The Internal Geography of Firms *

Dominick Bartelme
University of Michigan

Oren Ziv
Michigan State University

December 2020

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 center employ significantly less workers than closer plants within the same firm, and that this effect is attenuated in large firms. These findings suggest that large firms face lower costs of geographic expansion.

JEL codes: R32,L25

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

*We thank C. Luke Watson for excellent research assistance as well as Camillo Acosta-Mejia, Andrei Levchenko, Martha Stinson, Junfu Zhang, Jens Wrona and numerous seminar and conference participants for their helpful comments. Any opinions and conclusions expressed herein are those of the author(s) and do not necessarily represent the views of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed. Contact: dbartelm@umich.edu, orenziv@msu.edu.

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.

First, 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 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

[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 et al., 2015; Alcacer and Delgado, 2016; Charnoz et al., 2018; Atalay et al., 2019). These papers generally find that distance to management and/or other production units reduces plant performance.

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, where we show that on average firms that add plants expand their geographic footprint by substantially less than the baseline industry location patterns alone would predict.

Second, and perhaps 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 firms 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 geographic 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 different 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.

Third, we examine the relationship between internal distance and plant performance. The existing literature has documented that of various frictions in moving goods, people or ideas across space that impede coordination between distant plants and impair plant performance (Giroud, 2013; Kalnins and Lafontaine, 2013; Eichholtz et al., 2015; Alcacer and Delgado, 2016; Charnoz et al., 2018; Atalay et al., 2019). 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 center is negative and sizable for small firms but significantly smaller in magnitude for large firms. The differences are economically significant: applying the distance penalty faced by small firms to large firms while holding their plant locations fixed results in a loss to firm employment of 14%, while the analogous exercise for small firms yields an employment gain of 8%.

This is primarily a descriptive paper, and we do not take a strong stand on the underlying mechanisms. One appealing and parsimonious explanation that can rationalize all of the empirical patterns that we find is that firms face costs of geographic dispersion, but that large firms face lower costs than smaller firms. Costs of dispersion rationalize the observed average clustering, while lower costs for larger firms would explain both the diminished clustering and lower distance-size elasticity we find in large firms. These cost differences could be primitive, or they might be the result of different investments made by firms with different productivity levels. This explanation also resonates with the previous literature, cited above, that has documented the significance of internal distance costs. However, 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 estimating a quantitative structural model, and is challenging because we lack a tractable theory of plant location with many alternatives when locations are substitutes (Jia, 2008; Tintelnot, 2017; Arkolakis and Eckert, 2017; Antras et al., 2017). This exercise is beyond the scope of this paper.

Our paper contributes to the empirical literature on plant location in multi-unit firms. Our finding of geographical clustering in the cross-section has antecedents in the work of Henderson and Ono (2008) on headquarter location of manufactures and Behrens and Sharunova (2015) on Canadian manufacturing firms, which finds spatial clustering within these firms and examines the contribution of input-output linkages to this pattern. Our time series results on the geographic expansion paths of firms is related to industry studies such as Holmes (2011), who ties the gradual spatial expansion of Walmart to supply chain management. Our contribution generalizes the results from this literature and suggest that these costs exercise a powerful albeit heterogeneous quantitative influence on firm geography across the economy. To the extent that the costs are responsive to policies such as transportation and communication infrastructure investment, there may be substantial long-run efficiency gains from the spatial reorganization of firms. Hsieh and Rossi-Hansberg (2019) and Rossi-Hansberg et al. (2018) consider the extensive margin expansion of large firms into local new cities but do not consider the geographic footprint of these firms.

Our paper is also related to work on the international geography of multinational activity (see Antràs and Yeaple (2014) for a review). Our results that, on average, domestic firms cluster their plants geographically resembles the fact that multinational expansion within the firm is also geographically and culturally concentrated (e.g. Yeaple (2009); see Alfaro and Chen (2018) for a review). Concentration is linked to trade and communication costs which reduce productivity in distant affiliates(Keller and Yeaple, 2013; Bahar, 2020; Gumpert, 2018).Irarrazabal et al. (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 findings suggest that the some of the same economic forces shape the patterns of both multinational and domestic expansion of firms. However, 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 the pattern of multinational expansion.

2 Data and Measurement

Measuring Dispersion 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 $c = (c_x, c_y)$.[2] This provides a computationally tractable measure of these establishments' dispersion in a way that 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."[3]

We also measure distance to headquarters using Equation 1, replacing d_{nc} with plant's distance to the firm 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 combine both measures when investigating the within-firm distribution of plant locations and plant characteristics in Section 4.

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.[4] 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

[2] 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.

[3] 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.

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

(defined below) from the preceding Census. We find approximately 44,000 such firms.[5]

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.

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.[6] In the main body of the paper, we report results pooling all firms and for the manufacturing sector. Results for business servies, 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.

[5]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 more than one employee in subsequent Census waves.

[6]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

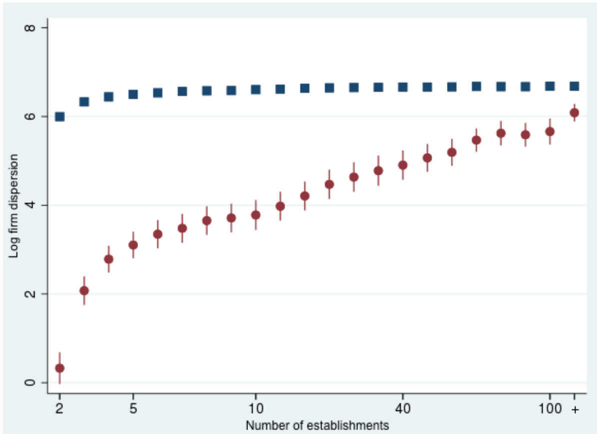
This section documents several empirical patterns in the domestic location decisions of multi-plant firms in the United States. Small and medium-size firms exhibit clustering their establishments relative to a counterfactual benchmark in which firms choose plant locations that replicate industry location 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

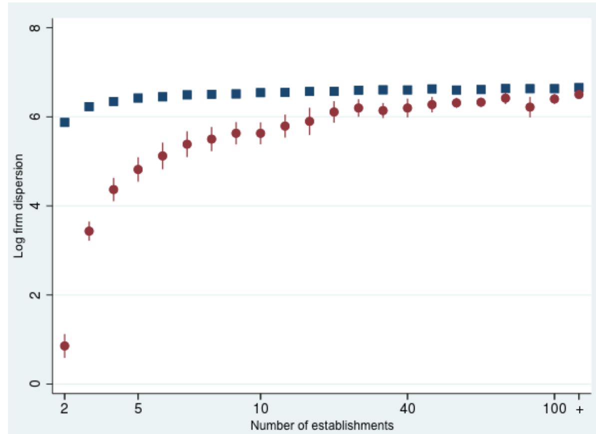
We first examine the cross-sectional patterns in within-firm plant location decisions. Figure 1a plots the average log dispersion by size class (in terms of the number of plants) for all firms in our sample with 2 or more plants, with standard errors clustered at the industry level, while Figure 1b plots 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.

