, $\Omega_1(d, HQ)$ decreases (the denominator grows) while $\Omega_2(d, HQ)$ increases, so both components reduce the bracketed term. With $\kappa_{HH} \geq 0$, the net marginal benefit strictly falls in HQ , implying strict concavity and a unique interior maximizer. $\square$

Note also that, by Lemma 1, $L^\dagger$ meets $\widehat{L}$ continuously at $HQ^{\text{crit}}(d)$, and κ is smooth. Hence $\pi_b(d, HQ)$ is continuous in (d, HQ) , and the optimal value $\pi_b^*(d) \equiv \pi_b(d, HQ^*(d))$ is continuous in d . Whenever $HQ^*(d)$ hits either $HQ^{\text{crit}}(d)$ or $\overline{HQ}$, $\pi_b^*(d)$ is continuous but may not be differentiable. The envelope theorem and all derivatives below apply anywhere away from those points. Monotonicity of $HQ^*(d)$ in d is established in the proof of Proposition 2 (see Section C7).

When d is such that $HQ^*(d)$ crosses either $HQ^{\text{crit}}(d)$ or $\overline{HQ}$, $\pi_b^*(d)$ remains continuous but may exhibit a kink; all derivative statements are understood away from these points.

C3 Employment Comparative Statics

With the headquarters policy in hand, we can now fully characterize how equilibrium branch employment responds to distance. Let $L^*(d) \equiv L^\dagger(HQ^*(d); d)$ denote employment obtained by evaluating the manager's reaction at the optimal $HQ^*(d)$. Totally differentiating and using the fact that $L^\dagger$ depends on d only through $\gamma(d)$ yields

$$\frac{dL^*(d)}{dd} = \frac{\partial L^\dagger}{\partial HQ} \frac{\partial HQ^*}{\partial d} + \frac{\partial L^\dagger}{\partial \gamma} \gamma'(d), \quad (12)$$

where all partial derivatives of $L^\dagger$ are evaluated at $(HQ^*(d); d)$.

Aligned-incentives region ($HQ^*(d) \geq HQ^{\text{crit}}(d)$). When incentives are aligned and there is no over-hiring, the optimal labor choice is $L^\dagger = \widehat{L} = (\alpha/w)^{1/(1-\alpha)} HQ$, so

$$\frac{\partial L^\dagger}{\partial HQ} = (\alpha/w)^{\frac{1}{1-\alpha}} > 0, \quad \frac{\partial L^\dagger}{\partial \gamma} = 0,$$

and equation (12) reduces to

$$\frac{dL^*(d)}{dd} = (\alpha/w)^{\frac{1}{1-\alpha}} \frac{dHQ^*(d)}{dd}.$$

Thus, in this region, employment moves one-for-one (up to the constant factor) with headquarter inputs. In particular, by Proposition 2 (as proved in Section C7), $dHQ^*/dd \leq 0$, so $L^*(d)$ is weakly decreasing in d on this region.

Over-hiring region ($HQ^*(d) < HQ^{\text{crit}}(d)$). On this region the manager has incentives to over-hire. Two channels operate in equation (12). First the *moral hazard* channel ($\partial L^\dagger / \partial \gamma$) $\gamma'(d)$, which is weakly positive because $\partial L^\dagger / \partial \gamma > 0$ and $\gamma'(d) \geq 0$. Second, the *HQ* channel ($\partial L^\dagger / \partial HQ$) (dHQ^*/dd), whose sign depends on how the manager's desired L reacts to marginal HQ time.

By Lemma 1, $\partial L^\dagger / \partial HQ$ on the over-hiring region is single-peaked in HQ and changes sign once. Evaluated at the equilibrium input $HQ^*(d)$, we arrive at the sign test

$$\frac{\partial L^\dagger}{\partial HQ}(HQ^*(d); d) \geq 0 \iff \delta HQ^*(d) \leq \frac{1-\alpha}{\alpha} [\beta w - \gamma(d)]. \quad (13)$$

Define $\bar{d}$ as any distance at which equality holds in (13) (when such a distance exists). We call distances with $\partial L^\dagger / \partial HQ > 0$ *productivity-dominated* (marginal HQ mainly raises the marginal product of labor), and those with $\partial L^\dagger / \partial HQ < 0$ *discipline-dominated* (marginal HQ mainly disciplines over-hiring). Uniqueness of $\bar{d}$, when it exists, is established in Proposition 3 (see proof in Section C7).

