When Distance Shrinks:
The Effects of Competitor Proximity on Firm Survival

Jasmina Chauvin*

August, 2019

Manuscript under review at *Management Science*.

Abstract

How does collocating with firms in one's industry affect firm performance? The existing evidence is mixed. I introduce a novel empirical approach that exploits changes in *effective* collocation driven by road improvements. Using comprehensive firm-level data from Brazil, I find that the effects of collocation differ starkly depending on the spatial scope of industries' product markets. In locally traded manufacturing industries, collocation leads to increased exit among the smallest firms but improved survival prospects for the largest. When collocation increases, firms react strategically, relocating and switching their primary industry in ways that lower exposure to competitors. In contrast, when collocation increases in nationally traded industries, survival rates improve for firms of all sizes and fewer firms relocate. While existing research focuses on the average effects of collocation, the findings of this study suggest that collocation intensifies both spatial competition and agglomeration forces, and that its effect on firm performance depends on the relative strength of these mechanisms in different industries.

*Georgetown University, McDonough School of Business. Email: jasmina.chauvin@georgetown.edu. I am grateful to Juan Alcácer, Laura Alfaro, Joel Baum, Kirill Borusyak, Jeffry Frieden, Marco Giarratana, Dan Gross, William Kerr, Tarun Khanna, Hong Luo, Brian McCann, Santiago Mingo, Frank Neffke, Haris Tabakovic, Neil Thakral and Claudia Steinwender for valuable comments. Special thanks to Newton de Castro and Eduardo Haddad for sharing data and to the Harvard Center for Geographic Analysis for technical support. This project received funding from Harvard Business School and the David Rockefeller Center for Latin American Studies.

1 Introduction

Given the high concentration of economic activity in cities and clusters, famously documented by Jacobs (1969), Porter (1990), and more recently by Ellison and Glaeser (1997), an extensive literature in multiple research fields has developed to explain the performance effects of collocation. However, a conclusive answer has proven difficult to establish. While studies of the manufacturing sector as a whole have generally found positive but small average effects of collocation on firm-level productivity (e.g., Martin, Mayer and Mayneris 2011), industry specific studies have yielded a range of results, including positive (Chung and Kalnins 2001), negative (Baum and Mezias 1992, Sorenson and Audia 2000), mixed (Henderson 2003, Beaudry and Swann 2009), and null (Buenstorf and Klepper 2009). Reviews of the literature have referred to all conclusions as "tentative" (McCann and Folta 2008, p. 540). This study introduces two novel features, one addressing causality and the other firm- and industry-level heterogeneity, to shed new light on the performance effects of collocation.

One key challenge existing research on collocation has faced is that, theoretically, spatial proximity between firms may be a double-edged sword. Well-known theories from urban economics predict that collocated firms generate agglomeration economies, which are positive externalities that arise from denser local pools of workers, inputs, ideas (Marshall 1920), or demand (Krugman 1991). Seminal theories in industrial organization, on the other hand, predict that proximity increases competition, eroding firms' ability to spatially differentiate and charge markups over marginal cost (Hotelling 1929, Salop 1979, Vogel 2008). While empirical research in these two fields has evolved separately, in practice, firms may be affected by both, agglomeration and spatial competition effects, when collocating. Consistent with this argument, Baum and Haveman (1997) and Alcacer (2006) find evidence that firms appear to balance agglomeration and competition forces when choosing where to locate. However, it remains unclear to what extent spatial competition versus agglomeration subsequently shape the performance of collocated firms.

I argue that the performance effect of collocation, which is effectively a *net* effect of both agglomeration and spatial competition forces, should vary predictably, depending on the spatial scope of an industry's product markets. In some industries, for example, bottled beverages, semi-processed foods, wood products, and primary metal products, competition is highly localized. Firms in these industries ship their output at a median distance of 250 miles or less according to data from the U.S. Commodity Flow Survey (Figure 2). For firms in such "locally traded" industries, the adverse competitive effects that collocation generates are likely to overshadow any positive agglomeration benefits. Meanwhile, for firms in industries that tend to serve national or international markets, for example, chemicals, apparel, and electronics, spatial competition forces are weak. In such "nationally traded" industries, the net effect of collocation will more likely reflect the positive agglomeration spillovers that the industry generates.

I test the argument that the spatial scope of product markets shapes the effect of collocation on firm performance by leveraging the insight illustrated in Combes *et al.* (2012) that the two underlying mechanisms — spatial competition and agglomeration — carry different predictions for the outcomes of firms at different points of the firm productivity distribution.[1] Specifically, in heterogeneous-firm models, such as Melitz (2003) and Syverson (2004), increased competition leads to a process of selection and reallocation, i.e., increased exit of the least productive firms from a market and reallocation of their market power to the most productive firms. If we envision a distribution of firms ordered by initial productivity, this effectively results in a left-truncation of the distribution. Meanwhile, agglomeration leads to improved performance for all collocated firms, effectively a rightward shift of the firm productivity distribution. Leveraging this insight and proxying productivity with firm size, I test the prediction that in locally traded industries collocation increases exit among the smallest firms but improves the survival rates of the largest firms (selection and reallocation), while in nationally traded industries, collocation improves the survival rates of firms of all sizes (rightward shift).[2]

On the empirical side, two traditional sources of endogeneity have plagued existing research

[1] Testing this argument directly would require observing the extent of agglomeration spillovers and degree of spatial competition across multiple industries. The data required for such an exercise rarely exists even for single industries and a single mechanism at a time.

[2] Indeed, in nationally traded industries, small firms may benefit from collocation proportionally more than large firms, i.e., a larger rightward shift. The possibility that small firms benefit more from agglomeration than large ones has received repeated support in the agglomeration literature, most recently in Faggio, Silva and Strange (2017).

— unobserved heterogeneity and simultaneity. Because smaller, less capable firms are more likely to collocate than larger, more capable firms, as was shown by Shaver and Flyer (2000) and Alcacer (2006), a simple relationship between collocation and performance may appear negative due to unobserved heterogeneity. In recent years, research has largely addressed this problem through the use of panel data, where identification proceeds from *changes* in collocation around a focal firm. However, the simultaneity problem remains. Entry and exit of other firms, which is the source of variation in standard collocation measures, are likely correlated with factors that also affect the focal firm's performance, for example, local demand shocks. According to Combes and Gobillon (2015), the standard solution to this problem calls for a compelling time-varying instrument for collocation, which has proven difficult to find.

I propose a different solution to this challenge by introducing the notion of *effective* collocation, and by leveraging a source of exogenous variation in it driven by road upgrades. While standard measures of collocation are constructed as raw or distance-weighted counts of firms in a focal firm's radius, effective collocation weights firms by their travel time, rather than geographic distance, to the focal firm. This choice has two distinct advantages. First, it yields a more micro-founded proxy of the mechanisms underlying agglomeration and spatial competition because it measures more closely the actual probability of interaction between firms.[3] More importantly, it can generate variation in collocation from changes in travel times among a fixed set of firms, rather than from potentially endogenous firm entry and exit patterns.

The empirical setting is Brazil, where the government invested roughly US$35 billion in road upgrades during 2007–2014 under a national investment program.[4] I collect detailed geospatial data on the condition of the federal and state road network before and after the program and combine it with data on the location and activities of all formal sector manufacturing firms from a high-quality administrative dataset. I find that the investments

[3]A number of recent studies provide direct causal evidence of the effects of roads on the likelihood of knowledge flows (Agrawal, Galasso and Oettl 2017), labor flows (Morten and Oliveira 2016), the quality of supplier matches (Bernard, Moxnes and Saito forthcoming), and price-cost markups (Asturias, García-Santana and Ramos forthcoming).

[4]Appendix Figure 4 illustrates the spatial distribution of the road investments under the *Programa de Aceleração do Crescimento*, (PAC).

lowered travel times across the set of Brazilian municipalities included in the analysis by 5.2 percent on average.

I exploit the changes in travel times induced by road upgrades for exogenous variation in effective collocation between a firm and its nearby competitors during the 2007—2014 period. Firms that saw travel times to competitors' locations fall effectively became more collocated. To address the possibility that road upgrades may have targeted certain locations or industries, the identification strategy relies on an extensive set of municipality and industry fixed effects, which net out any overall correlation between the extent of road upgrades and subsequent performance of firms in each industry and municipality.[5] The residual variation, then, which I exploit stems only from differences in how a road upgrade affects *different* firms in the *same* municipality. This variation, in turn, is driven solely from differences in the pre-existing location patterns of firms' nearby competitors relative to the road upgrades.

To illustrate the intuition behind the identification strategy, consider, for example, a municipality that receives an upgrade on a road leading westward. An industrial bakery whose competitors happened to lie to the west would see a larger increase in collocation than a water bottler whose competitors happened to lie to the east. The identifying assumption behind the identification strategy is that the differences in the pre-existing location patterns of firms' nearby competitors relative to the road upgrades are exogenous to the focal firm's performance, conditional on municipality and industry fixed effects. In support of this assumption, I find that conditional on controls, the change collocation induced by road upgrades is uncorrelated with a range of firm-level characteristics, including size, average wages, exporting or importing status, and firms' contributions to political campaigns.

The estimation results find strong evidence of heterogenous effects of collocation in locally and nationally traded industries. Using the Ellison and Glaeser (1997) index to define which industries are locally traded (present everywhere) versus nationally traded (geographically concentrated), I find that in locally traded industries, doubling collocation reduces the

[5]Conversations with the government secretariat in charge of the PAC road investment program suggest that while the initial project choice was largely determined by the existence of feasibility studies and capacity constraints on the existing network, other considerations also entered investment decisions over time. These included the desire to distribute investments geographically so that all states received some investment, more upgrades to a priori less developed parts of the country, as well as the needs of largest exporting sectors, such as and corn.

seven-year survival probability by 14.1 percentage points for the smallest firms but increases it by 2.6 percentage points for the largest firms (over a seven-year survival rate of 46.2 and 64.1 percent, respectively). This first result aligns with the prediction of competition-driven selection in locally traded industries. In contrast, the results in nationally traded industries imply that doubling collocation increases firms' survival probability by 14.9 percentage points on average (over a seven-year survival rate of 55.4 percent) with no significant differences across firm size categories. Finally, a placebo test on an intermediate set of industries shows no statistically significant effects of collocation on firm survival.

To inquire further into how firms reacted to the external shock, I decompose the main survival effect into cases of full exits, geographic relocations, and switches to a different primary industry. I find evidence of all three adjustment mechanisms for firms in locally traded industries. Firms that saw a greater increase in collocation were more likely to relocate and/or to switch to a different industry. Furthermore, the firms which switched their primary industry or relocated tended to move *away* from competitors, decreasing collocation in their destination relative to their origin. These actions are consistent with the literature on competitive response and repositioning, for example, Figueiredo and Silverman (2007) and Wang and Shaver (2014). In contrast, I find evidence that firms which saw a greater increase in collocation in nationally traded industries became less likely to relocate. When firms relocated or switched industry, they tended to move *toward* other firms, increasing collocation in their destination relative to their origin. These findings are consistent with agglomeration-seeking location strategies (Chung and Alcacer 2002, Kalnins and Chung 2004).

I perform several robustness checks on these results to consider alternative explanations. Roads can have effects on firm performance through channels that are distinct from spatial competition and agglomeration mechanisms, for example, by increasing firms' access to previously unserved markets, reducing the price of inputs, or increasing competition from new firm entry. I address each of these possibilities and find that the original conclusions remain unchanged.

This study makes two main contributions to the firm location literature. First, it shows that an observable industry-level characteristic, the spatial scope of product markets, shapes the effects of collocation on firm performance. I document that the effects of collocating differ starkly for firms in industries that compete for customers locally and firms that compete in broader geographic markets. Specifically, in locally traded industries collocation hurts some firms. It reduces the survival rate among the smallest firms, with some evidence that it increases it among the largest firms. In nationally traded industries, collocation improves firms' survival rates. The effects of collocation on a range of intermediate industries are unclear. These findings help to reconcile seemingly conflicting results of existing studies, which observe positive, negative, and null effects in different empirical contexts. They also suggest that studies which aggregate data from many heterogenous industries, such as those of the manufacturing sector as a whole, may find small or null effects depending on the composition of firms and industries in the sample. However, these muted average effects may hide significant heterogeneity.