How different is this pattern from the geographic clustering we observe in plant location across firms? That is, 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 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 $\tilde{f}_{it} \in \tilde{F}_t$ with the same

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.

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}.[7] 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 subgroups, 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,

[7]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.

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.

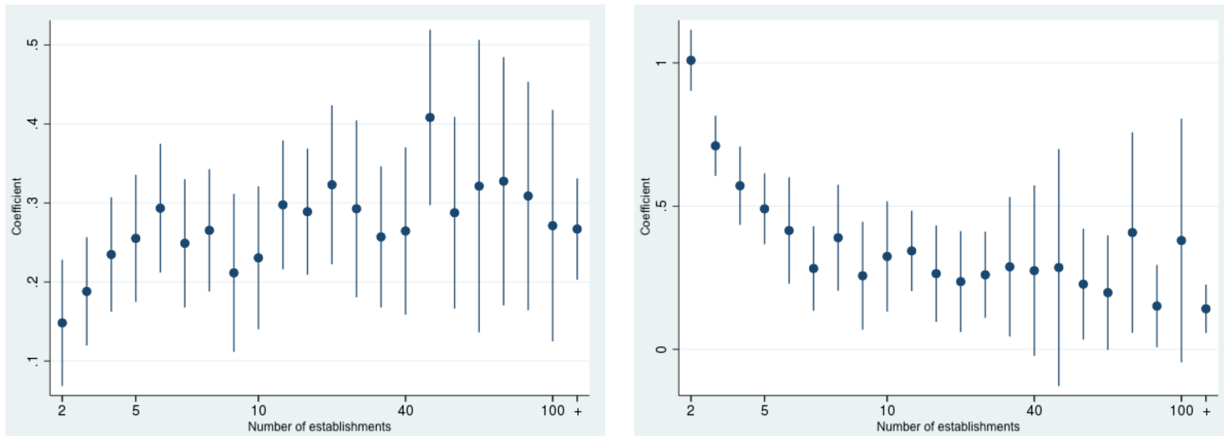
These results are robust to a number of specification changes and alternative measurements, all reported and discussed in more detail in Appendix B.3. We show that results are similar when using distance to headquarters as the measure of dispersion and when sorting firms by employment ventile rather than by the number of plants in the firm.[8] 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 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

[8]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 syntehtic 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.

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^c_{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.[9]

To address these issues, we proceed as follows. For each Census year[10] and initial size 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.[11] 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.

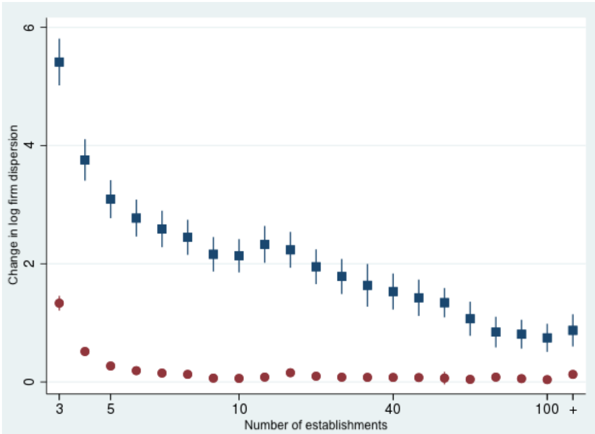
The circles in Figures 3a and 3b plot the results for all firms and manufacturing, respectively. In both groups, small and medium-sized firms tend to increase their average dispersion when they expand, at a rate decreasing in firm size. For the smallest firms,

[9]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.

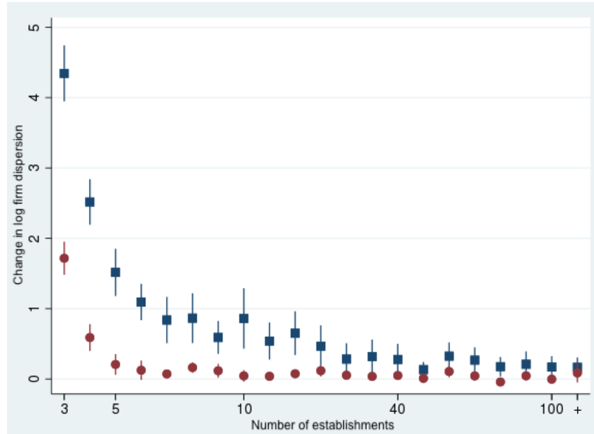
[10]We use Census waves because establishment birth years between Census waves are imputed for multi-unit firms in the LBD. The total number of establishments born between Census waves is not. We thank Martha Stinson for alerting us to this feature of the data.

[11]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.

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.

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.

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 $\tilde{f}_{it}$ 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.[12] We then compute the average increase in log dispersion across years,

[12] 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.

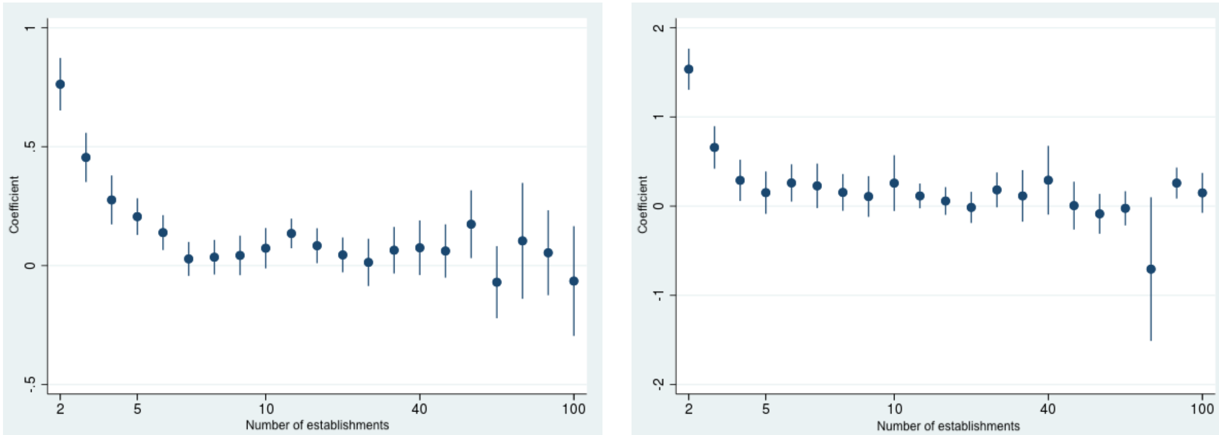
within size class, for the synthetic baseline in the same way as we did for the data.

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 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 sub-sample 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

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.