Productivity-dominated regime ($\partial L^\dagger / \partial HQ > 0$; equivalently $d < \bar{d}$ when $\bar{d}$ exists). Here the HQ force in equation (12) is non-positive because $dHQ^*/dd \leq 0$ (Proposition 2). Intuitively, increasing distance lowers HQ^* and, by complementarity in production, lowers the marginal product of labor and thus the manager's desired L . Private benefits from over-hiring are, in turn, weakly positive. When the HQ (productivity/monitoring) effect dominates the moral hazard effect, $dL^*/dd \leq 0$, so better connectivity (lower d) raises employment.

Discipline-dominated regime ($\partial L^\dagger / \partial HQ < 0$; equivalently $d > \bar{d}$ when $\bar{d}$ exists). Here the role of HQ in equation (12) is strictly positive: reducing distance raises HQ^* , which tightens discipline and *reduces* the manager's desired L . Together with the weakly positive benefits from over-hiring, this implies $dL^*/dd > 0$: better connectivity lowers employment.

In sum, close to headquarters, an extra unit of HQ inputs mostly boosts production, so managers want more workers when HQ rises. Farther away from headquarters, the same HQ inputs mainly works as monitoring, so extra HQ curbs over-hiring and lowers employment. The cutoff $\bar{d}$ is where this marginal role flips.

C4 Serving the Market from Headquarters

When the firm serves the market from headquarters, it sets $HQ = \overline{HQ}$ and chooses L to maximize

$$\pi_{\text{HQ}}(d, L) = \frac{L^\alpha \overline{HQ}^{1-\alpha}}{\tau(d)} - wL.$$

Since coordination costs $\kappa(\cdot)$ do not apply in this mode and the production term is increasing in HQ , the capacity constraint binds, so the firm chooses $HQ = \overline{HQ}$.

The first-order condition in L gives

$$\alpha L_{\text{HQ}}^*(d)^{\alpha-1} \frac{\overline{HQ}^{1-\alpha}}{\tau(d)} = w \implies L_{\text{HQ}}^*(d) = \left(\frac{\alpha}{w \tau(d)} \right)^{\frac{1}{1-\alpha}} \overline{HQ}.$$

Using the envelope theorem,

$$\frac{d \pi_{\text{HQ}}^*(d)}{dd} = \frac{\partial \pi_{\text{HQ}}(d, L)}{\partial d} \Big|_{L=L_{\text{HQ}}^*(d)} = - \frac{L_{\text{HQ}}^*(d)^\alpha \overline{HQ}^{1-\alpha}}{\tau(d)^2} \tau'(d) < 0.$$

Thus, optimized profits from serving the market directly are strictly decreasing in distance (via $\tau'(d) > 0$).

C5 Mode Choice: Three Channels and a Sufficient Condition

Let $\Delta\pi(d) \equiv \pi_b^*(d) - \pi_{\text{HQ}}^*(d)$ denote the branch-profit premium. Differentiating and applying the envelope theorem,

$$\frac{d \Delta\pi(d)}{dd} = \underbrace{- \frac{\partial \kappa(d, HQ^*(d))}{\partial d}}_{\text{internal coordination}} + \underbrace{(1 - \beta) \Psi(d) \frac{d\gamma(d)}{dd}}_{\text{moral hazard } (\leq 0)} + \underbrace{[L_{\text{HQ}}^*(d)]^\alpha \overline{HQ}^{1-\alpha} \frac{\tau'(d)}{\tau(d)^2}}_{\text{client-service } (> 0)}, \quad (14)$$

where the moral-hazard envelope term $\Psi(d)$ is defined as:

$$\Psi(d) \equiv \frac{\partial}{\partial \gamma} \{ [L^\dagger(HQ^*(d); d)]^\alpha HQ^*(d)^{1-\alpha} - w L^\dagger(HQ^*(d); d) \} \leq 0.$$

Note $\Psi(d) \leq 0$ on the over-hiring region because $\alpha L^{\alpha-1} HQ^{1-\alpha} - w = -(\gamma(d) - \delta HQ)/\beta < 0$ by the manager's FOC and $\partial L^\dagger / \partial \gamma > 0$. In the aligned-incentive region $HQ^*(d) \geq HQ^{\text{crit}}(d)$ we have $L^\dagger = \hat{L}$, which does not depend on γ . Hence, $\partial L^\dagger / \partial \gamma = 0$ and therefore $\Psi(d) = 0$.

The first two terms in equation (14) push against branch operation as distance rises (coordination costs and a stronger over-hiring incentive); the last term pushes toward branch operation (serving from HQ becomes more costly due to client-facing frictions).

