The Internal Geography of Firms*

Dominick Bartelme
University of Michigan

Oren Ziv
Michigan State University

March 2022

Abstract

We document that plants belonging to small and mid-sized firms are geographically concentrated, while large firms are much more dispersed. These differences are sizable; firms with 2 plants have a dispersion that is 5 log points lower than predicted by industry location patterns, while the corresponding figure is less than 2 log points for firms with 40 plants and less than a half for firms with 100 or more plants. These patterns are qualitatively robust across industries, time periods, and alternative specifications. We also find that plants that are farther from the firm headquarters employ significantly less workers than closer plants within the same firm, and that this effect is attenuated in large firms. We interpret these findings through the lens of a model of plant location that suggests that large firms face lower costs of geographic expansion.

JEL codes: R32,L25

Keywords: economic geography, agglomeration, firm location, firm growth

---

*Bartelme: dbartelm@umich.edu. Ziv: orenziv@msu.edu. We thank C. Luke Watson for excellent research assistance as well as Camillo Acosta-Mejia, Andrei Levchenko, Jeffrey Lin, Martha Stinson, Junfu Zhang, Jens Wrona and numerous seminar and conference participants for their helpful comments. The paper draws on confidential data from the US Census Bureau, which requires the following disclaimer: "Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the US Census Bureau. This research was performed at a Federal Statistical Research Data Center. All results have been reviewed to ensure that no confidential information is disclosed."

1 Introduction

An important decision facing large firms is where to locate their establishments, both in absolute terms and in relation to one another. Geographic dispersion allows firms to maximize access to markets and local inputs, but can also generate costs for the firm, such as transportation and time costs associated with shipping intermediate inputs between plants or management and coordination costs that increase with distance. An established literature documents some of these costs and their effects on plant performance in various settings.[1] However, we know far less about the quantitative importance of these forces in determining the locations of firms' constituent plants.

This paper uses U.S. Census microdata to document several empirical regularities in the plant location decisions of firms within the continental United States. We focus on the location of a firm's plants relative to one another, rather than the particular regions in which firms locate. Using both cross-sectional and time series approaches, we show that small and medium-sized firms tend to cluster their establishments closely in space to a far greater extent than suggested by industry-level agglomeration and co-agglomeration patterns. However, this clustering is much less evident for large firms, whose location decisions appear to "defy gravity." We also show that measures of plant performance decline with plant distance from the center of the firm more strongly for small and medium sized firms relative to large firms. One potential explanation for these results is that the costs of geographic expansion are lower for large firms.

We show that on average firms tend to cluster their establishments closely in space. For each observed firm we construct a counterfactual firm with the same number of plants in the same industries, but that chooses plant locations randomly according to industry agglomeration and co-agglomeration patterns without regard for the internal distance

[1]Several papers investigate the link between proximity (via trade, travel and communication costs) and plant performance, especially (Giroud, 2013; Kalnins and Lafontaine, 2013; Eichholtz, Holtermans, and Yönder, 2015; Alcacer and Delgado, 2016; Charnoz, Lelarge, and Trevien, 2018; Atalay et al., 2019). These papers generally find that distance to management and/or other production units reduces plant performance.

between the plants. We assess the degree of firm clustering by comparing the distance between the plants of the actual firms to those of the counterfactual firms. The magnitude of this difference can be substantial; for example, two randomly chosen plants in the same 6-digit NAICS industry are expected to be roughly 800 miles away from each other, while the average distance between plants belonging to a single firm with 2 plants in the same industry is about 4 miles (a gap of about 5 log points). This pattern is also evident in the time series: firms that add plants expand their geographic footprint by substantially less than the baseline industry location patterns alone would predict.

More surprisingly, we show that this within-firm spatial clustering is strongly heterogeneous across the firm size distribution. Large firms are much less spatially clustered than small and medium firms, relative to the level of clustering predicted by their own matching counterfactual firms. Moreover, large firms tend to add establishments in a way that is close to indistinguishable from the manner in which the counterfactual firms add establishments. For example, relative to firms with two plants, firms with 40 plants have less than half the percent difference between actual and counterfactual geographic footprint (5 log points vs less than 2 log points), and firms with over 100 plants close the gap by roughly 90%. The correlation between size and dispersion also holds *within* size classes based on the number of plants: a doubling of firm employment is associated with between a 20% and 100% increase in dispersion compared to similar firms with the same number of plants. It also holds with respect to future growth; firms that will add plants in the next five years are initially more dispersed than observationally similar firms that do not add plants over that time period, with the difference being larger for smaller firms. These conclusions are robust across industries and time periods, as well as to numerous alternative samples, measurements, hold for other distance metrics, including distance to headquarters and alternative baseline comparison groups and controls.

To help interpret these findings, we model the firm's decision of where to locate plants around a headquarters. Firms pay fixed costs to open new establishments producing a distinct varieties, and receive a location-specific productivity shifter that decreases with distance from the headquarters. When these productivity penalties are common across all firms, the model predicts that more productive firms will have more plants, and firms with more plants will be more dispersed, consistent with our findings above. However, the observed within-firm distance-productivity gradient will be less steep for smaller firms, who far-away plants are differentially positively selected. The gradient will be steeper for smaller firms only if large firms face smaller costs related to distance.

We then examine the empirical relationship between internal distance and plant performance. We test whether plant performance declines with distance in the same way for large and small firms. We find that the elasticity of plant employment with respect to distance to the firm's headquarters is negative and sizable for small firms but significantly smaller in magnitude for large firms. Our model suggests an appealing and parsimonious explanation rationalizing all of the empirical patterns; firms face costs of geographic dispersion, but large firms face lower costs. These cost differences could be primitive, or they might be the result of different investments made by firms with different productivity levels.[2]

Our results support and extend some key results from the literature on plant location in multi-unit firms. Our finding of geographical clustering in the cross-section are consistent with prior results on headquarters locations Henderson and Ono (2008). Other recent work examines the contribution of input-output linkages Behrens and Sharunova (2015) and managerial structure Antoni, Gumpert, and Steimer (2022) to firms' spatial clustering. Our time series results generalize the findings of Holmes (2011), who ties the gradual spatial expansion of Walmart to supply chain management. In contemporaneous work,

[2]A number of other mechanisms such as differences in the spatial correlation of demand or in internal cannibalization effects (e.g. Tintelnot (2017)) in combination with sizable trade costs could also be contributing to the observed differences in location patterns across the firm size distribution. Evaluating the contribution of these alternative mechanisms requires developing and estimating a tractable quantitative structural model of plant location with many alternatives, a difficult problem on which there has been some recent progress (Arkolakis, Eckert, and Shi, 2021; Oberfield et al., 2020). This exercise is beyond the scope of this paper.

(Acosta and Lyngemark, 2021) find expanding firms also are less geographically compact. Motivated by the agglomeration literature (Ellison and Glaeser, 1997; Duranton and Overman, 2005), we develop a novel formal test for clustering, which compares firms' location decision to firm-specific random baselines that preserves the firm's industrial structure. Previous work has documented that investment (Giroud, 2013) and size (Kalnins and Lafontaine, 2013; Eichholtz, Holtermans, and Yönder, 2015) decline with distance within firms, and that establishment survival rates decline with distance (Alcacer and Delgado, 2016). This literature highlights management and communication costs as a source for these patterns (Charnoz, Lelarge, and Trevien, 2018; Antoni, Gumpert, and Steimer, 2022).

Our primary and most distinctive contribution to this literature is to show that these results are driven by the smallest multi-units, and location decisions of and the impact of distance on larger firms are quantitatively and qualitatively distinct. With respect to the cross-sectional clustering of firms, we show that for the largest firms, clustering is nearly indistinguishable from the random baseline. With respect to our findings within the firm, we again find that distance elasticities of productivity within the firm are present for smaller firms, but are 60% to 70% lower for larger firms. These new findings suggest that either the geographic forces at work in or their equilibrium impact on the largest firms is different.[3] The idea that distance costs are lower for the most productive (and in equilibrium largest) firms has the potential to explain findings in the multi-establishment geography literature, for example why distant establishments have more management layers (Antoni, Gumpert, and Steimer, 2022), while at the same time firms centralize management as they add more distant establishments (Acosta and Lyngemark, 2021).

The possibility that larger firms benefit from lower distance costs also has implications for the geography of multinational activity. In this literature, concentration is linked to management- and communication-correlated distance costs within the firm, which

[3]The literature on the geography of large firms is substantially slimmer. Hsieh and Rossi-Hansberg (2019) and Rossi-Hansberg, Sarte, and Trachter (2018) consider the extensive margin expansion of large firms into local new cities but do not consider the geographic footprint of these firms. In contemporaneous work, Oberfield et al. (2020) find that the largest firms sort into denser locations within cities.

reduce productivity in distant affiliates (Keller and Yeaple, 2013; Bahar, 2020; Gumpert, 2018).Irarrazabal, Moxnes, and Opromolla (2013) and Antràs and Yeaple (2014) find that affiliate sales fall with distance to the home country, a result that we find applies to domestic plants that are further from the firm center as well. These clustering patterns are traditionally explained with a concentration-proximity trade-off (e.g. Yeaple (2009); see Alfaro and Chen (2018) for a review), which typically takes the form of a variable-profit vs. fixed cost trade-off where distance costs are symmetric for large and small firms. In our model, a similar trade-off can generate clustering but *does not* yield within-firm distance elasticities differentials that are consistent with our empirical findings, while heterogeneous distance costs do. Overall, the heterogeneity that we find in location decisions and spatial productivity penalties across large and small firms has not yet been explored in the literature on multinational location decisions. Our findings point to the potential existence and significance of this heterogeneity in explaining the pattern of multinational expansion.

2 Data and Measurement

Measuring Dispersion For firms with no headquarters identifiable in our data, our principal measure of the geographic footprint of the firm is the log average number of miles between firm i's constituent plants and the geographic centroid of the firm, given by

$$\log(Dispersion_i) = \log\left(\frac{\sum_{n=1}^{N} d_{nc}}{N}\right), \tag{1}$$

where d_{nc} is the distance between plant n and firm centroid $c = (c_x, c_y)$.[4] This provides a computationally tractable measure of these establishments' dispersion in a way that

[4]We approximate the coordinates of the centroid as the average latitude and longitude of all the establishments in the firm, given by

$$c_x = \frac{\sum_{n \in N} \mathbf{x}_n}{N}, \; c_y = \frac{\sum_{n \in N} \mathbf{y}_n}{N}$$

where $\mathbf{x}_n$ and $\mathbf{y}_n$ are coordinates of the nth point.

includes multilateral resistance – taking into account each establishment's relation to all others in the firm. We refer to this measure as the "dispersion" of the firm and refer to the centroid as the "center of the firm."[5]

We also measure distance to headquarters using Equation 1, replacing d_{nc} with plant's distance to the firm headquarters or regional headquarters. Headquarters information is missing for most firms. We provide results using distance to headquarters as robustness checks of our main results in Section 3, and explicitly use regional headquarters in addition to distance to centroid when investigating the within-firm distribution of plant locations and plant characteristics in Section 5.

To operationalize our measures, we use the firm identifiers and establishment zip codes to first calculate the geographic center of the firm and identify headquarters.

Main Sample Our main sample uses data from multi-unit firms (firms with two or more establishments) observed in any 5-year Economic Census between 1992 and 2012 where all establishments in the firm have positive sales and more than one employee, and where the firm exists in at least two zip codes so that an internal firm distance can be measured. We remove extreme outliers: any establishment above the top 0.05 percentile in sales, employment, payroll, and (for manufacturing establishments) value added.[6] Our sample does not include foreign establishments of U.S. firms, although it does include the U.S. establishments of foreign firms.

Section 3 also uses a sub-sample of these firms that expanded between Census years. We find expanding firms by comparing the number of establishments in each firms in each Census wave, isolating firm-year observations where the firm moved one size class (defined below) from the preceding Census. We find approximately 44,000 such firms.[7]

[5]This measure is conceptually and practically related to the (log) mean of the distribution of bilateral plant distances, which is used by Behrens and Sharunova (2015) and similar to that of Duranton and Overman (2005). Appendix B shows that the two measures are highly correlated in practice.

[6]All results are robust to eliminating all of these cuts as well as to restricting our sample to establishments with greater than five employees.

[7]Because a significant fraction of establishments report zero employment in their first year, we include in this sub-sample establishments who in their first year report fewer than two employees but that do report

Size classes To explore heterogeneity in our findings by firm size, we group firms into 22 size categories: groups 1 through 9 for firms with between 2 and 10 establishments respectively, next grouping by 5s for firms with between 11 and 40 establishments, then by 10s for firms with up to 100 establishments, and a final group for firms with more than 100 establishments. These are chosen in part to satisfy Census disclosure requirements. All results are robust to using the number of establishments as a continuous measure.

Sectors Distance forces and their quantitative effects on firms vary by sector. We assign firms in our main sample into five sectors: manufacturing, business services, wholesale and retail, finance and insurance, and all others.[8] In the main body of the paper, we report results pooling all firms and for the manufacturing sector. Results for business services, also tradable, are similar to those for manufactures, while other sectors are well represented by results pooling all firms. Appendix B breaks out main results by sector.

Firm and establishment measures Key outcome variables in Section 3 will be employment and sales per worker, which are observed or calculated at the establishment and firm levels. Additional results in appendix B use value added, value added per worker, and TFPr. All measures are from Economic Censuses and the Longitudinal Business Databaes (LBD). As a control, we use age group (0-5, 5-15, and 15+ years) for establishments and firms (measured as the maximum age group of the firms' establishments). Appendix A provides further description and summary statistics for our samples and the measures we employ in our main analysis.

---

more than one employee in subsequent Census waves.

[8]Firms may have establishments coded in multiple sectors. To isolate firms' modal sector, we first follow Fort and Klimek (2016) to assign consistent NAICS codes to establishments before 1997. We then calculate each sectors' share of the firms, weighing establishments by their sales. Manufacturing firms are then defined as firms where the modal dollar of revenue is generated at a manufacturing establishment. Weighing establishments by employment does not change classifications significantly and does not change results. As a further robustness check, we divide firms into firm-sector groupings and repeat our empirical exercise on these sub-firm units. Our results are robust to this cut of the data as well.

3 The Spatial Pattern of Firm Location Decisions

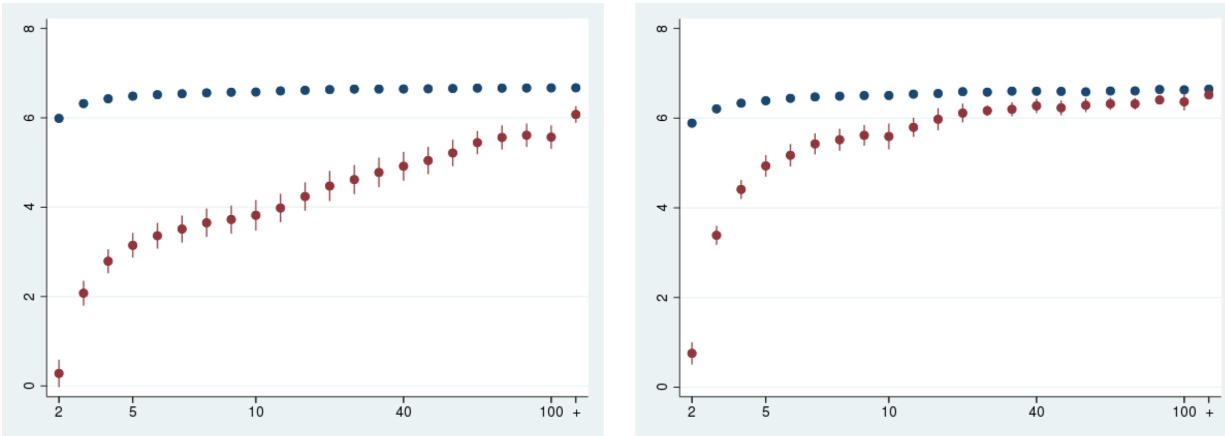
This section documents several empirical patterns in the domestic location decisions of multi-plant firms in the United States. Small and medium-size firms exhibit clustering their establishments relative to a counterfactual benchmark in which firms choose plant locations that replicate industry location patterns. However, larger firms exhibit substantially less clustering. Both of these patterns holds in the time series: small and medium-size firms that add establishments choose locations that are close to their existing establishments, while large firms that add establishments do so to a lesser degree.

3.1 Firm Dispersion in the Cross-Section

We first examine the cross-sectional patterns in within-firm plant location decisions. Figure 1a plots the average log dispersion by size class (in terms of the number of plants) for all firms in our sample with 2 or more plants, with standard errors clustered at the industry level, while Figure 1b plots the means for the manufacturing-only sub-sample. Small firms are highly geographically clustered: for firms with only two plants, the average distance to the firm center is less than two miles. Geographic dispersion rises steadily with the number of plants: firms with 41-45 plants have an average distance from the centroid of about 150 miles. A similar pattern holds for the manufacturing only sub-sample, although manufacturing firms are significantly more dispersed than the average firm along the entire size class distribution.