Second, the study develops and illustrates a novel empirical approach in this literature, relying on plausibly exogenous changes in *effective* collocation from changes in travel times. Recently, a number of studies have causally evaluated the effects of road availability on regional outcomes (e.g., Duranton and Turner 2012, Faber 2014, Donaldson 2018). My research design is the first to my knowledge to exploit variation in *firm-level* exposure to roads in order study the effects of collocation on firm performance.[6] The surge in new technologies and investments in mobility currently taking place provides an opportunity to adapt the approach developed here to other settings to study how changes in the costs and patterns of mobility affect firms' strategic interactions.

2 Theoretical Background

Empirical studies of collocation are frequently rooted in urban economics frameworks in which collocation increases firm productivity though positive supply-side spillovers. In a standard Cobb-Douglas production function, where quantity produced is a function of capital, labor, and total factor productivity (TFP), greater collocation can increase TFP.

[6] In the existing literature, the current study is most similar to Holl (2016), who provides one of the only firm-level studies of the effects of road investments but does not use roads as a shock to collocation.

The theoretical mechanisms behind this hypothesized effect date back to Marshall (1920) and include positive externalities generated when specialized input suppliers establish themselves to serve collocated firms, workers invest in industry-specific skills, and knowledge spillovers become more likely. Collocated firms experience greater productivity as a result of the increased quality and availability of inputs, workers, and knowledge.

While theoretical models of agglomeration cast predictions on TFP, research has often tested agglomeration theories on broader empirical outcomes. Typically, firm-level profitability or survival is highlighted in the case of management research, whereas regional employment, dynamism, and diversification are the focus of economic geography and urban economics research. The reasons for this are twofold. First, TFP estimation is fraught with challenges, including limited availability of plant-level data on input and output quantities, unobserved differences in input and output quality, as well as endogenous firm investment decisions. Second, frequently scholars are fundamentally interested in the effects of collocation on broader economic outcomes. While these can be affected by TFP, they are not necessarily affected linearly, or even in the same direction, because of competitive and general equilibrium effects.

However, in moving to broader outcomes, research has not always accounted for all of the theoretical mechanisms through which spatial proximity between firms can affect performance, and in particular, demand-side mechanisms like those described in industrial organization models of spatial competition (Hotelling 1929, Salop 1979). In Hotelling's framework, distance creates a barrier to competition and sustains positive profits. The effect here of greater proximity of firms in the same industry would be to intensify price competition and erode such profits. When firms are of heterogeneous productivity as in Melitz (2003), Syverson (2004), and Vogel (2008), heightened competition has precise predictions for firm survival: it leads to greater exit of the least productive firms and to reallocation of market power to the most productive firms. At the level of regions or local markets, the result is a lower density of low-quality firms, i.e., a left-truncation of the firm quality distribution.

A different class of models predict demand side benefits from collocating, through reduced customer search costs (Fischer and Harrington 1996). This positive mechanism is especially relevant in business-to-consumer (B2C) industries in which product heterogeneity is high

Table 1: *Theoretical Effects of Collocation*

Direction of effect	**Supply-side**	**Demand-side**
Positive	Agglomeration externalities	Reduced search costs
Negative	Congestion and input rivalry	Spatial competition

and quality hard to ascertain. This can explain why we often observe diamond retailers, shoe stores, and vegetable sellers clustered on the same street block. Search costs have also shown to be relevant for the collocation of hotel establishments (Baum and Haveman 1997). It is less likely to apply to manufacturing industries, many of which are business-to-business (B2B) intermediate good providers.

Finally, there may also be negative externalities on the supply side when collocated firms generate congestion and competition over limited inputs, for example, by raising the cost of local real estate and pushing up the wages of specialized workers. Firms can also be harmed if some of their unique and valuable resources, such as specialized workers, technologies, and IP, become more likely to spill over to competitors.

Considering the theoretical frameworks jointly (Table 1), there is the possibility that, in any industry, collocation can act as both a positive force on productivity and on demand, and as a negative force on input- and output prices. What, then, is the effect of increased collocation on an aggregate outcome like firm revenue, profitability, or survival? In other words, which of the underlying theoretical forces is likely to be dominant? While this is ultimately an empirical question, we can consider: under which conditions agglomeration benefits on the production side are likely to be the more dominant? Under which conditions agglomeration externalities are likely small, relative to competitive price effects?

One characteristic of manufacturing industries that can shed light on the relative importance of spatial competition versus agglomeration mechanisms is the spatial scope of the industry's product markets. While nearly all manufacturing industries are considered tradeable (Delgado, Porter and Stern 2015), in reality the spatial scope of markets for different products differ. For example, most electronics, apparel, and advanced machinery products are shipped nationally or internationally. Meanwhile, food, beverages, wood products, and metal products tend to be shipped in a radius of 250 miles or less (Figure 2). Rather than considering all manufacturing industries as tradeable, it is therefore more appropriate to consider nationally traded" and "locally traded" industries *within* manufacturing.[7] Firms in locally traded industries interact in product markets with a significantly lower geographic scope than firms in nationally traded industries.[8]

Why would the effect of collocation on firm performance vary by industry for locally and nationally traded industries? Even without taking into account the extent of agglomeration externalities in these industries, it is clear that for locally traded industries spatial competition is likely to be relatively more important than for nationally traded industries. Because these firms compete over local demand, their pricing strategy is sensitive to the actions of nearby competitors. A greater number of collocated competitors should lead to a lower average and less dispersion in prices. When prices fall and firms are of heterogeneous quality, marginal firms who were just breaking even risk greater exit. As the weakest firms exit, their market share reallocates to more productive surviving firms. These theoretical mechanisms in locally traded industries are testable with the following hypotheses:

HYPOTHESIS 1a: In locally traded industries, an increase in collocation will lead to increased exit of the least productive firms.

HYPOTHESIS 1b: In locally traded industries, an increase in collocation will improve the survival prospects of the most productive firms.

Meanwhile, in industries that compete nationally, sensitivity of prices to collocation in a particular market is likely to be negligible and, therefore, demand-side competition effects are limited. To the extent that firms generate supply-side benefits by collocating, greater collocation should increase firms' survival rates. If, as prior literature suggests, weaker firms benefit from positive agglomeration externalities more than the most productive firms (Shaver and Flyer 2000, Alcacer 2006), they will see a disproportionally larger positive effect. These insights lead to two additional hypotheses that are testable in nationally traded industries:

HYPOTHESIS 2a: In nationally traded industries, an increase in collocation will improve

[7] Note, the classic work by Chandra and Thompson (2000) makes a similar argument, pointing out that road investments should give rise to spatial competition forces in what they termed "regionally traded" goods. While their article considers retail and services industries as falling in this regionally traded category, I argue that many manufacturing industries also constitute locally traded goods.

[8] While most industries are also traded internationally, I refer to nationally traded rather than internationally traded industries to indicate the most likely geographic scope of the average firm in my dataset, which are single-plant manufacturing firms in Brazil. We know that even in the Unites States, exporters are concentrated among a small percentage of (mostly multi-unit) firms (Bernard, Jensen, Redding and Schott 2007).

the survival prospects of collocated firms.

HYPOTHESIS 2b: In nationally traded industries, an increase in collocation will improve the survival prospects of the least productive firms most.

Next, I describe the identification strategy and institutional setting designed to empirically test these predictions.

3 Identification Strategy and Institutional Context

The following section describes the identification strategy and institutional context. I discuss advantages and assumptions inherent in using road upgrades for exogenous variation in effective collocation. Finally, I describe the data sources and construction of key variables.

3.1 Identification Strategy

Empirical studies of collocation typically measure collocation by counting the number of firms (or their total employment) in an industry and defined geographic area. The choice of geographic area is usually an administrative unit (e.g., state, labor market region, county, or ZIP Code) or a pre-defined spatial distance band (e.g., 10 miles, 50 miles). In some cases, researchers further depreciate the count of each firm by its geographic distance to each focal firm (e.g., Sorenson and Audia 2000, Rosenthal and Strange 2003, Haveman and Rider 2014) to account for the fact that the mechanisms underlying collocation are sensitive to distance. Though no consensus exists on what the best choice of geographic area is (Singh and Marx 2013), a good guiding principle is to choose the spatial scope that corresponds to the likely scope of the economic interactions of interest (Alcacer and Zhao 2016).

I define collocation following these established principles but incorporate the travel time, rather than geographic distance, between firm locations in the measure. Specifically, I define collocation of firm i in industry s and location m as

$$CoLoc_i^{sm} = \sum_{k \in M} \frac{x_j^{sk}}{\tau_{mk}}, \tag{1}$$

where x_j^{sk} is the object being counted (i.e., firm j in industry s and location k) and τ_{mk}

is the road travel time between locations m and k.[9] M is the geographic area over which collocation is defined.

Incorporating travel time into the collocation measure offers a number of advantages. First, it allows me to construct a location-specific empirical definition of geographic area over which to assess collocation, which should reduce measurement error in existing collocation metrics that rely on uniform distance bands or administrative boundaries. Considering the nature of the economic interactions that are the focus of this study — localized competition, input-sourcing, labor-pooling, and knowledge-sharing — I define the geographic area that corresponds to all of the destinations that can be reached in four hours from each focal firm. The choice of four hours is motivated after considering that localized interactions (e.g., same day delivery of a good, travel to work) are likely constrained to occur within one working day, including the return trip.[10] Travel times also enable me to depreciate each firm more precisely by its effective distance to the focal firm. While in advanced economies, straight-line distance and travel costs may be highly correlated, this is less the case in emerging markets. Even if two firms are geographically proximate, if the travel possibilities between them are limited, they are unlikely to generate competitive pressures or agglomeration externalities.

Beyond the measurement improvements, arguably the most important advantage of the proposed collocation measure is that it creates a source of variation in collocation that is independent of firm location choices: changes in travel times. As the Introduction highlights, it has proven difficult to address identification challenges in research on collocation and firm performance. One challenge is selection bias, created by the fact that firms of different quality sort themselves in to more or less collocated places, thus creating an endogenous relationship between collocation and firm performance. Both Shaver and Flyer (2000) and Alcacer (2006) find that lower quality firms are more likely to collocate, arguably because they are more dependent on external agglomeration benefits. Firms can also use location choice to affect

[9] In this study, locations correspond to the more than 5,500 Brazilian municipalities. Because I do not observe within-municipality travel times, I set the within-municipality travel time equal to 15 minutes when counting other firms in the same municipality-industry as the focal firm. This choice is motivated by the average size of a municipality in Brazil (roughly 1,500 square kilometers) and the expected travel time between two randomly chosen points in a square of that size.

[10] This procedure naturally yields a smaller geographic area around the Amazonian city of Manaus than around the well-connected metropolis of Sao Paulo (Figure 3).

competitive dynamics, as Ghemawat and Thomas (2008) show in the global cement industry.

Panel data as in Baum and Mezias (1992), Sorenson and Audia (2000), and Henderson (2003) can partially address selection by identifying effects of collocation from *changes* in collocation around the focal firm after its entry. Nonetheless, this approach suffers from endogeneity if unobserved factors (e.g., a boom in local demand) affect both the dynamics of firm entry/exit and the performance of the focal firm. Addressing this second concern calls for instruments or natural experiments that affect collocation. However, beyond one-off empirical settings (e.g., the fall of the Berlin Wall as in Ahlfeldt *et al.* (2015)), such instruments have proven difficult to find.

By leveraging changes in travel times between firm locations as an exogenous source of variation in collocation, this study illustrates a new empirical strategy which is potentially more broadly applicable. The rationale behind using changes in travel times is grounded in the idea that the actual costs and patterns of mobility, rather than distance *per se*, affect the underlying economic mechanisms that generate competition and agglomeration spillovers. Indeed, recent evidence shows that transportation infrastructure investments can have powerful effects on the extent of supplier-buyer relationships (Bernard, Moxnes and Saito forthcoming), patterns of worker mobility (Morten and Oliveira 2016), and the flows of knowledge that support innovation (Agrawal, Galasso and Oettl 2017). Changes in the costs of spatial interactions can also affect competitive interactions between firms as Asturias, García-Santana and Ramos (forthcoming) and Haveman and Rider (2014) show. The key to the identification strategy is that travel times provide the only source of variation in collocation, and that the measure does not incorporate firm decisions regarding entry and exit.