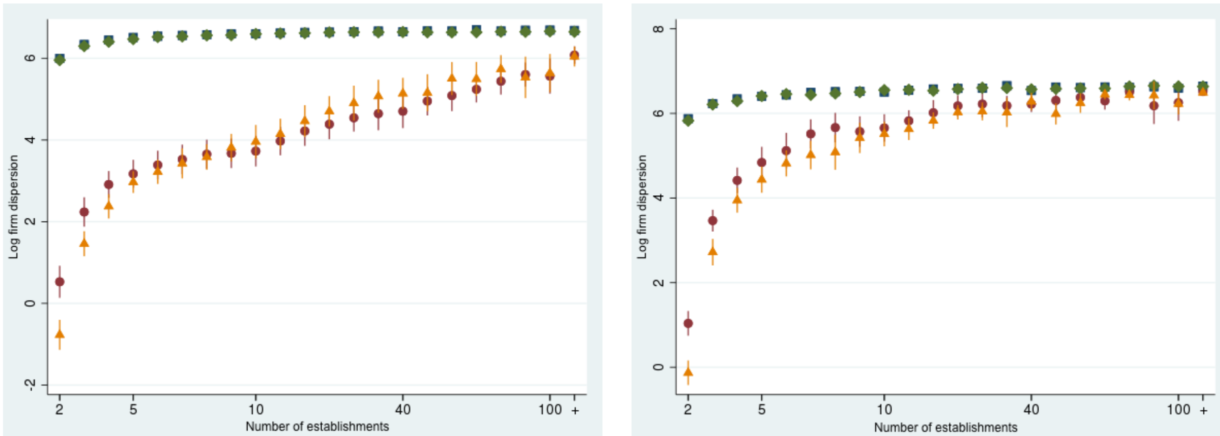
example, the difference in average log dispersion between 2 and 3 plant manufacturing firms in the cross-section is greater than 2 log points, while the average firm with 2 plants that we observe adding a third plant increases its dispersion by substantially less than 2 log points. The discrepancy is 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 random benchmark than comparable firms in the same size class that do not grow. The result-

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.

ing 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 1976 are still active in 2012, major changes in location patterns should be at least somewhat reflected in differences in cross-sectional log dispersions over this time period. There is clear evidence for modest 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. This rising dispersion can also help explain the difference between the cross-sectional and time series patterns, especially at the low end, since the gaps in dispersion between size classes tend to narrow over time.

3.3 Discussion

Our findings can be summarized in 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 (i.e. the "proximity-concentration" tradeoff, see Antràs and Yeaple (2014) for a review). 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 (Antras et al., 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 et al., 2015; Alcacer and Delgado, 2016; Charnoz et al., 2018;

Atalay et al., 2019). Within this paradigm, our findings suggest that these costs of dispersion may have pervasive and significant effects on plant location choices for smaller firms across all sectors of the economy. It is important to note 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 these alternative mechanisms would be interesting in and of themselves, evaluating their relative contribution to the high level of clustering we find requires estimating a quantitative structural model and is challenging because we lack a tractable theory of plant location with many alternatives when locations are substitutes (Jia, 2008; Tintelnot, 2017; Arkolakis and Eckert, 2017; Antras et al., 2017).

However, larger, more productive firms appear to be less affected, either because the costs of dispersion are lower for larger firms or because the benefits of dispersion are correspondingly greater for those firms. It could be that larger, more productive firms invest in lowering the costs of dispersion, or they might be "born" with lower costs of dispersion which in turn contribute to their measured productivity. Rigorously differentiating between these explanations (or quantifying their relative importance) using purely empirical methods would require 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.

Instead, we focus on testing a particular implication of the hypothesis that large firms face lower costs of dispersion: if this is the case, we would expect that plants that are far from the geographic center of large firms would be more productive (relative to other plants in the same firm) than correspondingly placed plants in small firms. That is, the "productivity penalty" paid by more distant plants should be higher for small firms than for large. We investigate this hypothesis in the following section.

4 The distance-productivity relationship within the firm

We examine the within-firm relationship between plant characteristics, chiefly as employment and sales per worker, and their distance from the geographic center of the firm. Firm size variables (such as employment) have often been used as a measure of productivity, based on the theoretical link between the two variables (e.g. (Hopenhayn, 1992; Melitz, 2003)). Although the link between sales per worker and a welfare-relevant notion of productivity is more tenuous, sales per worker is often used as a productivity measure in non-manufacturing industries, e.g. (Haltiwanger et al., 2007). In general, these variables jointly yield a robust qualitative measure of firm productivity since, all else equal, plants exhibiting both higher employment and higher or the same 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 link relative profitability to the relative Marshallian surplus generated by the firms.[13] These measures have their limits, especially as quantitative indicators, but involve far fewer conceptual and practical difficulties than computing measures of revenue TFP, especially for non-manufacturing firms. We use these variables in our main tables and report all robustness checks in Appendix B.3.

We run regressions of the form

$$y_{jt} = \alpha_{ikt} + \beta \ln dist_{ijt} + \theta \cdot \mathbf{1}_{10+} \ln dist_{ijt} + \gamma \mathbf{x}_{jt} + \epsilon_{jt}, \tag{4}$$

where y_{jt} is an establishment j characteristic like employment or sales per worker, $dist_{ijt}$ is a measure of plant j's geographic position within the firm i (e.g. distance to firm centroid), $\mathbf{1}_{10+}$ is a dummy that equals 1 if firm i has 10 or more plants,[14] α_{ikt} is a firm-year-(plant industry) dummy and x_{jt} is a set of year-(plant state)-(plant industry) - (plant age group) dummies interacted with $\mathbf{1}_{10+}$.[15]

Our predictions, based on the location patterns documented in the previous section, are that a) more distant plants suffer a productivity penalty relative to more central

[13]We provide a brief discussion of the issues in Appendix C.
[14]We report some results for more detailed size class breakdowns in Appendix B.3.
[15]Results are robust to using counties instead of states as the geographic unit.

