Plants in Space

Ezra Oberfield

Princeton University

Esteban Rossi-Hansberg

University of Chicago

Pierre-Daniel Sarte

Federal Reserve Bank of Richmond

Nicholas Trachter

Federal Reserve Bank of Richmond

To decide the number, size, and location of its plants, a firm balances the benefit of delivering goods from multiple plants with the cost of setting up and managing these plants and the potential for cannibalization among them. Modeling the decisions of heterogeneous firms in an economy with a vast number of distinct locations involves a large combinatorial problem. Using insights from discrete geometry, we study a tractable limit case of this problem in which these forces operate at a local level. Our analysis delivers predictions on sorting across space. Compared with less productive firms, productive firms place more plants in dense, high-rent locations and fewer plants in markets with low density and low rents. We present evidence consistent

We thank Milena Almagro, Fernando Alvarez, Donald Davis, Felix Tintelnot, Harald Uhlig, and Conor Walsh, as well as participants at numerous seminars and conferences for their feedback. We thank Eric LaRose, Reiko Laski, James Lee, Samira Gholami, Sara Ho,

Electronically published February 7, 2024

Journal of Political Economy, volume 132, number 3, March 2024.

https://doi.org/10.1086/726907

with these and several other predictions, using US establishment-level data.

I. Introduction

Delivering products and services to locations where consumers can easily access them involves complex decisions about where to locate plants and how large these plants should be. Having too few establishments is costly because it increases the distance to consumers. Having too many involves large span-of-control and fixed costs, as well as plants that cannibalize each other's customers. Understanding how these trade-offs play out for firms with different characteristics in an economy consisting of many local markets that differ in demand and production costs is complex. Perhaps because of the difficulty of the problem, little is known about the solution to this fundamental problem of how to organize production. The sorting of firms in space determines not only the profitability of firms but also consumers' surplus as well as the characteristics of individual locations. In this paper, we study this core component of a firm's production problem, provide a methodology that simplifies it significantly, and contrast its implications with the data.

Consider the case of Starbucks, which operated around 14,000 stores in 2019 in locations across the United States. Of course, not all Starbucks are equal in size, not all locations in the United States have a Starbucks, and the distance between neighboring Starbucks stores in a location differs across space. Simply put, there is a lot of variation across space in how individual stores are arranged. This variation is naturally related to the spatial distribution of population density, wages, and other characteristics. For example, figure 1 shows the location of Starbucks' establishments in three markets: Princeton, NJ; Richmond, VA; and New York City. Clearly, the number of establishments, as well as the distance between them, varies across these cities. Even within New York, the number of establishments is much larger, and the distance between them is much shorter, in the densest parts of Manhattan. What are the general characteristics of establishment location decisions? Clearly, density matters, but the scale of establishments is by no means constant in space. The average plant employment of Starbucks in New York is more than 23% higher than that in Richmond.

Casual evidence and introspection might suggest that firms simply sell in the densest markets, with the marginal market determined by a firm's productivity. A closer look, however, reveals a more nuanced pattern. Figure 2 provides a simple example. Walgreens and Rite Aid are pharmacies

and Suren Tavakalov for outstanding research assistance. The views expressed herein are those of the authors and do not necessarily represent the views of the Federal Reserve Bank of Richmond or the Federal Reserve System. This paper was edited by Harald Uhlig.

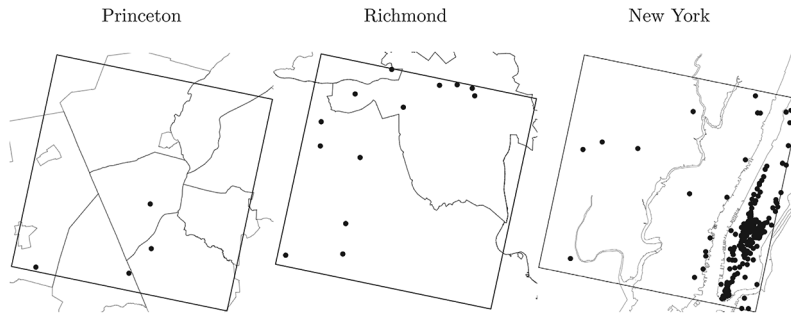

FIG. 1.—Density of Starbucks locations in a 12 × 12-mile area in Princeton, Richmond, and New York City.

that operate nationally, but Walgreens's total employment is larger, and it has more establishments. The figure shows that, in fact, both pharmacy chains tend to have more establishments in denser locations. However, Rite Aid has more stores than Walgreens in less dense locations. Is this

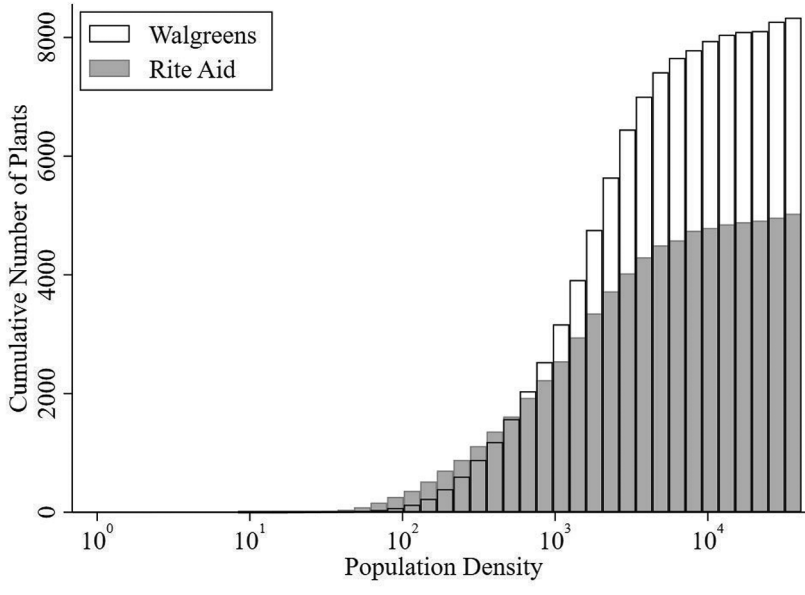

FIG. 2.—Sorting: Walgreens versus Rite Aid. For Walgreens and Rite Aid, this graph plots the cumulative number of establishments in locations with no more than the given population density. Population density is measured as the population density of the 6 × 6-mile square in which the establishment is located, using data from the 2010 decennial census, taken from Manson et al. (2021).

form of sorting across locations a general feature of the solution to the location problem or of the data?

More generally, we aim to provide insights into two main questions: First, which firms set up plants in which locations? Second, what determines the scale and location of production? Answering these questions requires us to think about plants and firms as distinct, albeit related, economic entities. In particular, we set up an economy with a continuum of heterogeneous locations. These locations have different productivities and amenities that determine, in general equilibrium, the distribution of population density, wages, and residential and business rents. They also determine, given the form local competition takes in a location, residual demand for a firm's product. We focus primarily on the problem of a firm that takes these distributions as given and needs to decide whether and how to serve each consumer. The firm decides where to set up production plants, how large each plant should be, and from which plant to serve each of its customers. We assume that firms face iceberg transport costs. Setting up a plant entails a fixed cost that depends on local land rents. The productivity of a plant depends on local characteristics as well as a firm-specific component that decreases with the total number of plants the firm operates. In other words, increasing a firm's span of control by adding plants implies a management cost that lowers its productivity. The main trade-off faced by the firm, therefore, is to reduce transport costs by setting up more plants close to consumers versus setting up fewer plants to economize on fixed costs, augment productivity by lowering its span of control, and reduce cannibalization between plants. Ours is a standard setup of this canonical firm decision problem.

Solving this production problem when the set of potential locations is large and heterogeneous involves a large and challenging combinatorial optimization problem. Our contribution is to focus on a limit formulation of the problem in which the firm chooses a density of plants, rather than a discrete set. The firm's problem then becomes one of calculus of variations, which is simpler to solve. Crucially, in the limit we propose, all relevant trade-offs described above remain active. Specifically, we study a limit of the problem in which fixed and span-of-control managerial costs become small while transportation costs become large. In this limit economy, the problem of the firm becomes amenable to an analytical characterization, making it easy to transparently characterize its implications.

In characterizing the solution to the firm's problem in the limit economy, we apply insights from discrete geometry, exploiting the sum-of-moments theorem by Fejes Tóth (1953). The theorem provides the optimal way to arrange plants across space when economic activity is uniform and the number of plants is large. When space is one-dimensional, plants should be located at the centers of equal-length intervals. In two dimensions, the result states that plants should be located at the centers of catchment areas

given by hexagons arranged so as to cover all locations. The intuition for this result is that, among all polygons with which one can construct a uniform tessellation, the hexagon is the closest to a circle.[1] A circle minimizes the average distance from a plant located at its center to its customers. However, unlike hexagons, circles cannot be used repeatedly to form a tessellation. We extend the theorem to an environment where economic activity is heterogeneous across space. Specifically, customers are not necessarily uniformly distributed across space, while plant costs and productivities also differ across locations.

Apart from being obviously important for practical applications, our extension of the theorem allows us to study sorting patterns, namely, the many-to-many matching between heterogeneous firms and heterogeneous locations. It helps us understand examples such as Starbucks, or Walgreens and Rite Aid, for which the number of establishments changes with customer density but at different rates and with different ranges of locations that vary with the firm's aggregate scale. In general, our theory delivers testable implications on sorting patterns across firms for industries that are well approximated by our limiting economy—industries for which catchment areas are small. First, more productive firms set up more plants in denser, high-rent locations than less productive firms. Perhaps more surprising is that they also set up fewer plants in markets with lower density and rents. For highly productive firms, the span-of-control cost of managing additional plants in low-density locations outweighs the benefits of low rents. In contrast, for less productive firms, low local rents are more attractive, since they manage fewer plants and so span-of-control costs are less relevant. Second, among firms that operate the same number of plants in a location, the plants from the more productive firm are larger in that location. Highly productive firms obtain large variable profits from each plant but limit the number of plants in a location because of the span-of-control cost they impose on the whole firm. Hence, when they set up the same number of plants as a less productive firm in a location, they choose to make these plants larger. In the final section of the paper, we present evidence, using the National Establishment Time Series (NETS) dataset, that corroborates these and other predictions of our theoretical analysis.

One of the advantages of using our limit economy to analyze the firm production problem is that the simplicity of the solution allows us to embed the problem into an equilibrium setup. To illustrate this without adding too much additional structure, we embed the firm's problem into a

[1] A tessellation is an arrangement of shapes, especially of polygons, in a repeated pattern without gaps or overlaps.

small industry that does not affect local characteristics.[2] We show that, in equilibrium, the local industry price index is a function of local productivity and a weighted sum of the productivity of all firms present in the location, with weights that depend on each firm's local footprint, that determines the cost of delivering goods to customers. These are the only local characteristics that determine equilibrium outcomes. Analyzing realistic quantitative equilibrium counterfactuals is not the focus of our study, but the method we develop to make the firm's problem tractable can be readily used to do so. We provide an algorithm to compute an industry equilibrium of our economy and illustrate the effect of improvements, in a single industry, in the technology to manage a firm's span of control and the technology to transport goods.

Canonical models of firm dynamics (i.e., Jovanovic 1982 and Hopenhayn 1992) make no clear distinction between a plant and a firm. However, mounting evidence points to the importance of considering plants and firms as different but related entities. Rossi-Hansberg and Wright (2007) highlight large differences between the size distributions of enterprises and establishments. In addition, Rossi-Hansberg, Sarte, and Trachter (2021) show evidence of diverging trends in market concentration at the national and local levels resulting from the expansion of the largest firms into new markets. Hsieh and Rossi-Hansberg (2022) show that industries with large increases in national market concentration also saw their top firms expand their operations geographically through the opening of new plants in smaller markets. Further, Aghion et al. (2019) observe that the average number of plants per firm has risen considerably in the United States, and Aghion et al. (2019) and Cao et al. (2019) provide evidence that growth through the opening of new plants has been a key margin of firm employment growth since 1990.