Specifically, I define change in collocation for firm i from time t_0 to time t_1 as

$$\Delta\, CoLoc^{sm}_{i,t_0-t_1} = \sum_{k\in M_{t_1}} \frac{x^{sk}_{j,t_0}}{\tau_{mk,t_1}} - \sum_{k\in M_{t_0}} \frac{x^{sk}_{j,t_0}}{\tau_{mk,t_0}}, \tag{2}$$

which sums over only the incumbent firms of time t_0 in the focal firm's four-hour radius (including any which did not survive) and weights each by its respective travel time to the focal firm at times t_1 and t_0, respectively. Thus, the only sources of variation in this measure come from (i) changes in the weight of each firm in the summation, τ_{mk}, and/or (ii) changes in the geographic region, M, over which collocation for the focal firm is calculated.[11] This variation should affect the underlying mechanisms of competition and agglomeration (relevance), but it does not suffer from the same selection bias and unobserved heterogeneity as traditional approaches (exogenous to firm location choice).

However, the identification does rely on the assumption that changes in collocation induced by travel times are *a priori* exogenous to the performance of the firms in the sample. To address the validity of this assumption, I next discuss it and the institutional context in detail.

3.2 Institutional Context

The empirical setting of the study is Brazil during 2007–2014. Following years of positive economic growth, in 2007 the government launched a national investment program — the *Programa de Aceleração do Crescimento* (PAC) — which, among others, invested more than 70 billion reals (roughly US$35 billion) into upgrades for the federal road network.[12] This program provides an excellent setting for the study, not only because it was economically significant, but also because most economic activity in Brazil relies on roads. Based on recent firm-level surveys, more than 99 percent rely on roads; other modes of transport (e.g., rail, waterways) account for a smaller share of traffic than in similar countries (Savaris *et al.* 2013).[13] At the same time, the condition of the road network in Brazil is extremely poor, and poses a critical constraint to economic activity.[14]

The stated aim of the PAC road investment program was to relieve key bottlenecks in the

[11]The metric will ultimately be expressed as a difference of logged values because this has the benefit of approximating percentage differences for small enough changes. I add 0.1 to the raw measure in order to include observations that have a zero measure of collocation in any of the years.

[12]Source: http://www.pac.gov.br/. The figures refer to Phases I and II of the program, which took place during 2007–2014.

[13]Information of the surveyed firms was accessed from https://www.epl.gov.br/perfil-de-embarcadores-e-servicos-demandados.

[14]In 2015, Brazil ranked 123rd out of 144 countries on the World Economic Forum's "Quality of overall infrastructure" index, well behind China (51st) and India (74th). Surveyed executives cite an inadequate supply of infrastructure as the fourth most problematic factor for business, after tax rates, restrictive labor regulations, and corruption.

federal road network. Rather than building new roads, the program focused on improving existing roads, either by upgrading the road surface type (e.g., duplicating a one-lane highway, paving a dirt road) or improving the surface condition to enhance performance and capacity (e.g., re-paving deteriorated stretches and potholes). The program was highly decentralized, with more than 250 different projects taking place across different parts of the country (Figure 4). Overall, as Tables **??** and **??** in the Appendix show that about 30 percent of the total federal road network saw some upgrade.

A potential threat to identification is the possibility that road upgrades targeted specific industries, regions, or firms in ways correlated with their future performance. Interviews with policymakers in Brazil indicate that special interests were less likely to have shaped the program in its initial years because project choice was frequently constrained by the existence of feasibility studies, which can take several years to develop. Thus, the government tended to allocate funds to projects which were shovel-ready. However, in later stages, industrial lobbies became more involved in the project selection process, especially the large agricultural exporters of soy, corn, and sugar. Additionally, over time, regions and municipalities with worse road conditions at the outset were more likely to receive investments due to equity concerns.

I tailor the empirical strategy to be robust to the possibility that investments targeted certain industries or regions based on their expected performance. To do so, I control for the full set of industry and municipality fixed effects in all empirical analyses. If industries and regions more likely to receive road investments systematically fared better (or worse) in the period under study, that trend will be absorbed into the fixed effects. Note that the inclusion of the fixed effects places a very high bar on the empirical analysis, as the fixed effects for the roughly 250 different industries and more than 3,700 municipalities represented in the sample absorb three-quarters of the variation in the change in collocation measure.[15]

The remaining variation, which is used for identification, stems from differences in how much different firms in the same municipality were affected. With the municipality and industry fixed effects in place, this variation is generated not primarily from the placement of

[15]Specifically, the OLS regression of change in collocation on industry and municipality fixed effects has an R^2 of 0.75.

the road upgrades, but rather from the pre-existing locations of a firm's competitors relative to the road upgrades. This variation — differences in preexisting location patterns of a firm's local competitors relative to the roads — is assumed to be plausibly exogenous to the focal firm's performance. As a check on this empirical design, I will show that, controlling for municipality and industry fixed effects, the resulting change in collocation is uncorrelated to a series of firm-level characteristics. This includes firm size, age, average wage levels, importing and exporting status, and the amount of funds that the firm contributed to political campaigns. Before doing so, I discuss the sources of the geographic and firm-level data.

4 Data and Measurement

Travel Times. I collect data from multiple sources to generate a unique dataset of estimated travel times between all pairs of Brazilian municipalities in 2007 and 2014. While new platforms such as Google Maps have made it possible to collect travel time estimates in real-time, historical sources of comparable data allow researchers to observe how the costs of mobility *change* are rare. I combine three sources of data in the analysis: (i) geo-referenced maps of the Brazilian federal and state road network from the Brazilian Ministry of Transportation; (ii) annual indicators the surface type (e.g., duplicated, paved, dirt) for each roughly five-kilometer segment of each federal road from the National Department of Transportation Infrastructure (DNIT); (iii) annual indicators of the condition (e.g., excellent, good, poor) of each road segment from the Brazilian National Transport Confederation (CNT), and independent, private-sector industry association.[16] I match the annual surface type and condition data to the geo-referenced map to construct a representation of the road network in 2007 (at the outset of the PAC program) and in 2014 (after completion of PAC Phases I and II). Finally, I assign a velocity to each road segment as a function of its surface type and condition (Table **??**).

With the spatial representation of the road network in place, I apply Dijkstra's algorithm to the network to calculate optimal routes between all pairs of municipalities. For each

[16]The Data Appendix describes the data sources and steps taken to process the data in detail. The use of independently audited road condition and quality indicators, rather than funds allocated or spent, is an advantage, especially in a setting like Brazil where corruption and misuse of public funds is a concern.

possible municipality origin and destination pair, the algorithm flexibly searches for the route that minimizes travel time.[17] I separately apply the algorithm to the 2007 and 2014 representation of the road network, thus generating an origin-destination (OD) matrix of minimum travel times for each of the two years. I perform numerous checks using data from Google Maps to ensure that the estimates are representative of real-world travel times (see the Data Appendix for a detailed discussion).

I find that over the entire network, the road investments decreased travel times by 5.2 percent on average. Given the size of the federal and state road network (over 280,000 kilometers), this is a large effect, and comparable to other programs recently evaluated (e.g., Gibbons *et al.* 2019). Figure 1 illustrates how travel times change in the spatial range used in this study corresponding to a four-hour area around any municipality. While the average decrease in travel times in these local markets is two percent, there is significant variation, with some municipalities seeing local travel times fall by 15 percent or more. Other municipalities see net deteriorations in road conditions, leading to an increase in local travel times.

Firm-Level Data. Detailed firm-level data comes from the *Relação Anual de Informações Sociais* (RAIS), an employer-employee matched administrative dataset maintained by the Brazilian Ministry of Labor (MTE). Because all formal firms are required to report RAIS (otherwise, employees forego a federally mandated 13th salary), RAIS is effectively an annual census of all formal-sector employers, one which has been increasingly used for academic research, especially in the field of international trade (e.g., Helpman *et al.* 2017). While RAIS is reported at the worker contract level, I aggregate the data at the firm level using each firm's unique tax identifier (CNPJ).[18]

I restrict the sample to manufacturing firms with a minimum size of three workers during 2007–2014.[19] I drop industries which may be directly affected by road construction, industries

[17] Solving the algorithm is a functionality available in the ArcGIS Network Analyst optimal routing tool.

[18] To the extent that an establishment is associated with more than one record in the same year (e.g., due to inconsistencies in field values for industry, legal form, municipality, etc.), I aggregate these records and associate them with the characteristics featured in the largest number of employee contracts. I drop any establishments with invalid IDs, CEI entities (these are multi-jurisdiction entities primarily associated with the construction sector), and any establishments that were not a business entity during the entire period.

[19] Manufacturing firms are identified with the *Classificação Nacional de Atividades Econômicas* (CNAE) version 1.0 industry code in the 15–37 range. For more detail, see

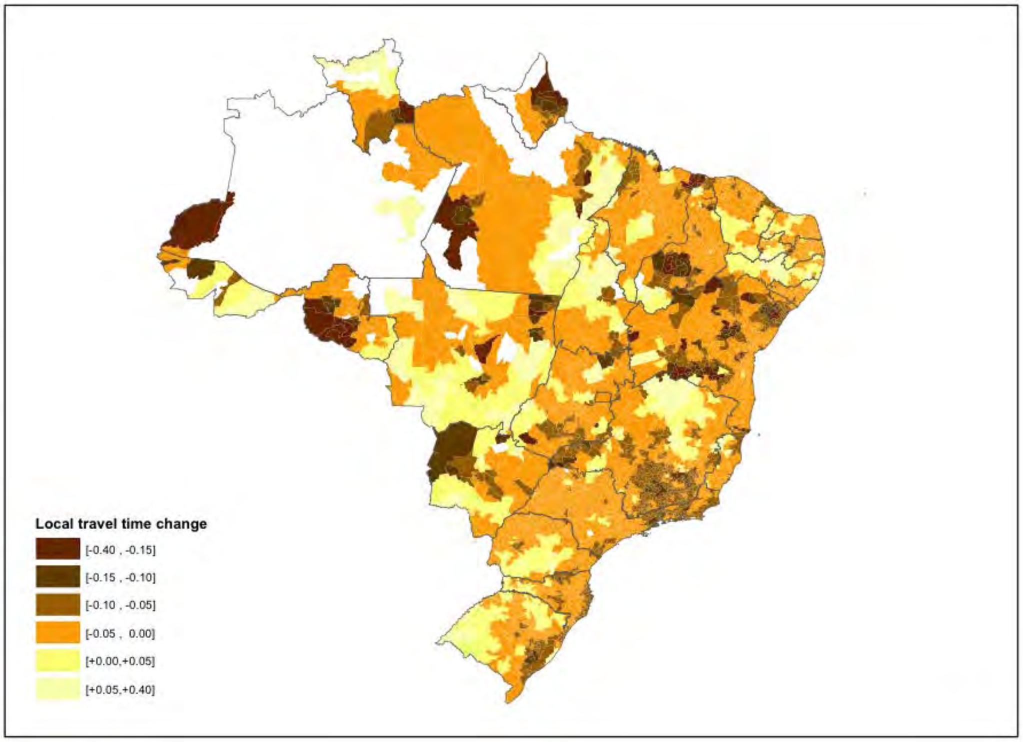

Figure 1: *Change in Local Travel Times Between Brazilian Municipalities, 2007–2014*

Notes. Author's calculations based on data from the Ministry of Transport of Brazil, the National Department of Transportation Infrastructure (DNIT), and the Brazilian National Transport Confederation (CNT). The map shows the total percentage change in the average time of traveling from each municipality to each other municipality in its local market. A local market is defined as any municipality that can be reached in four hours or less traveling by federal and state roads in 2007. Areas not shaded (white) are municipalities whose capital lies more than 50 kilometers away from a state or federal road. These municipalities are excluded from the analysis.

with fewer than ten firms active in 2007, and three industry categories that are too general to be meaningful for identifying competition or agglomeration forces.[20] In the analyses of firm performance (but not in the measures of collocation), I also exclude all establishments of multi-unit firms because for such firms, changes in internal travel times may trigger additional adjustment mechanisms beyond spatial competition and agglomeration, for example, plant exit due to internal spatial reorganization (Alcacer and Delgado 2016) or changes in entry and exit patterns driven by strategic interactions across markets as in Ghemawat and Thomas (2008).[21] I also exclude all manufacturing companies with state ownership to lessen concerns around the special treatment of such companies. The final sample consists of nearly 130,000 single-plant manufacturing firms.