Table 1: Employment and geography within the firm

Variable	ln employment	ln employment	ln employment
		(A) All Firms	
ln miles to centroid	-0.118		-0.029
	(0.018)		(0.010)
1_{10+} x ln miles to cent.	0.085		0.034
	(0.013)		(0.012)
ln miles HQ		-0.027	-0.026
		(0.005)	(0.006)
1_{10+} x ln dist HQ		0.001	-0.001
		(0.003)	(0.003)
N	6,415,000	2,105,000	2,105,000
		(B) Manufacturing	
ln miles to centroid	-0.149		-0.173
	(0.025)		(0.066)
1_{10+} x ln miles to centr.	0.097		0.162
	(0.023)		(0.085)
ln miles HQ		-0.020	-0.014
		(0.008)	(0.007)
1_{10+} x ln miles HQ		-0.016	-0.022
		(0.007)	(0.008)
N	270,000	80,000	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 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.

plants in the same firm, but b) this effect is weaker for larger firms. This prediction flows from a "gravity of knowledge" model of the costs of dispersion (e.g. Keller and Yeaple (2013); Ramondo and Rodríguez-Clare (2013); Arkolakis et al. (2018)), where productivity decays with distance from the center. Our empirical test is based on comparing the relative productivity of similar plants located at similar distances from their respective firm centers, for small and large firms. Mapping this test to the theory would be straightforward if plant locations were randomly chosen, but in practice we have non-random selection of plant locations. If distance-based productivity penalties are stronger for

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 to centroid	-0.005		0.002
	(0.002)		(0.006)
1_{10+} x ln miles to cent.	0.004		0.010
	(0.004)		(0.007)
ln miles HQ		-0.004	-0.004
		(0.001)	(0.002)
1_{10+} x Ln dist HQ		-0.004	-0.004
		(0.001)	(0.002)
N	6,415,000	2,105,000	2,105,000
		(B) Manufacturing	
ln miles to centroid	-0.013		-0.036
	(0.014)		(0.055)
1_{10+} x ln miles to centr.	0.008		0.061
	(0.014)		(0.052)
ln miles HQ		-0.007	-0.006
		(0.005)	(0.007)
1_{10+} x ln miles HQ		0.003	0.002
		(0.005)	(0.007)
N	270,000	80,000	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 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.

small firms, then small firms will only select a distant site if it has a higher-than-average fundamental productivity draw for that location. Thus if we observe a small firm choosing the same plant location (relative to its center) as a large firm, we can infer that on average the small firm is likely to have a relatively higher fundamental productivity at that site. This selection effect implies any observed larger distance penalty for small firms will tend to be an understatement of the true *ceteris paribus* differential.

Table 1 reports the results of estimating equation (4) using establishment employment as the outcome variable. The specification in the first column uses distance from plant centroid as the independent variable of interest. Using our full sample in panel (A), we find plant-level employment declines sharply with distance from the firm centroid for small firms, with a sharply estimated elasticity of around -0.12. However, for larger firms the penalty is much lower, with an elasticity estimate of about -0.03. Qualitatively and quantitatively similar patterns are observed in the manufacturing subsample in panel (B), although the dramatic decrease in the sample size raises standard errors somewhat.

The second column uses plant distance to headquarters as the independent variable instead of distance to firm centroid. Here our sample drops significantly, as few firms have unique declared headquarter establishments. Interestingly, the results for all firms imply that large and small firms have similar employment penalties for distance to headquarters (elasticity about -0.03) while the results for manufacturing even suggest that large firms may have slightly larger penalties for distance to headquarters. The third column includes both variables and yields similar results to the independent specifications: small firms face sizable distance-to-centroid penalties and distance-to-headquarters penalties, whereas large firms only face the distance-to-headquarters penalty. These results are robust to using other plant-level outcomes related to the scale of production, such as sales or value added.

Table 2 reports results for the same regressions using sales per worker as the outcome variable. Broadly speaking the point estimates are qualitatively in line with what we find for employment, although the estimated elasticities are much smaller and the standard errors much larger. Sales per worker does not appear to be robustly related to the firm's internal geography in any quantitatively significant way. This conclusion is robust to other indicators of input productivity (which we have for manufacturing only), like value added per worker and total factor productivity.

To summarize, for small firms establishments further away from the center of the firm are significantly smaller, with similar sales per worker as closer establishments. These facts jointly imply that further away establishments are less profitable and less productive productive in many common theoretical frameworks.[16] This observed dis-

[16]For example, with general cost functions and CES demand. See Appendix C for a brief discussion.

tance penalty in productivity could be one explanation for the high geographic clustering of plants that we documented for small firms in the previous sections. The observed penalty due to distance to centroid is significantly smaller for larger firms, which is a potential explanation for our second finding that larger firms tend to be more geographically dispersed dispersed. One caveat to this interpretation is that these are correlations based on observed location choices rather than random assignment; however, as discussed above, we would expect the non-random selection of plants to narrow rather than magnify the differences in observed distance penalties. Another wrinkle is that the observed headquarters penalty tend to be similar if not bigger for large firms. However, as we show below, quantitatively the smaller distance-to-centroid penalties for large firms dominate the slightly larger distance-to-headquarters penalties.

We use the estimates from the last column of Table 1 to give a rough calculation of the magnitude of the differences in observed distance penalties – to the firm centroid as well as to headquarters – on employment in small and large firms. We generate predicted employment for establishments of firms in each group by keeping each establishment's actual distance to centroid and headquarters, but imposing the distance penalties observed in the other group. Using the large firm penalties applied to small firms results in an average increase in firm employment of 8% over the values actually observed. Using the small firm penalties applied to large firms results in an average decrease of 14% in firm employment.[17] Of course, these calculations are not reflective of true counterfactuals; even if we accept the estimates as true *ceteris paribus* productivity penalties, firms can also adjust their location decisions in response to a change in costs. What these calculations illustrate is that the overall distance penalties incurred by small firms are quantitatively large compared to those of large firms.

A number of mechanisms may be behind the observed productivity penalties. Of these, increased cost of headquarter services is likely best captured by distance to headquarters, for which we find no evidence of a differential burden on small firms. This is consistent with an empirical literature that finds significant costs associated with

[17]These differences are even larger in manufacturing, with small firms growing by 43% and large firms shrinking by 78%. However, the standard errors are much larger in manufacturing due to the smaller sample size.

distance-to-headquarters even in large multinational firms (Keller and Yeaple, 2013), although the cross-country context is somewhat different than ours. The differential effect on centroid may reflect differential distance costs not captured by headquarter distance, such as minimizing distance along supply chains (Holmes, 2011; Atalay et al., 2019) and other costs of coordination among units.

5 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. They also raise questions about the nature and size of these costs to dispersion. Both the costs and the benefits of dispersion might be small, or they might both be large. Costs could be "constants of nature" that are largely invariant to policies such as regulations and infrastructure, or they might be quite responsive to the right policy changes. 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 et al. (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. It therefore seems important for future research to explore the size of the costs and what policy levers they might respond to.

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. However, despite recent developments that apply when locations are complements (see e.g. (Jia, 2008; Tintelnot, 2017; Arkolakis and Eckert, 2017; Antras et al., 2017)), modeling multiple discrete choices with general interdependence remains challenging.

References

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

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

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

_ and Stephen R Yeaple, "Multinational firms and the structure of international trade," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 55–130.

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

Arkolakis, Costas and Fabian Eckert, "Combinatorial Discrete Choice," Technical Report, Yale University 2017.

_ , Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and production in the global economy," *American Economic Review*, 2018, *108* (8), 2128–73.

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

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

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

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

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

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

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

Foster, Lucia, John Haltiwanger, and Chad Syverson, "Reallocation, firm turnover, and efficiency: selection on productivity or profitability?," *American Economic Review*, 2008, *98* (1), 394–425.