The distinction between firms and plants has been more prevalent in the international trade literature, given the interest in multinational production and export platforms. Examples of papers in this literature include Ramondo and Rodríguez-Clare (2013), Ramondo (2014), and Tintelnot (2017), among many others. Ramondo and Rodríguez-Clare (2013) and Arkolakis et al. (2018) use a probabilistic structure that is remarkably tractable. Each plant's technology is constant returns to scale, so decisions on how to serve each market are independent across markets. Introducing fixed costs of setting up plants or span-of-control costs would render the model intractable. In general, the firm's plant location problem can be split into two parts: an inner problem of establishing the set of locations serviced by each of a firm's plants, their "catchment areas," and an outer

[2] Setting up the model in full general equilibrium with many industries and labor mobility is straightforward but requires making assumptions on a number of economic fundamentals that are unrelated to our main focus.

problem of determining the location and number of plants. Tintelnot (2017) introduces fixed costs and solves the inner problem by assuming that firms sell a continuum of products and that plant productivities follow an extreme value distribution, which implies that plants sell to all locations. This smooths out the firm's objective function but still requires solving the combinatorial problem of where to set up plants and how many. All frameworks in this literature either solve the combinatorial problem with only a few countries or assume away fixed costs of setting up new plants. Methodologically, our main contribution is to solve the outer problem and characterize it analytically in a limiting case where all costs remain relevant. In contrast to the multinational literature, which has focused mostly on manufacturing, the particular limiting economy we study is likely to be a better approximation of industries with high transport costs, where the number of plants per firm is large. Our main substantive contribution is to incorporate span-of-control costs into the firm's problem and to characterize the resulting sorting patterns. Most models in the literature have the feature that the less profitable markets are reached only by the more productive firms. In contrast, our environment is one in which it is the less productive firms that locate in the more marginal markets.

The industrial organization literature has also analyzed how individual firms set up distribution networks in space. Seminal papers include Jia (2008) and Holmes (2011). Importantly, many of these frameworks study cases where opening stores in one location increases the marginal value of opening stores in other locations, the so-called supermodular cases.[3] The lack of cannibalization across plants makes these cases somewhat easier to handle. On the contrary, cases where cannibalization is prevalent, so that setting up new plants reduces the value of other plants, cannot easily be solved except for algorithms that, at worst, evaluate all possible combinations. Recently, Hu and Shi (2019) and Arkolakis, Eckert, and Shi (2023) have developed algorithms to solve these types of "submodular" problems more efficiently by iteratively pruning the choice set, but doing so for large numbers of locations remains a challenge. Furthermore, the purely numerical nature of essentially all this literature implies that few general insights have been obtained. Our analytical approach has the advantage of providing a set of general implications that we can contrast with micro data.

There is a large, active literature in operations research studying the facility location problem. The classic Weber problem of placing a single plant to serve many destinations at minimal cost (Weber 1909) was generalized to study the placement of multiple plants by Stollsteimer (1961) and Balinski (1965). There are many versions of the problem. One approach, used by much of the economics literature, studies the problem with a finite

[3] Holmes (2011) assumes submodularity but does not solve the model; he estimates parameters using moment inequalities.

set of discrete locations. The alternative, which we follow, models space as continuous. Another distinction is whether or not there are limits on each plant's output (these are known as the "capacitated" or "uncapacitated" facility location problems). Given the complexity of the problem, the literature has focused on numerical algorithms that deliver approximate solutions in polynomial time.[4]

We provide a characterization of the matching of heterogeneous firms with multiple plants to heterogeneous locations. Nocke (2006), Gaubert (2018), and Ziv (2019) study the assignment of single-plant firms to heterogeneous locations. Behrens, Duranton, and Robert-Nicoud (2014), Eeckhout, Pinheiro, and Schmidheiny (2014), Diamond (2016), Davis and Dingel (2019), and Bilal and Rossi-Hansberg (2021) study the assignment of workers to heterogeneous locations. None of these papers, however, address sorting when the agent, in our case a firm, can choose many locations concurrently.[5]

The rest of the paper is organized as follows. Section II presents the problem of the firm, proposes and studies the limit problem, and derives our main results. Section III embeds heterogeneous firms solving the production problem with multiple plants into an industry equilibrium. It also presents numerical examples that illustrate the effect of changes in the efficiency of span-of-control and transport costs. Section IV contrasts some of the main implications of our solution with panel data of firms and establishments. Section V concludes. An appendix, available online, includes all technical derivations, presents additional robustness results and data constructions details, and describes the numerical algorithm.

II. The Multiplant Firm Problem

We consider the problem of a firm deciding how to serve customers located in a unit square, $\mathcal{S} = [0, 1]^d \subset \mathbb{R}^d$, where $d \in \{1, 2\}$ is the dimension of the space.[6] Each location $s \in \mathcal{S}$ is characterized by an exogenous

[4] These problems have been shown to be NP-hard in both one and two dimensions (Fowler, Paterson, and Tanimoto 1981). It has been shown that, unless $P = \text{NP}$, there is a bound on the performance of such algorithms: they cannot guarantee a solution that is better than 1.463 times the actual minimal cost (Guha and Khuller 1999; Sviridenko 2002). Algorithms that deliver performance close to this bound have been recently proposed by Byrka and Aardal (2010) and Li (2013).

[5] Empirically, assessing sorting patterns when each plant is a stand-alone unit is difficult because of the reflection problem. In particular, one can observe whether plants in denser locations are larger, but it is not clear whether that is due to sorting or to the impact of being in a dense location. In our setting, with firms that operate many units in different locations, we can exploit leave-out strategies to argue that there is clear evidence of positive assortative matching.

[6] Our results can be easily generalized to a Euclidean space $\mathcal{S}$ that is closed, bounded, and Jordan measurable. While it seems intuitive that our results could be extended beyond two-dimensional space, doing so would require a resolution of the Gersho conjecture (Gersho 1979), which remains open for three or more dimensions.

productivity level B_s, as well as local equilibrium characteristics that firms take as given, namely, the residual demand function, $D_s(\cdot)$, the wage rate, W_s, and the commercial rent, R_s.

There is a set of firms, $j \in J$. Each firm produces a unique variety. A firm is characterized by its productivity, q_j. It chooses a finite set of locations $O_j \subset \mathcal{S}$ in which to set up plants. Conditional on operating a plant at location o, production requires only local labor, which is employed at wage W_o. A plant's productivity is the product of a local component, B_o, and a firm component, $Z(q_j, N_j)$. The firm component is increasing in a standard idiosyncratic productivity level q_j and decreasing in the firm's total number of plants $N_j = |O_j|$. The latter captures the productivity cost of increasing the firm's span of control. We also assume that $Z(q, 0) < \infty$. In sum, if a firm operates a total of N_j plants, its productivity in location $o \in O_j$ is $B_o Z(q_j, N_j)$. Each plant takes up ξ units of commercial real estate, with rental cost R_s per unit of space. Trade between any two locations incurs an iceberg shipping cost. For one unit of a good to arrive at distance δ, $T(\delta) \geq 1$ units must be shipped. We normalize $T(0) = 1$ and assume that $T(\delta)$ is strictly increasing, satisfies the triangle inequality, and diverges as $\delta \to \infty$.

We posit a market structure in which firms transport goods to households and choose a separate price for its good at each destination.[7] Firms will serve customers in the least costly possible way. Thus, the cost of delivering one unit of good j from a plant in o to a consumer in s is $W_o T(\delta_{so})/B_o Z(q_j, N_j)$, where δ_{so} denotes the distance between s and o. Let $\Lambda_{js}(O_j) \equiv \min_{o \in O_j}(W_o T(\delta_{so})/B_o Z(q_j, N_j))$ be j's minimal cost of delivering one unit of good j to a consumer in location s. Let p_{js} be the price charged by j to consumers in s. Then, if $D_s(p_{js})$ is the residual demand for variety j in location s, the optimal price maximizes

$$\max_{p_{js}} D_s(p_{js})(p_{js} - \Lambda_{js}).$$

The problem above can lead to pricing rules where markups depend on local characteristics. To simplify the problem, we abstract from spatial variation in markups and assume the following about the residual demand function.

ASSUMPTION 1. Residual demand satisfies $D_s(p_{js}) = D_s p_{js}^{-\varepsilon}$, where D_s subsumes all determinants of local demand, including the local price index.

[7] An alternative market structure—one in which households incur shipping costs and choose which plant to purchase from and each firm chooses a separate price at each of its plants—is equally natural. If households have isoelastic residual demand, as we assume below, the two market structures yield the same revenue and employment for each plant and the same consumption and expenditures for each consumer. We focus on the market structure in which firms incur shipping costs because it is simpler to state and work with.

Assumption 1 is satisfied in the standard case with monopolistic competition and CES (constant elasticity of substitution) preferences with elasticity of substitution across varieties given by ε. Then, as usual, $p_{js} = [\varepsilon/(\varepsilon - 1)]\Lambda_{js}$.

Firm j's profit can be expressed as

$$\pi_j = \max_{O_j}\left\{\int_s \max_{p_{js}} D_s p_{js}^{-\varepsilon}\left(p_{js} - \Lambda_{js}(O_j)\right) ds - \sum_{o\in O_j} R_o\xi\right\}, \tag{1}$$

or, using the expression for j's price,

$$\pi_j = \max_{O_j}\left\{Z(q_j, N_j)^{\varepsilon-1}\int_s D_s \max_{o\in O_j}\left\{b_o T(\delta_{so})^{1-\varepsilon}\right\} ds - \sum_{o\in O_j} R_o\xi\right\}, \tag{2}$$

where $b_o \equiv [(\varepsilon - 1)^{\varepsilon-1}/\varepsilon^{\varepsilon}](B_o/W_o)^{\varepsilon-1}$ summarizes local productivity and wages.

A. *The Catchment Area of a Plant*

The catchment area of a plant in location o is formed by locations s to which the firm sells goods from the plant in o. Formally, the catchment area of a plant in location o is

$$\left\{s \in \mathcal{S} \text{ for which } o = \arg\max_{\tilde{o}\in O_j}\left\{b_{\tilde{o}} T(\delta_{s\tilde{o}})^{1-\varepsilon}\right\}\right\}. \tag{3}$$

A plant's catchment area can be empty if its cost, relative to other nearby plants, is high enough.

Note that, once plants are placed in locations O_j, the catchment area of each plant depends only on transportation costs and on the production cost of locations where plants are placed. When space is one-dimensional, $d = 1$, and transport costs rise sufficiently fast with distance, catchment areas are simply a set of nonoverlapping intervals covering $\mathcal{S}$, such that the cost of servicing costumers at the boundary is identical for both adjacent plants. Therefore, the size of the catchment area of a particular plant is decreasing in the plant's cost and increasing in the cost of adjacent plants. When transport costs do not rise fast with distance, the catchment area of a plant can be the union of disjoint segments. With two-dimensional space, catchment areas can be substantially more complicated. In any of these cases, it is straightforward, although computationally costly, to solve numerically for catchment areas.[8] Nevertheless, we show that when we

[8] This problem is equivalent to constructing a weighted Voronoi diagram. Tintelnot (2017) approaches this problem by assuming that plants sell a continuum of goods produced with productivities drawn from a distribution with infinite support. The implication is that the catchment areas of all plants overlap and cover the whole space.

incorporate the decision of where to place plants, the optimal choice leads to a structure of catchment areas that greatly simplifies the problem. In particular, in the limiting case that we study, local catchment areas are always characterized by intervals ($d = 1$) or hexagons ($d = 2$).

Examples in one and two dimensions.—This section illustrates the role of the production cost in determining catchment areas. We first focus on the simpler case of $d = 1$ with $\mathcal{S} = [0, 1]$, so that catchment areas partition the unit interval. We show that nonconvex catchment areas can arise when transport costs do not rise quickly with distance. We then turn to an environment with two dimensions where $\mathcal{S}$ is the unit square. We show examples of how changes in fundamentals across locations affect catchment areas. For each exercise, we set $\varepsilon = 2$ and solve two cases with different distributions for b_o across space.[9]

Figure 3 presents the one-dimensional case. We consider the problem of a firm that (perhaps suboptimally) places five plants at regular intervals across space.[10] We assume that transport costs are given by $T(\delta_{so}) = 1 + 0.75\delta_{so}$, where δ_{so} is the Euclidean distance between s and o. The left-hand panel presents the resulting catchment areas when production costs are the same across all locations—that is, $b_o = 1 \; \forall \; o \in O_j$—while the right-hand panel shows the resulting catchment areas when the third location is 14.5% more productive than its counterpart in the left-hand panel. In both cases, catchment areas are characterized by a collection of segments. When all plants face the same cost, catchment areas are equally sized line segments. When one plant, in this example plant 3, faces a lower cost, catchment areas vary in size and need not be convex. More productive locations have larger catchment areas, as evidenced by the spatial expansion of plant 3's catchment area. Note that plant 3's catchment area depends on its own cost, its location, and the cost of its neighbors—plants 2 and 4—as well as the location and cost of its neighbors' neighbors—plants 1 and 5. In other words, designing the catchment area of plant 3 requires understanding this plant's interactions with all other plants.[11] Note also that the nonconvexity in the catchment area of plant 3 is possible because transport costs do not rise quickly with distance.