However, this clustering is meaningful to the extent to which this pattern is distinct from the geographic clustering we observe in plant location across firms. That is, we should ask: how tightly clustered are a firm's plants relative to the clustering behavior of a "similar" group of plants outside the firm? To answer this question, we construct a matching set of synthetic firms with plant locations drawn randomly from the set of "similar" plants, and use the dispersion observed in the synthetic firms as a baseline for

Figure 1: Cross-Sectional Firm Dispersion



(a) All firms (b) Manufacturing

Notes: Red circles plot the average log mean establishment distance from firm centroid for each firm size category. Blue squares plot the corresponding measure for synthetically constructed firms. All standard errors are clustered by firm modal 4-digit industry.

comparison to the data. For each 6-digit NAICS code k and year t, and each firm i with 2 or more plants, we group all the plants in the sample that belong to industry k at time t to create sets of plants P_{kt} with the observed geographic locations and a set of firms F_t with their observed number of plants, their industries and their locations. For each actual firm in $f_{it} \in F_t$, we create a synthetic firm $f_{it} \in F_t$ with the same number of plants in the same industries, but with the geographic locations of its plants randomly chosen, with equal probability and with replacement, from the relevant set P_{kt}.[9] We then compute the average log distance from the firm's center by size class for the synthetic firms, exactly as we did with the actual firms, and compare the results. This procedure controls for any industry agglomeration and co-agglomeration patterns in the data, allowing us to separate the within-firm component of geographic clustering from the across-firm components.

Figures 1a and 1b plot the results from the synthetic baseline alongside the actual data. In contrast to the pattern of rising dispersion found in the data, the average dispersion in the synthetic baseline is relatively flat across size classes and similar for both sub-groups,

[9]Drawing with replacement yields a lower expected value of the synthetic firm's dispersion. We discuss robustness to alternative sampling procedures in Appendix B.3, including procedures which minimize and eliminate self-sampling.

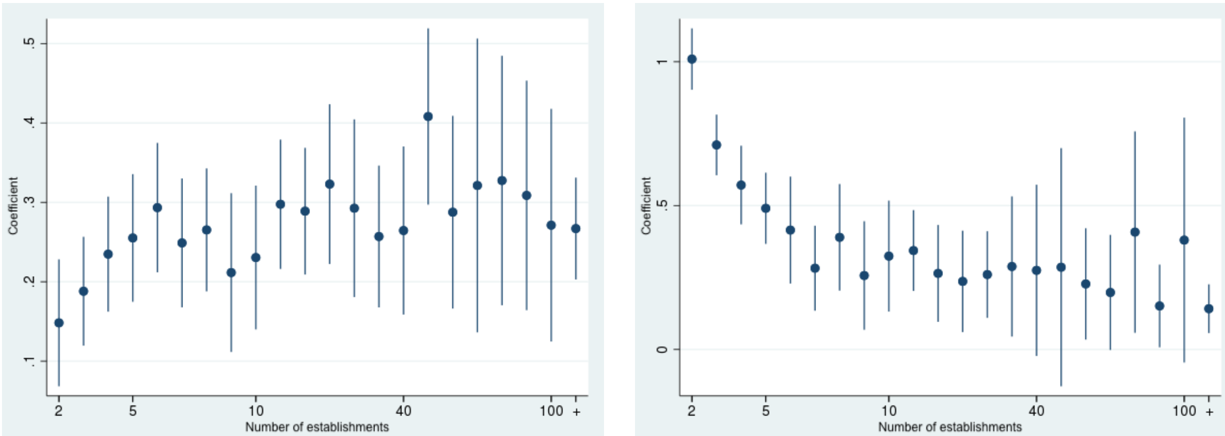
averaging about 500 miles. For small firms the baseline dispersion is far greater than the actual dispersion: the average 2-plant firm has a dispersion that is about 5 log points lower than the baseline. This gap shrinks for medium-sized firms, with actual dispersion being about 1 log point (for manufacturing firms) or 2.5 log points (for all firms) lower than baseline for firms with 10 plants. Small and medium-sized firms, therefore, are highly geographically clustered relative to the typical group of plants in their industries. In contrast, the gap between the random baseline and the actual almost entirely disappears for manufacturing firms with more than 20 plants and for the very largest non-manufacturing firms. Large firms "defy gravity" in their location choices, especially in manufacturing.

These results are robust to a number of specification changes and alternative measurements, all reported and discussed in more detail in Appendix B.3. We show that results are similar when using distance to headquarters as the measure of dispersion and when sorting firms by employment ventile rather than by the number of plants in the firm.[10] We also show that these results are not driven by differences in industry composition or other observable firm characteristics across the size categories, by running regressions of log dispersion on size-class dummies (to identify average dispersion by size class) while controlling for industry-time dummies and other firm characteristics. We discuss the specification and results in more detail in Appendix B.3, but the bottom line is that it makes little qualitative or quantitative difference for our results.

Another concern is that, while our synthetic baseline controls for the industry composition of each firm's plants, there may be other omitted plant characteristics that drive agglomeration patterns both within and across firms. In a robustness check, we use the Census of Manufactures to match plants based on the full set of products produced and inputs used in production when constructing the synthetic firms and find similar results. Finally, we check to make sure that self-sampling (generated by drawing with replacement) is not biasing the baseline downwards. We generate alternative baselines including ones

[10]The correlation between the number of plants and firm employment is 0.70.

Figure 2: Dispersion and Firm Size, by Number of Establishments



(a) All firms

(b) Manufacturing

Notes: Blue circles plot the coefficients from regression of difference between log firm dispersion and log dispersion of synthetic firm against log firm employment controlling for industry-by-year-by-age group fixed effects for each firm size category. All standard errors are clustered by firm modal 4-digit industry.

drawn without replacement, from more aggregated industries, and where establishments in the baseline cannot be drawn from the original firm, and show that these alternative procedures are barely distinguishable from those reported in Figures 1a and 1b.

So far, we have shown that larger firms, whether measured by the number of plants or employment, are more dispersed both in absolute terms and relative to the baseline defined by industry location patterns. We investigate whether this pattern continues to hold *within* the size classes defined by the number of plants, by running regressions of the form

$$y_{it} - y^s_{it} = \beta_z \ln emp_{it} + \gamma_z \mathbf{x}_{it} + \epsilon_{it}, \tag{2}$$

where y_{it} is actual firm dispersion, y^s_{it} is the predicted baseline firm dispersion, emp_{it} is firm employment and $\mathbf{x}_{it}$ is a group of industry-year-age group dummies. We estimate the models separately for each size class z, and plot the resulting β_z coefficients in Figures 2a and 2b. Holding fixed the number of establishments in a firm, firms with higher employment choose to disperse their establishments more than firms with lower employment in the same industry and age group. This increased dispersion for larger firms is sizable, averaging between 0.2 and 0.4 log points, and holds broadly across size classes.

Taken together, these results imply that size and dispersion are intimately related along all dimensions of size.

3.2 Spatial Growth Patterns of Firms

We next investigate the spatial pattern in which firms add new establishments as they expand. When a firm decides to add a plant, how does the location of its existing plants influence the placement of the new plant? Are firms with n establishments more agglomerated than what we would find if a firm with $n-1$ establishments randomly chose a location for its n^{th} establishment? Figures 1a and 1b already show that, on average, larger firms are more dispersed, and one might think that the spatial pattern of firm growth is already revealed by differencing the coefficients across size classes. However, as we explain below, there are three issues with this approach.

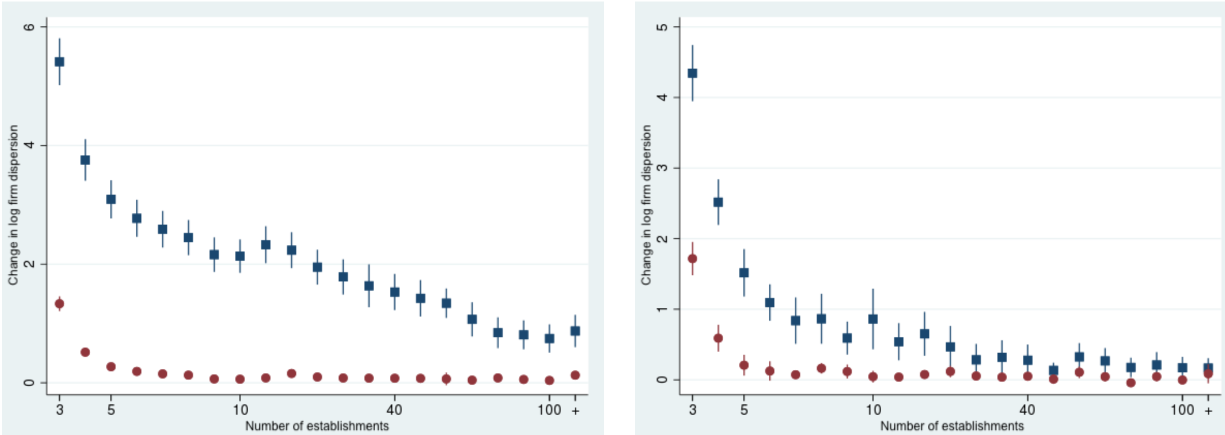
First, the set of firms that expand by a size class are not randomly selected from the firms in the previous size class. Expanding firms may be systematically different in their initial location choices from the firms that do not expand. Second, since plants tend to be quite long-lived, the cross-sectional results in Figures 1a and 1b reflect location choices made in both the recent and more distant past. Given the sweeping technological and economic changes of the past 50+ years, it would hardly be surprising to find that the location choices made by firms today are systematically different than those of the past. Third, the synthetic baseline used in the cross-section is not useful for studying spatial growth patterns. We would like to compare the change in dispersion associated with the actual location choice made by an expanding firm to the change associated with a counterfactual choice, *taking its existing plant locations as given*. There is no way to do that with the information in Figures 1a and 1b.[11]

To address these issues, we proceed as follows. For each Census year[12] and initial size

[11] The synthetic baselines in Figures 1a and 1b show that, in a world where all plant location choices are random, firm expansion is not associated with any economically significant change in dispersion.

[12] We use Census waves because establishment birth years between Census waves are imputed for multi-unit firms in the LBD. The total number of establishments born between Census waves is not. We thank

Figure 3: Time Series Firm Dispersion



(a) All firms (b) Manufacturing

Notes: Red circles plot growth in log average establishment distance to firm centroid for firms observed moving up one firm size category between 5-year Census waves from 1992-2012. Blue squares plot the measure for synthetic expansions. All standard errors are clustered by firm modal 4-digit industry.

class, we find every firm that moved up by exactly one size class relative to the previous Census year and compute the change in its log dispersion.[13] We then average across firms and years, by size class, to get the average change in log dispersion for this subset of expanding firms, for each size class.

The circles in Figures 3a and 3b plot the results for all firms and manufacturing, respectively. In both groups, small and medium-sized firms tend to increase their average dispersion when they expand, at a rate decreasing in firm size. For the smallest firms, these increases are quantitatively large: the third plant tends to increase firm dispersion by about 1.5 log points in both groups. The growth in dispersion tails off rapidly as firms continue adding plants, settling on between 0.2 and 0 log points after the 7$^{\text{th}}$ plant or so. These results are qualitatively consistent with the patterns found in the cross-section, although for most size classes growing firms tend to increase dispersion by somewhat less than what differencing the cross-sectional estimates would imply.

Martha Stinson for alerting us to this feature of the data.

[13] We observe some firms that simultaneously open and close multiple establishments. Both choice of entry and exit can have a meaningful impact on the firm's footprint. We use the net change the number of establishments to classify firms as having moved up one size class. Appendix B.3 discusses the fuller matrix of transitions including negative transitions.

How do these expansion patterns compare to those that would be observed if firms did not consider the location of their existing plants when expanding? We construct a new synthetic baseline by starting with the same underlying sample in Figures 3a and 3b, the set F_{zt} of firms that have z plants in Census year t and had $z-1$ plants in Census year $t-1$, where z refers to the size class, for each year and size class. For each firm f_{it} in this set, we construct a synthetic firm f_{it} by keeping the continuing plants of f_{it}, dropping the new plants and replacing each one with a plant drawn randomly (with replacement) from the set of plants P_{kt} in the same 6-digit NAICS code that were active in that year.[14] We then compute the average increase in log dispersion across years, within size class, for the synthetic baseline in the same way as we did for the data.

The results of the synthetic baseline are plotted using squares in Figures 3a and 3b. While small and medium firms in both groups disperse the most as they expand, the increase in their dispersion as predicted by the random location model is substantially greater than what is actually observed in the data. This implies that even as small and medium-sized firms tend to grow outward from their centers, they do so in stages, growing in space at a pace slower than predicted by the random location model. The gap between the random location model and the data decreases with firm size, much more rapidly for manufacturing firms, largely due to decline in the baseline dispersion growth rate. This in turn reflects the fact that larger firms are initially much closer to the baseline dispersion.

These results are robust to alternative specifications, measurements and samples, all reported and described in more detail in Appendix B.3. As in the cross-sectional analysis, we show that using distance to headquarters and including industry and firm controls yield qualitatively similar results. We also report results using alternative constructions of the synthetic baseline including alternative treatments of mergers, restricting the baseline to be selected from newer establishments only, and additional specifications. Finally, we

[14]From the set of expanding firms, some subset had also closed establishments between years. These choices also impact the footprint of the firm, so to simulate the new footprint of the firm with randomized location choices, we also randomize the establishments which closed from the firms' base-year locations.

examined firms that moved more than one size class between Census waves. There are fewer of these firms and describing the full transition matrix is prohibitive because of small sample sizes. At the same time, the heterogeneity in dispersion between size classes suggests pooling results across different size class changes could be misleading. What we can clear for release is that, for the sample that includes all sectors, the confidence intervals of the data and synthetic baseline did not overlap for any transition between size classes with fewer than 25 establishments. The same statement applies to the manufacturing subsample for transitions across up to three size classes with fewer than 10 establishments. These results are consistent with those reported in Figures 3a and 3b.

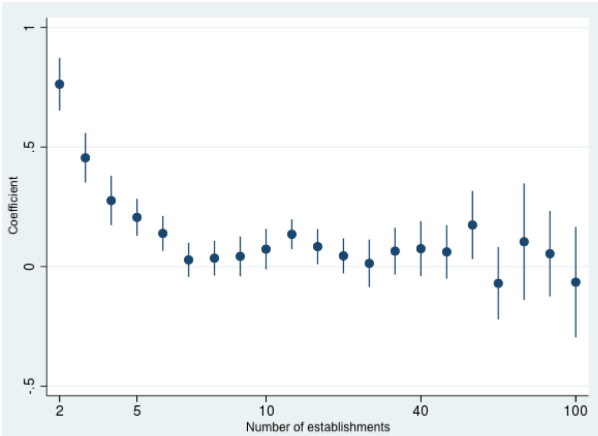
A comparison of the cross-sectional and time series results in Figures 1 and 2 reveals that the average growth rate in dispersion in firms that we observe adding a plant is generally less than the corresponding difference in dispersion between size classes. For example, the difference in average log dispersion between 2 and 3 plant manufacturing firms in the cross-section is greater than 2 log points, while the average firm with 2 plants that we observe adding a third plant increases its dispersion by substantially less than 2 log points. The discrepancy is systematic, although not always quantitatively large, and tends to diminish for larger firms. Two potential explanations suggest themselves: the non-random selection of firms into plant growth and the influence of historical plant location decisions that were made under different constraints than firms face today.

We first examine the role of selection by running the following regression separately for each size class z:

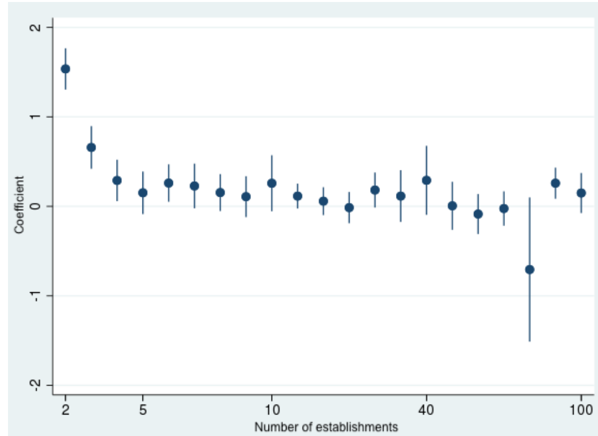
$$y_{it} - y_{it}^{s} = \beta_z \mathbf{1}_{z,t+1} + \gamma_z \mathbf{x}_{it} + \epsilon_{it}, \tag{3}$$

where y_{it} is actual firm dispersion, y_{it}^{c} is the predicted baseline firm dispersion, $\mathbf{1}_{i,t+1}$ is an indicator for whether or not that firm added any number of plants (on net) in the succeeding 5 year period and $\mathbf{x}_{it}$ is a group of industry-year-age group dummies. A positive β_z indicates that firms that grow are initially more dispersed relative to the

Figure 4: Selection into Growth



(a) All firms



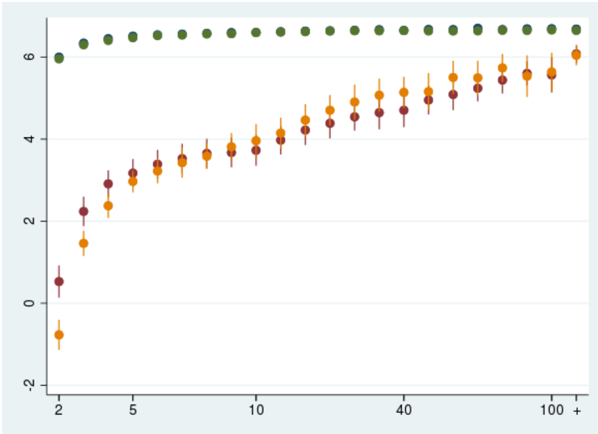
(b) Manufacturing

Notes: Blue circles plot coefficient results of difference between log observed and syntehtic dispersions on pre-expansion employment, controlling for industry-year-age group fixed effects for each size category. All standard errors are clustered by firm modal 4-digit industry.