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

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

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

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

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

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

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

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

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

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

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

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

Ramondo, Natalia and Andrés Rodríguez-Clare, “Trade, multinational production, and the gains from openness,” *Journal of Political Economy*, 2013, *121* (2), 273–322.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter, “Diverging trends in national and local concentration,” Technical Report, National Bureau of Economic Research 2018.

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

Yeaple, Stephen Ross, “Firm heterogeneity and the structure of U.S. multinational activity,” *Journal of International Economics*, 2009, *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 4 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 headquarters. Roughly one quarter of firms have 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). 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 dispersions. We then computed both measures and obtained a correlation of 0.97 between them.

B Additional Results

B.1 Results for Additional Sectors

We divided non-manufacturing firms into four subsamples 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.

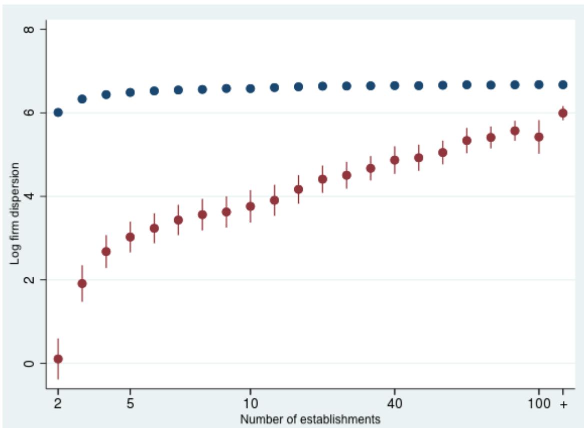
Broadly speaking, clustering patterns both in the cross section and time series for Finance and Insurance, Wholesale and Retail, and Other all follow the patterns we find for All Firms panels in the main text. Clustering patterns for Business Services appear more similar to those in manufacturing in 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

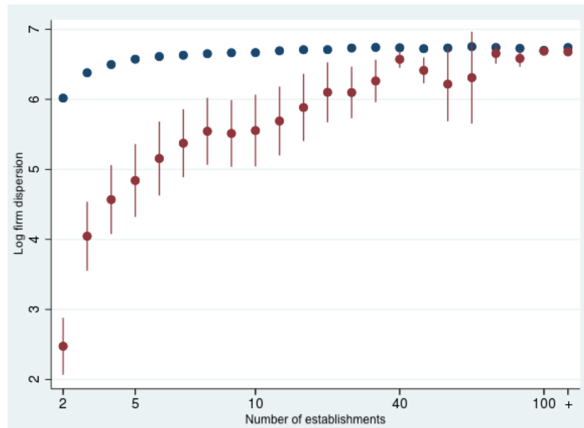
Here we report results using alternative measures of firm distance and alternative firm categorizations.

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 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

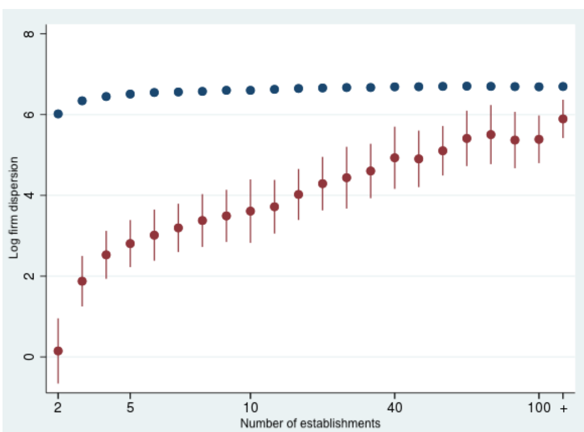
Figure 2a: Cross-Sectional Firm Dispersion, Other Sectors



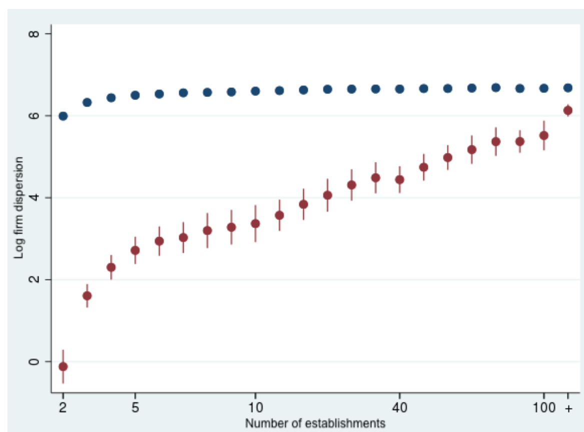
(a) Retail and wholesale firms



(c) Business services firms



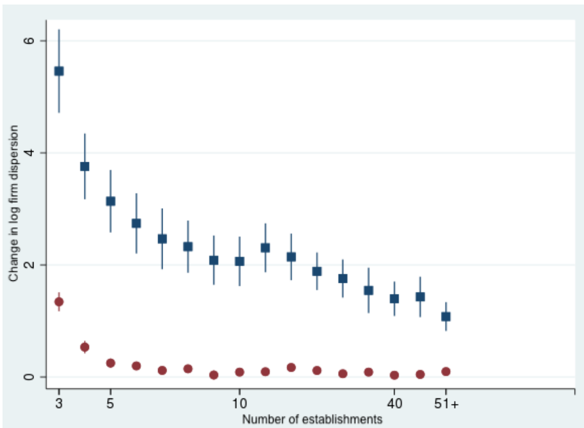
(b) Finance and insurance firms



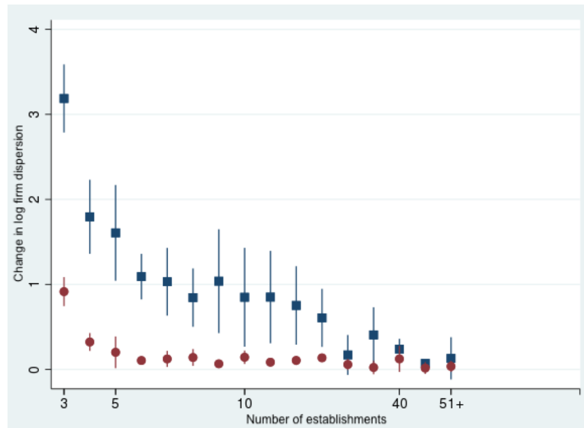
(d) Other firms

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.