In the RAIS panel, I observe a firm's primary industry, its location (municipality), the size of its workforce, and wages paid. By observing entry and exit, I can deduce a firm's age and its survival from one year to the next. I define the variable *Survive* as an indicator taking a positive value if a 2007 incumbent firm continues to exist in the same municipality and four-digit industry in 2014 and a value of zero otherwise. Therefore, in the baseline analyses of survival, firms that exit a market but relocate, along with firms that exit an industry but enter a different one, have a *Survive* value equal to zero. Based on each firm's unique CNPJ identifier, I also match RAIS to a list all importing and exporting firms published by the Brazilian Secretariat of Foreign Trade (SECEX). Finally, I collect data on monetary donations given by firms to any political parties or candidates for the years 2006, 2008, 2010, and 2012, published by the Brazilian Superior Electoral Tribunal (TSE).

Locally and Nationally Traded Industries. In the absence of data on internal trade that shows the distance of firms' shipments, I follow prior studies (e.g., Delgado, Porter and Stern 2015, Mian and Sufi 2014) and classify manufacturing industries following the principle that nationally traded industries concentrate production in a few locations and ship

http://biblioteca.ibge.gov.br/visualizacao/livros/liv2314.pdf.

[20] The industries excluded are: manufacturing of concrete (2630), stone (2691), cement (2691), construction bulldozers (2953), construction equipment (2995), earthmoving and paving equipment (2954), heavy military equipment manufacturing (2972), and three industries relating to chemical products, machine repair, and motor vehicle parts containing "not elsewhere specified" in the industry name.

[21] Multi-unit firms account for seven percent of manufacturing establishments in Brazil. While they are omitted from the analysis for the reasons above, the results are robust to their inclusion.

their product to more distant markets, whereas locally traded or non-traded industries are found everywhere. To implement this principle, I calculate the industry-level Ellison-Glaeser index (Ellison and Glaeser 1997), using data on industry location patterns in Brazil from RAIS. This well-known measure of spatial concentration counts the share of an industry's employment in a geographic unit and compares it to the share of the industry's employment nationally, adjusting for the industry Herfindahl index to account for concentrations that are due to the industry simply having few firms.[22] Industries with an index value above zero exhibit greater geographic concentration than economic activity overall and those with index values below zero exhibit less concentration than economic activity overall.

I use the cutoffs proposed by Ellison and Glaeser and categorize industries with index values above 0.05 as nationally traded and those below 0.02 as locally traded. This procedure yields 52,496 firms in locally traded and 21,930 firms in nationally traded industries. As expected, the industries classified as locally traded feature food, beverages, construction materials, and other low value-to-weight products. Nationally traded industries tend to feature more advanced manufacturers (Table 10).

Table 7 provides descriptive statistics for the firms in the sample. Overall, the firms represent 252 unique manufacturing industries and are located in 3,724 different municipalities. The average firm is eight years old and employs 23 full-time workers. Five percent of the firms in the sample export and 4.8 percent import in any given year. Just over half of all firms, 55.6 percent, survive during the 2007–2014 period, a survival rate similar to the rate of survival of U.S. manufacturing firms. Calculating change in collocation per Equation 2, I find that the average change in collocation is roughly 3.6 percent.

Tables 8 and 9 present the analyses that serve as a check on the identifying assumption, that the change in collocation is uncorrelated with firm-level characteristics, conditional on controls. These models regress change in collocation on each of firm size, firm age categories, the firm's average annual worker wage, exporting and importing dummies, and the survival rate of firms in the industry-municipality in the seven-year period preceding the

[22] The geographic units that I use to calculate the index are Brazilian regional urban divisions (*Divisão Urbano Regional - Regioes Imediatas de Articulação Urbana*), defined by the Brazilian Statistical Institute (IBGE). They are exhaustive of Brazil's land mass, centered around urban areas, and comparable to U.S. economic areas. https://ww2.ibge.gov.br/home/geociencias/geografia/default_divisao_urbano_regional.shtm.

PAC program, 2000–2007. The results provide support for the identifying assumption. The estimated coefficients on the large majority of the variables hover close to zero and are not statistically significant. Only firm age has some predictive power for the sample of firms in locally traded industries, suggesting collocation increased marginally more among older firms, although the size of the effect is very small. Overall, these results suggest that after conditioning on municipality and industry fixed effects, firm characteristics that are plausibly correlated with firms receiving preferential government treatment are not predictive of the *residual variation* in collocation.

5 Empirical Specification and Results

Here I present the empirical model motivated by the Hypotheses and the choices of control variables. I present and interpret the results of estimating the empirical model. Finally, this section presents the results of several robustness test on the main result as well as several extensions, which consider whether firms strategically react to changes in collocation.

5.1 Empirical Specification

To test the predictions of the effects of collocation on firm survival, I estimate the following empirical model,

$$Y_{i,07-14} = \beta \Delta CoLoc^{sm}_{i,07-14} + \gamma X_{i,07} + \theta Z^{sm}_{07} + ind + muni + \varepsilon_i, \tag{3}$$

where Y_i is a binary variable indicating firm i's survival over the seven-year period from 2007–2014. $\Delta\, CoLoc^{sm}_{i,07-14}$ is the log change in collocation for firm i in industry s in municipality m, and β is the main coefficient of interest, which estimates the average effect of change in collocation on the probability of firm survival. $X_{i,07}$ is a vector of firm-level controls and Z^{sm}_{07} a vector of industry-municipality-level controls, both measured in the baseline year. ind and $muni$ are industry and municipality fixed effects, and ε_i is a random, normally-distributed error term.

In order to test the theoretical prediction of heterogeneous effects, I next estimate a set of models per Equation (4), which introduce interaction terms for the change in collocation and proxies for productivity that are based on firm size in 2007. A strong, positive relationship between firm size and productivity of manufacturing firms is well-documented, for example in Haltiwanger, Lane and Spletzer (1999). The first proxy that I use is a measure of firm size calculated as the number of full-time workers employed at the firm in the baseline year, $Size_{07}$. As other proxies, I calculate different quantiles of firm size for each focal firm relative to all other firms in its industry, as well as relative only to firms in its industry and local market, $Size_{q,07}$. In models that include interactions of collocation with firm size, β is the effect of collocation for firms in the omitted category (usually the smallest firms), while β_qs estimate the effects for firms in the larger size categories. I estimate the model as a linear probability model using ordinary least squares due to the large number of fixed effects and check that the estimated probabilities rarely fall outside the meaningful [0,1] interval.[23]

$$Y_{i,07-14} = \beta \Delta CoLoc^{sm}_{i,07-14} + \sum_q \beta_q \Delta CoLoc^{sm}_{i,07-14} \cdot Size_{q,07} + \gamma X_{i,07} + \theta Z^{sm}_{07} + ind + muni + \varepsilon_i. \tag{4}$$

Beyond the variables of interest, the model includes a number of firm-level and municipality-industry-level controls:

Firm-level controls. All regressions control for the baseline level of a firm's collocation in 2007. First, the baseline level may be correlated with unobserved firm characteristics due to better firms sorting into more or less competitive locations, or because more or less competitive locations yield more productive firms (e.g., through selection). Second, the baseline level of collocation is mechanically correlated with the change in collocation because of convergence effects (relative changes are smaller from a larger baseline). Also, I include the set of controls analyzed in Tables (8) and (9) also in the baseline models to control for any potential residual correlation between firm characteristics and collocation.

Municipality-industry-level controls. At the level of each municipality-industry, I control

[23] A logit model, with more than 3,500 fixed effects, is unable to converge. In robustness checks, I have performed the analysis using a logit model, leaving out the municipality fixed effects. The direction and significance of the estimated coefficients is in line with the regressions performed using the LPM. These analyses are available from the author.

for a count of firms in the focal firm's four-hour market, but outside its industry. In the agglomeration literature, overall density of economic activity, or "urbanization economies," can also generate positive agglomeration externalities (Jacobs 1969). I also control for the level of competition in the market-industry measured as the inverse of the standard Herfindahl index in order to control for any correlation between the baseline levels of competition in the market and subsequent trends in firm survival.

5.2 Results

Column (1) in Table 2 shows the results of estimating Equation 3 in locally traded industries. The beta coefficient, which estimates the average effect of change in collocation on firm survival, is large and negative but not statistically significant, potentially indicating a large degree of heterogeneity. Taken at face value, the coefficient size suggests that a doubling in collocation leads to a 9.1 percentage point lower probability of survival over the seven-year period, an elasticity of roughly 1/10.

The next several columns directly test for firm-level heterogeneity by interacting change in collocation and various measures of relative firm size per Equation 4. In Column (2), which interacts change in collocation with a continuous measure of firm size (log workers), the coefficient on change in collocation is now large, negative, and significant; the interaction effect is positive and highly significant. Specifically, the estimated coefficients indicate that for the smallest firms, doubling the level of collocation leads to a roughly 19 percentage point decrease in the probability of survival over seven years. This result provides the first evidence that the effect of collocation is heterogeneous in firm size in locally traded industries.

Columns (3)–(5) report the results of models that interact change in collocation with increasingly finer quantiles of a firm's size. In each case the coefficient on the change in collocation, which measures the effect for the smallest firms (omitted category), is large, negative, and significant. Its size implies that a doubling of collocation reduces the survival probability of the smallest firms between 12 and 16 percentage points, depending on the specification. Meanwhile, the coefficients for the larger size quantiles are positive, significant, and roughly equal in size, suggesting a net zero effect of collocation on survival for the larger firms.

Column (6), where the quantiles are measured relative only to other firms in the focal firm's industry and market, provides evidence that the survival rates of the largest firms in the market actually increase as collocation increases. The marginal effects of this model suggest that doubling collocation reduces the survival probability of the smallest firms by 14.1 percentage points. The coefficient on the interaction term for the firms in the largest size quartile is 0.167 and statistically significant at the 5 percent level, suggesting that the survival rate of these firms *increases*, by 2.6 percentage points.

The coefficients of other variables in the model conform to expectations. Firm size, age, and average worker wage are positively associated with survival. The coefficient on firm size suggests that a doubling in firm size is associated with a roughly 6 percentage point higher survival rate. Older firms have higher survival rates than younger firms. Firms paying higher average wages also have a higher probability of surviving while exporters and importers are, somewhat surprisingly, no more likely to survive than firms that do not trade. Baseline levels of also have no significant correlation with firm survival in locally traded industries. The coefficients on urbanization economies appear as positive and significant, suggesting greater density of economic activity outside the own industry is associated with a higher survival rate, while the coefficients on the level of competition in the industry-market are either small and positive or not statistically significant.

Overall, the results in locally traded industries point to significant, negative effects of increased collocation on the smallest firms, lending support for *Hypothesis 1a*. The results in Column (6) also lend some support for *Hypothesis 1b*, showing small, positive effects of increased collocation for the largest firms in locally traded industries. Combined, the results are consistent with the predictions of selection and reallocation.

Table 3 shows the results from estimating the same models in the sample of firms in nationally traded industries. Immediately, a key difference appears: the average effect of collocation on survival for firms in nationally traded industries is positive and statistically significant. The size of the coefficient implies that a doubling of collocation leads to a 14.9 percentage point higher survival rate for firms in nationally traded industries.