Similar logic applies for the two-dimensional case. Figure 4 presents catchment areas when nine plants are arbitrarily placed in a regular grid.[12] For these exercises we set transportation costs equal to $T(\delta_{so}) = 1 + \delta_{so}$. As in the one-dimensional case, we assume that there is no variation in

[9] The choice $\varepsilon = 2$ is not essential to generate these examples.

[10] Plants locations are $o \in \{1/10, 3/10, 5/10, 7/10, 9/10\}$.

[11] One can show that if production costs are similar across space and if trade costs rise sufficiently fast with distance, a plant's catchment area depends only on its cost and that of its direct neighbors.

[12] In the figure, plants are located at (1/6, 1/6), (1/6, 1/2), (1/6, 5/6), (1/2, 1/6), (1/2, 1/2), (1/2, 5/6), (5/6, 1/6), (5/6, 1/2), and (5/6, 5/6).

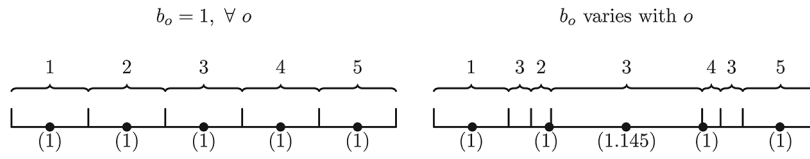

FIG. 3.—Catchment areas in one dimension. The figure presents the catchment areas, as defined in equation (3), for the one-dimensional case where five plants are located in a line. The left-hand plot presents the case where all locations are equally productive, while the right-hand plot presents a case where plant 3's location is more productive. Each dot in each plot corresponds to the location of a plant. The number in parentheses under each dot corresponds to the value of b_o for that dot. The number above each brace indicates which plant serves that location, i.e. the catchment area of a particular plant.

economic fundamentals, so $b_o = 1\ \forall\ o \in O_j$, in the left-hand panel. The right-hand panel presents the results when we increase the costs of the location in the top-left corner by setting productivity to $b_o = 0.85$, and we reduce the costs of the central and bottom-right corner locations by letting $b_o = 1.2$.

As figure 4 shows, when production costs are constant across production locations, catchment areas are all equally sized, are all convex, and are all polygons. However, when production costs vary, the catchment areas can take different shapes and sizes, can be nonconvex, and are not polygons.

While the one-dimensional case is simpler in the sense that catchment areas are, at most, a combination of disjoint intervals, both the one- and two-dimensional cases share the same basic features and complexities: local characteristics affect the size of catchment areas, and nonconvex catchment areas can arise as a combination of transport cost varying slowly with

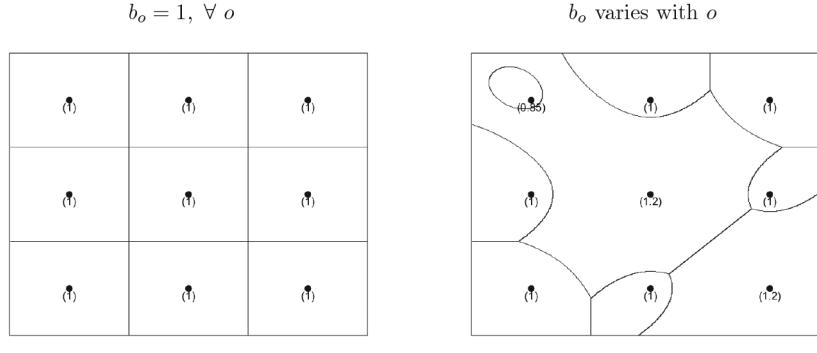

FIG. 4.—Catchment areas in two dimensions. The figure present the catchment areas, as defined in equation (3), for the case where nine plants are located in the square space. Each dot in each plot corresponds to the location of a plant. The number in parentheses next to each dot corresponds to the value for b_o for that dot.

distance and specifics about the productivity of a location, the productivity and location of its neighbors, the productivity and location of the neighbors' neighbors, and so on.

The difficulty of these problems may seem daunting, particularly once we introduce more locations and richer heterogeneity and especially when we incorporate the outer problem of how many plants to use and where to place them. Perhaps surprisingly, in the limit economy we study below, solving this outer problem of optimal plant locations imposes structure that leads to a simple characterization of the inner problem of solving for catchment areas in both one and two dimensions. The only case where we can characterize the solution to the full problem without relying on the limit economy is the one-dimensional case with uniform locations. We turn to that problem first.

B. *A Simple Special Case in One Dimension*

We now turn to the outer maximization problem of determining how many plants to set up and where to do so. We start by considering one special case for which the solution to this outer problem is straightforward. If space is one-dimensional and economic activity is uniform across locations, the optimal configuration is to have plants equally spaced, so that catchment areas are equal in length. In this case, the firm's profit function can be expressed as

$$\pi_j = \sup_N xZ\left(q_j, N\right)^{\varepsilon-1} NG(1/N) - RN,$$

where $x \equiv [(\varepsilon - 1)^{\varepsilon-1}/\varepsilon^{\varepsilon}]D(B/W)^{\varepsilon-1}$ and $G(u) = \int_{-u/2}^{u/2} T(|\delta|)^{1-\varepsilon}\, d\delta$. The variable x combines the demand, D, facing the plant at each location with the cost of effective labor, W/B, into a measure of local profitability, and $G(u)$ is the integral of the function $T(\cdot)^{1-\varepsilon}$ over all distances between the origin and points of a line segment of length u centered at the origin.

We refer to the function $\kappa(N) \equiv NG(1/N)$ as the *efficiency of distribution*. It represents the fraction of the value of sales a firm retains after subtracting the cost of optimally transporting the goods to consumers from its N plants. Figure 5 provides a graphical representation of the function $G(\cdot)$ and the implied function $\kappa(\cdot)$. The figure presents an example with four plants, $N = 4$, with catchment areas of length $1/N$. It plots the function $T(\delta)^{1-\varepsilon}$, where δ is the distance from the plant to the customer's location. For customers at the plant's location, $\delta = 0$, there is no loss from transport costs. For customers farther away, profits are reduced by a factor of $T(\delta)^{1-\varepsilon}$. The shaded area is $G(1/N)$, the fraction of profit the firm gets from a plant's catchment area relative to that in a world with no transport costs. The efficiency of distribution, $\kappa(N) = NG(1/N)$, is the fraction of profit the firm gets per unit of space, relative to that in a world with no transport costs.

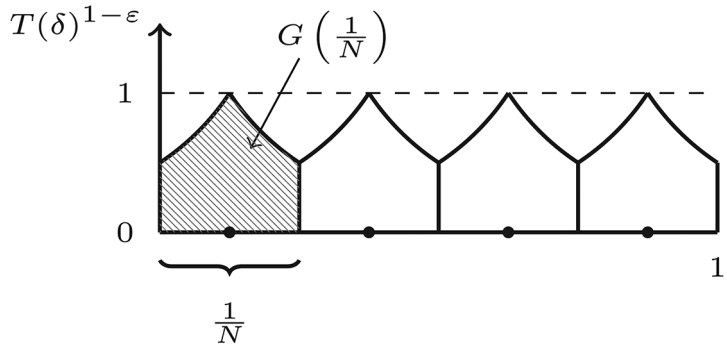

FIG. 5.—One-dimensional representation of the efficiency of distribution, $\kappa(N) \equiv NG(1/N)$.

In two dimensions, there is no closed-form solution for the exact placement of plants, even when economic activity is uniform across space. Nevertheless, there is a known upper bound for the profit a firm can attain, also based on the strategy of placing plants in a regular pattern across space. We discuss this strategy further in section II.D.

As discussed above, we aim to venture beyond these special cases in which economic activity is uniform, so that we can discuss how a firm's local footprint varies with local economic conditions or how firms sort across space. Thus, we now propose a reformulation of this problem that can be tractably studied, while still preserving its main features and trade-offs.

C. *A Tractable Limit*

We propose a tractable limit of the firm's problem in which the number of plants per firm grows large so that the firm is essentially choosing a density of plants, rather than a discrete number. In particular, we study a limit in which the space that plants take up grows small, trade costs grow large, and the productivity cost for having many plants grows small. We take limits at carefully chosen rates so the problem is well behaved in the limit. Specifically, for some $\Delta > 0$, let

$$\xi^{\Delta} = \Delta^{d},$$

$$T^{\Delta}(\delta) = t\left(\frac{\delta}{\Delta}\right), \quad \text{and}$$

$$Z^{\Delta}(q, N) = z(q, \Delta^{d} N).$$

for $d = 1, 2$. We study the limit as $\Delta \to 0$.

We want to study a limit in which the key trade-offs between the fixed and managerial span-of-control costs of setting up plants and the cost

of reaching consumers remain relevant, a limit in which plants continue to potentially cannibalize each other's customers. Thus, as Δ declines and the cost of adding plants falls, we increase transport costs. In two-dimensional space, $d = 2$, ξ^{Δ} and Z^{Δ} depend on the square of Δ, since space is two-dimensional, while, in contrast, transport costs are a function of distance, which is one-dimensional.[13] The following proposition describes the profits of the firm in this limit.

PROPOSITION 1. Suppose that R_s, D_s, and B_s/W_s are continuous functions of s. Then, in the limit as $\Delta \to 0$, the profits of firm j satisfy

$$\pi_j = \sup_{n:\mathcal{S}\to\mathbb{R}^+} \int_s \left[x_s z\left(q_j, \int n_{\tilde{s}} d\tilde{s}\right)^{\varepsilon-1} n_s g(1/n_s) - R_s n_s \right] ds,$$

where $x_s \equiv [(\varepsilon-1)^{\varepsilon-1}/\varepsilon^{\varepsilon}] D_s (B_s/W_s)^{\varepsilon-1}$. In one dimension, $g(u)$ is the integral of the function $t(\cdot)^{1-\varepsilon}$ over all distances between the origin and points of a line segment of length u centered at the origin. In two dimensions, $g(u)$ is the integral of the function $t(\cdot)^{1-\varepsilon}$ over all distances between the origin and points of a regular hexagon of area u centered at the origin.[14]

Proposition 1 shows that in the limit, the firm's problem is one of calculus of variations, which, as we show below, is much easier to analyze. As before, the variable x_s combines both local demand facing the plant, D_s, and local cost of effective labor, W_s/B_s, into a measure of local profitability. Hence, in the limit, a location's characteristics can be fully summarized by two variables: local rent, R_s, and local profitability, x_s. In contrast, before taking the limit, the relevant features of a location were infinite-dimensional, comprised of the local effective wage and demand in surrounding locations. Note also that, in the limit, aside from the span-of-control considerations, the problem is completely separable across locations. An important implication of this last result is that the problems in one and two dimensions

[13] There is a natural analogy to the continuous time limit of discrete-time portfolio choice problems. In those models, as the length of a period shrinks to zero, the amount of risk must grow without bound, so that there is a meaningful amount of risk to compare across assets. The key, as in our setup, is that the speed at which risk grows is the same as the speed at which the period shrinks. Thus, the relative importance of risk per period unit remains constant. In the appropriate limit, the value of assets follows a Brownian motion.

[14] In one dimension, $g(u) = \int_{-u/2}^{u/2} t(|\delta|)^{1-\varepsilon}\, d\delta$. In two dimensions,

$$g(u) = \int_0^{\sqrt{3^{-3/2}2u}} t(\delta)^{1-\varepsilon} 2\pi\delta\varpi\left(\frac{\delta}{\sqrt{3^{-3/2}2u}}\right) d\delta,$$

where $\varpi(r)$ is the fraction of circle with radius r that intersects with the interior of a regular hexagon with side length 1. As we show in app. A.1.2, simple trigonometric arguments yield

$$\varpi(\delta) = \begin{cases} 1 & 0 \le \delta \le \sqrt{3}/2, \\ 1 - \dfrac{6}{\pi}\cos^{-1}\left(\dfrac{\sqrt{3}/2}{\delta}\right) & \sqrt{3}/2 \le \delta \le 1. \end{cases}$$

end up being incredibly similar. The only difference is that when solving the problem in one dimension, the function $g(u)$ requires integrating over a line segment, while in two dimensions it requires integrating over a hexagon.

Before presenting a sketch of the proof of proposition 1, we go back to the simpler case where all locations are identical that we analyzed in section II.B for one-dimensional space. We then proceed to sketch the proof of proposition 1 and characterize the solution to the profit maximization problem. Most formal proofs are relegated to the appendix unless explicitly stated.