A sufficient condition for the client-service channel to dominate is:

$$\frac{\tau'(d)}{\tau(d)^2} > \frac{\frac{\partial \kappa(d, HQ^*(d))}{\partial d} - (1 - \beta) \Psi(d) \frac{d\gamma(d)}{dd}}{[L_{HQ}^*(d)]^\alpha \overline{HQ}^{1-\alpha}}.$$

The left-hand side is the (absolute) client–service slope that makes serving from HQ worse as distance rises. The right-hand side is the branch’s own distance penalty: coordination costs grow with d , and distance also worsens the agency problem (the manager is more tempted to over-hire), both of which make branches less attractive. When the left exceeds the right, $d\Delta\pi/dd > 0$, so improving connectivity (lower d) tilts the firm *away* from branches.

C6 Entry Cutoff

Define the optimized value function

$$\Pi(d) \equiv \max \{ \pi_b^*(d), \pi_{HQ}^*(d), 0 \}.$$

From equation (11) and Lemma 1, the branch profit schedule satisfies

$$\frac{d \pi_b^*(d)}{dd} = - \frac{\partial \kappa(d, HQ^*(d))}{\partial d} + (1 - \beta) \Psi(d) \frac{d\gamma(d)}{dd} < 0,$$

because $\partial \kappa / \partial d > 0$, $\Psi(d) \leq 0$, and $d\gamma/dd \geq 0$.

From Section C4, the HQ profits schedule satisfies

$$\frac{d \pi_{HQ}^*(d)}{dd} = - [L_{HQ}^*(d)]^\alpha \overline{HQ}^{1-\alpha} \frac{\tau'(d)}{\tau(d)^2} < 0,$$

since $\tau'(d) > 0$.

Branch profits move smoothly with distance: $L^\dagger$ meets $\hat{L}$ without a jump at $HQ^{\text{crit}}(d)$ and κ is smooth, so $\pi_b^*(d)$ has no discontinuities. The HQ profit schedule $\pi_{HQ}^*(d)$ is smooth in d as well. It follows that $\Pi(d) = \max \{ \pi_b^*(d), \pi_{HQ}^*(d), 0 \}$ is continuous. Moreover, on any range where $\Pi(d) > 0$ it coincides with either $\pi_b^*(d)$ or $\pi_{HQ}^*(d)$, each strictly decreasing, so $\Pi(d)$ is strictly decreasing on those ranges.

Assume that entry is profitable at short distance (e.g., $\pi_b^*(0) > 0$ or $\pi_{HQ}^*(0) > 0$). Then there exists a unique cutoff

$$d^{\text{entry}} \equiv \min \{ d \geq 0 : \Pi(d) \leq 0 \},$$

such that $\Pi(d) > 0$ for $d < d^{\text{entry}}$ and $\Pi(d) = 0$ for $d \geq d^{\text{entry}}$. This establishes the existence/uniqueness and the monotone decline of optimized profits with distance stated in Proposition 1; a formal proof is provided in Section C7.

Intuitively, both operating modes deteriorate with distance: branches because internal coordination becomes costlier (and agency worsens), and HQ service because client-facing activities become harder. Taking the better of the two modes yields a value function that falls with distance until it hits zero exactly once, which pins down the market-entry threshold.

C7 Proofs of Propositions

Proof of Proposition 1

Define $\Pi(d) = \max\{\pi_b^*(d), \pi_{\text{HQ}}^*(d), 0\}$. From equation (11) and Lemma 1,

$$\frac{d\pi_b^*(d)}{dd} = -\frac{\partial\kappa(d, HQ^*(d))}{\partial d} + (1-\beta)\Psi(d)\frac{d\gamma(d)}{dd} < 0,$$

since $\partial\kappa/\partial d > 0$, $\Psi(d) \leq 0$, and $d\gamma/dd \geq 0$.

From Section C4,

$$\frac{d\pi_{\text{HQ}}^*(d)}{dd} = -[L_{\text{HQ}}^*(d)]^\alpha \overline{HQ}^{1-\alpha} \frac{\tau'(d)}{\tau(d)^2} < 0,$$

since $\tau'(d) > 0$. Both schedules are continuous in d (Lemma 1 and smoothness of κ, τ), hence $\Pi(d)$ is continuous and strictly decreasing wherever $\Pi(d) > 0$.