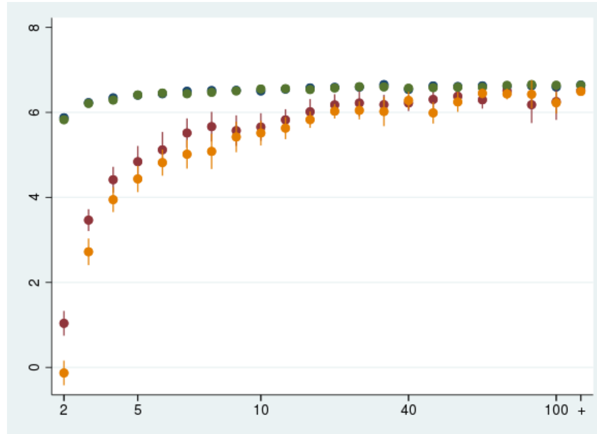
random benchmark than comparable firms in the same size class that do not grow. The resulting β_z coefficients are plotted in Figures 4a and 4b. Firms that grow are generally more dispersed initially than stagnating firms of the same size, with the difference being especially pronounced for smaller firms: for example, 2-plant manufacturing firms that grow have an average initial distance between their first 2 plants that is about 1.5 log points greater than the average distance between the plants of comparable stagnating manufacturing firms. This difference in initial dispersion between growing and stagnating firms diminishes or disappears for larger firms. Thus the pattern of positive selection of growing firms on initial dispersion can help explain the discrepancy between the cross-section and the time series: growing firms are already much more dispersed than the average firm in their size class. This finding can also be interpreted as further support for the relationship between size and dispersion: firms that will be larger *in the future* are already more dispersed *today*.

We next examine how plant location patterns have changed over time. Figures 5a and 5b plot the cross-section of firm dispersion by size class (analogous to Figures 1a and 1b) for 1976 (the orange triangles) and 2012 (the red circles). While some plants that were active in

Figure 5: Firm Dispersion in 1976 and 2012



(a) All firms



(b) Manufacturing

Notes: Red circles plot the average log mean establishment distance from firm centroid for each firm size category for firms in the 2012 Economic Census. Orange triangles plot the corresponding measure for firms in the 1976 Economic Census. Blue squares and green diamonds plot the same measures of synthetic firms in 2012 and 1976 respectively. All standard errors are clustered by firm modal 4-digit industry.

1976 are still active in 2012, major changes in location patterns should be at least somewhat reflected in differences in cross-sectional log dispersion over this time period. There is modest evidence for increases in the geographic dispersion of firms over time, especially for the smallest size categories and for manufacturing firms. However, the cross-sectional patterns are qualitatively similar in the two time periods. Acosta and Lyngemark (2021) find firms have increased in size over this same period. These findings are consistent with the pattern in Figure 5 when considered in tandem with our findings on selection: while the general pattern that larger firms are less clustered has stayed consistent over time, larger firms select into growth, and the overall increase in firm geographic size has come at the extensive margin as firms move up size categories.

4 Theoretical Framework

4.1 Discussion

Our findings so far can be summarized by two broad conclusions. First, firms have a strong tendency to cluster their establishments in space, over and above what would be expected based on industry location patterns alone. This tendency can be observed in both the cross-section and the time series, in different sectors to varying degrees, and in recent data as well as in the past. Second, the tendency to cluster is very heterogeneous in firm size, however measured: large firms are much less spatially clustered than small firms. Again, we observe this tendency broadly across sectors and time periods, in the cross-section and the time series, and in the extensive margin of plants and the intensive margin of employment. We even observe that firms that will grow in the near future are already more dispersed than observationally similar firms that will not.

How can we explain these patterns? The classic frameworks of location choice, many of which are developed in the literature on multinational firms, have firms trading off the benefits of geographic dispersion against the costs. The benefits of dispersion are most often conceptualized as proximity to consumers so as to minimize trade costs, although other benefits such as access to geographically dispersed factor supplies or technologies are also relevant (Antràs and Yeaple, 2014; Antras, Fort, and Tintelnot, 2017; Antràs and De Gortari, 2017). The costs of dispersion are generally thought to be supply side frictions in moving goods, people or ideas across space (via trade, travel and communication costs) that impede coordination between far-flung production units (Keller and Yeaple, 2013; Giroud, 2013; Kalnins and Lafontaine, 2013; Eichholtz, Holtermans, and Yönder, 2015; Alcacer and Delgado, 2016; Charnoz, Lelarge, and Trevien, 2018; Atalay et al., 2019). Our finding of different location patterns for large and small firms suggest that either the benefits of dispersion are larger or the costs of dispersion are smaller for large firms relative to small ones.

A complete empirical and quantitative accounting for the sources of the differences in observed location patterns across firms is beyond the scope of this paper; while such an exercise would be interesting and valuable there are a number of practical challenges.[15] In-

[15]Estimating and quantifying the various sources of differences requires estimating a quantitative model of firm location choice with a large number of alternatives. This is generally an intractable problem, although methods have been developed for special cases and there is ongoing work to extend their scope (Jia, 2008; Tintelnot, 2017; Arkolakis, Eckert, and Shi, 2021; Antras, Fort, and Tintelnot, 2017). In addition, one would

stead, we focus on a particular hypothesis: large firms face a lower "productivity penalty" than small firms when choosing more distant plant locations. We develop a simple model of plant location in which firms are able to produce additional varieties by opening plants in new locations, at the cost of paying a common fixed cost. For each potential location, they receive a variable productivity shifter than is decreasing with the plant's distance to the firm's headquarters. We show that when the productivity penalty is common across all firms, the model predicts that more productive firms will have more plants and that firms with more plants will be more dispersed, consistent with our findings above. However, with a common productivity-penalty, *measured* plant-level productivity or employment at more productive firms will be *more* sensitive to distance from the firm's headquarters, due to the endogenous selection of plant locations. The model can generate the opposite result, that large firms have smaller (in absolute value) measured elasticity of plant productivity with respect to distance, if large firms pay a smaller productivity penalty than small firms. We then test this prediction by implementing a version of the model-implied regression on our data in the subsequent section.

Appendix C discusses several extensions: allowing wages and productivity shifters to be location specific, endogenizing the headquarters location decision, and introducing some aspects of vertical differentiation into the model.

4.2 Setup

There are a finite (but large) number of discrete locations I. Each location i is *ex ante* identical. Workers are mobile and choose a location i in which to supply labor inelastically and consume. They have Cobb-Douglas preferences over goods and non-traded housing,

like detailed data on the inner workings of the firm over a substantial period of time (to capture long-run relocation effects) and quasi-experimental variation in firm productivity and coordination frictions, neither of which are available in our data.

with α being the traded goods expenditure share

$$U(i) = C(i)^{\alpha} H(i)^{1-\alpha}. \tag{4}$$

Workers have CES preferences over varieties of traded goods which are produced by firms at their (potentially multiple) plants such that

$$C_i = \left[\int_{\omega \in \Omega} q_i(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma}{\sigma-1}}, \quad \text{and} \quad q_i(\omega) = \left[\sum_{n=1}^{N(\omega)} q_{i,n}(\omega)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}}, \tag{5}$$

where $N(\omega)$ is the set of products produced by firm ω, and $q_{i,n}(\omega)$ is the quantity of firm ω's variety n sold at location i.

In order to focus our exposition on the within-firm location decision, we make the simplifying assumptions that goods are traded costlessly and housing is supplied at a constant marginal cost wP_H at all locations. This ensures the price index at each location, $P_i = \mathbf{P}$, and the equilibrium wage does not vary by location, $w_i = w$. We present the general equilibrium as well as some extensions to this framework in Appendix C.

At each location, an unbounded mass of *ex ante* identical potential entrants must pay a fixed cost wF_e to enter and build a headquarters. If a firm enters, it draws a core firm productivity $z(\omega)$ and a structural elasticity of productivity with respect to distance from headquarters $\gamma(\omega)$ from a joint distribution, as well as a vector of location-specific i.i.d. productivity multipliers $\epsilon(\omega)$ drawn from distribution with a finite expected value, with each entry corresponding to productivity in every location. It can then produce a single variety at its headquarters location h at constant marginal cost $w/z(\omega)$. For each non-headquarters location $n \neq h$, the firm has the option of paying an additional fixed cost wf_e, denominated in labor and paid at the headquarters, in order to produce a new variety at n at constant marginal cost

$$mc_n(\omega) = \frac{w}{z(\omega)\epsilon_n(\omega)} \cdot d_{h,n}^{\gamma(\omega)}. \tag{6}$$

Here $d_{h,n}$ is the distance between location n and the headquarters h.[16]

4.3 Prices, Profits, and Location Ranking

The problem of a firm that has paid the entry cost is to choose the optimal number and locations of plants and production levels to maximize profits. Since the headquarters entry cost is sunk and enables the firm to produce at the headquarters with no additional fixed cost, CES demand guarantees that the firm will always produce a positive amount at the headquarters location. Firms can rank all potential plants in terms of their potential profitability. This allows us to break the firm's problem up into two steps. First, for a firm with $N(\omega)$ plants, we compute the operating profits from producing at the headquarters and its $N(\omega) - 1$ additional plants. Second, we find the value for $N(\omega)$ for which the additional operating profit from opening an additional plant is insufficient to cover the additional fixed cost, which identifies the optimal number of plants.

Let location 1 denote the headquarters location of a firm ω. If the firm additionally chooses to operate the $N(\omega) - 1$ highest productivity non-headquarters plants, it solves

$$\max_{q_1(\omega), q_2(\omega), \ldots, q_{N(\omega)}} \sum_{n=1}^{N(\omega)} p_n(\omega) q_n(\omega) - \sum_{n=1}^{N(\omega)} mc_n(\omega) q_n(\omega). \tag{7}$$

The demand curve for each variety is given by

$$q_n(\omega) = \left(\frac{p_n(\omega)}{\mathbf{P}}\right)^{-\sigma} \cdot \frac{wL}{\mathbf{P}}, \tag{8}$$

where $\mathbf{P} = \left(\int_\Omega \sum_{i=1}^{N(\omega)} p_i(\omega)^{1-\sigma} \, d\omega\right)^{\frac{1}{1-\sigma}}$. Given the standard constant markup rule, operating

[16]We focus on distance to headquarters rather than distances between pairs of non-headquarters plants, because it is in line with most of the literature and also much more tractable. When productivity at each plant depends on the spatial distribution of every other plant in the firm, the location decision becomes a difficult combinatorial problem.

profits (gross of all fixed costs) are given by

$$\pi(\omega) \propto \sum_{n=1}^{N} mc_n(\omega)^{1-\sigma}. \tag{9}$$

Profits are increasing in $z(\omega)$ for any given number of plants $N(\omega)$, and also increasing in the number of plants. It is also apparent that the firm can indeed rank each potential plant in order of its potential contribution to operating profits, and that this ranking is inversely proportional to the marginal cost at n. The change in the variable profit from adding an additional plant in location $N(\omega) + 1$ is $mc_{N(\omega)+1}^{1-\sigma}$. This is falling in $N(\omega)$ due to the fact that the plant-level productivity is falling in the rank $R_n(\omega)$ of that location in terms of contribution to profits. Firm ω will stop expanding when $\Delta_{N+1}\pi(\omega) \leq wf_e \leq \Delta_N \pi(\omega)$.

4.4 Main Results

More productive firms have more plants Firms enter at n if variable profit is higher than the fixed cost of entry. It follows that firm ω enters location n if an only if

$$\epsilon_n(\omega) \geq \xi \frac{d_{h,n}^{\gamma(\omega)}}{z(\omega)}, \tag{10}$$

where ξ is a constant (from the firm's perspective) defined in Appendix C. Letting $\bar{\epsilon}_n(\omega)$ be the cutoff productivity, the probability of a firm with core productivity z and distance elasticity $\gamma(\omega)$ entering location n is $1 - G(\bar{\epsilon}_n(\omega))$ where G is the cumulative distribution function of ϵ. It is immediate that plants are less likely to be placed at greater distances when $\gamma(\omega)$ is larger, are equally likely to be placed at any distance when $\gamma(\omega) = 0$, and that more productive firms have a higher expected number of plants if $Corr(\gamma(\omega), z(\omega)) \leq 0$.

Larger firms are more dispersed Assume that all firms have the same distance elasticity γ. The expected average distance to headquarters of a firm with N plants is

$$ED_N = E\left[\frac{\sum_{n=1}^{I} \mathbf{1}_n(\omega) \cdot d_{h,n}}{N} | N\right], \tag{11}$$

where $\mathbf{1}_n(\omega)$ is an indicator that equals 1 if location n is chosen and zero otherwise. Since firms always enter locations in order of decreasing productivity (or increasing marginal cost), we can write this as

$$ED_N = \frac{\sum_{n=1}^{I} Pr(R_n(\omega)) \leq N) \cdot d_{h,n}}{N}.$$

This is due to the key property that the distribution of the rank $R_n(\omega)$ does not depend on the number of plants when all firms have the same distance elasticity. Furthermore, for firms headquartered at h,

$$ED_{N+1} - ED_N = \frac{\sum_{n=1}^{I} \left[\eta_n \cdot d_{h,n}\right]}{N+1}. \tag{12}$$

where $\eta_n = Pr(R_n(\omega)) \leq N+1) - \frac{N+1}{N} Pr(R_n(\omega)) \leq N)$. η_n is monotonically increasing in distance. Furthermore, because $\sum_{n+1} \eta_n$ (unweighted by $d_{h,n}$) is zero, we have that $ED_{N+1} - ED_N > 0$ by Chebyshev's sum inequality.[17] Therefore, firms with more plants have a higher expected distance to headquarters. Since expected distance to headquarquarters is decreasing in γ, this would remain true if firms with higher N have systematically lower distance elasticities, or as long as $Corr(\gamma(\omega), z(\omega)) \leq 0$.

Distance-productivity elasticities and selection Again assume that all firms have the same distance elasticity γ. The observed log productivity of firm ω's plant in location n,

[17]Note that this is distinct from Chebyshev's inequality.

$prod_n(\omega) \equiv \frac{w}{mc_n(\omega)}$ is

$$\log(prod_n(\omega)) = \log(z(\omega)) - \gamma(\omega)\log d_{h,n} + \epsilon_n(\omega). \tag{13}$$

We define the average elasticity of realized plant productivity with respect to distance for firms with N plants as the coefficient on $d_{h,n}$ in the regression

$$\log(prod_{n,\omega}) = \alpha_\omega + \beta_N \log d_{h,n} + \widetilde{\epsilon}_n(\omega) \tag{14}$$

where α_ω is a firm fixed effect, $\widetilde{\epsilon_n}(\omega)$ has mean zero within each firm and we have a random sample of firms all with exactly N plants. Note that $\beta_N \geq -\gamma$ since $E[\epsilon_n(\omega)(d_{h,n})] \geq 0$. This is due to selection: since firms are systematically less likely to locate in more distant locations, we expect plants at further locations to have drawn higher idiosyncratic productivity $\epsilon_n(\omega)$. In Appendix C, we show that the bias of β_N away from $-\gamma$ (towards zero) is smaller for larger, more productive firms and hence $\beta_N \geq \beta_{N+1}, \forall N$. Furthermore, this extends easily to other plant outcomes such as employment or total sales, since these vary proportionally with productivity.[18]. We would need to have $Cov(\gamma(\omega), z(\omega)) < 0$ in order for the model to generate a smaller (in absolute value) realized distance elasticity for larger firms.