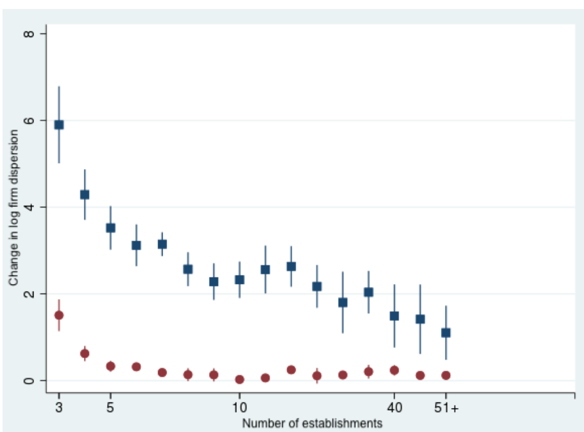
Figure 3a: Time Series Firm Dispersion, Other Sectors



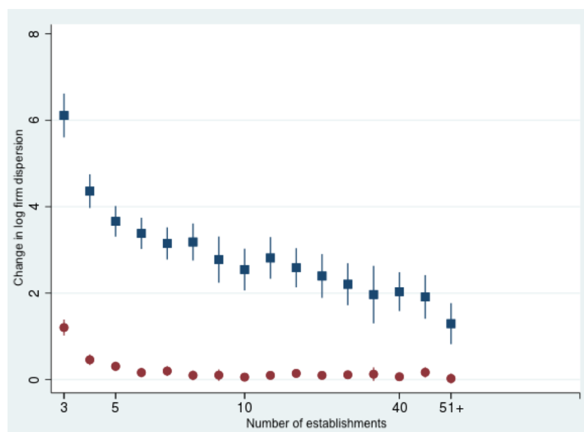
(a) Retail and wholesale firms



(c) Business services firms



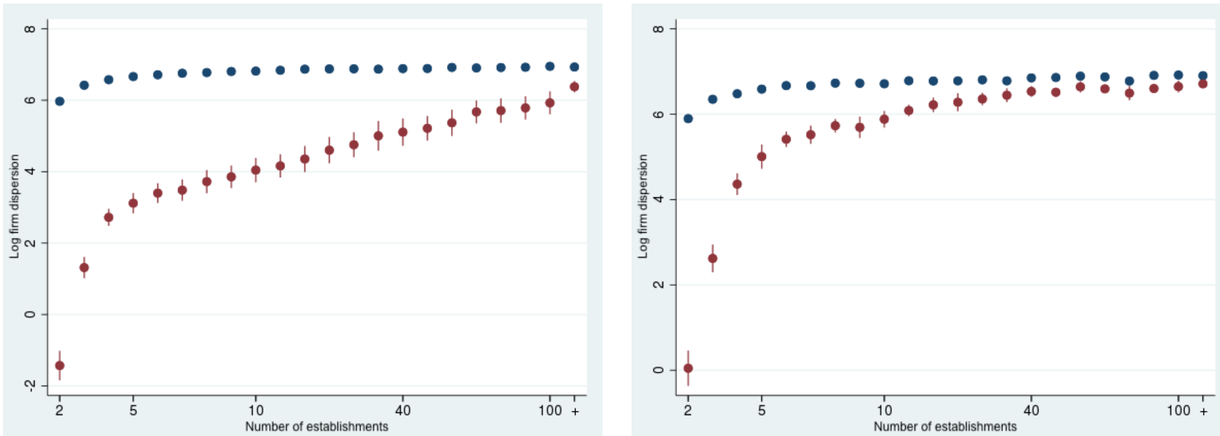
(b) Finance and insurance firms



(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.

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.

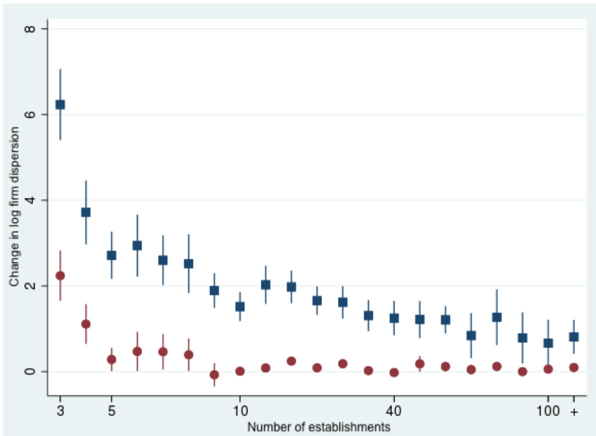
Alternative firm size measures Next, we replicate our main cross sectional results using an alternative firm categorization scheme. We assign firms to employment ventiles and replicate the procedure in Figure 1. While the curvatures of the resulting figures are slightly different, all qualitative results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

B.3 Alternative Synthetic Controls

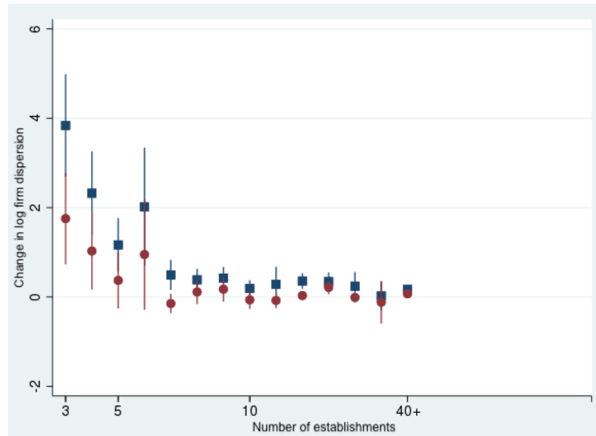
Alternative sampling procedures and self-sampling corrections In principle, several choices could affect the construction of our synthetic baseline and impact results. First, we choose to sample with replacement as a conservative measure, this may generate the concern that our finding that the largest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

Figure 5a: Time Series, Distance to Headquarters



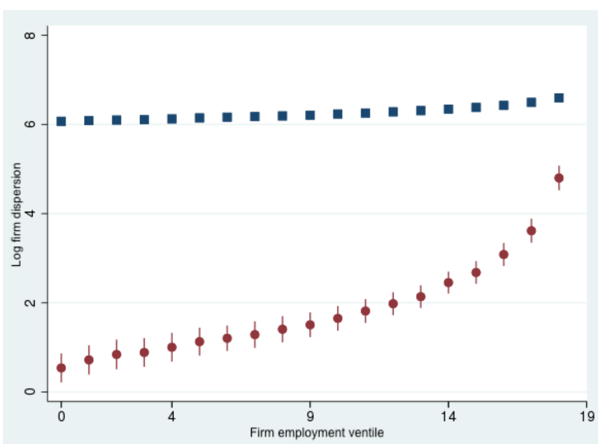
(a) All firms



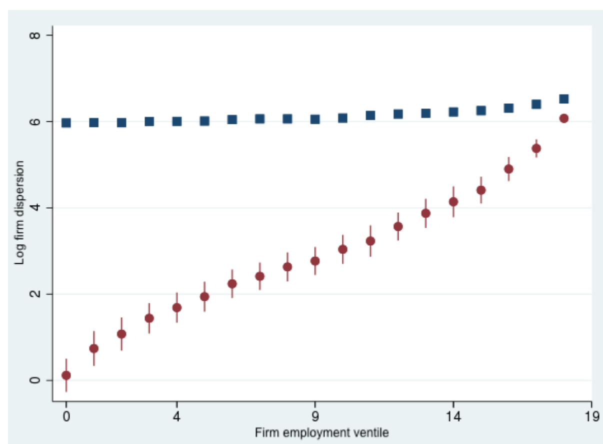
(b) Manufacturing

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.