Since $\Pi(d)$ takes, at each distance d , the larger of $\pi_b^*(d)$, $\pi_{\text{HQ}}^*(d)$, and 0, it is continuous and strictly decreasing wherever it is positive. Assume entry is profitable at short distance (e.g., $\Pi(0) > 0$) and becomes unprofitable for sufficiently large distance (i.e., there exists D with $\Pi(D) \leq 0$). By continuity, there exists a smallest distance $d^{\text{entry}} \geq 0$ such that $\Pi(d^{\text{entry}}) = 0$ and $\Pi(d) > 0$ for all $d < d^{\text{entry}}$. Uniqueness holds because Π cannot increase on any interval where it is positive and therefore can cross zero at most once. Moreover, on the region where $\Pi(d) > 0$ it coincides with either $\pi_b^*(d)$ or $\pi_{\text{HQ}}^*(d)$, each strictly decreasing, so $\frac{d\Pi(d)}{dd} < 0$ throughout that region.

□

Proof of Proposition 2

When an interior solution exists, $HQ^*(d) = \widetilde{HQ}(d)$ is implicitly defined by the first-order condition in equation (11). Totally differentiating this expression, and using the implicit

function theorem yields

$$\frac{d\widetilde{HQ}(d)}{dd} = - \frac{\frac{\partial^2 \kappa(d, HQ)}{\partial d \partial HQ} - (1 - \beta) \frac{\partial B(d, HQ)}{\partial d}}{\frac{\partial^2 \kappa(d, HQ)}{\partial HQ^2} - (1 - \beta) \frac{\partial B(d, HQ)}{\partial HQ}} \Big|_{HQ = \widetilde{HQ}(d)},$$

where $B(d, HQ)$ denotes the bracketed benefit term on the right-hand side of (11).

By strict convexity, the denominator is strictly positive. By $\partial^2 \kappa / \partial d \partial HQ \geq 0$ and Assumption 1, the numerator is weakly positive, hence $d\widetilde{HQ}/dd \leq 0$. If the HQ capacity constraint binds, $HQ^*(d) = \overline{HQ}$ and hence $dHQ^*/dd = 0$. Putting the regions together, the optimizer is weakly decreasing in distance: in the interior it falls with d , and when the cap binds it remains constant, yielding $dHQ^*/dd \leq 0$ for all d . $\square$

Proof of Proposition 3

Conditional on operating a branch, equilibrium employment is $L^*(d) = L^\dagger(HQ^*(d); d)$ and $\frac{dL^*(d)}{dd}$ is given by equation (12).

Aligned-incentive regions. If $HQ^*(d) \geq HQ^{\text{crit}}(d)$, then $L^\dagger = \widehat{L} = (\alpha/w)^{1/(1-\alpha)}HQ$ so $\partial L^\dagger / \partial HQ = (\alpha/w)^{1/(1-\alpha)} > 0$ and $\partial L^\dagger / \partial \gamma = 0$. By Proposition 2, $dHQ^*/dd \leq 0$, which implies that $dL^*/dd \leq 0$.

Over-hiring region. If $HQ^*(d) < HQ^{\text{crit}}(d)$, then $\partial L^\dagger / \partial \gamma > 0$ (Lemma 1) and thus the second term in equation (12) is weakly positive. The sign of the first term is governed by the expression in equation (13). Denote by $\bar{d}$ any distance where equality in that expression holds (if it exists). For $d < \bar{d}$ (productivity-dominated), $\partial L^\dagger / \partial HQ > 0$ and, by Proposition 2, $dHQ^*/dd \leq 0$, so the first term is non-positive; whenever this effect dominates the private benefits from over-hiring term, and $dL^*/dd \leq 0$. For $d > \bar{d}$ (discipline-dominated), $\partial L^\dagger / \partial HQ < 0$ so the first term is strictly positive (lower d raises HQ^* and reduces the manager's desired L), and together with the non-negative term reflecting the incentives to over-hire, we obtain $dL^*/dd > 0$.

Finally, the uniqueness of $\bar{d}$ can be established by rewriting the sign condition—using the optimality condition in equation (11) evaluated at $HQ^*(d)$ —as a comparison of two functions of d . The left-hand side is weakly increasing in d when $\frac{\partial^2 \kappa(d, HQ)}{\partial d \partial HQ} \geq 0$, while the right-hand side is weakly decreasing in d when $\frac{d\gamma(d)}{dd} \geq 0$. Hence the two functions can intersect at most once. When the firm does not operate a branch, employment coincides with $L_{\text{HQ}}^*(d)$ from Section C4, which declines strictly with d . $\square$