5 The Distance-Productivity Relationship Within the Firm

We now turn to examining the distance-productivity relationship within the firm and how it changes with firm size, which the previous section identified as the key set of moments that differentiates the common-productivity-penalty model from the model in which large firms face smaller productivity penalties. We identify the nearest regional headquarters for each establishment in our sample and test for differences in the elasticity of plant-level outcomes with respect to distance to headquarters across firms with different

[18]Note that sales per worker equal a constant markup over wages and are hence invariant to distance to headquarters

numbers of establishments.[19] We also examine elasticities with respect to distance to firm centroid, which is available for a much larger number of firms and may also reflect frictions that operate between non-headquarters plants, such as transportation costs along supply chains. Our chief plant-level outcome measure is employment, which maps one-to-one with physical productivity in our model and is more widely available and better measured than physical productivity. We examine additional outcomes and controls after presenting our main results.

Our ideal specification compares the relative productivity of plants across two firms at exactly the same locations producing exactly the same products. We rely on a set of fixed effects to replicate this ideal as closely as possible. First, we control for firm by year by the plant's 6-digit NAICS industry dummies, which means that we compare—within firms—plants producing exactly the same modal product to the 6-digit level. Second, we add county-industry-year controls to account for any industry-specific demand or productivity differences in space that could be correlated with distance to headquarters as well as plant outcomes. These geographic controls may also reduce the scope for the impact of sectoral specialization on tasks within the firm-industry (Duranton and Puga, 2005). Third, we add industry-year-age group fixed effects to account for industry-specific differences in plant vintage that could potentially be correlated with distances and outcomes. Our specifications take the form

$$y_{jt} = \alpha_{ikt} + \beta \ln dist_{ijt} + \theta \cdot \mathbf{1}_{10+} \ln dist_{ijt} + \gamma \mathbf{x}_{jt} + \epsilon_{jt}, \tag{15}$$

where y_{jt} is an establishment j characteristic, $dist_{ijt}$ is a measure of plant j's geographic position within the firm i, $\mathbf{1}_{10+}$ is a dummy that equals 1 if firm i has 10 or more plants,[20] α_{ikt} is a firm-year-(plant) industry dummy and x_{jt} is a set of year-(plant) county-(plant)

[19]The distinction between regional and global headquarters to be empirical relevant for our elasticity estimation. The decision of larger firms to generate multiple regional headquarters, which is beyond the scope of our model, may be a mechanism that mitigates overall distance costs, and so focusing on distance to regional headquarters should yield a more conservative estimate of the overall difference in distance costs.

[20]We report some results for more detailed size class breakdowns in Appendix B.3.

industry - (plant) age group dummies interacted with $\mathbf{1}_{10+}$.[21]

Table 1: Employment and Distance Within the Firm

Variable	ln Employment	ln Employment	ln Employment
		(A) All Firms	
ln miles HQ	-0.039		-0.040
	(0.008)		(0.008)
1_{10+} x ln dist HQ	0.024		0.027
	(0.007)		(0.006)
ln miles to centroid		-0.119	0.024
		(0.021)	(0.011)
1_{10+} x ln miles to cent.		0.083	-0.021
		(0.015)	(0.014)
N	2,105,000	6,415,000	2,105,000
		(B) Manufacturing	
ln miles HQ	-0.125		-0.122
	(0.037)		(0.063)
1_{10+} x ln miles HQ	0.089		0.088
	(0.042)		(0.061)
ln miles to centroid		-0.163	-0.025
		(0.038)	(0.161)
1_{10+} x ln miles to centr.		0.132	-0.017
		(0.044)	(0.158)
N	80,000	270,000	80,000

Notes: The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. All regressions include 6-digit industry-firm-year fixed effects as well as 6-digit industry-year-county and 6-digit industry-year-age group fixed effects. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.

Table 1 reports the results of estimating equation (15) using establishment employment as the outcome variable. Panel (A) uses plants in all industry and the first column focuses on the subset of plants with information on distance to closest headquarters. We find plant-level employment declines sharply with distance from the nearest headquarters with a sharply estimated elasticity of around -0.039. However, for larger firms the penalty

[21] Results are robust to using states instead of counties as the geographic unit.

is much lower, with an elasticity estimate of about -0.015. Similar patterns are observed in the manufacturing sub-sample in panel (B), although both point estimates and standard errors are quantitatively larger and the sample is much smaller.

The second column uses plant distance to centroid, enabling us to use our full sample. In both panels the distance elasticity is significantly attenuated for large firms, as in the distance to headquarters specification. The third column includes both variables. Interestingly, the coefficients on distance to headquarters – both for large and small firms – appear largely the same as in the first column, while the coefficients on distance to centroid are now quantitatively much smaller and more noisily estimated. We interpret these results as modest evidence that distance-to-headquarters captures much of the relevant variation in the data, providing some support our theoretical focus on this variable.

The findings of Table 1 are supportive of the hypotheses that small firms face significant productivity penalties from locating plants further from headquarters, and that large firms face much smaller penalties for doing so. While the estimates in Table 1 do not utilize quasi-experimental evidence and hence cannot be taken to identify the true productivity penalties, our model implies that the true productivity penalties should be even larger than found in Table 1 due to non-random selection of plant locations. The model with smaller productivity penalties for large firms can thus account for all the empirical patterns documented in this paper.

In the remainder of this section we briefly describe some additional exercises that examine the robustness of our results. The main concern is that while employment maps one-to-one with physical productivity in our model and is often used as a measure of productivity based on the theoretical link between the two variables (e.g. (Hopenhayn, 1992; Melitz, 2003)), it may no longer be a sufficient statistic for productivity in more complex environments such as those with multiple factors of production, increasing marginal costs or departures from CES demand. A related concern is that employment comparisons may be misleading when production functions differ across plants in the firm, perhaps due to vertical and/or functional specialization.

We address the first set of concerns, to the extent that we are able, by considering sales per worker as an additional outcome measure. Sales per worker is often used as a productivity measure in non-manufacturing industries, e.g. (Haltiwanger, Lane, and Spletzer, 2007), although the link between sales per worker and welfare-relevant notions of productivity is complex. For example, in the model of Section 4 sales per worker are equalized across plants regardless of productivity differences. While sales per worker alone is generally not informative regarding productivity, we demonstrate in Appendix C.2 that jointly considering employment and sales per worker in a more general environment with multiple factors of production and increasing marginal costs yield an increase in information relative to considering just employment in isolation. Plants exhibiting both higher employment and weakly higher sales per worker must be more profitable for the firm; relative profitability, in turn, is a measure of relative productivity in demand systems (such as CES) that rightly link relative profitability to the relative Marshallian surplus generated by the firms.

Table 2 reports results for the same regressions using sales per worker as the outcome variable. When using the all firm sample in Panel (A), we find little consistent evidence that sales per worker varies systematically with distance to headquarters or distance to centroid. When we examine manufacturing establishments in Panel (B) we find some modest evidence that distance penalties in sales per worker are smaller for larger firms, although standard errors are fairly large and the results are sensitive to the details of the specification (e.g. state vs county -by-industry-year fixed effects). All in all, sales per worker does not appear to be robustly related to the firm's internal geography in any quantitatively significant way. In conjunction with our results on the size-distance gradient, we take these results as further evidence that the productivity of plants further away from the headquarters location is lower, while less so for larger firms.

The second set of concerns is that vertical relationships or functional specialization

Table 2: Sales Per Worker and Distance Within the Firm

Variable	ln Sales/Worker	ln Sales/Worker	ln Sales/Worker
		(A) All Firms	
ln miles HQ	0.005		0.007
	(0.004)		(0.004)
1_{10+} x Ln dist HQ	-0.003		-0.004
	(0.004)		(0.005)
ln miles to centroid		0.001	-0.004
		(0.003)	(0.007)
1_{10+} x ln miles to cent.		0.004	0.008
		(0.007)	(0.007)
N	2,105,000	6,415,000	2,105,000
		(B) Manufacturing	
ln miles HQ	-0.095		-0.082
	(0.026)		(0.037)
1_{10+} x ln miles HQ	0.081		0.067
	(0.023)		(0.038)
ln miles to centroid		0.021	-0.045
		(0.012)	(0.054)
1_{10+} x ln miles to centr.		-0.019	0.071
		(0.008)	(0.070)
N	80,000	270,000	80,000

Notes: The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. All regressions include 6-digit industry-firm-year fixed effects as well as 6-digit industry-year-county and 6-digit industry-year-age group fixed effects. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.

across plants could both alter the employment-productivity relationship across plants and systematically vary with distance to headquarters. For example, employment gradients could be driven by differences in capital intensity or managerial intensity. We provide additional size measures in Appendix B.3, including establishment sales and, for manufactures, production workers, and we also control for and then directly test for evidence of differences in factor intensity or tasks. We don't find strong or consistent evidence that our findings are driven by other correlated geographic patterns, which we attribute to the saturated fixed effects in our specification that ensure a high degree of functional similarity across the plants that are being compared.

6 Conclusion

We have shown that small and medium sized firms tend to cluster their plants in space, while larger firms do so to a much smaller degree. We have also documented that more distant plants are smaller than closer plants within small firms, but this pattern is weaker or absent for plants belonging to larger firms. These findings suggest that spatial dispersion is costly for small firms, and less so for large. However, we caution that we do not rule out alternative explanations, such as differences in the spatial correlation of demand or in internal cannibalization effects (e.g. Tintelnot (2017)), which in combination with sizable trade costs could also be contributing to the observed differences in location patterns across the firm size distribution.

While we treat these cost differences as exogenous characteristic, an interesting question may be whether they are the result of endogenous investment decisions made by the larger firms. Moreover, distance costs could be policy responsive. Figures 5a and 5b shows that the cross-sectional patterns of firm dispersion have changed modestly since 1976, especially for the smallest firms. Along with the evidence from changes in travel costs in Giroud (2013) and Charnoz, Lelarge, and Trevien (2018) this suggests that the costs are at least somewhat mutable. If such costs are both large and responsive to policy, there is scope for policies aimed at reducing these costs to generate significant welfare gains through the spatial reorganization of firms.

Firm clustering impacts industry agglomeration Bartelme and Ziv (2021), and the differential clustering of larger establishments may have consequences for the distribution of economic activity. While beyond the scope of this paper, a natural direction for future research is to estimate a structural model of plant location decisions in general equilibrium. This exercise could be useful for learning both about the mechanisms underlying our findings as well as the aggregate implications for the location of economic activity and welfare.

References

Acosta, Camilo and Ditte Håkonsson Lyngemark. 2021. "The internal spatial organization of firms: Evidence from Denmark." *Journal of Urban Economics* 124:103366.

Alcacer, Juan and Mercedes Delgado. 2016. "Spatial organization of firms and location choices through the value chain." *Management Science* 62 (11):3213–3234.

Alfaro, Laura and Maggie Xiaoyang Chen. 2018. "Transportation cost and the geography of foreign investment." In *Handbook of International Trade and Transportation*. Edward Elgar Publishing.

Antoni, Manfred, Anna Gumpert, and Henrike Steimer. 2022. "Firm organization with multiple establishments." .

Antràs, Pol and Alonso De Gortari. 2017. "On the geography of global value chains." Tech. rep., National Bureau of Economic Research.

Antras, Pol, Teresa C Fort, and Felix Tintelnot. 2017. "The margins of global sourcing: Theory and evidence from us firms." *American Economic Review* 107 (9):2514–64.

Antràs, Pol and Stephen R Yeaple. 2014. "Multinational firms and the structure of international trade." In *Handbook of international economics*, vol. 4. Elsevier, 55–130.

Arkolakis, Costas, Fabian Eckert, and Rowan Shi. 2021. "Combinatorial Discrete Choice." Tech. rep., Yale University.

Atalay, Enghin, Ali Hortaçsu, Mary Jialin Li, and Chad Syverson. 2019. "How wide is the firm border?" *The Quarterly Journal of Economics* 134 (4):1845–1882.

Bahar, Dany. 2020. "The hardships of long distance relationships: time zone proximity and the location of MNC's knowledge-intensive activities." *Journal of International Economics* :103311.

Bartelme, Dominick and Oren Ziv. 2021. "JUE insight: Firms and industry agglomeration." *Journal of Urban Economics* :103372.

Behrens, Kristian and Vera Sharunova. 2015. "Inter-and intra-firm linkages: Evidence from micro-geographic location patterns." *CEPR Discussion Papers* (10921).

Charnoz, Pauline, Claire Lelarge, and Corentin Trevien. 2018. "Communication costs and the internal organisation of multi-plant businesses: Evidence from the impact of the french high-speed rail." *The Economic Journal* 128 (610):949–994.

Duranton, Gilles and Henry G Overman. 2005. "Testing for localization using micro-geographic data." *The Review of Economic Studies* 72 (4):1077–1106.

Duranton, Gilles and Diego Puga. 2005. "From sectoral to functional urban specialisation." *Journal of urban Economics* 57 (2):343–370.

Eichholtz, Piet, Rogier Holtermans, and Erkan Yönder. 2015. "The economic effects of owner distance and local property management in US office markets." *Journal of Economic Geography* 16 (4):781–803.

Ellison, Glenn and Edward L Glaeser. 1997. "Geographic concentration in US manufacturing industries: a dartboard approach." *Journal of political economy* 105 (5):889–927.

Fort, Teresa C and Shawn D Klimek. 2016. "The effects of industry classification changes on US employment composition." Tech. rep., Tuck School at Dartmouth.

Giroud, Xavier. 2013. "Proximity and investment: Evidence from plant-level data." *The Quarterly Journal of Economics* 128 (2):861–915.

Gumpert, Anna. 2018. "The organization of knowledge in multinational firms." *Journal of the European Economic Association* 16 (6):1929–1976.

Haltiwanger, John C, Julia I Lane, and James R Spletzer. 2007. "Wages, productivity, and the dynamic interaction of businesses and workers." *Labour Economics* 14 (3):575–602.

Henderson, J Vernon and Yukako Ono. 2008. "Where do manufacturing firms locate their headquarters?" *Journal of Urban Economics* 63 (2):431–450.

Holmes, Thomas J. 2011. "The Diffusion of Wal-Mart and Economies of Density." *Econometrica* 79 (1):253–302.

Hopenhayn, Hugo A. 1992. "Entry, exit, and firm dynamics in long run equilibrium." *Econometrica: Journal of the Econometric Society* :1127–1150.

Hsieh, Chang-Tai and Esteban Rossi-Hansberg. 2019. "The industrial revolution in services." Tech. rep., National Bureau of Economic Research.

Irarrazabal, Alfonso, Andreas Moxnes, and Luca David Opromolla. 2013. "The margins of multinational production and the role of intrafirm trade." *Journal of Political Economy* 121 (1):74–126.

Jia, Panle. 2008. "What happens when Wal-Mart comes to town: An empirical analysis of the discount retailing industry." *Econometrica* 76 (6):1263–1316.

Kalnins, Arturs and Francine Lafontaine. 2013. "Too far away? The effect of distance to headquarters on business establishment performance." *American Economic Journal: Microeconomics* 5 (3):157–79.

Keller, Wolfgang and Stephen Ross Yeaple. 2013. "The Gravity of Knowledge." *American Economic Review* 103 (4):1414–44.

Melitz, Marc J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *Econometrica* 71 (6):1695–1725.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter. 2020. "Plants in space." Tech. rep., National Bureau of Economic Research.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2018. "Diverging trends in national and local concentration." Tech. rep., National Bureau of Economic Research.

Tintelnot, Felix. 2017. "Global production with export platforms." *The Quarterly Journal of Economics* 132 (1):157–209.

Yeaple, Stephen Ross. 2009. "Firm heterogeneity and the structure of U.S. multinational activity." *Journal of International Economics* 78 (2):206–215.

A Data and Measurement Appendix

A.1 Sample

Figure 1 uses the average distance between establishments in each firm and the firm's centroid as a measure of the firm's geographic dispersion, and compares this measure in the data to that of a synthetic control constructed as described in Section 3. Panel A of Table A1 summarizes these measures for the 650,000 firm-by-year observations in this sample as well as the 49,000 firm-year observations in the sub-sample of manufacturing firms. Panel A also reports, for both all and the sub-sample of manufacturing firms, firm log employment and firm log output per worker.

Our time series analysis uses a subset of the firms in our main sample which can be observed moving up exactly one firm size category between Economic Censuses. This yields 44,000 firm-year observations from the full sample and 4,900 firm-year observations from the manufacturing firm sub-sample, or around 10% of each initial sample. Panel A reports changes in real and synthetic firm data for these sub-samples.

Section 5 examines relationships between establishments' log miles to the firm center and establishment log employment or sales per worker. Summary statistics for the 6,413,000 establishment-year observations in the full sample and 270,000 establishment-year observations for the sub-sample of manufacturing establishments are reported in Panel B.

Some analyses also use distance to regional headquarters. Roughly one quarter of firms have at least one identifiable headquarters based on establishment NAICS codes. Some previous studies have imputed plant headquarters from payroll information (e.g. Giroud (2013)). However, we choose to examine only the subsample of firms with an establishment with the headquarters code (551114) because payroll is linked to outcomes of interest (such as sales per worker). For each plant, we calculate closest of the firms' plants with the 551114 NAICS code and designate that the plant's closest regional headquarters.