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

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

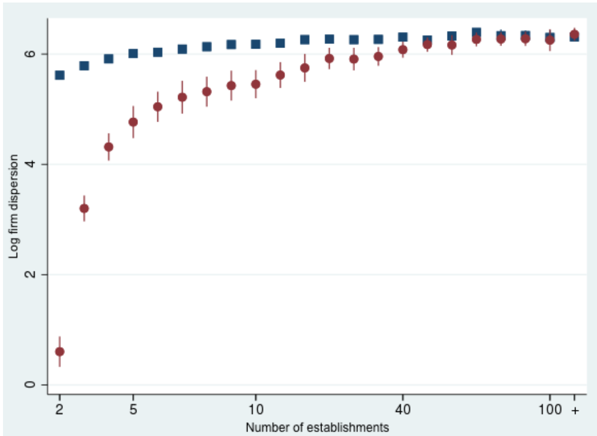
Using the matched plants, we recompute the firm distance to centroid as our new synthetic baseline. We repeat this for the vector of outputs $\vec{S_i} = [t_1, t_2, ..., t_O]$ and report the results below in Appendix Figure 6a.

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.

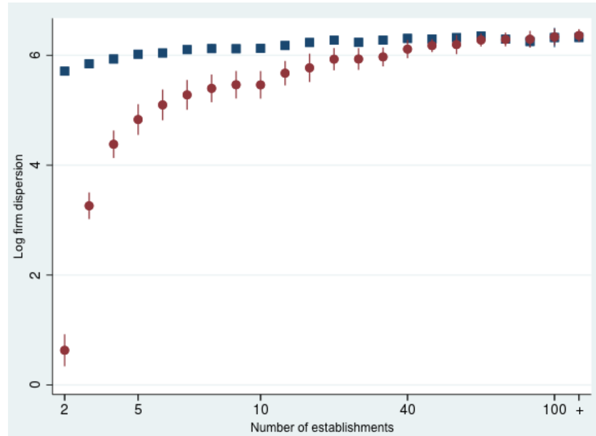
Results using manufacturing establishments matched on input and output usage by firm size category do not pass disclosure avoidance requirements and cannot be reported.

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 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.

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.

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.

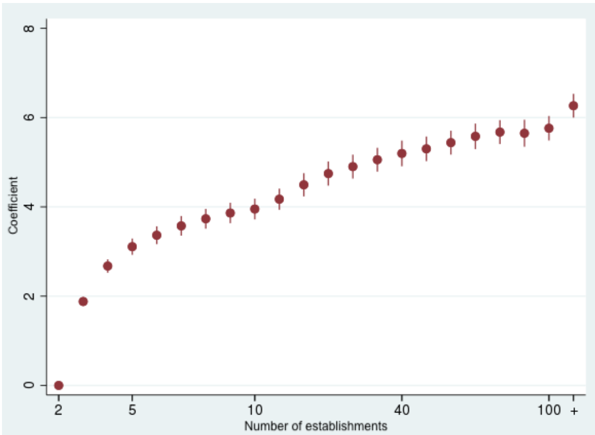
This specification increases the standard errors for each bin and reduces the slope across size bins as compared to Figure 1, but the findings remain qualitatively and quantitatively similar. We conclude that the differences between bins in Figure 1 are in part due to composition differences but are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

B.5 Additional robustness checks

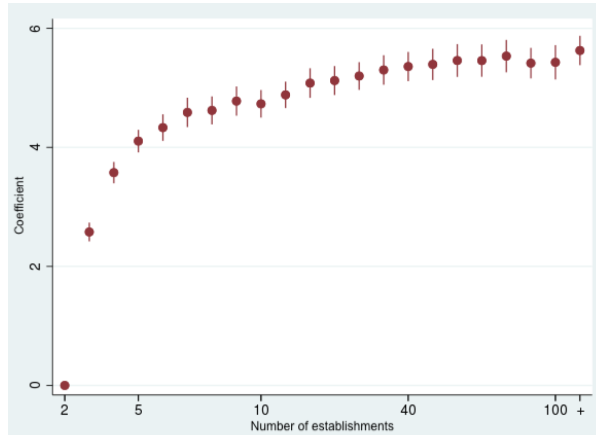
Transitions down and up by more than one category Our main time series specification looks at firms that grow, moving up one size category. Alternatively, we can look at firms that move down one size category. Results are qualitatively similar for these firms.

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.

Figure 8a: Cross Section With Controls



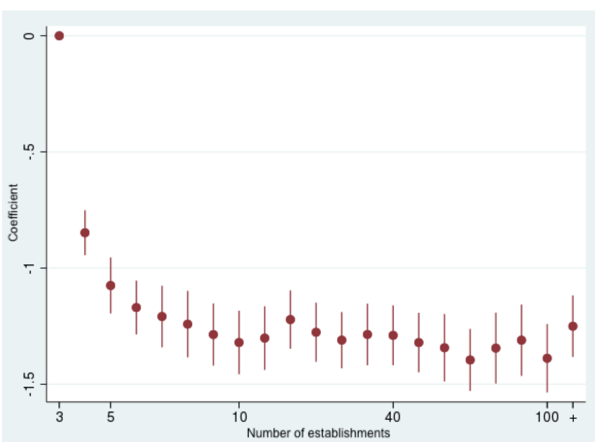
(a) All firms



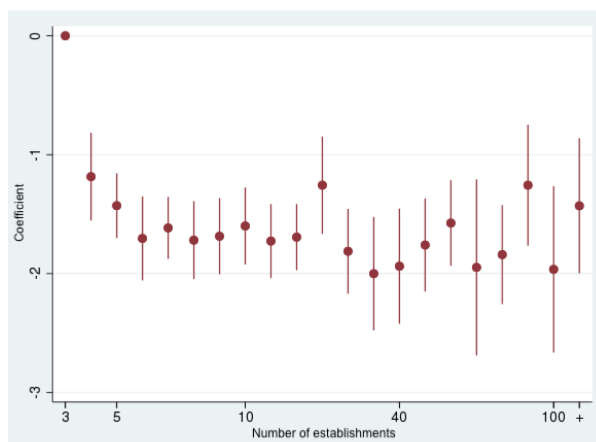
(b) Manufacturing

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.

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.