When change in collocation is interacted with firm size in Column (2), the positive and statistically significant effect for the smallest firms remains. The coefficient on the interaction

Table 2: *Effects of Change in Collocation on Firm Survival for Locally Traded Industries*

Dependent variable: *Survive (0/1)*	Baseline (1)	Interacted with size (2)	2 Quantiles in industry (3)	3 Quantiles in industry (4)	4 Quantiles in industry (5)	4 Quantiles in ind-mkt (6)
$\Delta\, CoLoc_{07-14}$	-0.091	-0.194***	-0.124**	-0.149**	-0.161**	-0.141**
	(0.059)	(0.068)	(0.062)	(0.065)	(0.069)	(0.071)
$\Delta\, CoLoc_{07-14}$ x Firm size (log workers)		0.056***				
		(0.021)				
$\Delta\, CoLoc_{07-14}$ x 2nd quantile			0.113**	0.119*	0.069	0.033
			(0.050)	(0.061)	(0.074)	(0.078)
$\Delta\, CoLoc_{07-14}$ x 3rd quantile				0.115*	0.133*	0.053
				(0.066)	(0.076)	(0.084)
$\Delta\, CoLoc_{07-14}$ x 4th quantile					0.159**	0.167**
					(0.076)	(0.079)
Firm level controls (2007):						
Firm size (log workers)	0.059***	0.057***				
	(0.002)	(0.002)				
Firm in 2nd quantile (dummy)			0.113**	0.082***	0.072***	0.073***
			(0.050)	(0.006)	(0.007)	(0.007)
Firm in 3rd quantile (dummy)				0.136***	0.115***	0.119***
				(0.007)	(0.009)	(0.008)
Firm in 4th quantile (dummy)					0.157***	0.152***
					(0.008)	(0.008)
Firm born prior to 1993 (dummy)	0.089***	0.090***	0.098***	0.095***	0.094***	0.095***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Firm born during 1993-2000 (dummy)	0.058***	0.059***	0.061***	0.059***	0.059***	0.060***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.006)	(0.005)
Average annual worker wage (log reals)	0.029***	0.029***	0.038***	0.034***	0.033***	0.033***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Exporter (dummy)	-0.003	-0.003	0.023	0.019	0.017	0.018
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
Importer (dummy)	0.004	0.005	0.047***	0.042***	0.039***	0.041***
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Municipality-Industry Controls (2007):						
Baseline collocation (log)	0.006	0.006	0.005	0.005	0.005	0.003
	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.006)
Urbanization economies (log)	0.137*	0.136*	0.151**	0.144**	0.146**	0.119
	(0.073)	(0.074)	(0.072)	(0.073)	(0.073)	(0.074)
Herfindahl index, inverse (log)	0.003***	0.004***	-0.001	0.000	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Firms	52,496	52,496	52,496	52,496	52,496	52,496
Industries	73	73	73	73	73	73
Municipalities	2,982	2,982	2,982	2,982	2,982	2,982
R-squared	0.103	0.104	0.097	0.100	0.100	0.100
Industry FE	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y

Notes. The table presents results of the linear probability model estimating the likelihood of firm survival during 2007–2014 as a function of changes in collocation and baseline controls. *Survive* is an indicator variable taking the value one if a 2007 incumbent firm continues to exist in the same municipality and four-digit industry in 2014, and a value of zero otherwise. Identification comes from pre-period industry location patterns. Robust standard errors, clustered at the industry-municipality level, in parentheses. The omitted categories are the smallest firms and the youngest firms (born after 2000). * $p < .10$, ** $p < .05$, *** $p < .01$.

effect is negative but not statistically significant, providing weak evidence that the positive effect is smaller for larger firms. This conclusion is again confirmed in Columns (3)–(6) where, across the specifications, the results continue to show a positive and significant effect of increased collocation on the survival of the smallest firms, and negative but not statistically significant coefficients on the interaction terms. Overall, these results support *Hypothesis 2a* and the conclusion that collocation has positive effects on the survival of firms in nationally traded industries. There is only weak evidence for *Hypothesis 2b*, that small firms benefit from collocation more than large firms.

The coefficients on all other control variables are very similar, both in direction and in magnitude, to those estimated for firms in locally traded industries except for one key difference: the baseline level of collocation. Higher baseline levels of collocation are positively associated with firm survival among firms in nationally traded industries. But the coefficient on the level is much smaller than the coefficient on the change. Specifically, it suggests that a doubling of collocation is associated with a 1.5 percentage point higher firm survival rate. This effect size indeed falls within the range typically found in agglomeration studies, which tend to be within the 1–3 percent range. Meanwhile, the causal estimates using the change in collocation suggest that the actual effects of increased collocation are substantially larger, at least in the short to medium term. Overall, the contrast between the two is consistent with the notion that levels of collocation confound causal effects of collocation with selection of lower quality firms into collocated areas, leading to downward-biased estimates.

Finally, Table 4 presents the result of estimating the parallel models in the sample of firms that are between the two cutoff points I use to classify industries as nationally or locally traded (the "neither" set of industries). Here, changes in collocation have no significant effects on firm survival, consistent with the notion that for firms in this set of industries, no theoretical mechanism clearly dominates. Rather, the opposing mechanisms may offset one another.

5.3 Robustness of Main Results

A key challenge lies in distinguishing the effects of increased collocation with competitors from other potential effects of improved roads. For example, besides changing the nature of

Table 3: *Effects of Change in Collocation on Firm Survival for Nationally Traded Industries*

Dependent variable: *Survive (0/1)*	Baseline (1)	Interacted with size (2)	2 Quantiles in industry (3)	3 Quantiles in industry (4)	4 Quantiles in industry (5)	4 Quantiles in ind-mkt (6)
$\Delta\, CoLoc_{07-14}$	0.149**	0.217**	0.201**	0.188**	0.178*	0.184*
	(0.076)	(0.110)	(0.083)	(0.092)	(0.103)	(0.106)
$\Delta\, CoLoc_{07-14}$ x Firm size (log workers)		-0.030				
		(0.038)				
$\Delta\, CoLoc_{07-14}$ x 2nd quantile			-0.106	0.035	0.063	-0.063
			(0.097)	(0.118)	(0.112)	(0.116)
$\Delta\, CoLoc_{07-14}$ x 3rd quantile				-0.121	-0.030	-0.023
				(0.113)	(0.134)	(0.134)
$\Delta\, CoLoc_{07-14}$ x 4th quantile					-0.133	-0.122
					(0.134)	(0.145)
Firm level controls (2007):						
Firm size (log workers)	0.054***	0.055***				
	(0.004)	(0.004)				
Firm in 2nd quantile (dummy)			0.108**	0.107***	0.089***	0.078***
			(0.009)	(0.010)	(0.011)	(0.011)
Firm in 3rd quantile (dummy)				0.152***	0.147***	0.139***
				(0.011)	(0.011)	(0.011)
Firm in 4th quantile (dummy)					0.169***	0.153***
					(0.013)	(0.013)
Firm born prior to 1993 (dummy)	0.083***	0.083***	0.089***	0.086***	0.086***	0.087***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Firm born during 1993-2000 (dummy)	0.043***	0.043***	0.046***	0.044***	0.044***	0.045***
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Average annual worker wage (log reals)	0.022***	0.022***	0.030***	0.026***	0.025***	0.026***
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Exporter (dummy)	0.001	0.001	0.026*	0.022	0.022	0.022
	(0.015)	(0.015)	(0.014)	(0.014)	(0.014)	(0.014)
Importer (dummy)	0.019	0.020	0.047***	0.043***	0.044***	0.046***
	(0.015)	(0.015)	(0.014)	(0.014)	(0.015)	(0.015)
Municipality-Industry Controls (2007):						
Baseline collocation (log)	0.014**	0.014**	0.016**	0.014**	0.014**	0.015**
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Urbanization economies (log)	0.092*	0.091*	0.094*	0.093*	0.090	0.085
	(0.055)	(0.055)	(0.057)	(0.056)	(0.056)	(0.055)
Herfindahl index, inverse (log)	0.001	0.001	-0.003*	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Firms	21,930	21,930	21,930	21,930	21,930	21,930
Industries	99	99	99	99	99	99
Municipalities	2,302	2,302	2,302	2,302	2,302	2,302
R-squared	0.103	0.104	0.097	0.100	0.100	0.100
Industry FE	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y

Notes. The table presents results of the linear probability model estimating the likelihood of firm survival during 2007–2014 as a function of changes in collocation and baseline controls. *Survive* is an indicator variable taking the value one if a 2007 incumbent firm continues to exist in the same municipality and four-digit industry in 2014, and a value of zero otherwise. Identification comes from pre-period industry location patterns. Robust standard errors, clustered at the industry-municipality level, in parentheses. The omitted categories are the smallest firms and the youngest firms (born after 2000). * $p < .10$, ** $p < .05$, *** $p < .01$.

Table 4: *Effects of Change in Collocation on Firm Survival for "Neither" Industries*

Dependent variable: *Survive (0/1)*	Baseline (1)	Interacted with size (2)	2 Quantiles in industry (3)	3 Quantiles in industry (4)	4 Quantiles in industry (5)	4 Quantiles in ind-mkt (6)
$\Delta\, CoLoc_{07-14}$	0.010	0.007	0.007	0.007	0.008	0.008
	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)	(0.015)
$\Delta\, CoLoc_{07-14}$ x Firm size (log workers)		0.021				
		(0.019)				
$\Delta\, CoLoc_{07-14}$ x 2nd quantile			0.087	0.035	-0.027	-0.011
			(0.055)	(0.065)	(0.066)	(0.066)
$\Delta\, CoLoc_{07-14}$ x 3rd quantile				0.107*	0.070	0.097
				(0.065)	(0.071)	(0.074)
$\Delta\, CoLoc_{07-14}$ x 4th quantile					0.081	0.109
					(0.077)	(0.079)
Firms	52,555	52,555	52,555	52,555	52,555	52,555
Industries	80	80	80	80	80	80
Municipalities	2,860	2,860	2,860	2,860	2,860	2,860
R-squared	0.105	0.105	0.105	0.105	0.106	0.106
Industry FE	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y

Notes. The table presents results of the linear probability model estimating the likelihood of firm survival during 2007–2014 as a function of changes in collocation and baseline controls. *Survive* is an indicator variable taking the value one if a 2007 incumbent firm continues to exist in the same municipality and four-digit industry in 2014, and a value of zero otherwise. Identification comes from pre-period industry location patterns. Robust standard errors, clustered at the industry-municipality level, in parentheses. The omitted categories are the smallest firms and the youngest firms (born after 2000). All controls are included. * $p < .10$, ** $p < .05$, *** $p < .01$.

spatial competition and affecting agglomeration economies, roads can allow firms to access previously unserved customers, they can affect the prices of inputs, and lead to the increased entry of new firms. In this section I test the robustness of the main findings to controls that address these alternative channels though which road upgrades may affect firm performance.

Establishing the effect that improved roads may have on firms through the demand channel, i.e., by enabling better access to markets and customers independent of any spatial competition effects, has been the focus of number of recent studies (e.g., Donaldson and Hornbeck 2016). The standard empirical approach to measuring changes in market access is to calculate the change in total GDP or population accessible to a firm. In line with this literature, I create a variable named *Change in local market access*, which is the change in the travel-time weighted sum of the gross domestic product (GDP) of all municipalities that lie within a firm's four-hour market from 2007 to 2014. As in the measure of change in collocation, variation in market access comes from changes in travel time between municipalities and changes in the market area accessible in four-hours (and not the values of GDP, which are kept at their 2007 values). But unlike the change in collocation measure, market access varies at the level of municipalities, not firms or industries, and, hence, can only be included in a model that excludes municipality fixed effects.