Table A1 provides summary statistics for these firms as well.

Table A1: Summary Statistics

(A) Firm-level Variables

Variable	Obs	Mean	St. Dev.
All Firms			
Number of establishments	650,000	9.38	86.21
Log average miles to centroid	650,000	2.43	3.60
Log average miles to centroid, synthetic	650,000	6.51	6.77
Log employment	650,000	2.83	1.09
Change in log average miles to centroid	44,000	0.60	2.35
Change in log average miles to centroid, synthetic	44,000	3.27	3.70
Manufactures			
Number of establishments	49,000	8.51	42.13
Log average miles to centroid	49,000	3.65	4.08
Log average miles to centroid, synthetic	49,000	6.52	3.65
Log employment	49,000	3.98	1.06
Change in log average miles to centroid	4,900	0.90	3.14
Change in log average miles to centroid, synthetic	4,900	2.47	4.32

(B) Establishment-level Variables

Variable	Obs	Mean	St. Dev.
All Establishments			
Log miles to firm centroid	6,415,000	4.50	2.58
Log employment	6,415,000	2.62	1.25
Log sales per worker	6,415,000	4.80	1.13
Log miles to headquarters	2,105,000	5.8	2.13
Log employment, HQ sample	270,000	2.70	1.30
Log sales per worker, HQ sample	270,000	4.91	1.14
Manufactures			
Log miles to firm centroid	270,000	4.47	3.10
Log employment	270,000	4.05	1.43
Log sales per worker	270,000	5.32	0.96
Log miles to headquarters	80,000	4.47	3.10
Log employment, HQ sample	80,000	4.38	1.51
Log sales per worker, HQ sample	80,000	5.63	0.99

A.2 Average Bilateral Distance vs Distance to Centroid

Here we discuss how our measure of within-firm dispersion relates to another measure of dispersion, the log mean of the bilateral plant distances, used in Behrens and Sharunova (2015) and similar to the measure in Duranton and Overman (2005). For a firm f with N constituent plants indexed by i and j, and a distance measured in miles of d_{ij} between the plants, the average bilateral distance is

$$\text{Average Bilateral Distance} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}\mathbf{1}_{i\neq j}2\cdot d_{ij}}{N\cdot(N-1)}.$$

Our measure is:

$$\text{Distance to Centroid} = \frac{\sum_{i=1}^{N} d_{ic}}{N}$$

where d_{ic} is the distance between plant i and firm centroid $c = (c_x, c_y)$, which we approximate as the average latitude and longitude of all the establishments in the firm,

$$c_x = \frac{\sum_{j\in J}\mathbf{x}_j}{J},\ c_y = \frac{\sum_{j\in J}\mathbf{y}_j}{J},$$

where $\mathbf{x}_j$ and $\mathbf{y}_j$ are coordinates of the jth point.

Both measures are multilateral: a single plant's 'closeness' is a product of its relationship to all other plants. In a single dimension, these measures are identical. In two dimensions these measures can differ because the average distance to other establishments' locations is not identical to the distance to the average location of the plants in two dimensions. However, the two measures are highly correlated. To show this, we use the public Zipcode Business Patterns data from 2012 to generate fake firms with different number of establishments and different geographic dispersion. We then computed both measures and obtained a correlation of 0.97 between them.

B Additional Dispersion Results

B.1 Results for Additional Sectors

We divided non-manufacturing firms into four sub-samples using the same modal sales decision rule described in Section 2 of the main text. Business services firms are firms with modal sales generated at establishments coded as belonging in NAICS sector 53. Retail and Wholesale firms are firms with modal sales generated at establishments in NAICS sectors 42, 44, or 45. Finance and insurance is similarly sector 52. Our final sector is the residual category.

Below, we replicate Figures 1 and 3 for these sectors. Due to small sample sizes, we are unable to release the time series graphs for each sector individually. Instead, we pool firms with 51 or more establishments into a single category.

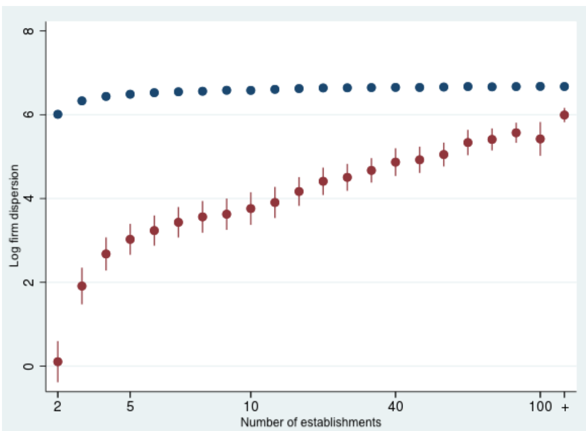
Broadly speaking, clustering patterns both in the cross section and time series for Finance and Insurance, Wholesale and Retail, and Other all follow the patterns we find for All Firms panels in the main text. Clustering patterns for Business Services appear more similar to those in manufacturing in that the gaps between the actual and baseline dispersions tend to fall faster than in other industries. This could be connected to to tradability; cannibalization of sales across plants is theoretically a larger concern for these industries, and thus all else equal we should expect to see greater dispersion for tradable industries.

B.2 Results Using Alternative Measures

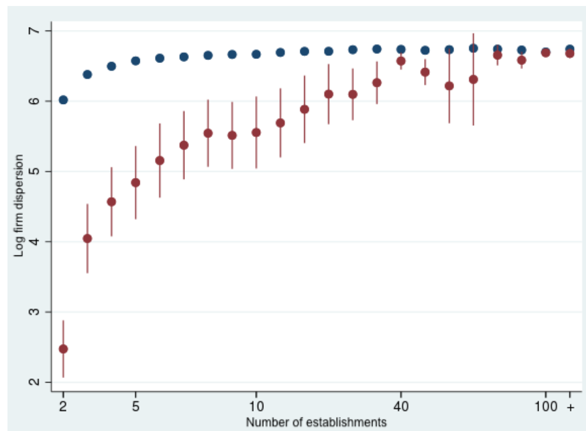
Here we report results using alternative measures of firm distance and alternative firm categorizations.

Distance to headquarters First, we re-plot our main cross sectional and time series results measuring firms by their establishments' average distance to headquarters. Because only

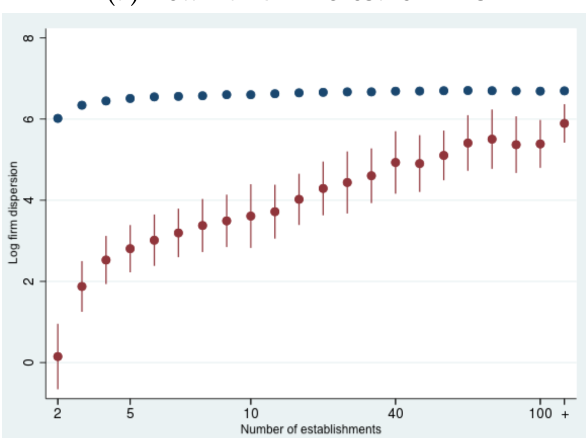
Figure 2a: Cross-Sectional Firm Dispersion, Other Sectors



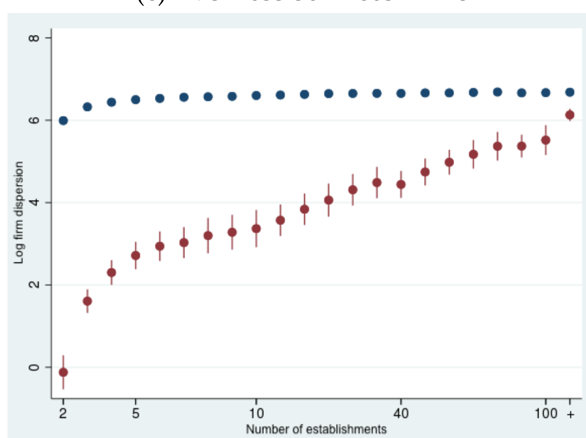
(a) Retail and wholesale firms



(c) Business services firms



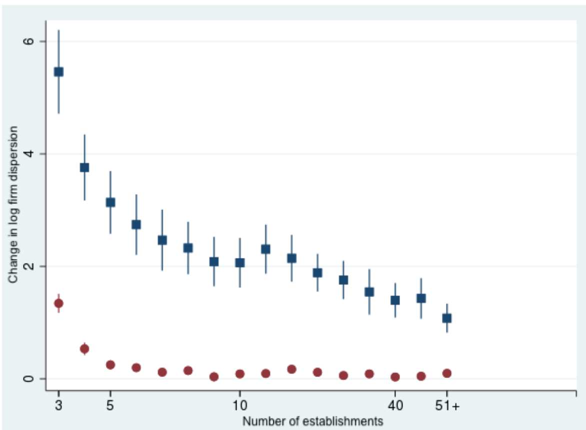
(b) Finance and insurance firms



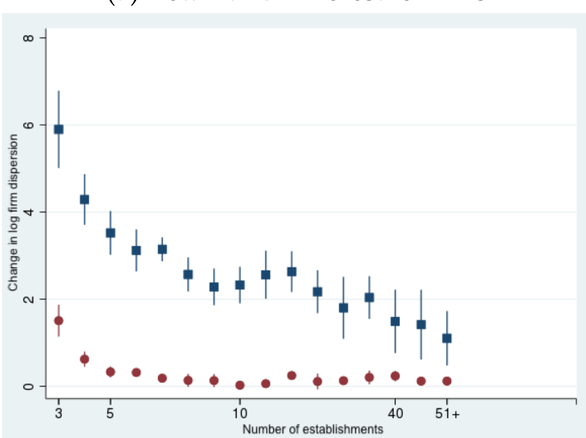
(d) Other firms

Notes: Blue circles plot the coefficients from a regression of the difference between log firm dispersion and log dispersion of synthetic firm on log firm employment, controlling for industry-by-year-by-age group fixed effects for each firm size category. Standard errors are clustered by firm modal 4-digit industry.

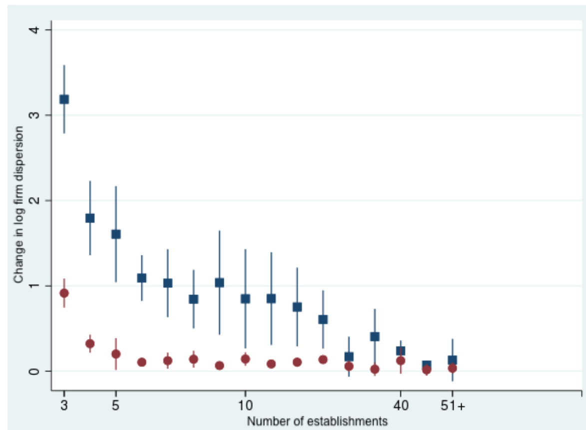
Figure 3a: Time Series Firm Dispersion, Other Sectors



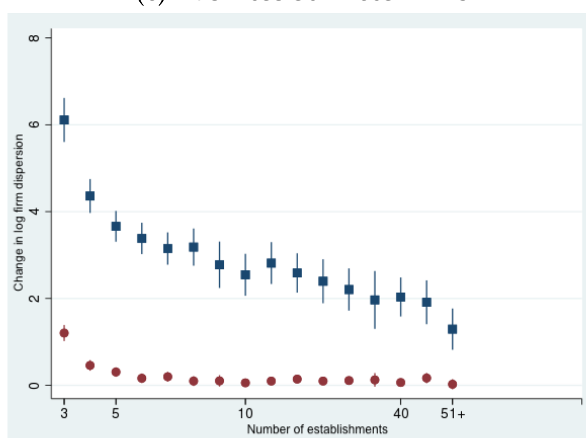
(a) Retail and wholesale firms



(b) Finance and insurance firms



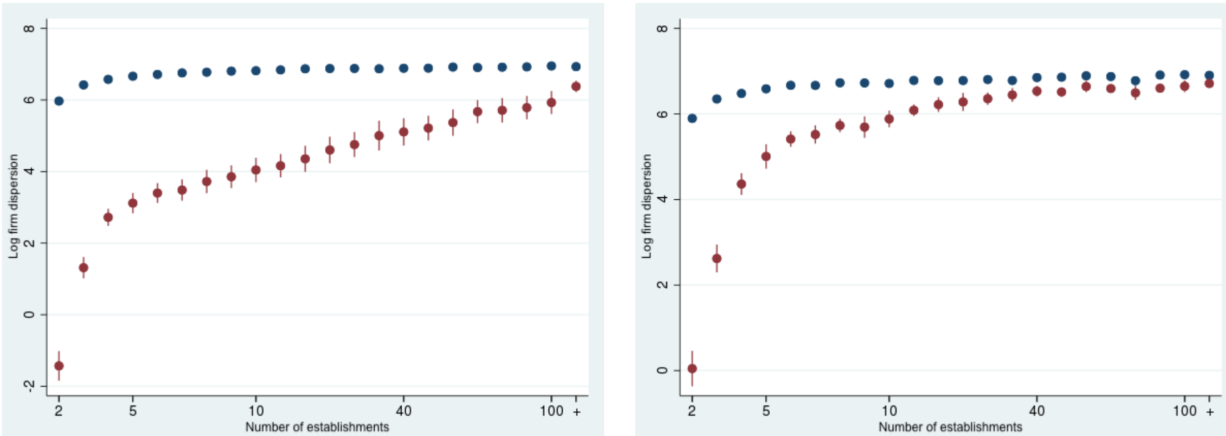
(c) Business services firms



(d) Other firms

Notes: Red circles plot growth in log average establishment distance to firm centroid for firms observed moving up one firm size category between 5-year Census waves from 1992-2012. Blue squares plot the measure for synthetic expansions. Standard errors are clustered by firm modal 4-digit industry.

Figure 4a: Cross-Sectional Firm Dispersion: Distance to Headquarters



(a) All firms (b) Manufacturing

Note: Red circles plot the average log mean establishment distance from firm's headquarters for each firm size category. Blue squares plot the corresponding measure for synthetically constructed firms. Standard errors are clustered by firm modal 4-digit industry.

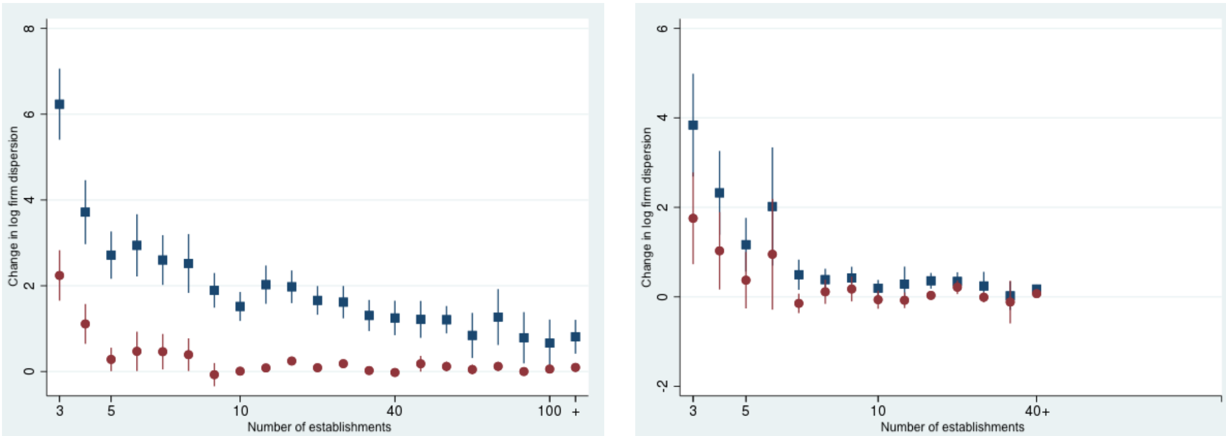
about a quarter of our sample contains headquarters information, some establishment size categories cannot be cleared for publication in the manufacturing sector. As a result, we pool manufacturing firms with 40 or more establishments into a single size category. Our results both in the cross section and time series are similar to our results in Figures 1 and 3 in the main text.

Alternative firm size measures Next, we replicate our main cross sectional results using an alternative firm categorization scheme. We assign firms to employment ventiles and replicate the procedure in Figure 1. While the curvatures of the resulting figures are slightly different, all qualitative results remain. Results are similar for ventile assignments based on sales or sales per worker.