B.6 Within-firm comparisons: alternative groupings and outcome variables

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

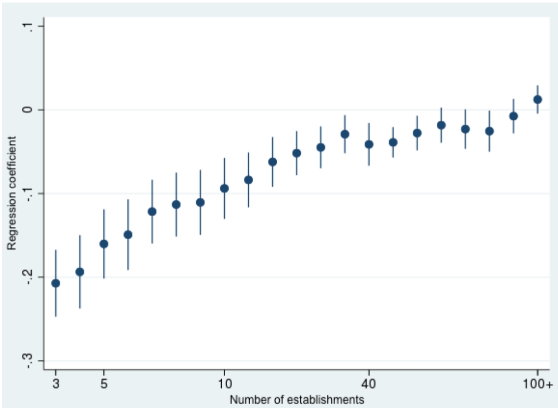
$$y_{ijt} = \alpha_{itj} + \beta_z \ln dist_{ijt} + \gamma_z \mathbf{x}_{jt} + \epsilon_{ijt},$$

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.

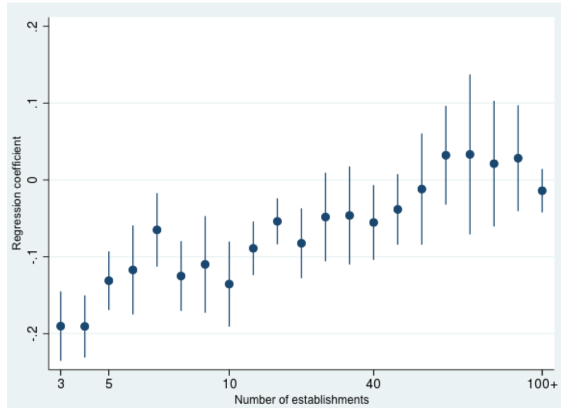
In Panel (A), the pattern is distance penalty becomes gradually less negative as firm size increases. For manufacturing the estimates are noisier and no longer monotonically increasing, but jointly maintain an upward trend.

Within-firm results using alternative measures In this section, we replicate results in Table 1 and 2 using alternative establishment measures of size and productivity. We rerun specification 4 using sales as an outcome variable and report outcomes in the tables below. First, Table A2 reports results using sales, in place of employment in Table 1. In both the full sample and the subsample of manufacturing establishments, distance penalties and differential effects on establishments of large firms closely mirror the patterns reported in 1, qualitatively as well as in magnitude. Next, in Table A3 we rerun results using value added, another measure of establishment size. These measures are only available for the manufacturing subsample. Point estimates here also closely

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 4, 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.

match results using sales and employment.

Finally, we use two other measures of productivity available for the subsample of manufacturing establishments: value added per worker and TFPr, which we measure as in Foster et al. (2008). Results are reported in Tables A4 and A5, and are comparable to those in the main text in Table 2, in panel (B), which looks at sales per worker of manufacturing establishments in particular. Results using these measures are slightly smaller in magnitude and as noisy as those in Table 2. While some signs change direction, these results are consistent with findings in Section 4 that there are no clear establishment distance penalties observable using productivity outcomes.

Table A2: Sales and geography within the firm

Variable	ln sales	ln sales	ln sales
		(A) All Firms	
ln miles to centroid	-0.123		-0.027
	(0.018)		(0.014)
1_{10+} x ln miles to cent.	0.089		0.044
	(0.015)		(0.016)
ln miles HQ		-0.031	-0.030
		(0.006)	(0.006)
1_{10+} x ln dist HQ		-0.003	-0.005
		(0.003)	(0.003)
N	6,415,000	2,105,000	2,105,000
		(B) Manufacturing	
ln miles to centroid	-0.162		-0.209
	(0.018)		(0.088)
1_{10+} x ln miles to centr.	0.104		0.224
	(0.025)		(0.110)
ln miles HQ		-0.028	-0.020
		(0.009)	(0.010)
1_{10+} x ln miles HQ		-0.012	-0.021
		(0.010)	(0.012)
N	270,000	80,000	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 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.

Table A3: Value added and and geography within the firm, Manufacturing Only

Variable	ln value added	ln value added	ln value added
ln miles to centroid	-0.151		-0.159
	(0.024)		(0.104)
1_{10+} x ln miles to centr.	0.078		0.156
	(0.031)		(0.132)
ln miles HQ		-0.028	-0.022
		(0.011)	(0.012)
1_{10+} x ln miles HQ		-0.001	-0.007
		(0.009)	(0.012)
N	270,000	80,000	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 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.

Table A4: Value added per worker and geography within the firm, Manufacturing Only

Variable	ln va / worker	ln va / worker	ln va / worker
ln miles to centroid	-0.008		-0.012
	(0.011)		(0.052)
1_{10+} x ln miles to centr.	0.001		0.046
	(0.013)		(0.050)
ln miles HQ		-0.002	-0.017
		(0.004)	(0.005)
1_{10+} x ln miles HQ		0.006	0.004
		(0.004)	(0.005)
N	270,000	80,000	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 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.

Table A5: Total Factor Productivity and geography within the firm, Manufacturing Only

Variable	ln TFPr	ln TFPr	ln TFPr
ln miles to centroid	-0.001		-0.009
	(0.006)		(0.025)
1_{10+} x ln miles to centr.	-0.006		-0.002
	(0.006)		(0.027)
ln miles HQ		0.002	0.002
		(0.002)	(0.002)
1_{10+} x ln miles HQ		0.001	0.001
		(0.002)	(0.003)
N	270,000	80,000	80,000

Notes: Dependent variable is establishment TFPr calculated in Foster et al. (2008). 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.

C 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.

$$p_i(q_i) = a_i q_i^{-\frac{1}{\sigma}}, \;\; q_i = b_i f(\mathbf{z}_i), \;\; \forall i \in J, \tag{5}$$

with f differentiable and $\sigma > 1$.[18] We assume that firms face common factor prices $\mathbf{w}$. Under these assumptions, we have

$$c_i(\mathbf{w}, q_i/b_i) = c_j(\mathbf{w}, q_j/b_j) = c(\mathbf{w}, \tilde{q}), \;\; \forall \mathbf{w}, \; i, j \in J \tag{6}$$

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

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

$$\frac{\partial c(\mathbf{w}, \tilde{q})}{\partial w_\ell} > 0, \;\; \forall \mathbf{w}, \tilde{q}. \tag{7}$$

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

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

$$\max_{\ell_i^*} \frac{\mu_i \cdot \tilde{q}(\ell_i^*)^{1-\frac{1}{\sigma}}}{\ell_i^*} \cdot \ell_i^* - c(\tilde{q}(\ell_i^*)). \tag{8}$$

[18] 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.

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

$$Prod_i = \int_o^{\ell_i^*} \left[\frac{\mu_i \cdot \tilde{q}(\ell)^{1-\frac{1}{\sigma}}}{\ell} - c'(\tilde{q}(\ell))\tilde{q}'(\ell) \right] d\ell. \tag{9}$$

We are now ready to examine the relationship between total sales, sales per worker, 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 $c(\tilde{q}(\ell^*))$. 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.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.[19] 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

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

employment together are better indicators of productivity than either alone.

C.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 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.