Table 5 shows the results of controlling for change in market access. Here we can compare the baseline model (Column 1) to a model that excludes the municipality fixed effects (Column 2) and then adds a control for the change in local market access (Column 3). For the locally traded industries, both the removal of the municipality fixed effects and the inclusion of the additional control do little to change the main result of a strong negative effect of increased collocation on the smallest firms. The only difference is that the coefficient for the largest firms is now smaller in magnitude and no longer statistically significant. In nationally traded industries, however, the parallel models in Columns (7) and (8) are sensitive to the removal of the municipality fixed effects, but do not appear sensitive to the inclusion of the market access control. The sensitivity of the main effect to the removal of the municipality fixed effects is consistent with the possibility that regions with *worse* outcome trends in traded sectors were more likely to receive road investments, i.e., the possibility that the government directed more investment to lagging regions. Such potential biases emphasize the importance

Table 5: *Robustness Checks*

	Locally traded industries					Nationally traded industries				
	Baseline	Without Muni FE	Local mkt. control	Entry control	Factor prices	Baseline	Without Muni FE	Local mkt. control	Entry control	Factor prices
Dependent variable:	*Survive*	*Survive*	*Survive*	*Survive*	*Wage growth*	*Survive*	*Survive*	*Survive*	*Survive*	*Wage growth*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$\Delta\, CoLoc_{07-14}$	-0.161**	-0.115**	-0.141**	-0.168**	0.028	0.178*	0.051	0.060	0.175*	-0.163
	(0.069)	(0.054)	(0.061)	(0.070)	(0.381)	(0.103)	(0.084)	(0.087)	(0.103)	(0.468)
Δ *Local mkt. access*			0.026					-0.014		
			(0.030)					(0.045)		
$Entry\, rate_{07-14}$				-0.018***					-0.011*	
				(0.003)					(0.006)	
$\Delta\, CoLoc_{07-14}$ x 2nd quantile	0.069	0.036	0.034	0.069	-0.424	0.063	0.020	0.020	0.062	-1.018
	(0.074)	(0.070)	(0.071)	(0.074)	(0.469)	(0.112)	(0.110)	(0.110)	(0.112)	(0.730)
$\Delta\, CoLoc_{07-14}$ x 3rd quantile	0.133*	0.092	0.089	0.135*	-0.174	-0.030	-0.091	-0.091	-0.029	0.014
	(0.076)	(0.077)	(0.077)	(0.076)	(0.414)	(0.134)	(0.116)	(0.117)	(0.134)	(0.505)
$\Delta\, CoLoc_{07-14}$ x 4th quantile	0.159**	0.114	0.111	0.160**	-0.356	-0.133	-0.129	-0.129	-0.131	0.008
	(0.076)	(0.072)	(0.073)	(0.076)	(0.487)	(0.134)	(0.121)	(0.121)	(0.134)	(0.531)
Firms	52,496	52,985	52,985	52,496	28,870	21,930	22,575	22,575	21,930	11,849
R-squared	0.100	0.054	0.054	0.101	0.128	0.152	0.071	0.071	0.152	0.186
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Municipality FE	Y	N	N	Y	Y	Y	N	N	Y	Y
State FE	N	Y	Y	N	N	N	Y	Y	N	N

of the municipality controls.

Columns (4) and (9) address the sensitivity of the results to controlling for the entry of new firms. It may be that road improvements attracted new entrants and that both the competitive and agglomeration effects experienced by incumbent firms are a result of the change in entry dynamics. To test this possibility, I control for the 2007–2014 municipality-industry entry rate, calculated as the number of new entrants during 2007–2014 as a share of the 2007 incumbent firms. The results in Columns (4) and (9) provide evidence that higher entry has a statistically significant negative association with the survival probability for incumbents, suggestive of increased competition. However, the inclusion of the entry control does not impact the main effect of the change in collocation.

Finally, Columns (5) and (10) look for evidence of increased factor prices, by replacing firm survival with a new dependent variable, which is the growth in the average wage of the firm from 2007 to 2014. Note that this analysis can only be performed on the set of surviving firms. The results do not show evidence that the shock to collocation affected wage trends in incumbent firms, offering little evidence for an input prices channel in the case of labor.

6 Additional Adjustments: Industry Switches and Relocations

To test whether firms exhibit other strategic responses to the unexpected changes in collocation, I investigate how much the reduction in the rate of survival, as measured, is due to firms exiting altogether, as opposed to exiting by relocating to a different municipality or switching to a different industry. A firm facing increased competition in its product and local market can respond by repositioning (Wang and Shaver 2014), for example, by changing its location or switching products. Given that they face the more serious competitive threats, I expect that the smallest firms in locally traded industries would be most likely to respond to increased collocation by relocating to a different municipality, or switching to a different product. Meanwhile, in nationally traded industries, increased collocation should make it less likely that a firm would relocate or switch industry, as the original choice of industry and location becomes more attractive due to increased agglomeration spillovers.

I define two additional dependent variables: *Switch* as an indicator taking positive value if an incumbent firm reports a different industry code in 2007 and in the last year that it is observed in the sample, and *Relocate* as an indicator taking positive value if an incumbent firm is located in a different municipality in the last year that it is observed.[24] The model is parallel to Equation 4, with the exception of an additional firm-level control for the last year that the firm is observed in the sample, i.e., the year of "exit" (or the year 2014, in the absence of exit). This control is important because time is the main predictor of product switches and relocations and, as we know from the prior analysis, changes in collocation affect the likelihood of firms surviving and thus remaining in the sample.

Table 6 shows the results of the analyses of industry-switches and relocations in locally and nationally traded industries. The differences across the two industry types are once again impressive and in line with theory. In locally traded industries, the results for the sample as a whole in Columns (1) and (3) show that a positive change in collocation increases the likelihood that a firm relocates and switches to a different industry, although the effects are not statistically significant at typical significance levels. The results become more pronounced, however, after introducing the interactions of collocation with firm size quartiles.

For the smallest firms, both the likelihood of relocating and switching to a different industry increase as collocation increases. Specifically, doubling collocation increases the likelihood of relocating by 3.8 percentage points, and the likelihood of industry switching by 7.4 percentage points among the smallest firms. Given the average likelihood of moves and product switches of 3.1 and 8.4 percent, respectively, this is roughly a doubling of the likelihood of these events. The coefficients on the interaction terms with size suggest that larger firms are less likely to relocate and switch industries, through the differences are not always statistically significant.

In nationally traded industries, on the other hand, the likelihood of relocating falls with an increase in collocation. For the smallest firms, doubling collocation decreases the probability of relocating by 7.4 percentage points, again, more than doubling the baseline probability

[24]Note that, given the definition of *Survive*, any firm moving or switching is part of the subset of firms defined as having not survived in their original location and industry. The main results on survival are robust to the exclusion of movers and switchers from the sample, i.e., full "exits."

of 4.1 percent. The coefficients on the interaction terms suggest that the effect is muted for large firms. The results on industry switches are ambiguous in the empirical results. While the coefficients are negative, suggesting lower propensity to switch industries, none are statistically significant at the standard thresholds.

A final difference becomes apparent when considering the collocation in the origins and destinations that precede and follow the relocation or industry switch. Figure 5 graphically depicts the average difference in the levels of collocation between the origin and destination for all firms that relocate and switch industry. We see that firms that relocate or switch industries in locally traded industries tend to move *away* from competitors, while in nationally traded industries, they move *toward* competitors. While this last piece of evidence is descriptive, it lends support for the main hypothesis that proximity to competitors plays a fundamentally different role in nationally traded and locally traded industries.

7 Conclusion

This study estimates the effect of collocating with firms in one's industry on firm survival, leveraging reductions in travel times between firm locations stemming from improved roads as exogenous variation in collocation. I find that in industries that compete for customers locally, increased collocation produces effects consistent with heightened competition: doubling collocation lowers the survival rate of the smallest firms by 14.1 percentage points while increasing it by 2.6 percentage points for the largest firms. Meanwhile, in industries that compete in national markets, increased collocation produces effects consistent with increased agglomeration spillovers. Doubling collocation increases firms' survival rate by 18.4 percentage points, with no significant differences across firm sizes. As further evidence consistent with increased competition in locally traded industries, I observe a higher propensity of firms to relocate to a different municipality or switch their primary product after being brought closer to competitors, and when they do so, they tend to evade competition. Meanwhile, in nationally traded industries I observe fewer relocations and when these occur, they are moves towards competitors.

Table 6: *Effects of Change in Collocation on Relocations and Product Switches*

	Locally traded				Nationally traded			
Dependent variable:	*Relocate*		*Switch*		*Relocate*		*Switch*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\Delta\, CoLoc_{07-14}$	0.027	0.038*	0.056	0.074*	-0.042	-0.074**	-0.065	-0.064
	(0.019)	(0.022)	(0.039)	(0.041)	(0.031)	(0.031)	(0.050)	(0.061)
$\Delta\, CoLoc_{07-14}$ x 2nd quantile		-0.037**		-0.053		0.040		-0.002
		(0.017)		(0.036)		(0.041)		(0.078)
$\Delta\, CoLoc_{07-14}$ x 3rd quantile		-0.001		-0.021		0.073*		0.006
		(0.021)		(0.039)		(0.044)		(0.070)
Firm in 2nd quantile		0.009***		0.027***		0.008**		0.027***
		(0.002)		(0.004)		(0.004)		(0.006)
Firm in 3rd quantile		0.015***		0.039***		0.018***		0.035***
		(0.002)		(0.004)		(0.005)		(0.008)
Firms	52,496	52,496	52,496	52,496	21,930	21,930	21,930	21,930
R-squared	0.077	0.078	0.113	0.115	0.114	0.115	0.152	0.154
Industry FE	Y	Y	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y	Y	Y

Notes. The table presents the long-differenced analysis of the linear probability model estimating the likelihood of relocations and product switches during the 2007-2014 period as a function of the changes in co-location and baseline controls. *Relocate* and *Switch* are indicator variables taking the value one if a 2007 incumbent firm is observed in a different municipality/ industry code in the last year that it appears in the data. Identification comes from pre-period industry location patterns. Robust standard errors, clustered at the industry-municipality level, in parentheses. The omitted category in the case of size are firms in the first size quantile (smallest) and firms born after 2000 in the case of firm age. All specifications include the full set of controls as well as a control for the last year firm is observed in the data. Some results are sensitive to the number of firm size categories included; the sharpest results (involving three size categories) are presented. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The findings of significant differences in the way that firms respond to collocation with competitors suggest that there is not a single answer to the question how proximity affects performance, but rather that studies need to be careful to consider both industry and firm-level heterogeneity. The estimates of the current study suggest that for some industries, benefits from collocation may be substantially larger than the typical ranges of agglomeration benefits estimated in prior studies. Meanwhile, other industries experience null or negative effects from collocating. This study shows that the tradability of an industry's product, which determines the spatial scope of the markets for the good, is a key statistic that determines the nature of the effects from collocation.

The ability of the study to detect effects on firm behavior based only on changes to the actual cost of mobility also carries the important implication that "space" is not a constant, but rather it is shaped by the costs and patterns of human mobility. This opens up opportunities for further inquiry on how changes in technologies and policies that affect the cost of mobility shape the competitive and collaborative interactions between firms.

References

Agrawal, A., Galasso, A. and Oettl, A. (2017). Roads and Innovation. *Review of Economics and Statistics*, **99** (3), 417–434.

Ahlfeldt, G. M., Redding, S. J., Sturm, D. M. and Wolf, N. (2015). The Economics of Density: Evidence From the Berlin Wall. *Econometrica*, **83** (6), 2127–2189.

Alcacer, J. (2006). Location Choices Across the Value Chain: How Activity and Capability Influence Collocation. *Management Science*, **52** (10), 1457–1471.

— and Delgado, M. (2016). Spatial organization of firms and location choices through the value chain. *Management Science*, **62** (11), 3213–3234.

— and Zhao, M. (2016). Zooming in: A Practical Manual for Identifying Geographic Clusters. *Strategic Management Journal*, **37** (1), 10–21.

Asturias, J., García-Santana, M. and Ramos, R. (forthcoming). Competition and the welfare gains from transportation infrastructure: Evidence from the Golden Quadrilateral of India. *Journal of the European Economic Association*.

Baum, J. A. C. and Haveman, H. A. (1997). Love Thy Neighbor? Differentiation and Agglomeration in the Manhattan Hotel Industry, 1898-1990. *Administrative Science Quarterly*, **42** (2), 304.

— and Mezias, S. J. (1992). Localized competition and organizational failure in the Manhattan hotel industry, 1898- 1990. *Administrative Science Quarterly*, **37** (4), 580–604.

Beaudry, C. and Swann, G. M. P. (2009). Firm growth in industrial clusters of the United Kingdom. *Small Business Economics*, **32** (4), 409–424.