B.3 Alternative Synthetic Controls

Alternative sampling procedures and self-sampling corrections In principle, several choices could affect the construction of our synthetic baseline and impact results. First, we choose to sample with replacement as a conservative measure, this may generate the

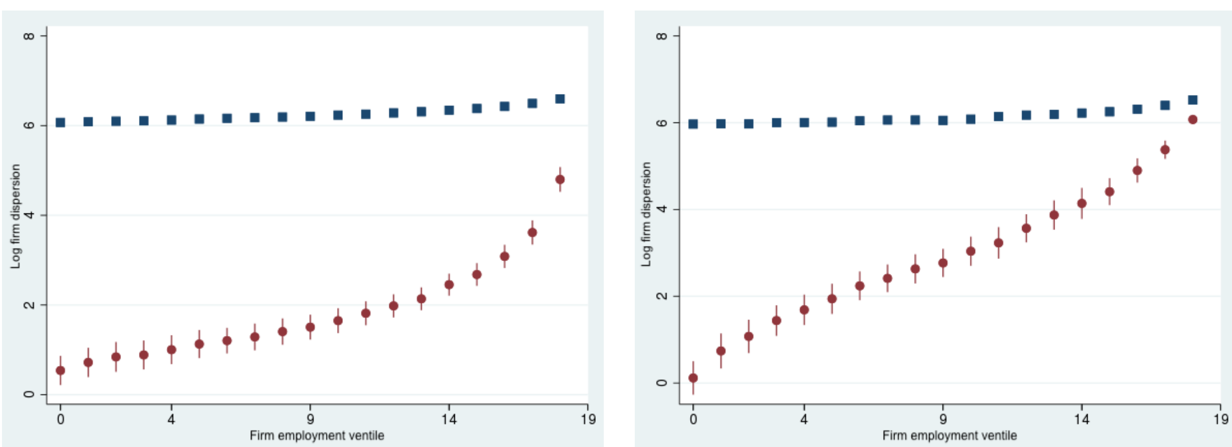
Figure 5a: Time Series, Distance to Headquarters



(a) All firms (b) Manufacturing

Notes: Red circles plot growth in log mean establishment distance from firm's headquarters for firms observed moving up one firm size category between 5-year Census waves from 1992-2012. Blue squares plot the measure for synthetic expansions. Standard errors are clustered by firm modal 4-digit industry.

Figure 6a: Cross-Sectional Firm Dispersion, by Firm Employment Ventile



(a) All firms (b) Manufacturing

Note: Red circles plot the average log mean establishment distance from firm centroid for each firm employment ventile. Blue squares plot the corresponding measure for synthetically constructed firms in each ventile. Standard errors are clustered by firm modal 4-digit industry.

concern that our finding that the largest establishments are not clustered is an artifact of this decision. However, in practice, sampling without replacement does not meaningfully impact the synthetic baseline or any result.

Second, because we wish to sample from a set of meaningfully similar plants, we sample from fine 6-digit industries. However, within some such industries, several large firms dominate. As a result, some plants could be sampled from the same firm. In particular, one concern is that our result that the largest firms are not clustered could be an artifact of drawing synthetic firms by and large from the original firm's own establishments. We address this concern using three alternative baseline selection criteria: (1) We sample from a larger pool of 4-digit NAICS codes, where a single firm is less likely to dominate (2) We resample to ensure no firm's establishments are drawn from their own plants and (3) a combination of (1) and (2). These alternative sampling methods have very small impacts on the baseline and do not appreciably affect any results.

Matching firms on inputs and products A key assumption in Figure 1 is that for each firm, the set of establishments selected for the synthetic baseline are identical to the firm's actual establishments in all but location. This assumption may be violated if there are systematic differences between establishments within the same 6-digit industry. For instance, establishments in the same industry code may sell the same product but differ in their production processes. Or they may have the same modal product but one may be a multi-product firm while the other produces only a single product. If these differences affect the establishment's location, the null will reflect any corresponding geographic differences in production in addition to the underlying cross-firm dispersion within industries.

To address this, we use the product and material files in the Census of Manufactures to match establishments within the same 6-digit industry based on their entire set of products sold and, separately, the entire set of inputs used in production. To perform such a match, we first have to measure the similarity between establishments based on the entire set of their product sales and inputs used. However, there is no obvious way to compare the similarity of two distinct inputs used or products sold. Our approach is to consider each input, and separately, each product, as a distinct dimension. Of I possible inputs and P possible products, each plant i can then be described by an I and P dimensional vector of inputs $\vec{T}_i = [t_1, t_2, ..., t_I]$ and outputs $\vec{S}_i = [t_1, t_2, ..., t_O]$, respectively.

To measure the similarity of the two vectors of inputs, we use the angle of similarity defined as

$$sim_{T_i,T_j} = \frac{\vec{T}_i \cdot \vec{T}_j}{\|\vec{T}_i\| \|\vec{T}_j\|}.$$

We then use sim_{T_i,T_j} as a sampling weight. For each plant i, we draw a matched plant from the same 6-digit industry with N plants such that the probability of drawing any particular plant j as plant i's match is

$$Pr(i,j) = \frac{sim_{T_i,T_j}}{\sum_{k=1}^{N} sim_{T_i,T_k}}.$$

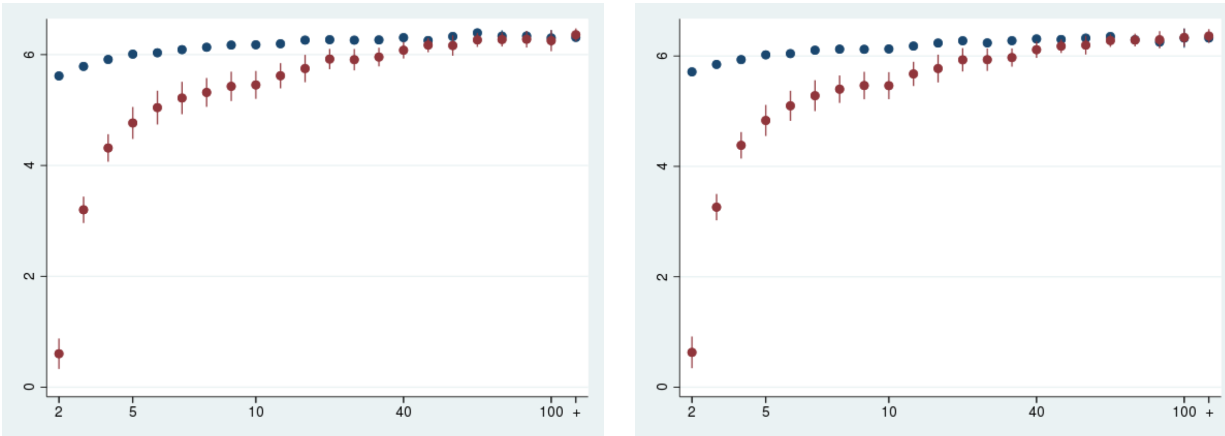
Using the matched plants, we recompute the firm distance to centroid as our new synthetic baseline. We repeat this for the vector of outputs $\vec{S}_i = [t_1, t_2, ..., t_O]$ and report the results below in Appendix Figure 6a.

Synthetic baseline firms constructed using this approach do show somewhat less dispersion across the board. However, the qualitative results of Fact 1 are unchanged: firms on average are clustered and smaller firms are significantly more clustered relative to their baselines.

Results using manufacturing establishments matched on input and output usage by firm size category do not pass disclosure avoidance requirements and cannot be reported.

Time series alternatives To form the synthetic control, we match new establishments with randomly selected establishments in the same 6-digit industry. Changes in the spatial distribution of firms over time, which are documented in Figure 3, may mean we are selecting here from a systematically different set of establishments: those born within our

Figure 7a: Synthetic Controls Using Angle of Product and Input Similarity



(a) Product mix (b) Input mix

Note: Red circles plot the average log mean establishment distance from firm centroid for each firm size category. Blue squares plot the corresponding measure for synthetically constructed firms where existing firms' plants are matched to and replaced with new plants based on plant output and input mix, respectively. Standard errors are clustered by firm modal 4-digit industry.

time frame and those born before it. Ideally, we would address this by selecting establishments for the null that are both of the same age and industry. However, these constraints quickly reduce the full set of establishments from which to choose. An intermediate approach matches new establishments with those in the same industry that are born within a specific window. For instance, we can match new establishments with establishments that are five of fewer years of age. Our results are robust to such restrictions.

Our time series results are also robust to different treatments of growth resulting from merger activity. In our main results, we do not differentiate between newly born establishments and establishments that enter the firm as the result of a merger. Our synthetic baseline randomizes the location of both of these groups of new establishments. Alternatively, we can run our specification randomizing only new births. When computed in this method, the synthetic control is closer to the real data but similarly preserves economically and statistically meaningful growth above what is observed in real data for the same size categories as in our main specification.

B.4 Robustness to industry controls.

Figure 1 compares, within each size bin, the observed dispersion to the synthetically constructed null. By construction, the null accounts for the characteristics of the firm, including industry composition. However, the cross-group pattern of decreasing dispersion in unconditional means may in part reflect compositional differences in the groups' characteristics.

Controlling for industry and other compositional differences ideally involves a saturated fixed-effect model. For example, detailed industry-by-year controls and age fixed effects would eliminate the possibility that such differences between bins accounts for the potential differences in means across bins. We can estimate

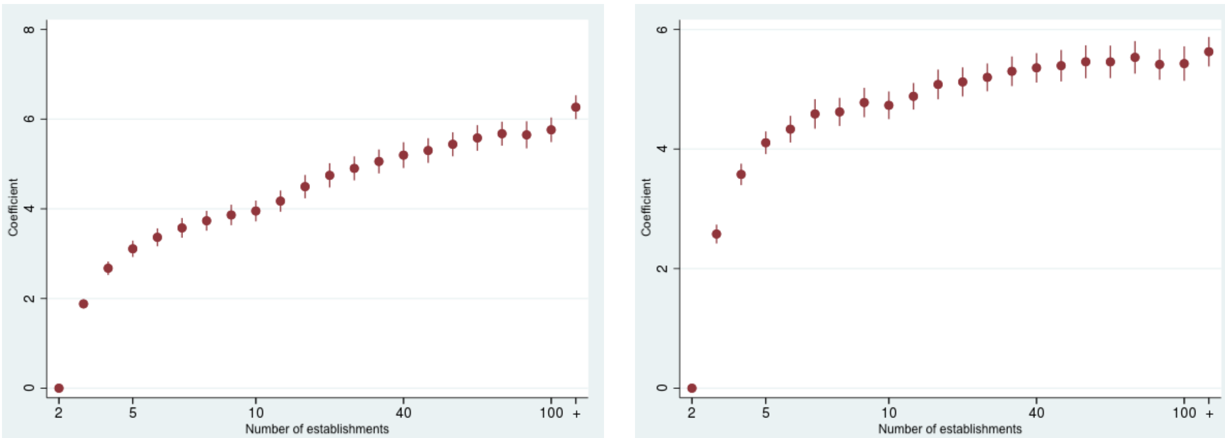
$$y_{jkt} = \alpha_{kt} + \beta_z \cdot \mathbf{1}_{estgroup=Z} + \alpha_a \cdot \mathbf{1}_{agegroup=a} + \epsilon_{jkt}$$

where y_{jkt} is the dispersion measure of firm j in industry k at time t, α_{kt} is an industry-year fixed effect, β_z is the conditional mean of firms for each establishment group in the set of groups z, and α_a is a fixed effect for firm age group. The differences between estimates β_z are now interpretable as the composition-adjusted differences in means.

However, that the overall level of the set of β_z is now indeterminate with the included fixed effects α_{kt} and α_a. Hence this specification is useful in that it allows us to look at whether differences between groups persist even after cleansing groups of industry composition differences, but is not useful for our original purpose, understanding the level differences between group means and synthetic controls.

This specification increases the standard errors for each bin and reduces the slope across size bins as compared to Figure 1, but the findings remain qualitatively and quantitatively similar. We conclude that the differences between bins in Figure 1 are in part due to composition differences but are robust to controlling for such differences.

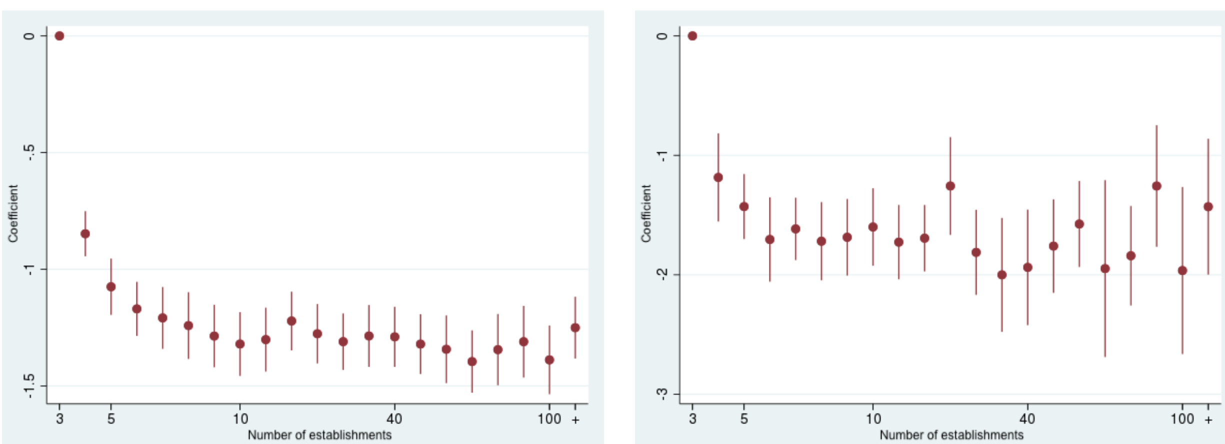
Figure 8a: Cross Section With Controls



(a) All firms (b) Manufacturing

Note: Red circles plot the coefficients for dummy variables for each firm size category from regression of log mean establishment distance from firm's centroid on dummies and industry-by-year-by-age-by-state controls. Standard errors are clustered by firm modal 4-digit industry.

Figure 9a: Time Series With Controls



(a) All firms (b) Manufactures

Note: Red circles plot the coefficients for dummy variables for each firm size category from regression of change of log mean establishment distance from firm's centroid on dummies and industry-by-year-by-age-by-state controls. Standard errors are clustered by firm modal 4-digit industry.

B.5 Additional robustness checks

Transitions down and up by more than one category Our main time series specification looks at firms that grow, moving up one size category. Alternatively, we can look at firms that move down one size category. Results are qualitatively similar for these firms.

Of course, firms may move more than one size category within 5-year windows. We can rerun our specification for the full transition matrix. Individual cells with few observations lose significance, but observed growth in dispersion is generally lower than synthetic baseline growth, and the difference is statistically significant for all growing firms for any transitions to under 25 establishments, as well as for transitions by up to three groups and to 10 establishments for manufacturing firms.

Other Robustness checks Results are robust to replacing interacted state fixed effects with interacted county fixed effects and replacing 4-digit industry controls with finer 6-digit controls. In order to rule out the possibility that our results are driven by extremely small establishments, we drop firms with five or fewer employees. None of our results are significantly affected by this change to the sample selection criteria. To ensure our results are not driven by assignments of firms to modal sectors, we also break firms into firm-sectors and repeat our analysis. Our results hold at the firm-sector level.

C Theoretical Appendix

C.1 Model Appendix

C.1.1 Plant location decisions and distance gradients

First, we demonstrate that when the distance penalty is constant across firms $\gamma\omega = \gamma$, the bias in the distance gradient attenuates with firm productivity ω.

Expression for sales For a plant from firm ω at location n, total revenue will be

$$r_n(\omega) = \left(\frac{p_n(\omega)^{1-\sigma}}{\mathbf{P}^{-\sigma}}\right) \cdot \frac{wL}{\mathbf{P}}, \tag{16}$$

where price is

$$p(\omega) = \frac{\sigma}{1-\sigma} \frac{w d_{h,n}^{\gamma(\omega)}}{\epsilon_n(\omega) \cdot z(\omega)}. \tag{17}$$

Rewriting revenue we have

$$r_n(\omega) = \rho \frac{d_{h,n}^{(1-\sigma)\gamma(\omega)}}{\epsilon_n(\omega)^{1-\sigma} \cdot z(\omega)^{1-\sigma}} \cdot \frac{w^{2-\sigma} L}{\mathbf{P}^{1-\sigma}}, \tag{18}$$

where ρ is a constant defined as $\rho = \frac{\sigma}{1-\sigma}^{1-\sigma}$.

Expression for employment The total labor demand from production at firm ω's plant at location n will be total quantity produced there divided by productivity.

$$l_n(\omega) = \frac{q(\omega) d_{h,n}(\omega)^{\gamma(\omega)}}{z(\omega)\epsilon_n(\omega)}, \tag{19}$$

which is

$$l_n(\omega) = \frac{d_{h,n}(\omega)^{(1-\sigma)\gamma(\omega)}}{z(\omega)^{1+\sigma}\epsilon_n(\omega)^{1+\sigma}} \frac{\rho w^{1-\sigma} L}{\mathbf{P}^{1-\sigma}}. \tag{20}$$

Note that $l_n(\omega)$ and $r_n(\omega)$ are a function of a plant's realized productivity. Specifically, the distance elasticity is the same for both and related to the distance-productivity elasticity by $\sigma - 1$.