D. Uniform Space

We begin by discussing the simpler case where all economic activity is uniform across locations. We analyzed the one-dimensional version of this problem in section II.B. Here we also discuss the case of two dimensions where there is no closed-form solution for the placement of plants. As before, assume that the local demand shifter for all locations is D, effective productivity is b, and commercial land rents are R. When space is homogeneous, there are some known results to the solution to the firm's problem of choosing where to locate its plants. In particular, if a firm places N plants, the firm's payoff will be no higher than $xz(q_j, N)^{\varepsilon-1}NG(|\mathcal{S}|/N) - RN$ where $|\mathcal{S}|$ is the length of $\mathcal{S}$ in one dimension and the area of $\mathcal{S}$ in two dimensions, and, as stated in proposition 1, $x \equiv Db$ and $G(u)$ is the integral of the transportation cost $T(\cdot)^{1-\varepsilon}$ over either a line segment of length u or a regular hexagon with area u centered at the origin, in each respective case. For the one-dimensional case, the optimality of catchment areas that are congruent line segments is trivial, as shown in the left-hand panel of figure 6: given all the symmetry built into the model, it is straightforward to set the catchment area of each plant to be $|\mathcal{S}|/N$, with plants placed at the center of each catchment area. In this case, as we argued in section II.B, $xz(q_j, N)^{\varepsilon-1}NG(|\mathcal{S}|/N) - RN$ is not only an upper bound but is also equal to the maximum payoff to a firm when it places N plants in a one-dimensional uniform space.

When space is two-dimensional, the result follows from the sum-of-moments theorem in Fejes Tóth (1953), one of the landmark results in discrete geometry.[15] That is, the nearly optimal policy is to have uniform catchment areas in the form of hexagons, with plants at the center of each hexagon. The right-hand panel of figure 6 shows an example of this solution. Why are hexagonal catchment areas optimal in two-dimensional

[15] While the appearance of hexagons as a result of the optimal configuration of economic activity in space is sometimes associated with Christaller (1933), the formal statement and proof are due to Fejes Tóth (1953).

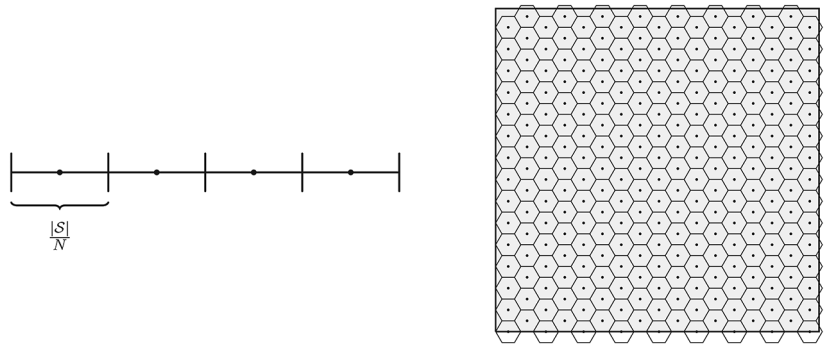

FIG. 6.—Filling out space. The left-hand panel shows a line of length $\mathcal{S}$ divided by $N = 4$ line segments of length $\mathcal{S}/N$. The right-hand panel shows a square of area $|\mathcal{S}|$ divided by N hexagons of area $|\mathcal{S}|/N$.

space? Jensen's inequality implies that it is optimal to have catchment areas of roughly the same area. Furthermore, optimality dictates that the shape of each catchment area should minimize the average distance from the center to the points in the catchment area. Among all shapes, a circle minimizes this average distance. However, one cannot form a tessellation with circles, as they would either overlap or leave empty spots. Among all polygons with which one can construct a uniform tessellation, the hexagon is closest to a circle. Note that this is an upper bound. As the right-hand panel in figure 6 shows, N disjoint uniform hexagons of size $|\mathcal{S}|/N$ generically do not fit exactly in the space $\mathcal{S}$.[16] It is straightforward to show that if N is large, that is, if $|\mathcal{S}|$ is large relative to the size of the catchment areas, then the boundary issue is quantitatively less relevant. In the "appropriate" limit, the upper bound is attained.

E. Heterogeneous Space

We are interested in understanding the location of a firm's plants in heterogeneous space. Section II.D provided important tools that we take advantage of for the general case with spatial heterogeneity. For the homogeneous-space case, we know how to construct the solution to the firm's problem for the $d = 1$ case for any number of plants N, and for the $d = 2$ case we know how to do it when the number of plants is "large." Interestingly, the limit that we explore allows us to apply both results. For the heterogeneous-space case, proposition 1 provides our key result. The proposition establishes that we can use a "large-N" limit to obtain a simple characterization

[16] Bollobás (1973) showed that the upper bound can be attained only if $\mathcal{S}$ is the union of N disjoint regular hexagons.

of the firm's optimization problem when space is heterogeneous. The key insight is that when economic activity is continuous over space, it is locally uniform. As a result, in the limiting economy, the solution for homogeneous space applies locally. The proposition states that, in the limit, the optimal policy is to place plants so that local catchment areas are uniform, infinitesimal intervals in one-dimensional space and hexagons in two-dimensional space. The variable n_s is the measure of plants in the neighborhood of s, so that $1/n_s$ is a measure of the size of the catchment areas.

Section II.D showed that, when economic activity is uniform, the solution for one-dimensional space is simpler than that for two dimensions. When economic activity is heterogeneous across space, the problem is considerably more complex for either one or two dimensions, as discussed in section II.A. The local logic that we exploit in this paper allows us to make substantive progress in both one- and two-dimensional problems with spatial heterogeneity. Once we know the solution for homogeneous space for the "large-N" limit, which we have from Fejes Tóth (1953), working with two-dimensional space is no more difficult than working in one dimension. For readability, we describe the proof when space is two-dimensional.

A sketch of the proof of proposition 1.—In economy Δ, firm j's profit is given by

$$\pi_j^{\Delta} = \max_{O_j}\left\{ Z^{\Delta}\left(q_j, N_j\right)^{\varepsilon-1}\int_s D_s \max_{o\in O_j}\left\{ b_o\, T^{\Delta}(\delta_{so})^{1-\varepsilon}\right\} ds - \sum_{o\in O_j} R_o \xi^{\Delta}\right\}.$$

The strategy is to create upper and lower bounds for firm j's profit. We start by dividing the space $\mathcal{S}$ into congruent squares with side length k, indexed by $i \in I^k$, denoted by $\mathcal{S}_i^k$ (for any k such that $1/k$ is an integer). For each Δ, k, we construct upper and lower bounds on firm j's profit, $\overline{\pi}_j^{k\Delta}$ and $\underline{\pi}_j^{k\Delta}$, such that

$$\underline{\pi}_j^{k\Delta} \le \pi_j^{\Delta} \le \overline{\pi}_j^{k\Delta}.$$

To construct the upper bound, we begin by considering a best-case scenario for each square by supposing that each square's highest demand, highest effective productivity, and lowest rent apply everywhere in the square. That is, for location s in square $\mathcal{S}_i^k$, we replace D_s, b_s, and R_s with $\bar{D}_s^k \equiv \sup_{\tilde{s}\in\mathcal{S}_i^k} D_{\tilde{s}}$, $\bar{b}_s^k \equiv \sup_{\tilde{s}\in\mathcal{S}_i^k} b_{\tilde{s}}$, $\underline{R}_s^k \equiv \inf_{\tilde{s}\in\mathcal{S}_i^k} R_{\tilde{s}}$. Similarly, to construct the lower bound, we consider a worst-case scenario for each square by supposing that each square's lowest demand, lowest effective productivity, and highest rent apply everywhere in the square.

In appendix G, we explore an example where we solve numerically for the upper and lower bounds $\overline{\pi}_j^{k\Delta}$ and $\underline{\pi}_j^{k\Delta}$. Intuitively, the example shows that the bounds get much tighter for small Δ's, when the chosen number of local plants is large.

We next use the sum-of-moments theorem separately for each square to give an upper bound on the profit the firm could attain from that square in that best-case scenario. As described in figure 6, if the firm chooses to place N_i plants in square $\mathcal{S}_i^k$, the upper bound corresponds to assigning to each of those plants a catchment area that is a regular hexagon with area $(k \times k)/N_i$.[17] To construct a lower bound for the worst-case scenario, we impose an ad hoc restriction on the firm's strategies so that all plants within $\mathcal{S}_i^k$ have catchment areas that are regular hexagons of the same size and are fully contained in the square $\mathcal{S}_i^k$; since regular hexagons do not form a tesselation of a square, not all customers in $\mathcal{S}_i^k$ are served by the firm in this suboptimal policy.

The second step is to fix an arbitrary k and study the limit as $\Delta \to 0$. Define the function $\kappa(n_s) \equiv n_s g(1/n_s)$. We prove that

$$\lim_{\Delta \to 0} \bar{\pi}_j^{k\Delta} \leq \bar{\pi}_j^k \equiv \sup_{\{n_s \geq 0\}} \int \left\{ \bar{D}_s^k \bar{b}_s^k z\left(q_j, \int n_{\tilde{s}}\, d\tilde{s}\right)^{\varepsilon-1} \kappa(n_s) - n_s \underline{R}_s^k \right\} ds$$

and that

$$\lim_{\Delta \to 0} \underline{\pi}_j^{k\Delta} \geq \underline{\pi}_j^k \equiv \sup_{\{n_s \geq 0\}} \int_s \left\{ \underline{D}_s^k \underline{b}_s^k z\left(q_j, \int n_{\tilde{s}}\, d\tilde{s}\right)^{\varepsilon-1} \kappa(n_s) - n_s \bar{R}_s^k \right\} ds.$$

If we define $\pi_j \equiv \lim_{\Delta \to 0} \pi_j^\Delta$ to be the firm's profit in the limiting economy, these, together with $\underline{\pi}_j^{k\Delta} \leq \pi_j^\Delta \leq \bar{\pi}_j^{k\Delta}$, imply that

$$\underline{\pi}_j^k \leq \pi_j \leq \bar{\pi}_j^k.$$

We obtain these results because, for any k, the economic features are uniform within each square $\mathcal{S}_i^k$ in both the best- and worst-case scenarios, so we can use results from discrete geometry to derive relatively simple expressions for the bounds. For the upper bound, the results imply that the catchment areas within each square $\mathcal{S}_i^k$ form a mesh with uniform regular hexagons. For the lower bound, the ad hoc restriction imposes that the catchment areas form a mesh with uniform regular hexagons.

The final step is to show that

$$\lim_{k \to 0} \bar{\pi}_j^k = \lim_{k \to 0} \underline{\pi}_j^k = \sup_{n_s \geq 0} \int_s \left\{ D_s b_s z\left(q_j, \int n_{\tilde{s}}\, d\tilde{s}\right)^{\varepsilon-1} \kappa(n_s) - R_s n_s \right\} ds,$$

[17] In constructing this upper bound, we compute, separately for each square $\mathcal{S}_i^k$, the profits the firm would earn from the plants in that square if it had no other plants in all of $\mathcal{S}$. We then add together the profits for all $k \times k$ squares. This double-counting of customers is fine because we are constructing an upper bound for profits, not the value of profit for any feasible policy. In any case, as $\Delta \to 0$, sales outside a plant's $k \times k$ square go to zero.

that is, that the limit can be interchanged with the supremum. Since $x_s = D_s b_s$, this proves proposition 1. The rest of the technical details of the proof are relegated to the appendix.

F. *Convergence of the Policy Function*

Proposition 1 established that the firm's profit function converges to a well-behaved limit. Here we establish further that the policy function, the choice of the measure of plants by location, converges as well.

PROPOSITION 2. Suppose that the problem in the limiting economy has a unique solution, n^*. For each Δ, let $O^{\Delta *}$ be a solution to the problem in economy Δ. Then for any ϵ, there is a $\bar{\Delta}$ small enough so that for any $\Delta < \bar{\Delta}$ and any Jordan-measurable set $\mathcal{A} \subseteq \mathcal{S}$,

$$\left| \Delta^{\mathrm{d}} \left| O^{\Delta *} \cap \mathcal{A} \right| - \int_{s \in \mathcal{A}} n_s^* \, ds \right| < \epsilon.$$

The proposition describes the precise sense in which the policy function converges. Consider any Jordan-measurable set of locations, $\mathcal{A}$.[18] As $\Delta \to 0$, the number of plants a firm places in a set of locations naturally rises, as rental costs fall and trade costs rise. Nevertheless, appropriately scaled by Δ^d, the number of plants placed in the set converges to a well-behaved limit that corresponds to the solution in the limit economy.[19] Hence, the proposition suggests a natural way to approach data on plant locations. Namely, rather than asking whether a firm placed a plant in a particular location, the proposition suggests looking at the number of plants a firm places in a contiguous area (e.g., a 12 × 12-mile square).