Bernard, A. B., Jensen, J. B., Redding, S. J. and Schott, P. K. (2007). Firms in International Trade. *Journal of Economic Perspectives*, **21** (3), 105–130.

—, Moxnes, A. and Saito, Y. U. (forthcoming). Production Networks, Geography and Firm Performance. *Journal of Political Economy*.

Buenstorf, G. and Klepper, S. (2009). Heritage and Agglomeration: The Akron Tyre Cluster Revisited. *Economic Journal*, **119** (537), 705–733.

Chandra, A. and Thompson, E. (2000). Does public infrastructure affect economic activity? *Regional Science and Urban Economics*, **30** (4), 457–490.

Chung, W. and Alcacer, J. (2002). Knowledge Seeking and Location Choice of Foreign Direct Investment in the United States. *Management Science*, **48** (12), 1534–1554.

— and Kalnins, A. (2001). Agglomeration Effects and Performance: A Test of the Texas Lodging Industry. *Strategic Management Journal*, **22** (10), 969–988.

Combes, P.-P., Duranton, G., Gobillon, L., Puga, D. and Roux, S. (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica*, **80** (6), 2543–2594.

— and GOBILLON, L. (2015). The Empirics of Agglomeration Economies. *Handbook of Regional and Urban Economics*, **5**, 247–348.

DELGADO, M., PORTER, M. E. and STERN, S. (2015). Defining clusters of related industries. *Journal of Economic Geography*, **16** (1), 1–38.

DONALDSON, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. *American Economic Review*, **108** (4-5), 899–934.

DURANTON, G. and TURNER, M. A. (2012). Urban Growth and Transportation. *Review of Economic Studies*, **79** (4), 1407–1440.

ELLISON, G. and GLAESER, E. L. (1997). Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach. *Journal of Political Economy*, **105** (5), 889.

FABER, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. *Review of Economic Studies*, **81** (3), 1046–1070.

FAGGIO, G., SILVA, O. and STRANGE, W. C. (2017). Heterogeneous Agglomeration. *Review of Economics and Statistics*, **99** (1), 80–94.

FIGUEIREDO, J. M. D. and SILVERMAN, B. S. (2007). Churn, Baby, Churn: Strategic Dynamics Among Dominant and Fringe Firms in a Segmented Industry. *Management Science*, **53** (4), 632–650.

FISCHER, J. and HARRINGTON, J. (1996). Product Variety and Firm Agglomeration. *The RAND Journal of Economics*, **27** (2), 281–309.

GHEMAWAT, P. and THOMAS, C. (2008). Strategic Interaction Across Countries and Multinational Agglomeration: An Application to the Cement Industry. *Management Science*, **54** (12), 1980–1996.

GIBBONS, S., LYYTIKÄINEN, T., OVERMAN, H. G. and SANCHIS-GUARNER, R. (2019). New road infrastructure: The effects on firms. *Journal of Urban Economics*, **110**, 35–50.

Haltiwanger, J. C., Lane, J. I. and Spletzer, J. R. (1999). Productivity Differences across Employers: The Roles of Employer Size, Age, and Human. *The American Economic Review*, **89** (2), 94–98.

Haveman, H. a. and Rider, C. I. (2014). The Spatial Scope of Competition and the Geographical Distribution of Entrepreneurship: Magazine Foundings and the U.S. Post Office. *Sociological Science*, **1**, 111–127.

Helpman, E., Itskhoki, O., Muendler, M.-A. and Redding, S. J. (2017). Trade and inequality: From theory to estimation. *The Review of Economic Studies*, **84** (1), 357–405.

Henderson, J. V. (2003). Marshall's scale economies. *Journal of Urban Economics*, **53** (1), 1–28.

Holl, A. (2016). Highways and productivity in manufacturing firms. *Journal of Urban Economics*, **93**, 131–151.

Hotelling, H. (1929). Stability in Competition. *The Economic Journal*, **39** (153), 41–57.

Jacobs, J. (1969). *The Economy of Cities*. New York: Random House.

Kalnins, A. and Chung, W. (2004). Resource-Seeking Agglomeration: A Study of Market Entry in the Lodging Industry. *Strategic Management Journal*, **25** (7), 689–699.

Krugman, P. (1991). *Geography and Trade*. Cambridge, MA: MIT Press.

Marshall, A. (1920). *Principles of Economics*. London: MacMillan.

Martin, P., Mayer, T. and Mayneris, F. (2011). Spatial concentration and plant-level productivity in France. *Journal of Urban Economics*, **69** (2), 182–195.

McCann, B. T. and Folta, T. B. (2008). Location Matters: Where We Have Been and Where We Might Go in Agglomeration Research. *Journal of Management*, **34** (3), 532–565.

Melitz, M. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica*, **71** (6), 1695–1725.

Mian, A. and Sufi, A. (2014). What Explains the 2007-2009 Drop in Employment? *Econometrica*, **82** (6), 2197–2223.

Morten, M. and Oliveira, J. (2016). The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City. *NBER Working Paper 22158*.

Porter, M. E. (1990). *The Competitive Advantage of Nations*. New York: The Free Press.

Rosenthal, S. S. and Strange, W. C. (2003). Geography, industrial organization, and agglomeration. *Review of Economics and Statistics*, **85** (2), 377–393.

Salop, S. C. (1979). Monopolistic Competition with Outside Goods. *The Bell Journal of Economics*, **10** (1), 141–156.

Savaris, B., Vinagre, F. and Magalhaes, D. (2013). *The Brazilian Infrastructure: It's "Now or Never"*. Tech. rep., Credit Suisse Securities Research and Analytics.

Shaver, M. J. and Flyer, F. (2000). Agglomeration Economies, Firm Heterogeneity, and Foreign Direct Investment in the United States. *Strategic Management Journal*, **21** (12), 1175–1193.

Singh, J. and Marx, M. (2013). Geographic Constraints on Knowledge Spillovers: Political Borders vs. Spatial Proximity. *Management Science*, **59** (9), 2056–2078.

Sorenson, O. and Audia, P. G. (2000). The Social Structure of Entrepreneurial Activity: Geographic Concentration of Footwear Production in the United States 1940-1989. *American Journal of Sociology*, **106** (2), 424–462.

Syverson, C. (2004). Market Structure and Productivity: A Concrete Example. *Journal of Political Economy*, **112** (6), 1181–1222.

Vogel, J. (2008). Spatial Competition with Heterogeneous Firms. *Journal of Political Economy*, **116** (3), 423–466.

Wang, R. and Shaver, M. (2014). Competition-Driven Repositioning. *Strategic Management Journal*, **1604**.

Table 7: *Firm Level Descriptive Statistics*

	Full sample n=129,325		Locally traded n=52,496		Nationally traded n=21,930	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Size (workers)	23.307	86.272	18.525	53.039	39.002	169.143
Age (years)	8.412	5.099	8.151	5.040	8.858	5.172
Avg. worker wage (reals)	8818.814	6080.362	7794.399	5053.723	9914.935	7295.826
Exporter	0.050	0.218	0.029	0.166	0.088	0.284
Importer	0.048	0.214	0.028	0.165	0.083	0.275
Survive	0.556	0.497	0.557	0.497	0.554	0.497
Relocate	0.037	0.188	0.031	0.173	0.045	0.207
Switch product	0.107	0.309	0.084	0.278	0.110	0.312
Level of collocation (2007)	20.468	51.863	34.101	75.281	5.210	7.258
Change in collocation	0.036	0.082	0.041	0.084	0.026	0.078
Ellison-Glaeser index	0.033	0.040	0.011	0.005	0.095	0.061
Urbanization economies	497.821	625.985	456.378	594.047	526.325	659.296
Herfindahl index, inverse	8.292	3.689	8.586	3.638	7.303	3.813

Table 8: *Predicting Change in Collocation with Firm-Level Characteristics, Locally Traded Industries*

Dependent variable: $\triangle Coloc_{07-14}$	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Number of workers (log)	-0.000						
	(0.000)						
Firm born prior to 1993 (dummy)		0.001**					
		(0.001)					
Firm born during 1993-2000 (dummy)		0.001**					
		(0.000)					
Average annual worker wage (log reals)			-0.000				
			(0.000)				
Exporter (dummy)				0.002			
				(0.001)			
Importer (dummy)					0.001		
					(0.001)		
Political donations (log reals)						-0.000	
						(0.000)	
Pre-period survival rate							-0.000
							(0.002)
Baseline collocation (log)	-0.015***	-0.015***	-0.015***	-0.015***	-0.015***	-0.015***	-0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Firms	52,496	52,496	52,496	52,496	52,496	52,496	52,496
R-squared	0.846	0.846	0.846	0.846	0.846	0.846	0.846
Industry FE	Y	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y	Y

Table 9: *Predicting Change in Collocation with Firm-Level Characteristics, Nationally Traded Industries*

Dependent variable: $\triangle Coloc_{07-14}$	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**
Number of workers (log)	-0.000						
	(0.000)						
Firm born prior to 1993 (dummy)		-0.001					
		(0.001)					
Firm born during 1993-2000 (dummy)		-0.001					
		(0.001)					
Average annual worker wage (log reals)			0.000				
			(0.000)				
Exporter (dummy)				0.001			
				(0.001)			
Importer (dummy)					-0.001		
					(0.001)		
Political donations (log reals)						-0.000	
						(0.000)	
Pre-period survival rate							-0.002
							(0.003)
Baseline collocation (log)	-0.018***	-0.018***	-0.018***	-0.018***	-0.018***	-0.018***	-0.018***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Firms	21,930	21,930	21,930	21,930	21,930	21,930	16,738
R-squared	0.692	0.692	0.692	0.692	0.692	0.692	0.754
Industry FE	Y	Y	Y	Y	Y	Y	Y
Municipality FE	Y	Y	Y	Y	Y	Y	Y

Table 10: *Fifteen Largest Industries (by Number of Firms in Sample)*

CNAE Code	Locally traded industries	CNAE Code	Nationally traded industries
1812	Apparel manufacturing, except underwear, sweaters, shirts	1931	Leather shoes
1581	Bakery products, confectionery and pastry manufacturing	2010	Wooden logs
2812	Manufacture of metal frames	2892	Artifacts manufacturing drawn
2222	Printing of educational material, and material for industrial or commercial uses	1779	Other textile products (knitting)
2842	Manufacturing locksmith articles, except frames	1939	Manufacture of footwear of other materials
2522	Manufacture of plastic packaging	2519	Manufacture of other rubber products
1589	Other food manufacturing	2833	Manufacturing metal stamped artifacts
2029	Manufacturing various wooden goods, straw, cork and braided materials, except furniture	2931	Manufacture of machinery and equipment for agriculture, poultry and obtaining animal products
1542	Dairy product manufacturing	2021	Plywood, laminated, pressed or sintered wood manufacturing
1813	Manufacture of professional apparel	2473	Manufacture of perfumes and cosmetics
2893	Manufacturing body shop items and metal articles for household use and personal	2940	Machine tool manufacturing
1749	Manufacture of other textile articles, including weaving	3210	Basic electronic material manufacturing
1543	Ice cream manufacturing	2481	Manufacture of paints, varnishes, enamels and lacquers
1584	Pasta manufacturing	2749	Other non-ferrous metals and alloy metallurgy
1921	Manufacture of bags, handbags, suitcases and other travel goods of any material	2221	Printing of newspapers, magazines and books

Source: Author's calculations based on RAIS data, year 2007. Definition of locally traded industries uses the Ellison-Glaeser index, with a value of 0.2 or less. Nationally traded industries based on the Ellison-Glaeser index, with a value of 0.5 or more.