Expression for profits Because the markup is a constant over marginal cost, the operating profit at location n for firm ω is positive when

$$\pi_n(\omega) = \frac{1}{\sigma} r(\omega) = \frac{\rho}{\sigma} \frac{d_{h,n}^{(1-\sigma)\gamma(\omega)}}{\epsilon_n(\omega)^{1-\sigma} \cdot z(\omega)^{1-\sigma}} \cdot \frac{w^{2-\sigma} L}{\mathbf{P}^{1-\sigma}} - w f_e > 0 \tag{21}$$

Expression for cutoff productivity Variety is placed at a location if it is profitable to do so.

Rearranging terms above to get a cutoff interms of epsilon:

$$\epsilon_n(\omega) > f^{\frac{1}{\sigma-1}} \frac{\sigma^{\frac{2-\sigma}{1-\sigma}}}{1-\sigma} \frac{wL^{\frac{1}{1-\sigma}}}{\mathbf{P}} \cdot z(\omega)^{-1} \cdot d_{h,n}^{\gamma(\omega)} \tag{22}$$

substituting for the constant terms we have the cutoff draw $\bar{\epsilon}_n$

$$\bar{\epsilon}_n = \xi \cdot \frac{d_{h,n}^{\gamma}(\omega)}{z(\omega)} \tag{23}$$

where $\xi = f^{\frac{1}{\sigma-1}} \frac{\sigma^{\frac{2-\sigma}{1-\sigma}}}{1-\sigma} \frac{wL^{\frac{1}{1-\sigma}}}{\mathbf{P}}$ is a constant from the firm's perspective. The cutoff draw is increasing in distance and decreasing in firm's core productivity. It may also be clarifying to note that for a given firm, the distance-adjusted minimum draw $\frac{\epsilon_n(\omega)}{d_{h,n}^{\gamma}(\omega)}$ is constant, and across all firms the minimum productivity draw (adjusting for both firm productivity and distance) is constant.

Differential selection by firm productivity The expected observed draw for a given firm at a given location post selection will be:

$$\widetilde{\epsilon_n} = \frac{\int_{\bar{\epsilon}_n}^{\infty} \epsilon_n dG(\epsilon_n)}{1 - G(\bar{\epsilon}_n)} \tag{24}$$

Note that $\widetilde{\epsilon_n}$ varies only by $\bar{\epsilon}_n$. That is, the average expected draw is only different across locations and firms by the cutoff productivity. This simplifies the analysis, as we know

$$\frac{\partial \bar{\epsilon}_n}{\partial d_{h,n}} = \frac{\partial}{\partial d_{h,n}} \cdot \frac{\xi d_{h,n}^{\gamma(\omega)}}{z(\omega)} = \frac{\gamma \xi d_{h,n}^{\gamma(\omega)-1}}{z(\omega)} > 0, \quad \text{and} \quad \frac{\partial \bar{\epsilon}_n}{\partial z(\omega)} = \frac{\partial}{\partial z(\omega)} \cdot \frac{\xi d_{h,n}^{\gamma(\omega)}}{z(\omega)} = \frac{-\xi d_{h,n}^{\gamma(\omega)}}{z(\omega)^2} > 0 \tag{25}$$

and furthermore,

$$\frac{\partial^2 \overline{\epsilon}_n}{\partial z(\omega)\partial d_{h,n}} = \frac{\partial}{\partial z(\omega)} \cdot \frac{\gamma \xi d_{h,n}^{\gamma(\omega)-1}}{z(\omega)} = \frac{-\gamma \xi d_{h,n}^{\gamma(\omega)-1}}{z(\omega)^2} < 0 \tag{26}$$

From these, it's clear that the corresponding comparative statics on $\widetilde{\epsilon_n}$ are of the same direction.

To illustrate, we can directly interrogate teh average draw when further imposing ϵ_n is drawn from a Pareto distribution with minimum 1 and shape parameter $\theta > 1$. So that $g(\overline{\epsilon}_n) = \theta \overline{\epsilon}_n^{-\theta-1}$, and $G(\overline{\epsilon}_n) = 1 - \overline{\epsilon}_n^{-\theta}$. Then,

$$\frac{\partial \overline{\epsilon}_n}{\partial z(\omega)} = \frac{-\xi d_{h,n}^{\gamma}(\omega)}{z(\omega)^2} = \frac{-\overline{\epsilon}_n}{z(\omega)} < 0 \tag{27}$$

The derivative of the average observed draw with respect to $z(\omega)$ is :

$$\frac{\partial \widetilde{\epsilon_n}}{\partial z(\omega)} = \frac{-g(\overline{\epsilon}_n)\frac{\xi d_{h,n}^{\gamma}(\omega)}{z(\omega)^2}}{1 - G(\overline{\epsilon}_n)} \cdot \left[\xi \frac{d_{h,n}^{\gamma}(\omega)}{z(\omega)} + \frac{\int_{\overline{\epsilon}_n}^{\infty} \epsilon_n g(\epsilon_n) d\epsilon_n}{1 - G(\overline{\epsilon}_n)}\right] = \frac{-g(\overline{\epsilon}_n)\frac{\overline{\epsilon}_n}{z(\omega)}}{1 - G(\overline{\epsilon}_n)} \cdot [\overline{\epsilon}_n + \widetilde{\epsilon_n}] < 0 \tag{28}$$

The derivative of the same with respect to d is :

$$\frac{\partial \widetilde{\epsilon_n}}{\partial d} = \frac{g(\overline{\epsilon}_n)\frac{\gamma(\omega)\xi d_{h,n}^{\gamma(\omega)-1}}{z(\omega)}}{1 - G(\overline{\epsilon}_n)} \cdot \left[\xi \frac{d_{h,n}^{\gamma}(\omega)}{z(\omega)} + \frac{\int_{\overline{\epsilon}_n}^{\infty} \epsilon_n g(\epsilon_n) d\epsilon_n}{1 - G(\overline{\epsilon}_n)}\right] = \frac{g(\overline{\epsilon}_n)\frac{\gamma(\omega)\overline{\epsilon_n}}{d_{h,n}}}{1 - G(\overline{\epsilon}_n)} \cdot [\overline{\epsilon}_n + \widetilde{\epsilon_n}] > 0 \tag{29}$$

This last derivative implies that there is increased selection at higher distance, i.e. plants at higher distances have on average higher idiosyncratic draws.

The derivative with respect to distance becomes:

$$\frac{\partial \widetilde{\epsilon_n}}{\partial d} = \frac{g(\overline{\epsilon}_n)\frac{\gamma(\omega)\overline{\epsilon_n}}{d_{h,n}}}{1 - G(\overline{\epsilon}_n)} \cdot [\overline{\epsilon}_n + \widetilde{\epsilon_n}] = \frac{\gamma(\omega)\overline{\epsilon_n}}{d_{h,n}} \frac{\theta \overline{\epsilon}_n^{-\theta-1}}{1 - (1 - \overline{\epsilon}_n^{-\theta})} \cdot [\overline{\epsilon}_n + \widetilde{\epsilon_n}] = \frac{\gamma(\omega)\theta}{d_{h,n}} \cdot [\overline{\epsilon}_n + \widetilde{\epsilon_n}] \tag{30}$$

Now we establish the cross-partial is positive, indicating selection is less severe for higher $z(\omega)$ firms:

$$\frac{\partial^2 \widetilde{\epsilon_n}}{\partial z(\omega)\partial d} = \frac{\gamma(\omega)\theta}{d_{h,n}} \cdot \left[\frac{\partial \overline{\epsilon}_n}{\partial z(\omega)} + \frac{\partial \widetilde{\epsilon_n}}{\partial z(\omega)}\right] < 0 \tag{31}$$

The first term is positive while both terms in parentheses are negative, from above. More productive firms have a lower selection gradient, in that the expected value of

Differential selection by plant count In our empirics, productivity is not directly observed, and instead we cut the sample by firms' plant count. Firms with a variety of underlying core productivities $z(\omega)$ will have the same realized number of plants due to differences in average draws. We now show that the differential selection follws that of average firm quality—that is, firms with more plants will on average also exhibit less attenuation in the distance elasticity.

First, consider two groups firms of size N=1 and N=2. The expected draw for a plant in group N will be

$$\widetilde{\epsilon_{n}}_{d,N} = \frac{\int_{\frac{\xi \cdot d_{h,n}^{\gamma(\omega)}}{z(\omega)}}^{\infty} \epsilon_n dG(\epsilon_n)}{1 - G(\frac{\xi \cdot d_{h,n}^{\gamma(\omega)}}{z(\omega)})} \tag{32}$$

Now, we are concerned with how $Cov(d_{h,n}^{\gamma}(\omega), \widetilde{\epsilon_n})$ is changing with N.

First, note that the threshold $\frac{\xi \cdot d_{h,n}^{\gamma(\omega)(z(\omega))}}{z(\omega)}$ and thus the expected value is only effected by N via $z(\omega)$. This is because the draws for each location are independent. In other words, conditional on $z(\omega)$, the value of N gives no additional information on the value of $\widetilde{\epsilon_n}$. However, the value of $z(\omega)$ will vary by N.

The expected difference between the expected draw for two plants at distance $d_1 < d_2$ from the two groups will be

$$\left(\widetilde{\epsilon_{n}}_{d_2,N=1} - \widetilde{\epsilon_{n}}_{d_1,N=1}\right) - \left(\widetilde{\epsilon_{n}}_{d_2,N=2} - \widetilde{\epsilon_{n}}_{d_1,N=2}\right) \tag{33}$$

$$= \int_{z(\omega)} (\widetilde{\epsilon_{n}}_{d_2}|_z - \widetilde{\epsilon_{n}}_{d_1}|_z) \cdot Pr(z = z(\omega)|N = 1)dz - \int_{z(\omega)} (\widetilde{\epsilon_{n}}_{d_2}|_z - \widetilde{\epsilon_{n}}_{d_1}|_z) \cdot Pr(z = z(\omega)|N = 2)dz \tag{34}$$

Notice that we drop the N subscripts on $\widetilde{\epsilon_n}$ as we are conditioning on $z(\omega)$ Now grouping terms:

$$= \int_{z(\omega)} (\widetilde{\epsilon_{nd_2}}|_z - \widetilde{\epsilon_{nd_1}}|_z) \cdot (Pr(z = z(\omega)|N = 1) - Pr(z = z(\omega)|N = 2))\, dz \tag{35}$$

Both terms in parentheses are decreasing in $z(\omega)$. We know the first is decreasing in $z(\omega)$ because the cross-partial $\frac{\partial^2 \widetilde{\epsilon_n}}{\partial z(\omega)\partial d} > 0$, and we know the second is true because more productive firms are more likely to have more plants. So, by Chebyshev Integral Inequality, the above is

$$> \int_{z(\omega)} (\widetilde{\epsilon_{nd_2}}|_z - \widetilde{\epsilon_{nd_1}}|_z)\, dz \cdot \left(\int_{z(\omega)} Pr(z = z(\omega)|N = 1) dz \int_{z(\omega)} -Pr(z = z(\omega)|N = 2) dz \right) dz \tag{36}$$

and since the probabilities are both equal to 1,

$$> \int_{z(\omega)} (\widetilde{\epsilon_{nd_2}}|_z - \widetilde{\epsilon_{nd_1}}|_z)\, dz \cdot (1 - 1)\, dz \tag{37}$$

So we know that for any two plants chosen from these two groupings of firms,

$$\left(\widetilde{\epsilon_{nd_2,N=1}} - \widetilde{\epsilon_{nd_1,N=1}}\right) > \left(\widetilde{\epsilon_{nd_2,N=2}} - \widetilde{\epsilon_{nd_1,N=2}}\right) \tag{38}$$

Finally, the above argument will hold when comparing any two groups of firms, thus will be true overall when we group firms by any single cutoff into high and low plant firms. Selection will be more severe for the grouping of firms with fewer plants.

C.1.2 General Equilibrium Conditions

Zero expected profit At each locations, firms of each type ω enter until the expected profit is zero.

$$E(\pi_n) = \pi(\omega) d\omega - E(N(\omega)) \cdot w f_e - w F_e = 0 \tag{39}$$

Housing market clearing condition At each location, labor employed in production, housing provision, or fixed cost provision for headquarters or locations spends a fixed share of income on a quantity of housing $q_{H,n}$:

$$q_{H,n} = \frac{(1-\alpha)w(1+P_H)}{P_H} \int_\omega \mathbf{1_n}(\omega)(l_n(\omega)+f_e) + \mathbf{1_1}(\omega)\cdot(l_n(\omega)+F_e)dG(\omega) \tag{40}$$

Labor market clearing For labor markets to clear, the total labor demanded by each firm at each location n, summed over all locations and firms, must equal to the total labor supply $\mathbf{L}$.

$$\mathbf{L} = (1+P_H)\int_\omega \sum_{n\in N(\omega)} \mathbf{1_n}(\omega)(l_n(\omega)+f_e) + \mathbf{1_1}(\omega)\cdot(l_n(\omega)+F_e)dG(\omega) \tag{41}$$

noting that all labor at any location i employed in production or fixed cost provision must be supplied housing at labor-denominated cost P_H.

Goods market clearing condition For the goods market to clear, the total wage bill must be equal to the total quantity consumed. Workers consumer a fixed share α of their income on goods.

$$\alpha w\mathbf{L} = \sum_{n\in I}\int_\omega p_n(\omega)q_n(\omega)d\omega \tag{42}$$

C.1.3 Extensions

Inelastic housing and local productivity shifters If housing is supplied inelastically rather than at a constant marginal cost, wages w_n will now be location specific and will adjust by location to preserve constant utility across locations. In particular, the local wage w_n will be determined according to the spatial equilibrium condition:

$$\Xi = log(w_n) - \alpha log\mathbf{P} - (1-\alpha)logP_{H,n} \tag{43}$$

where $P_{H,n}$ is the equilibrium price of housing at n, and the constant Ξ is $log(\overline{u}) - \alpha log(\alpha) - (1 - \alpha)log(1 - \alpha)$ where $\overline{u}$ equal to the prevailing utiliy across space. Higher real wages will make production more costly in more populous locations. This will act as a reduction in the realized productivity of plants at that location. Specifically, the cutoff productivity at each location will now be:

$$\overline{\epsilon}_n = \xi \cdot \frac{d^{\gamma}_{h,n}(\omega)}{z(\omega)} w_n \tag{44}$$

where ξ is now defined as before but without the wage.

Similarly, a location n can have a productivity shifter δ_n entering multiplicatively with the firm core productivity $z(\omega)$ for all firms at location n. Locations with higher δ_n will have a lower idiosyncratic threshold $\overline{\epsilon}$ in a manner isomorphic to the threshold determined by the location-specific wage w_n. Fully conditioning on the location-specific productivity shifters and wages, the analysis in the main text holds. This highlights the importance of controlling for plant location in our analysis.

Headquarters location decision An alternative approach to allowing entry at each location would be to allow firms to choose the location of their headquarters after entry. Firms draw an I-dimensional vector $z(\omega)$ and choose location n for headquarters plants with productivity $z_n(\omega)$. Firms choose headquarters locations then receive the location draw vector $\epsilon(\omega)$ as in the main text.

In this framework, firms trade off idiosyncratic core productivity attached to locations with those locations' productive potential. More central locations would in expectation be more productive. The relevant distribution of core productivity would be that ex-post of headquarters site selection, and can be non-degenerate at each location. The analysis in the main text follows. Specifically, while firm core productivity would be selected on locations' market potential, analysis conditional on firm productivity would be unaffected.

Vertical Differentiation It is possible to reinterpret the model in the main text as a model with vertical differentiation. We can consider an economy where consumers consume CES aggregates of varieties in each of K industries:

$$C_i = \prod_{k \in K} \left[\int_{\omega \in \Omega} q_{i,k}(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right]^{\frac{\sigma \delta_k}{\sigma-1}} \tag{45}$$

where industry consumption shares δ_k sum to 1.

Each firm can produce a maximum of one variety in each industry, and for each industry, firms receive a vector of draws $\epsilon_k(\omega)$. Firms choose how many industries to enter and where to locate their production for each industry.

The decision regarding the number of industries to enter in this model is isomorphic to the decision regarding the number of locations to enter in the main text, and all results follow with minor changes to the equilibrium conditions. As there are no trade costs in general, this setup eschews the effects of intra-firm trade cost present in vertically integrated firms and the subsequent forces.

C.2 Productivity Comparisons Across Firms and Plants

In this section we provide a simple example with CES demand and a general cost function to establish the firm-level relationship between total sales, sales per worker and a notion of productivity which encompasses the effects of both supply and demand shocks. We refer to the production unit as the "firm" below, then discuss how the analysis extends to plants within a firm.