A sketch of the proof of proposition 2.—We show uniform convergence of the policy function in two steps. First, we derive properties of a firm's limiting problem. We show that if the limiting problem has a unique solution, n^*, then for any $\varepsilon > 0$ there exists an $\eta > 0$ such that $n \in \mathcal{N}$ and

[18] Jordan-measurable sets are, loosely, those that are well approximated by finite unions of rectangles. These include all bounded convex sets but not all Lebesgue-measurable sets. A set is Jordan measurable if and only if its indicator function is Riemann integrable. Why is the theorem restricted to Jordan-measurable sets? The set $O^{\Delta *}$ is finite for any $\Delta > 0$ and thus always has Lebesgue measure zero. It is hard to rule out the possibility that a single set $\mathcal{A}$ with Lebesgue measure zero (e.g., points in the unit square with rational coordinates) contains $O^{\Delta *}$ for all Δ, in which case the Lebesgue integral $\int_{s \in \mathcal{A}} n_s^* \, ds$ would equal zero.

[19] We do not know whether there is a unique optimal policy function for economy Δ (and we have no way of checking). However, the theorem applies to any optimal policy functions $O^{\Delta *}$. It is easier to assess uniqueness for the limiting economy. We can divide the problem of the limiting economy into a one-dimensional problem of choosing the total measure of plants and a subproblem of choosing a spatial allocation of those plants. It is straightforward to show that conditional on N, there is a unique solution to the subproblem of placing the plants in space (up to sets of measure zero). We do not provide sufficient conditions to ensure uniqueness of the outer problem, but the fact that it is one-dimensional means that it is easy to verify uniqueness numerically.

$|\Pi(n) - \Pi(n^*)| < \eta$ imply $\int_{s\in\mathcal{S}}|n_s - n_s^*|\,ds < \varepsilon$, where $\bar{\mathcal{N}}$ is a space of functions with a uniform bound and $\Pi(n)$ is the profit the firm would obtain from following strategy n.

In the second step, we study the sequence of economies as $\Delta \to 0$. As in the proof of convergence of the value function in appendix A.2, we construct a sequence of bounds on the profit function that get tighter as $\Delta \to 0$. We show that for economy Δ, the optimal choice $O^{\Delta*}$ has a corresponding strategy in the limiting economy, $n^{\Delta*}$. As $\Delta \to 0$, the bounds get tighter and two things happen. First, $O^{\Delta*}$ gets close to $n^{\Delta*}$: over any Jordan-measurable set $\mathcal{A}$, $\Delta^d|O^{\Delta*} \cap \mathcal{A}|$ uniformly approaches $\int_{s\in\mathcal{A}} n_s^{\Delta*}\,ds$. Second, the corresponding strategy $n^{\Delta*}$ delivers a value in the limit economy close to optimum. This, along with the first step, implies that $n^{\Delta*}$ converges to n^*. Namely, we have uniform convergence of the policy function to n^*.

In constructing the strategy $n^{\Delta*}$, we use segments of length k or $k \times k$ squares to find upper and lower bounds, as in the proof of proposition 1. In particular, for any Δ and k we can construct a strategy in the limiting economy that maximizes profit subject to the restriction that the measure of plants on each segment/square corresponds to the number of plants of $O^{\Delta*}$ in the segment/square multiplied by Δ^d. In the proof of proposition 1, the key step was to take the limit as $\Delta \to 0$ for a given k and then take $k \to 0$. Here, the key trick is to choose a sequence of $k = K(\Delta)$ so that, as we take the limit as $\Delta \to 0$, the sequence $k = K(\Delta)$ also converges to zero (albeit more slowly than does Δ). As a result, for each Δ, we construct the strategy $n^{\Delta*}$ in the limit economy that maximizes profits, subject to the restriction that the measure of plants on each $K(\Delta)$ segment, or $K(\Delta) \times K(\Delta)$ square, corresponds to the number of plants of $O^{\Delta*}$ in the segment/square multiplied by Δ^d.

G. *The Local Efficiency of Distribution and Its Properties*

As we discussed above, we refer to the function $\kappa(n_s) \equiv n_s g(1/n_s)$ as the local efficiency of distribution in the neighborhood of s. Recall that $\kappa(n_s)$ represents the fraction of the value of local sales a firm retains after subtracting the cost of optimally transporting the goods to consumers from its n_s plants. The following lemma describes some useful properties of κ.[20]

LEMMA 3. $\kappa(n) \equiv ng(1/n)$ is strictly increasing and strictly concave and satisfies the following properties:

[20] When space is one-dimensional, we can prove a converse of lemma 3. If κ is twice continuously differentiable, strictly increasing, strictly concave, and satisfies $\kappa(0) = 0$, $\kappa'(0) \in (0, \infty)$, $\kappa''(0) = 0$, and $\lim_{n\to\infty} n[1 - \kappa(n)] \in (0, \infty)$, then there is a strictly increasing transport cost $t(\delta)$ that generates κ, namely, $t(\delta) = [1 + \int_{1/2\delta}^{\infty} \kappa''(x)x\,dx]^{1/(1-\varepsilon)}$. For example, $\kappa(n) = \tan^{-1}(\phi n)/(\pi/2)$ for some constant $\phi > 0$ is consistent with the trade cost $t(\delta) = ([1/(\pi/2)]\{\tan^{-1}(1/2\delta\phi) - 1/[2\delta\phi + (2\delta\phi)^{-1}]\})^{1/(1-\varepsilon)}$.

1. $\kappa(0) = 0$,
2. $\lim_{n\to\infty}\kappa(n) = 1$, and
3. $1 - \kappa(n) \asymp_{n\to\infty} n^{-1/d}$.

If transport costs satisfy $\lim_{\delta\to\infty}\delta^{-2d/(\varepsilon-1)}t(\delta) = \infty$, then

4. $\kappa''(0) = 0$, and
5. $\kappa'(0) < \infty$.

The first property says that with no plants revenues are zero; $\kappa(n)$ is increasing, since more plants imply that customers are, on average, closer to a plant. It is concave, since additional plants cannibalize existing plants, leading to diminishing gains from reducing transport costs. The second property states that as n grows infinitely large, catchment areas grow small and $\kappa(n)$ approaches an upper bound of 1; in the limit, additional plants provide no significant gains, and the economy becomes "saturated." The third property states that $\kappa(n)$ follows an asymptotic power law as n grows large. If, asymptotically, trade costs increase sufficiently fast with distance, we can provide a sharper characterization of the efficiency of distribution when n is small. The fourth property states that when the number of plants is small, local profits increase linearly in the number of plants. Put together, cannibalization is irrelevant for the first set of plants but becomes the dominant force when the number of plants grows large. Finally, the fifth property says that there is no Inada condition at $n = 0$. Hence, there can be locations s in which the firm places no plants, $n_s = 0$.[21]

H. *The Assignment of Plants to Locations*

Proposition 1 can be used to characterize how firms place their plants. As before, we assume that the firm takes as given the distribution of commercial rents, R_s, and the distribution of local profitability, x_s. The problem of choosing how many plants to have, N_j, and their distribution in space, $n_j : \mathcal{S} \to \mathbb{R}^+$, can be stated as

$$\sup_{N_j, n_j : \mathcal{S}\to\mathbb{R}^+} \int_s \left[x_s z(q_j, N_j)^{\varepsilon-1}\kappa(n_{js}) - R_s n_{js}\right] ds,$$

subject to

$$\int_s n_{js}\, ds \;\le\; N_j.$$

[21] In one dimension ($d = 1$), $\kappa'(0) = 2\int_0^\infty t(\delta)^{1-\varepsilon}\, d\delta < \infty$. In two dimensions ($d = 2$), $\kappa'(0) = \int_0^\infty t(\delta)^{1-\varepsilon} 2\pi\delta\, d\delta < \infty$.

Letting λ_j be the multiplier on the constraint, the first-order condition with respect to n_{js} is given by

$$x_s z_j^{\varepsilon-1}\kappa'(n_{js}) \leq R_s + \lambda_j, \text{ with equality if } n_{js} > 0, \tag{4}$$

where we use z_j as shorthand for $z(q_j, N_j)$. The first-order condition with respect to N_j is

$$\lambda_j = -\frac{d[z(q_j, N_j)^{\varepsilon-1}]}{dN_j}\int_s x_s\kappa(n_{js})\,ds. \tag{5}$$

The productivity of the firm declines with its span of control, as measured by the total number of plants, N_j. Hence, λ_j can be interpreted as the marginal span-of-control cost for the firm. It amounts to an additional shadow fixed cost, on top of the explicit fixed cost R_s, of operating one more plant in location s.

Since κ is strictly concave, equation (4) implies that n_{js} is increasing in x_s and z_j and decreasing in R_s and λ_j; $R_s + \lambda_j$ comprise a plant's effective fixed cost. Naturally, higher effective fixed costs induce the firm to operate fewer plants in a location. The firm trades off this effective fixed cost against the gains from increasing the efficiency of distribution: more plants implies that, on average, customers will be closer to the plants. Larger x_s or z_j imply larger gains from a reduction in average distance, inducing the firm to operate more plants in the location.

To characterize the solution to this problem, it is useful to make the following assumption on the productivity function $z(q, N)$.

ASSUMPTION 2. $z(q, N) = q\Xi(N)$, where Ξ is a log-concave function.

Using the first-order conditions in equation (4), together with this assumption, we can show that firms with higher endogenous productivity have higher marginal span-of-control costs of increasing the number of plants, λ_j, even relative to their firm-specific profitability, $z_j^{\varepsilon-1}$.[22]

LEMMA 4. Consider two firms with $z_1 < z_2$. Then, either $\lambda_1/z_1^{\varepsilon-1} < \lambda_2/z_2^{\varepsilon-1}$ or $N_1 = N_2 = 0$.

Proof. Since κ is concave, the density of plants is a decreasing function of $(R_s + \lambda_j)/x_s z_j^{\varepsilon-1}$. Suppose that $\lambda_1/z_1^{\varepsilon-1} \geq \lambda_2/z_2^{\varepsilon-1}$. Then, in every market $(R_s + \lambda_1)/x_s z_1^{\varepsilon-1} > (R_s + \lambda_2)/x_s z_2^{\varepsilon-1}$. Therefore, $n_{1s} \leq n_{2s}$, with a strict inequality whenever $n_{2s} > 0$. If $N_2 > 0$, then $N_2 > N_1$, and the log concavity of z with respect to N, along with $\kappa' > 0$, implies that

$$\frac{\lambda_1}{z_1^{\varepsilon-1}} = (\varepsilon - 1)\frac{-z_N(q_1, N_1)}{z(q_1, N_1)}\int_s x_s\kappa(n_{1s})\,ds < (\varepsilon - 1)\frac{-z_N(q_2, N_2)}{z(q_2, N_2)}\int_s x_s\kappa(n_{2s})\,ds = \frac{\lambda_2}{z_2^{\varepsilon-1}},$$

a contradiction. If $N_2 = 0$, then $N_1 = 0$. QED

[22] Log concavity ensures that $-z_N(q, N)/-z(q, N)$ is nondecreasing.

Our next result uses lemma 4 to prove that more productive firms set up relatively more plants in locations with higher rents.

PROPOSITION 5. Consider two firms with $z_1 < z_2$. Let $R^*(z_1, z_2)$ be the unique rent that satisfies

$$\frac{R^*(z_1, z_2) + \lambda_2}{R^*(z_1, z_2) + \lambda_1} = \frac{z_2^{\varepsilon-1}}{z_1^{\varepsilon-1}}.$$

Then, $R_s > R^*(z_1, z_2)$ implies that $n_{2s} \geq n_{1s}$, with strict inequality if $n_{2s} > 0$; $R_s < R^*(z_1, z_2)$ implies that $n_{1s} \geq n_{2s}$, with strict inequality if $n_{1s} > 0$; and $R_s = R^*(z_1, z_2)$ implies that $n_{1s} = n_{2s}$.