Table 11: *Table of Correlations of Key Variables*

	Variables	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	Firm size (workers)	1												
2	Age (years)	0.0952	1											
3	Avg. annual worker wage (reals)	0.1379	0.129	1										
4	Exporter	0.2708	0.123	0.2499	1									
5	Importer	0.2779	0.1201	0.3064	0.4848	1								
6	Survive	0.0446	0.1243	0.0213	0.0561	0.0602	1							
7	Relocate	0.0186	-0.035	0.0576	0.0218	0.0243	-0.1461	1						
8	Switch product	0.0273	0.0166	0.0514	0.0214	0.0312	-0.3122	0.0578	1					
9	Level of collocation (2007)	-0.0251	0.0067	0.0124	-0.0151	-0.0284	0.0074	-0.0162	-0.061	1				
10	Change in collocation	-0.0175	0.0068	-0.0541	-0.0294	-0.0263	-0.0141	-0.0113	0.0098	-0.0731	1			
11	Ellison-Glaeser index	0.086	0.0335	0.0917	0.0837	0.0822	-0.0071	0.0248	0.0203	-0.1342	-0.0555	1		
12	Urbanization economies	0.0206	0.1406	0.3207	0.1038	0.1257	0.0298	0.0508	0.014	0.4228	-0.1415	0.0344	1	
13	Herfindahl index, inverse	-0.1257	0.0044	-0.0538	-0.0776	-0.0989	0.0137	-0.0068	-0.0634	0.0742	0.0606	-0.1396	0.0435	1

Note: Author's calculations based on data from *Rela??o Annual de Informa??es Sociais* (RAIS), SECEX, and Ministry of Transport of Brazil.

8 Figures

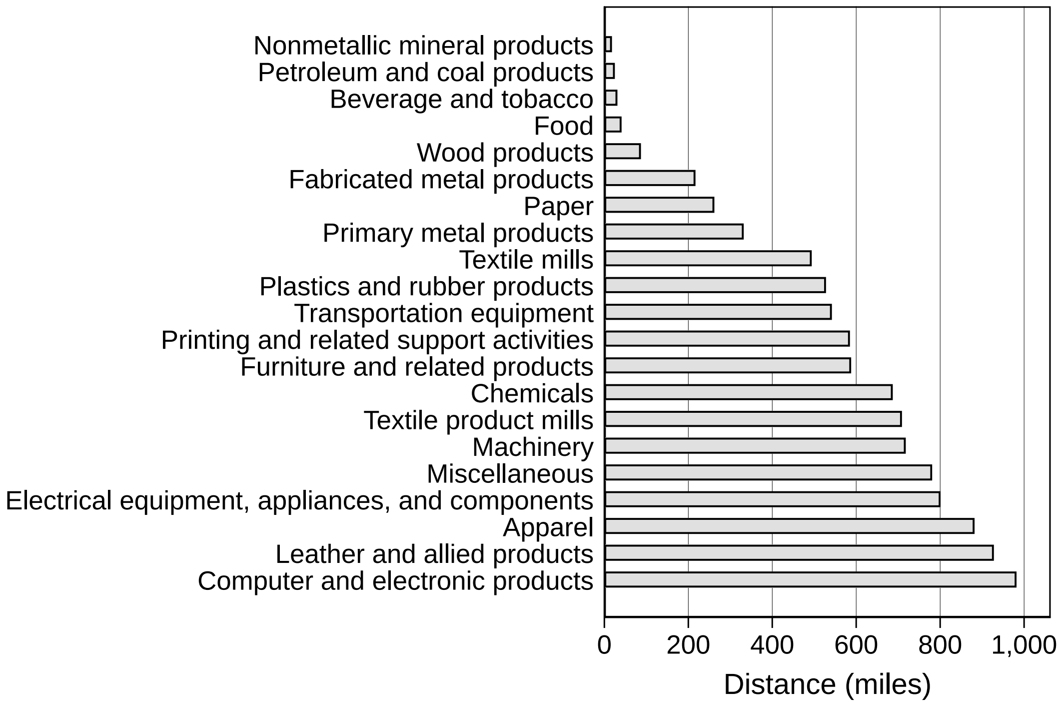

Figure 2: *Median Distance Goods Shipped, 2-digit NAICS sector*

Note: Author's calculations using the 2012 U.S. Commodity Flow Survey (CFS).

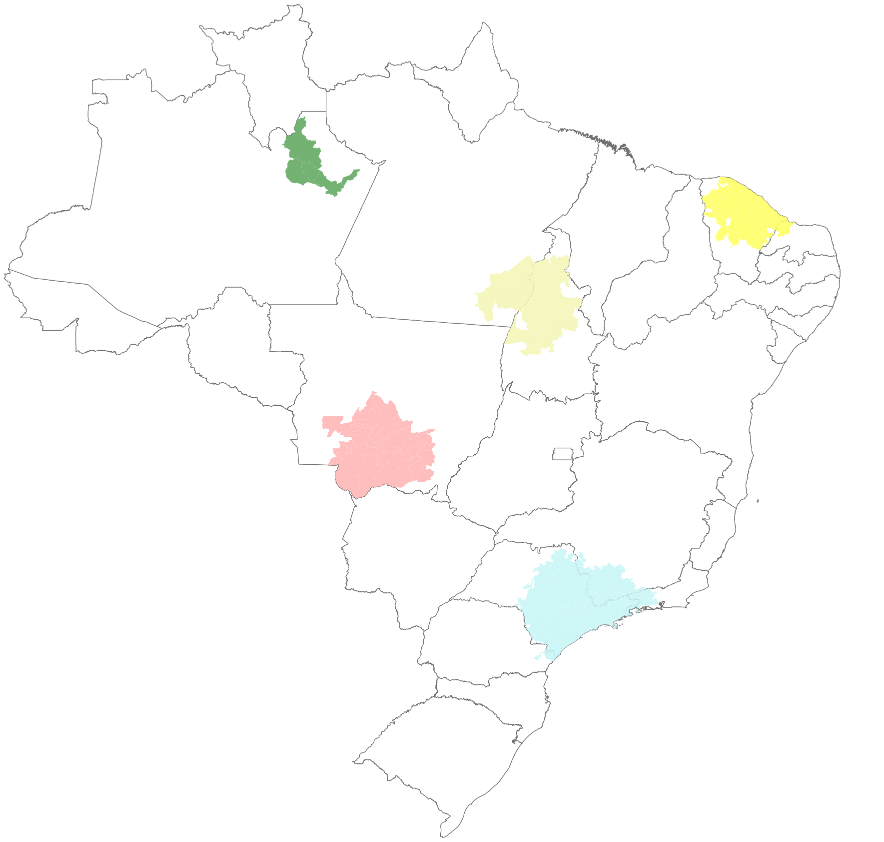

Figure 3: *Select Markets as Defined by Four Hours of Travel Time*

Note: Map based on author's calculations and data from the Ministry of Transport of Brazil, the National Department of Transportation Infrastructure (DNIT), and the Brazilian National Transport Confederation (CNT).

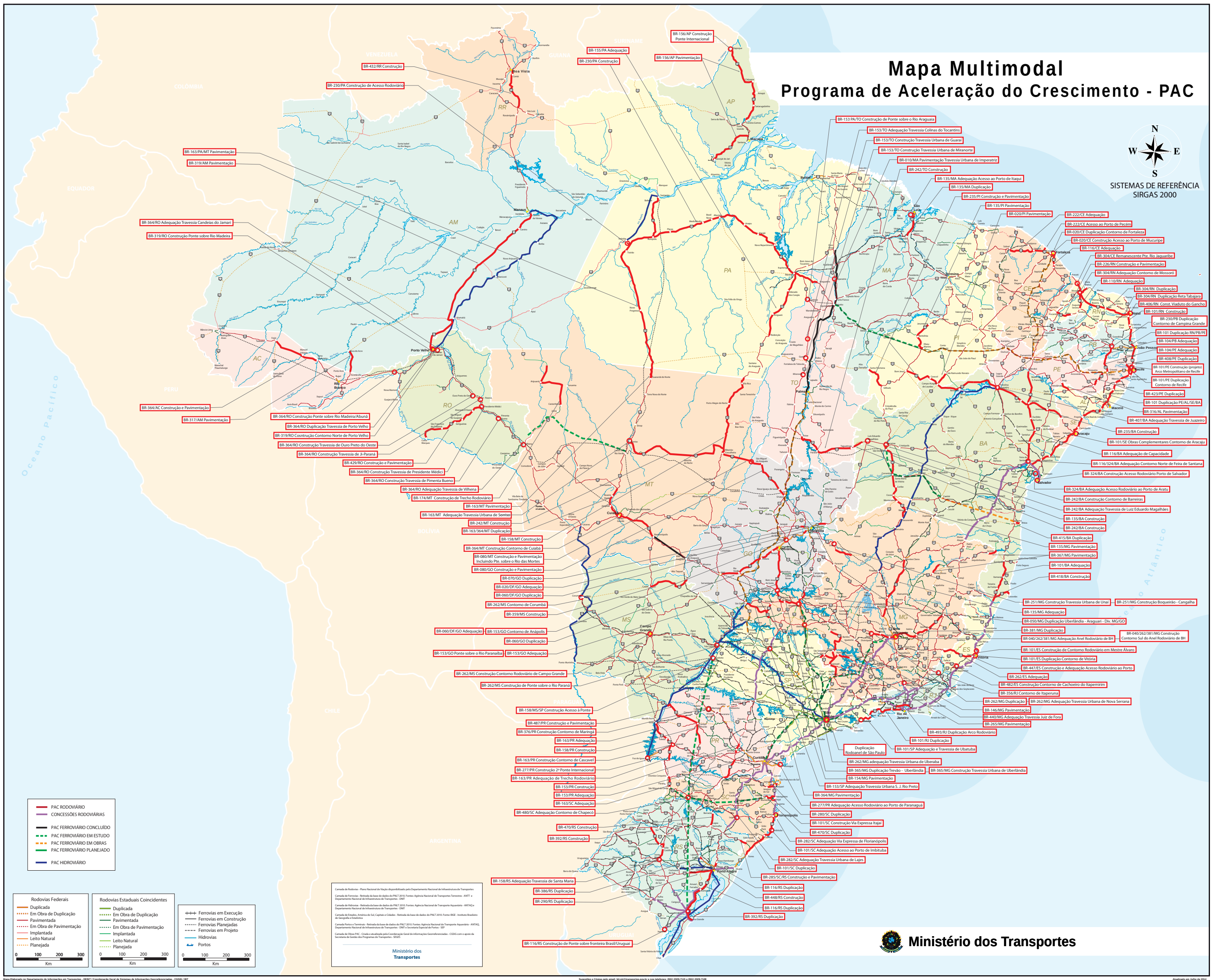

Figure 4: *Programa de Aceleração do Crescimento — Programmed Investments*

Source: Ministerio dos Transportes de Brasil.

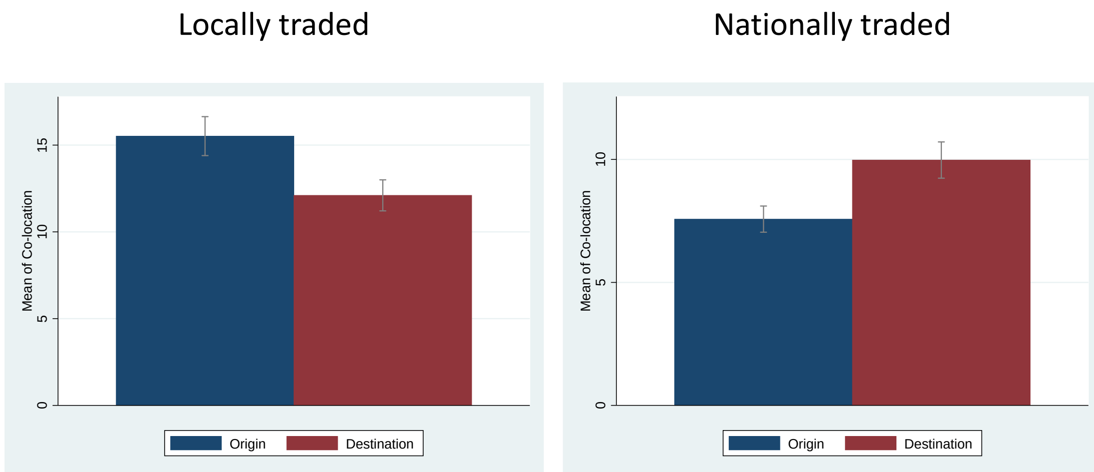

Figure 5: *Origin and Destination Collocation*

Note: Author's calculations.

9 Supplemental Data Appendix

Available from author upon request.