Consider a set of firms $i, j \in J$ that hire factors $\mathbf{z}$ on competitive markets and choose output quantity q in order to maximize profits. Production and demand functions across firms differ only by multiplicative constants, i.e.

$$p_i(q_i) = a_i q_i^{-\frac{1}{\sigma}}, \quad q_i = b_i f(\mathbf{z}_i), \quad \forall i \in J, \tag{46}$$

with f differentiable and $\sigma > 1$.[22] We assume that firms face common factor prices $\mathbf{w}$. Under these assumptions, we have

$$c_i(\mathbf{w}, q_i/b_i) = c_j(\mathbf{w}, q_j/b_j) = c(\mathbf{w}, q), \quad \forall \mathbf{w},\ i, j \in J \tag{47}$$

for each firm's cost function. That is, each firm faces the exact same cost of "producing" q, although each firm has a different mapping from q to actual quantity produced.

Letting ℓ be the quantity of labor hired, we further assume that

$$\frac{\partial c(\mathbf{w}, q)}{\partial w_\ell} > 0, \quad \forall \mathbf{w}, q. \tag{48}$$

By Shepard's lemma, this condition implies that each firm's conditional labor demand is strictly increasing. This allows us to associate each "quantity" q with a unique choice of labor input ℓ^*, so we can write $q(\ell^*)$, with $q'(\ell^*) > 0$.

Using the notation above, firm i's problem can be written as

$$\max_{\ell_i^*} \frac{\mu_i \cdot q(\ell_i^*)^{1-\frac{1}{\sigma}}}{\ell_i^*} \cdot \ell_i^* - c(q(\ell_i^*)). \tag{49}$$

where we have suppressed the dependence of both c and ℓ^* on the common factor prices. Only revenue per worker differs across firms, due to both supply and demand shocks, whose effects can be summarized by a single multiplicative parameter $\mu_i = a_i b_i^{1-\frac{1}{\sigma}}$ due to the assumption of CES demand. Firm productivity is then defined as the total Marshallian surplus generated by the firm, or

$$Prod_i = \int_o^{\ell_i^*} \left[\frac{\mu_i \cdot q(\ell)^{1-\frac{1}{\sigma}}}{\ell} - c'(q(\ell))q'(\ell) \right] d\ell. \tag{50}$$

We are now ready to examine the relationship between total sales, sales per worker,

[22]This analysis applies to multi-plant firms so long as we can aggregate the plant outputs into a quantity index with an associated price index that takes this form.

and productivity. Using the Envelope Theorem, it is easy to show that firms with higher μ can have higher or lower revenue per worker, depending on the shape of $c(q(\ell^*))$. For example, with constant marginal cost CES demand implies that all firms have the same revenue per worker regardless of the value of productivity μ_i. Thus, in this model, sales per worker is not very informative regarding firm productivity. Either total sales or total employment is a better indicator of firm productivity, since $\partial \ell_i^* / \partial \mu_i > 0$ and $\partial Prod_i / \partial \mu_i > 0$.

C.2.1 Departures from CES demand

We maintain the previous assumptions, but relax CES demand. In this setting, a firm with both (weakly) higher employment and (weakly) higher sales per worker (and hence weakly higher sales) must be (weakly) more profitable. Furthermore, neither sales per worker nor employment is sufficient to establish relative profitability on its own, so the two statistics together yield an increase in information. However, moving from firm profitability to Marshallian surplus cannot generally be done. While relative profitability can be inferred by comparing the height of the revenue per worker curve for a single employment level (the smaller firm's optimal employment), the relative consumer surplus depends on the global behavior of these curves as well as the marginal cost curve.[23] With both supply and demand shocks, this behavior can be complex. To the extent that relative profitability is informative about relative productivity, both sales per worker and employment together are better indicators of productivity than either alone.

C.2.2 Comparing plants within firms

The results above apply equally well to comparisons between plants in the same firm. Marginal revenue products of labor must be equalized across plants, but average revenue products need not be. Larger plants may have higher or lower average revenue products of labor, depending on the shape of the cost function. Under CES demand, total sales or

[23]CES demand allows us to infer global behavior from relative height at a single point, which makes it convenient for this problem.

employment is positively related to productivity and profitability.

Within-firm comparisons face the additional challenge of the possibility of joint production and/or shared inputs across plants. To the extent that this measurement error is random, it will tend to bias toward a null finding. If geographically more central plants tend to provide more shared inputs (e.g. headquarter services), that will tend to mechanically produce a negative relationship between centrality and sales per worker even in the absence of true productivity differences. Thus, even more caution is required in interpreting differences in sales per worker across plants than across firms.

D Additional Within-firm Results

Alternative firm groupings Our results are robust to moving the threshold for "large" firms in the vicinity of 10 firms. Below, we plot results from an alternative specification of the form

$$y_{ijt} = \alpha_{itj} + \beta_z \ln dist_{ijt} + \gamma_z \mathbf{x}_{jt} + \epsilon_{ijt},$$

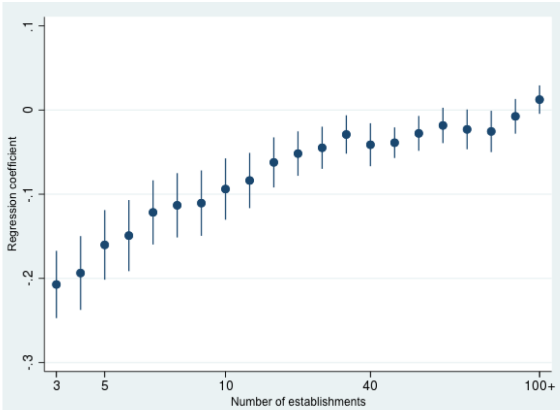
separately for each size class z, where y_{ijt} is an establishment j employment, $dist_{ijt}$ is a measure of the plant's distance to centroid, α_{it} is a firm-year-plant industry dummy and x_{ijt} is a set of year-plant industry - plant age group dummies.

In Panel (A), the pattern is distance penalty becomes gradually less negative as firm size increases. For manufacturing the estimates are noisier and no longer monotonically increasing, but jointly maintain an upward trend.

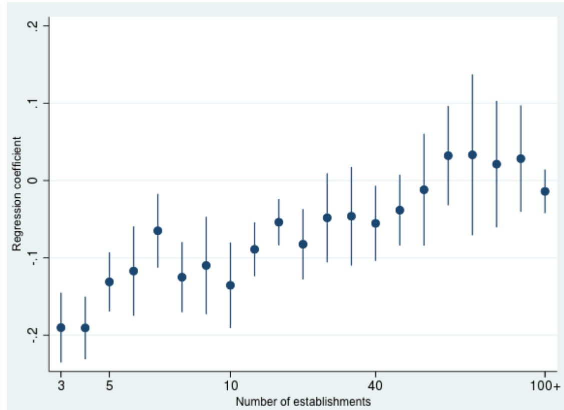
Within-firm results using alternative measures In this section, we probe for alternative explanations to the results of Table 1 and 2. First, we using alternative establishment measures of size. Next we explore the possibility that the nature of production, including factor intensities and managerial intensities are driving our within-firm result.

First, Table A2 reports results replicating the specification in Table 1 but using the

Figure 10a: Distance Penalty Within Firms



(a) Log employment, all firms



(b) Log employment, manufacturing

Note: Plotted points are firm-category specific slopes describing the relationship between establishments' log employment (top) or sales per worker (bottom) and log distance to firm centroid by firm size category as described by the specification in equation 15, which includes controls for a firm by (4-digit establishment) industry by year fixed effects as well as establishment age category fixed effects. All standard errors are clustered by establishment 4-digit industry.

natural log of sales as the outcome variable in place of employment. In both the full sample in Panel (A) and the sub-sample of manufacturing establishments in Panel (B), distance penalties and differential effects on establishments of large firms closely mirror the patterns reported in Table 1. On the one hand, because it conflates prices with quantities, sales is an imperfect measure of plant size. On the other hand, the majority of variation in price is likely either within firm-industry or location-industry, so that our fine controls on both of those dimensions likely strip most of the unwanted variance from these regressions. Broadly, the similarity between these results and those found in Table 1 reassure us that plants further away from the firm center are smaller.

Next, we directly test for other geographic patterns in production associated with plant size. Table A3 investigates the same specifications as in Table 1 using other outcomes related to production function technology. These variables exist only for the sub-sample of manufacturing plants. Panel (A) reports distance elasticity for the log of production workers. The distinction between production workers and total employment comes down to non-production workers, which are generally managerial workers. Thus production employment is a cleaner measure of employment which removes managerial services from labor inputs. The elasticities in all three columns of Panel (A) broadly reflect the results in Panel (B) of Table 1, indicating those results are not only driven by differences in managerial services.

Panel (B) reports distance elasticities for scope of production, measured as the log of the number of distinct products sold by the plant. While there is no clear relationship to distance to centroid, the number of products does fall with distance to regional headquarters for smaller firms only. On the one hand, larger plants are more likely to increase the number of products, and these results should be read as another substantiation of the results in Table 1. On the other hand, the existence of patterns in the number of products with distance indicates that output, and thus the production process, does indeed change in a way that is correlated with distance, despite our controls. We further investigate these differences in the final table in this section.

We are concerned that changes in factor intensities correlated with distance could drive the employment result in Table 1. First, we report that the results in Tables 1 and 2 are also robust to directly controlling for plant capital intensity. Table A4 further investigates the possibility of systematic differences of factor intensity across space, running the specification in Table 1 using factor intensities as an outcome. Panel (A) uses the natural log of the capital to labor ratio, while Panel (B) uses the natural log of the material input to labor ratio. Reading across both Panels, results in Column (1) indicate that capital intensity may be increasing in distance overall while material intensity is decreasing, although neither the coefficients nor their differential with those of large firms are significant. Furthermore, these patterns don't hold for distance to centroid in Column (2), and are highly sensitive to our fixed effect controls. Overall, we find no significant relationship between factor intensity and distance.

Finally, we test for evidence that tasks, and specifically managerial intensity, are correlated with distance in our specification. A common measure of managerial intensity in manufacturing is the percent of production workers. For the full sample of establishments, we lack task data on production vs non-production workers and we follow Giroud (2013) in using pay as a proxy for presence of managers, on the rationale that locations with higher managerial intensity will likely have higher average pay. Panel (A) uses the natural log of pay per worker as the right-hand side outcome in our specification for our main sample, while Panel (B) repeats the exercise on the sub-samples of manufacturing establishments. While pay, like sales, is a combination of prices and quantities, as in Table A2, our firm-industry and location-industry controls likely absorb a substantial amount of cross-plant price variation in similar worker types, leaving mostly variation due to differences in worker (and therefore task) types across plants. If differences in tasks and managerial intensity drove our employment elasticities, we would likely see pay per worker declining with distance for small firms especially. In both Panels (A) and (B), elasticities with respect to regional headquarters are small and insignificant. In Panel (A), there is a negative distance to centroid elasticity (Column 2) for small firms only, although this reverses in Column (3). Overall, we see very little consistent evidence for geographic patterns in average pay conditioning on our controls.

In Panel (C), we look for geographic patterns in the percent of production workers. The percent of production workers is more directly related to managerial intensity, but is only available for our sub-sample of manufacturing establishments. If managerial intensity declines with distance for small firms only, we should see a decline in the percent of production workers for those firms. Again, across all three columns, no such geographic patterns are evident. We conclude that, although our data on tasks within establishments is limited, we find—using our fine controls—no clear evidence of task or managerial intensity differences within the firm.

Table A2: Size and Distance Within Firms, Sales

		(A) All Firms: Sales	
Variable	ln Sales	ln Sales	ln Sales
ln miles HQ	0.034		-0.033
	(0.007)		(0.008)
1_{10+} x ln dist HQ	0.022		0.023
	(0.007)		(0.007)
ln miles to centroid		-0.119	-0.001
		(0.022)	(0.013)
1_{10+} x ln miles to cent.		0.079	0.013
		(0.018)	(0.012)
N	2,105,000	6,415,000	2,105,000
		(B) Manufacturing: Sales	
Variable	ln Sales	ln Sales	ln Sales
ln miles HQ	-0.220		-0.204
	(0.056)		(0.069)
1_{10+} x ln miles HQ	0.167		0.155
	(0.066)		(0.078)
ln miles to centroid		-0.142	-0.070
		(0.032)	(0.151)
1_{10+} x ln miles to centr.		0.112	0.055
		(0.039)	(0.153)
N	80,000	270,000	80,000

Notes: This table replicates Table 1 using the natural log of sales in place of employment. The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. All regressions include firm-industry-year fixed effects as well as industry-year-age group and industry-year-county fixed effects where industry controls are at the 6-digit level. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.

Table A3: Size and Distance Within Firms, Alternative Measures

	(A) Manufacturing: Production Workers		
Variable	ln Prod. Workers	ln Prod. Workers	ln Prod. Workers
ln miles HQ	-0.175		-0.191
	(0.067)		(0.065)
1_{10+} x ln miles HQ	0.139		0.160
	(0.075)		(0.070)
ln miles to centroid		-0.142	0.052
		(0.036)	(0.174)
1_{10+} x ln miles to centr.		0.097	-0.122
		(0.035)	(0.190)
N	80,000	270,000	80,000
	(A) Production Scope		
Variable	ln N Prods	ln N Prods	ln N Prods
ln miles HQ	-0.106		-0.091
	(0.025)		(0.038)
1_{10+} x ln miles HQ	0.118		0.102
	(0.026)		(0.038)
ln miles to centroid		-0.007	-0.061
		(0.008)	(0.076)
1_{10+} x ln miles to centr.		0.014	0.070
		(0.011)	(0.091)
N	80,000	270,000	80,000

Notes: Panel (A) repeats Table 1 Panel (B) using production workers rather than total plant employment. Panel (B) repeats Table 1 Panel (B) using total number of products. The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. All regressions include firm-industry-year fixed effects as well as industry-year-age group and industry-year-county fixed effects where industry controls are at the 6-digit level. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.

Table A4: Within-firm Elasticities, Additional Measures

		(A) Capital Intensity	
Variable	ln (K/L)	ln (K/L)	ln (K/L)
ln miles HQ	0.046		0.0278
	(0.062)		(0.069)
1_{10+} x ln miles HQ	-0.042		-0.025
	(0.067)		(0.075)
ln miles to centroid		-0.001	0.092
		(0.012)	(0.120)
1_{10+} x ln miles to centr.		0.018	-0.070
		(0.019)	(0.145)
N	80,000	270,000	80,000
		(B) Material Intensity	
Variable	ln (M/L)	ln (M/L)	ln (M/L)
ln miles HQ	-0.074		-0.061
	(0.050)		(0.058)
1_{10+} x ln miles HQ	0.076		0.061
	(0.049)		(0.054)
ln miles to centroid		0.018	-0.039
		(0.018)	(0.065)
1_{10+} x ln miles to centr.		-0.011	0.087
		(0.020)	(0.092)
N	80,000	270,000	80,000

Notes: The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. Panel (A) uses the log of the number of products produced in each plant as the dependent variable. Both panels repeat the exercise in Table 1 Panel (B) with different outcome measures. Panel (A) uses capital intensity, measured as the log of total value of plants' capital assets divided by total employment for the same. Panel (B) uses material intensity, measured as the log of value of materials used in production divided by total employment. All regressions include firm-industry-year fixed effects as well as industry-year-age group and industry-year-county fixed effects where industry controls are at the 6-digit level. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.

Table A5: Geography of Tasks Within the Firm

	(A) All Firms: Worker Pay		
Variable	ln Pay/Worker	ln Pay/Worker	ln Pay/Worker
ln miles HQ	0.005		0.003
	(0.004)		(0.003)
1_{10+} x Ln dist HQ	-0.001		-0.000
	(0.002)		(0.003)
ln miles to centroid		-0.013	0.067
		(0.003)	(0.005)
1_{10+} x ln miles to cent.		0.017	0.005
		(0.006)	(0.005)
N	2,105,000	6,415,000	2,105,000
	(B) Manufacturing: Worker Pay		
Variable	ln Pay/Worker	ln Pay/Worker	ln Pay/Worker
ln miles HQ	-0.013		-0.027
	(0.023)		(0.024)
1_{10+} x ln miles HQ	0.000		0.015
	(0.027)		(0.027)
ln miles to centroid		-0.005	0.066
		(0.010)	(0.030)
1_{10+} x ln miles to centr.		0.002	-0.070
		(0.007)	(0.033)
N	80,000	270,000	80,000
	(C) Manufacturing: Percent Production Workers		
Variable	% Prod Workers	% Prod Workers	% Prod Workers
ln miles HQ	-0.008		-0.007
	(0.016)		(0.018)
1_{10+} x ln miles HQ	0.010		0.009
	(0.014)		(0.016)
ln miles to centroid		0.006	-0.003
		(0.003)	(0.015)
1_{10+} x ln miles to cent.		-0.010	0.045
		(0.004)	(0.012)
N	80,000	270,000	80,000

Notes: Panels (A) and (B) replicate Table 1 using pay per worker rather than employment as the outcome variable. Panel (C) replicates Table 1 Panel (B) using plants' percent production workers as the outcome variable. The coefficient on distance is interacted with firm size indicator which is 1 for firms with ten or more establishments. All regressions include firm-year fixed effects as well as industry-year-age group fixed effects. Standard errors in parentheses are clustered by industry at the 4-digit NAICS level.