Proof. To start, $z_2 > z_1$ implies that $\lambda_2 > \lambda_1$ and $\lambda_2/z_2^{\varepsilon-1} > \lambda_1/z_1^{\varepsilon-1}$. Therefore, $(R + \lambda_2)/(R + \lambda_1) > 1$, so $(z_1^{\varepsilon-1}/z_2^{\varepsilon-1})[(R + \lambda_2)/(R + \lambda_1)]$ is strictly decreasing in R. Since $\lim_{R\to 0}(z_1^{\varepsilon-1}/z_2^{\varepsilon-1})[(R + \lambda_2)/(R + \lambda_1)] = (\lambda_2/z_2^{\varepsilon-1})/(\lambda_1/z_1^{\varepsilon-1}) > 1$ and $\lim_{R\to\infty}(z_1^{\varepsilon-1}/z_2^{\varepsilon-1})[(R + \lambda_2)/(R + \lambda_1)] = z_1^{\varepsilon-1}/z_2^{\varepsilon-1} < 1$, there is a unique R^* such that $(z_1^{\varepsilon-1}/z_2^{\varepsilon-1})[(R + \lambda_2)/(R + \lambda_1)] = 1$. If $R_s > R^*(z_1, z_2)$ and $n_{2s}, n_{1s} > 0$, then $\kappa'(n_{2s}) = (R_s + \lambda_2)/z_2^{\varepsilon-1}x_s < (R_s + \lambda_1)/z_1^{\varepsilon-1}x_s = \kappa'(n_{1s})$, and since κ' is decreasing, $n_{2s} > n_{1s}$. If $n_{2s} > 0$ and $n_{1s} = 0$, then of course $n_{2s} > n_{1s}$. If $n_{2s} = 0$, then $\kappa'(0) \leq (R_s + \lambda_2)/z_2^{\varepsilon-1}x_s < (R_s + \lambda_1)/z_1^{\varepsilon-1}x_s$, which implies that it is optimal for $n_{1s} = 0$. The argument for $R_s < R^*(z_1, z_2)$ is identical. The argument for $R = R^*(z_1, z_2)$ is trivial. QED

Proposition 5 states that for two firms with different productivities, there is a cutoff level of rent such that the firm with higher productivity places more plants in locations with higher rent and the firm with lower productivity places more plants in locations with lower rent. Thus, even while the two firms have overlapping footprints, there is a clear pattern of sorting. Figure 7 provides a graphical representation of this result.

The type of sorting implied by proposition 5 stands in sharp contrast to workhorse models of trade and multinational production in which the more marginal locations are reached by the most productive firms.[23] Here, it is the less productive firms that go to the lower-rent locations. Why the difference?

Consider first locations with the cutoff level of rent where two firms with different productivities choose to place the same number of plants. Firms balance the marginal profit from an additional plant, $x_s z_j^{\varepsilon-1}\kappa'(n_{js})$, against the effective fixed cost of a new plant, $R_s + \lambda_j$, which depends on the local rent and the productivity penalty arising from the larger span of control of managers. The higher-productivity firm earns more profit per plant in the location but chooses not to open more plants because of its higher marginal span-of-control costs (as shown in lemma 4). At locations with higher rent, the higher rent deters the lower-productivity firm from placing many

[23] See, e.g., Melitz (2003), Eaton and Kortum (2002), and Ramondo and Rodríguez-Clare (2013).

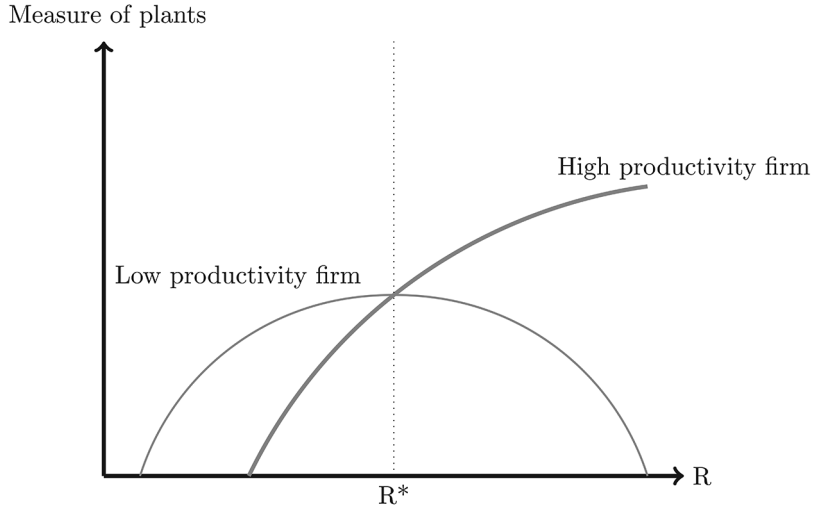

FIG. 7.—Location of plants of a high-productivity and a low-productivity firm.

plants, but it has a smaller impact on the higher-productivity firm's effective fixed costs. Hence, the higher-productivity firm places relatively more plants in these high-rent locations. Formally, since $d \ln(R_s + \lambda_j)/d \ln R_s$ is decreasing in λ_j, the effective fixed cost of setting up a plant rises proportionally less with rents for the high-productivity firm. A parallel argument implies that a lower rent induces the lower-productivity firm to place more plants, while the large marginal span-of-control cost of the high-productivity firm limits its presence.[24] Most models of plant location decisions in the literature do not feature span-of-control costs, and so this sorting implication is absent.[25]

The results above condition on firms with a positive density of plants in particular locations. Our next result shows that, for any given location, there is a productivity threshold such that firms with productivity below

[24] In the baseline model, a plant requires a fixed amount of space, and production uses only labor. Thus, a plant's fixed cost depends on the local rent, and its variable cost depends on the local wage. If both a plant's fixed cost and its production of output used (possibly different) bundles of labor and floor space, firms would still sort across locations. However, rather than sorting according to rent, they would sort according to the cost of the fixed-cost input bundle. That is, larger firms would place plants in locations in which the fixed-cost input bundle was more expensive.

[25] We can also show that the marginal efficiency of distribution of more productive firms is relatively smaller in higher-rent locations. Hence, in higher-rent locations, higher-productivity firms saturate the market relatively more, and the cannibalization between plants is larger. We relegate the formal statement of these additional results to app. A.4.

the threshold do not set up plants in that location. Under further restrictions on the span-of-control costs, there is another threshold such that firms with high enough productivity do not set up plants there either. That is, when all conditions are satisfied, only plants with productivities between these thresholds set up plants in a given location.

PROPOSITION 6. If $\lim_{\delta \to \infty}(\delta^d / t(\delta)^{\varepsilon-1}) = 0$, for any location s, there exists a productivity threshold $\underline{z}_s > 0$ such that $n_{js} = 0$ if $z_j < \underline{z}_s$. If $\lim_{z \to \infty}(\lambda_j / z^{\varepsilon-1}) = \infty$, then there exists an additional threshold $\bar{z}_s < \infty$ such that $n_{js} = 0$ if $z_j > \bar{z}_s$.

Our final result in this subsection refers to the total size of firms. The results above condition on a firm's productivity. However, empirically it is easier to condition on other firm observables, such as their total employment size or the total number of plants. We do not have a result that the total number of plants is increasing in firm productivity. Not only do firms sort their plants across locations, but their optimal plant size varies, depending on local characteristics. However, under particular parametric assumptions on a firm's productivity function, and if wages are constant across space, we can show that more productive firms employ more workers.[26] We let L_j denote the total number of workers of firm j.

LEMMA 7. Suppose that $z(q, N) = qe^{-N/\sigma}$ and that local wages are constant across locations at W. Consider two firms with $z_1 < z_2$; then, either $L_1 < L_2$ or $L_1 = L_2 = 0$.

III. Industry Equilibrium

We now proceed to embed the problem of the multiplant firm that we studied in the previous section into an industry equilibrium. We do so for a single "small" industry in the context of a full spatial equilibrium, of which we do not specify the details.[27] In particular, we are interested in how competition among firms interacts with the sorting of firms across space. After describing the industry equilibrium, we study two comparative statics: a relaxation of the span-of-control cost (perhaps driven by advances in information and communication technologies) and a reduction in transportation costs. We study the implications of these technological

[26] The assumption of equal local wages is consistent with the general equilibrium framework in sec. 3 of Oberfield et al. (2020).

[27] It is straightforward to embed the industry equilibrium into a full spatial equilibrium framework. This can be done in a number of ways. In one example, spelled out explicitly in Oberfield et al. (2020), each location is characterized by exogenous amenities in addition to productivity, people are freely mobile across locations, and land can be used for housing or commercial real estate. We can also accommodate further agglomeration and congestion forces or impediments to mobility. Of course, other alternative general equilibrium setups could work as well.

developments when the change occurs in that industry only. Hence, in the comparative-statics exercises, we keep rents and wages fixed, which implies that firms within an industry interact exclusively through the local industry price index. The exercises illustrate the relevance of transport costs and span-of-control costs in the limit economy and allow us to speak to the type of changes in sorting documented by Rossi-Hansberg, Sarte, and Trachter (2021) and Hsieh and Rossi-Hansberg (2022).[28]

There are $\mathcal{L}_s$ workers in location s with Cobb-Douglas preferences across industry aggregates from a unit continuum of industries indexed by $\omega \in [0, 1]$. Consistent with assumption 1, each industry aggregate is a Dixit-Stiglitz bundle of the varieties j produced by all firms in that industry, J_ω, with elasticity of substitution across varieties ε.[29] Thus, if $I_{s\omega}$ is the total expenditure on the industry aggregate for industry ω in locations s, the residual demand curve facing firm $j \in J_\omega$ is $I_{s\omega}P_{s\omega}^{\varepsilon-1}p_j^{-\varepsilon}$, where the price index for industry ω is $P_{s\omega} \equiv \left(\int_{j\in J_\omega} p_{js}^{1-\varepsilon}\, dj\right)^{1/(1-\varepsilon)}$. Aggregating across firms, we can characterize a location's industry price index.[30]

PROPOSITION 8. In the limit when $\Delta \to 0$, the local price index for industry ω is $P_{s\omega} = [\varepsilon/(\varepsilon - 1)](W_s/B_s\mathcal{Z}_{s\omega})$, where $\mathcal{Z}_{s\omega} \equiv \left(\int_{j\in J_\omega} z_j^{\varepsilon-1}\kappa(n_{js})dj\right)^{1/(\varepsilon-1)}$.

Note that industry productivity in a location, $\mathcal{Z}_{s\omega}$, is the CES aggregate of the firms' effective productivities, z_j, with the weight on a firm's productivity given by its local efficiency of distribution, $\kappa(n_{js})$. Then, the local price index is just the standard CES markup times the "aggregate" local marginal cost.[31]

A. *Numerical Illustration of an Industry Equilibrium*

To illustrate more concretely some of the equilibrium implications of our theory, we now specify all relevant functional forms and distributions and solve for an equilibrium of the model numerically. Our parameterization is intended to make the relevant forces visually clear and transparent.

Let transportation costs take the form $t(\delta; \phi) \equiv t(\delta/\sqrt{\phi})$, where ϕ indexes the efficiency of transportation (i.e., a higher ϕ implies lower trade

[28] We are well positioned to study this question relative to existing models of plant location that either have no span-of-control cost (Ramondo and Rodríguez-Clare 2013 and Tintelnot 2017) or limit a firm's location to a single plant (Gaubert 2018; Ziv 2019).

[29] For simplicity, we abstract from firm entry and hold the set of firms fixed in our comparative-statics exercises.

[30] In app. A.5, we determine other aggregate properties of the industry equilibrium when $\Delta \to 0$.

[31] We have not been able to show the existence or uniqueness of the type of equilibrium we desire. The main holdup is that we have not been able to show that the industry price index in the limiting economy is continuous. While we strongly suspect that this is the case—individual firms have incentives to place plants where other firms have not—we have no formal proof. We hope that future work can improve on this.

costs for a given distance traveled, δ). This implies that $\kappa(n; \phi) \equiv \kappa(\phi n)$.[32] We parameterize transportation costs as $t(\delta/\sqrt{\phi}) = e^{\delta/\sqrt{\phi}}$. We set $\phi = 0.04$. Firms' productivity is given by $z(q, N) = qe^{-N/\sigma}$, where σ indexes the efficiency of a firm's span of control (i.e., a higher σ implies a higher z for the same aggregate size of the firm, N). We set $\sigma = 1$. Finally, we posit a one-to-one mapping between locations' total expenditure on the industry I_s, and rent R_s; that these expenditures are distributed according to a truncated Pareto distribution; and that the distribution of firm productivities q_j is also given by a truncated Pareto distribution.[33]

We first describe the baseline industry equilibrium and then proceed to study comparative-static exercises with respect to σ and ϕ. Appendix F describes the numerical algorithm that solves the industry equilibrium. Figure 8 presents the distribution of plants, n_{js}, and sales, $(\varepsilon - 1)z_j^{\varepsilon-1}x_s\kappa(n_{js})$, for three representative firms: a firm with the lowest productivity, $q = 0.1$; a firm with intermediate productivity, $q = 1$; and a firm with the highest productivity, $q = 10$. As implied by proposition 5, for any pair of firms, there exists an income threshold (or, equivalently, a rent threshold, since rents are monotone in income) such that the more productive of the two firms sets up more plants above the threshold and fewer plants below. In our example, the most productive firm operates many plants in middle-income locations and fewer plants in very high-income or very low-income locations (the case of $q = 10$ in the left-hand panel of fig. 8). In fact, it operates no plants in the lowest-income locations. The logic should be clear; rents in high-income locations are high, which encourages high-productivity firms to economize on plants at the cost of having lower efficiency of distribution, $\kappa(n_{js})$. As shown in the right-hand panel of figure 8, they compensate with higher sales from each plant, which results in higher total sales. Low-income locations, in contrast, are less attractive to large firms, since their shadow cost of setting up an additional plant is high, given the productivity penalty that arises from their larger span of control (λ_j is increasing in q_j). Again, these firms compensate with higher sales from each plant. Firms with lower productivity then take advantage of low-rent locations, given their lower span of control and the lack of competition from top firms in those locations.

[32] Note that ϕ is defined so that it enters the function t as $\sqrt{\phi}$, while it enters the function κ linearly. The reason for the discrepancy is that the function κ is constructed from an integral over a two-dimensional space.

[33] Expenditure in location s, I_s, is distributed truncated Pareto, so that the measure of locations with income weakly less than I is $[1 - (I/\underline{I})^{-\chi_I}]/[1 - (\bar{I}/\underline{I})^{-\chi_I}]$, with $\underline{I} = 1$, $\bar{I} = 25$, and $\chi_I = 2$. We set the elasticity of substitution across varieties, ε, to 2. We assume that the distribution of fundamentals is such that the rent schedule in a location with income I_s is given by $R(I_s) = e^{\log(I_s)^2}$. There is a unit measure of firms, and the distribution of productivity is given by a truncated Pareto distribution so that the measure of firms with pure productivity no greater than q is $[1 - (q/\underline{q})^{-\chi_q}]/[1 - (\bar{q}/\underline{q})^{-\chi_q}]$, with $\underline{q} = 0.1$, $\bar{q} = 10$, and $\chi_q = 1.25$.

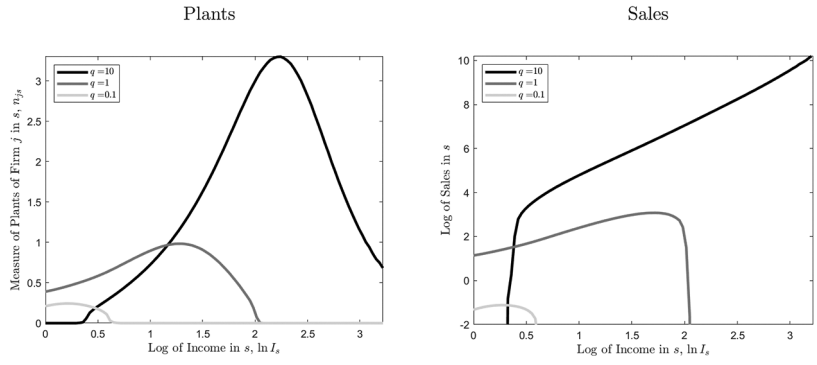

FIG. 8.—Sorting in industry equilibrium.

B. *Comparative Statics*

1. Improvements in an Industry's Span-of-Control Technology

Consider the effect of an improvement in the span-of-control technology captured by an increase in the parameter σ in the firm's productivity function, $z(q, N) = qe^{-N/\sigma}$. A better span-of-control technology increases firm productivity and lowers the shadow cost of adding new plants. This motivates firms to have more plants in more locations. In equilibrium, the additional entry leads to more local competition, through an increase in $\mathcal{Z}_s$ at all locations, which makes some firms shrink and others exit from some, or all, locations.

Figure 9 reproduces figure 8 (the solid lines computed for $\sigma = 1$) and compares it with findings for $\sigma = 3$ (dashed lines). In response to

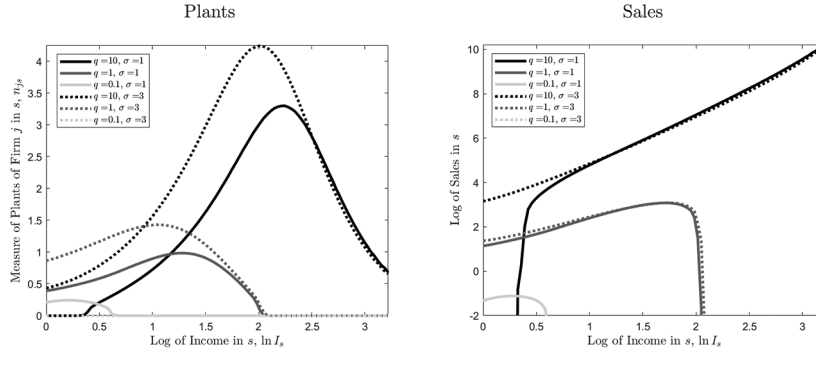

FIG. 9.—Span of control and sorting in an industry equilibrium.

the improvement in span-of-control technology, the top firm increases the measure of plants in low-income locations. It also reduces its presence slightly in the highest-income markets because of increased competition. The middle-productivity firm expands its presence in both lower- and higher-income locations. Holding fixed the actions of other firms, the lowest-productivity firm would benefit from the improved span-of-control technology as well. However, increased competition pushes it to exit all markets. The top firm not only enters lower-income markets but also, with improved span-of-control technology, ends up outselling the medium-productivity firm that already had a presence in those locations. The ability to manage a greater span of control, therefore, results in a net reallocation of sales from low- to high-productivity firms.

The left-hand panel in figure 10 shows how an improved span-of-control technology affects the shadow cost of additional plants, λ_j. As argued above, λ_j declines following the direct effect of the technological change. The effect is clearly magnified for high-productivity firms. These firms benefit most, since their better technology makes them want to expand more extensively in space and thus makes them benefit disproportionately from a technology that renders such an expansion less costly. The right-hand panel in figure 10 shows the effect of the span-of-control technology on local profitability, x_s. Increased competition lowers the local price index, particularly in low-income locations, which, in turn, lowers local profitability. These are the locations where top firms expand and where they now compete with lower-productivity firms. Although the total number of plants increases everywhere, low-income locations exhibit the largest increase in the number of plants.

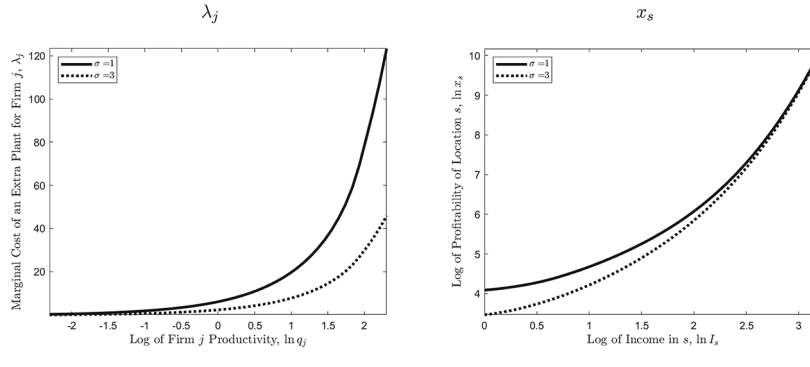

FIG. 10.—Effect of improvements in span-of-control technology on λ_j and x_s.

2. Improvements in an Industry's Transportation Costs

Consider the effects of an improvement in transportation technology captured by an increase in ϕ in the transportation cost function, $t(\delta) = e^{\delta/\sqrt{\phi}}$. An increase in transportation efficiency reduces the cost of reaching customers and so incentivizes firms to have fewer plants with larger catchment areas. Having fewer plants implies lower managerial cost associated with firms' span of control, which in turn increases productivity and induces firms to expand. The larger catchment areas effectively reduce the (fixed) rent costs of serving consumers in a location, which encourages the entry of all firms in more markets but particularly incentivizes the entry of less productive firms. Furthermore, lower transport costs imply more cannibalization among plants. This effect is particularly relevant for high-productivity firms, since they operate more plants. Hence, we expect improvements in transport efficiency to disproportionately benefit low-productivity firms. The incentives to enter more locations with fewer plants imply that competition at the local level increases everywhere, as reflected by an increase in $\mathcal{Z}_s$. This countervailing force reduces firms' sales in some locations.

Figure 11 shows the effect of an increase in ϕ from 0.04 to 0.4 on the mass of plants and sales of the representative high-, medium-, and low-productivity firms discussed above. The left-hand panel shows that all firms expand to new locations but also have fewer plants in most locations where they were already present. The top firm expands to low-income locations and now sells everywhere, while the medium- and low-productivity firms expand to higher-income locations. The increase in competition implies that profitability, x_s, declines almost uniformly across markets. As the right-hand side of figure 11 shows, increased competition implies that all

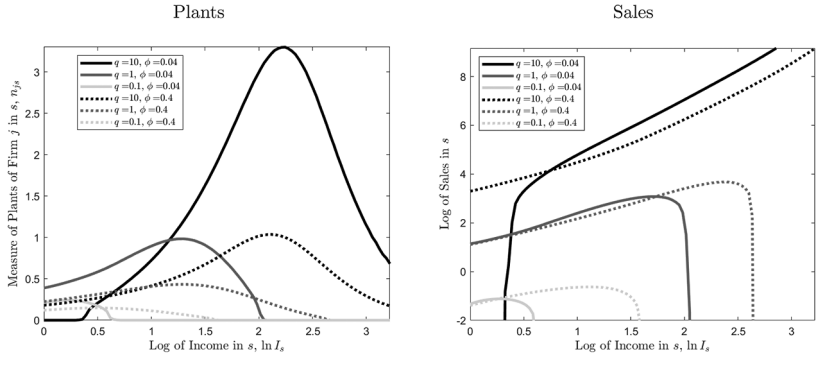

FIG. 11.—Transportation efficiency and sorting in an industry equilibrium.

three firms see their sales fall in many of the markets where they were operating.

While the improvement in transport technology leads firms to expand the range of locations in which they are active, they also have fewer plants. The effect on the total number of plants, therefore, is ambiguous. In this simulation, the total measure of plants falls in both low- and high-income markets. However, it increases in middle-income markets, as a large number of lower-productivity firms now choose to enter these markets. Overall, the improvement in transport costs favors low-productivity firms. Figure 12 shows that improvements in transportation technology lead total sales to increase for the lowest-productivity firms while total sales by the top firms decline.

IV. Empirical Evidence

Our theory provides a number of concrete implications about the location of plants in space for industries that can be approximated by our limit economy. In this section, we contrast its implications on sorting and the

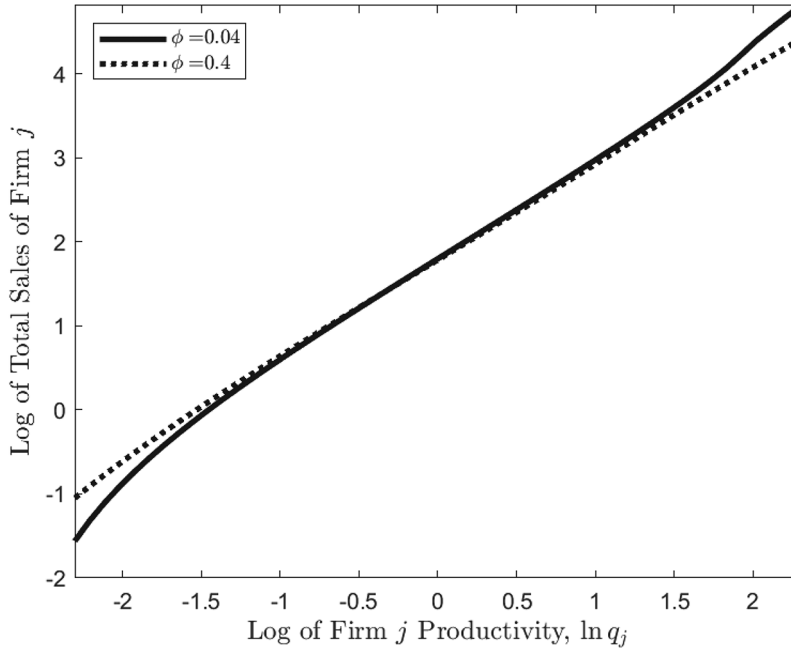

FIG. 12.—Effect of improvements in transportation efficiency on total firm sales.

role of span of control with US evidence. Our main source of data is NETS, which is provided by Walls & Associates. NETS provides yearly employment information for "lines of business," which we associate with plants in the theory and refer to as "plants" or "establishments" in the remainder of the paper.[34] For each establishment, we know its geographic coordinates, its industry classification, and its parent company.[35] We classify industries according to the SIC8 (8-digit standard industrial classification) industry classification, with over 18,000 distinct industries. We are interested in exploring how firms place their plants across space. To do so, we require a consistent definition of a "location." We follow Holmes and Lee (2010) and divide the continental United States into squares with side lengths of M miles. We present results for values of M ranging from 3 to 48 miles.

In order to contrast the theory's predictions with the data, we first need to map firm productivity and location characteristics to observable measures in the data. In lemma 7, we show that firm total national employment in a given industry is strictly increasing in its productivity. We can measure a firm's total employment directly in the data.[36] We can easily measure each location's population density in the data (since all locations are squares with the same area) and then use this metric to rank locations.

We have characterized the model's predictions for the limit economy. While the underlying forces we underscore—such as transport costs, span of control, and cannibalization—are likely relevant for any multiplant firm, the predictions of our theory are guaranteed to hold only in the limit economy. This limit is likely a better approximation for firms that set many plants across locations. Thus, when assessing the model's predictions, we

[34] The definition of a line of business is almost identical to the definition of an establishment or plant (which we use as synonyms). Although it is conceivable in principle that an establishment may contain one or more lines of business, in practice almost all plants have a single line of business. Thus, we refer to a line of business as a plant. For those cases where two lines of business are present in the same exact location, and thus in the same plant, each line of business is identified as a single plant.

[35] A more detailed description of NETS can be found in Rossi-Hansberg, Sarte, and Trachter (2021). We use only a cross section of NETS for 2014. Compared to census data, a cross section in NETS has an excess of very small firms, partly because it keeps track of nonemployee firms. Thus, we restrict our attention to firms with at least five employees. Crane and Decker (2019) observes that NETS has imputed employment data. Once we restrict to firms with at least five employees, the fraction of plants with nonimputed employment is 81.5%. In apps. C–E, we show that our empirical findings regarding sorting and span of control are robust to using only the nonimputed data.

[36] In our data, employment is better measured than revenue. Note also that, in our data, franchises are listed as separate firms. We hope that future research can explore how span-of-control considerations are affected or relaxed by contracting arrangements such as franchise agreements.

do so both using all industries and also restricting attention to industries in which plant catchment areas tend to be small, such as services.[37]

A. *Sorting in the Data*

A central and distinctive prediction of our model is that more productive firms sort toward "better" locations. Our main results related to sorting are presented in proposition 5. This proposition establishes that more productive firms set up relatively more plants in locations with higher land rents. We do not observe local rents in the NETS data. However, using alternative data sources, it is clear that there is a very tight positive relationship between rents and our ordering of locations using population density. Figure B.1, in appendix B (figs. B.1 and G.1 are available online), shows the relationship for zip codes and counties, using American Community Survey data to estimate densities and Zillow data to compute rents. Hence, in what follows, we use population density as a measure of the local characteristics on which firms sort.

As in the theory, let $\mathcal{L}_s$ denote population density in location s. The average density of the locations of a firm j, $\bar{L}_j$, is the average of the location density, $\mathcal{L}_s$, across all of the firm's plants. Once we compute $\bar{L}_j$, we subtract industry fixed effects. We use the residuals as our measurement of a firm's average density of its locations. The results below are robust to constructing $\bar{L}_j$ using a weighted average, where the weights are based on total employment in the plant's location.

Figure 13 shows that the relationship between $\ln \bar{L}_j$ and $\ln L_j$ is roughly linear when we restrict attention to firms with national size in an industry greater than 10 employees. Hence, in table 1 we estimate a linear relationship and show that, indeed, the relationship between a firm's log average location density and its log national employment size is positive and significant, after controlling for industry fixed effects. The table also presents a selected set of robustness checks. The implication of the theory holds robustly in the data: bigger firms sort toward dense locations. Column 2 of table 1 confirms the same finding when we look at a larger spatial resolution, $M = 48$, although the coefficient is smaller, probably because of spatial averaging across markets.

The implications we derived from the limit problem where transport costs are large and span-of-control and fixed costs are small should describe particularly well the behavior of firms that choose to set up many plants or that operate in industries in which trade costs are high. Column 3 of table 1 shows that the sorting pattern is indeed present and strongly significant when we limit the sample to firms with 100 or more plants, even

[37] In app. H, we show, for a numerical example, that the limit predictions derived from our model hold well for simulated plant locations when Δ is small but may fail when Δ is large and firms set up very few plants across space.

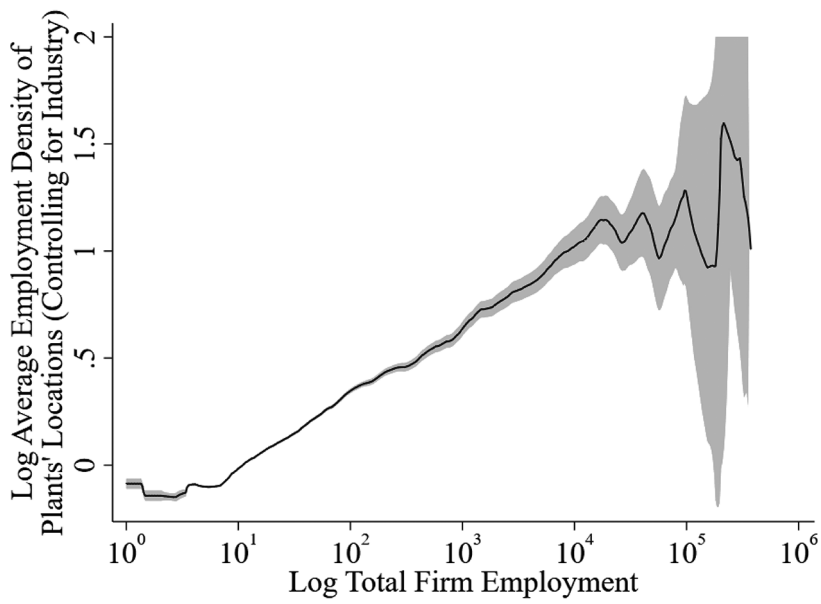

FIG. 13.—Sorting: firm size and local density. For each firm, we calculate the log of average employment density across all of the firm's plants (we use $M = 12$). Then, we subtract industry fixed effects and collect the residuals. Finally, we fit a kernel-weighted local regression of the residuals on the log of total firm employment. The regression shown in the figure uses a zero-degree polynomial (local-mean smoothing) and the bandwidth that minimizes the conditional weighted mean integrated squared error. The shaded area indicates the 95% confidence interval.

though it reduces the sample size tremendously. Column 4 restricts the sample to industries that have at least one firm with at least 100 plants and finds an even steeper positive relationship, which is again highly significant. Finally, the sorting pattern we have uncovered could arise from omitted characteristics of firms that are correlated with density. For example, if firms tend to set up plants where they are founded and denser locations incubate more productive firms, we would obtain a positive relationship between average firm density and total firm employment.[38] We address this concern by examining sorting patterns among firms with the same headquarters locations. As shown in column 5 of table 1, the positive relationship between total firm employment and the average employment density of the firm's plant locations is robust to the inclusion of fixed effects for the firm's headquarters locations for firms with more than 100 plants.

Table V, in appendix C (tables V–X are available online), shows many more variations of these results, using different thresholds and selection criteria. All of them show similar findings. In addition, we implement a

[38] See Walsh (2019) for a recent study of firm entry across locations.

TABLE 1
SORTING: FIRM SIZE AND LOCAL DENSITY

	$\ln \bar{L}_j$				
	Baseline		Firms with ≥100 Plants	Industries in Which Largest Firm Has ≥100 Plants	HQ FE, Firms with ≥100 Plants
	(1)	(2)	(3)	(4)	(5)
$\ln L_j$	.165***	.0952***	.146***	.172***	.0791**
	(.000975)	(.000848)	(.0249)	(.00164)	(.0350)
Observations	3,670,994	3,673,053	876	1,387,742	652
R^2	.139	.099	.384	.080	.664
SIC8 FE	Yes	Yes	Yes	Yes	Yes
HQ location FE	No	No	No	No	Yes
M	12	48	12	12	12

NOTE.—The table presents the results of regressing the log of the average employment density across all of the firm's plants on the log employment of the firm at the national level and industry fixed effects (FE). Column 1 presents the baseline results with $M = 12$. Column 2 shows the baseline results with $M = 48$. Column 3 restricts the analysis to firms with at least 100 plants. Column 4 restricts the analysis to industries where there is a firm with at least 100 plants. Column 5 adds the headquarters (HQ) location FE for each firm to the case where we restrict to firms with at least 100 plants. Robust standard errors are in parentheses.
** $p < .05$.
*** $p < .01$.

leave-out strategy to address the potential concern that the firm's presence could be driving local density. The resulting sorting is virtually identical. We also present results when we limit the sample to plants with nonimputed employment data, as well as using alternative weights.

Table 2 presents the estimated elasticity of firm average location density to firm national employment by major industrial sector. As discussed

TABLE 2
SORTING BY MAJOR INDUSTRY

	$\ln \bar{L}_j$				
	All	Manufacturing	Services	Retail Trade	FIRE
	(1)	(2)	(3)	(4)	(5)
$\ln L_j$	.165***	.0523***	.160***	.150***	.234***
	(.000975)	(.00272)	(.00153)	(.00247)	(.00320)
Observations	3,670,994	274,478	1,479,391	856,860	244,048
R^2	.139	.192	.097	.068	.122
SIC8 FE	Yes	Yes	Yes	Yes	Yes
M	12	12	12	12	12

NOTE.—The table presents the results of regressing the log of the average employment density of each location, weighted by the number of establishments of a particular firm operating in a particular industry in the location, on the log employment of the firm at the national level and industry fixed effects (FE). Column 1 presents the baseline case for all industries, and the rest of the columns present results by major sector. Robust standard errors are in parentheses.
*** $p < .01$.

above, the limit problem we study is likely a better approximation of the problem of firms that sell goods and services at short distances. Broad industry classifications do not provide an ideal grouping of industries according to their tradability. Nevertheless, it is probably the case that firms in industries within, say, retail trade sell services that are less tradable than firms in industries within manufacturing. Table 2 shows that the elasticity of firm average location density with respect to firm national employment is, in fact, smaller in manufacturing than in other sectors. We find the highest elasticities in FIRE (finance, insurance, and real estate) and services.

B. The Largest Firm in Town

We can also explore the implications of proposition 5 on sorting by looking at how the size of the firm with the largest number of plants in each location, $L_{j^*(s)}$, changes with population density, $\mathcal{L}_s$, where $j^*(s)$ is the identity of the firm that places the most plants in s. Sorting implies that, in locations with low population densities, low-productivity and smaller firms should place more plants than large firms. Table 3 shows that, in fact, the national size of the firm with the most plants in a location increases with population density, controlling for industry fixed effects. Column 1 presents our baseline case, column 2 the case with 48-mile-square resolution, column 3 when we restrict the sample to firms with at least 100 plants,

TABLE 3
NATIONAL SIZE OF THE LARGEST FIRM IN TOWN

	$\ln L_{j^*(s)}$			
	Baseline		Firms with ≥100 Plants	Industries in Which Largest Firm Has ≥100 Plants
	(1)	(2)	(3)	(4)
$\ln \mathcal{L}_s$	.395***	.594***	.131***	.516***
	(.00266)	(.00461)	(.00811)	(.00366)
Observations	1,984,474	1,006,305	211,517	616,248
R^2	.616	.644	.554	.600
SIC8 FE	Yes	Yes	Yes	Yes
M	12	48	12	12

NOTE.—The table presents the results of finding the log employment of the firm with the most plants in an industry and location and regressing its log total employment on the log density of the location and industry fixed effects (FE), weighted by each industry's total employment. In locations where multiple firms are tied for the highest number of plants, we take the average of the firm size. Column 1 presents the baseline results when $M = 12$, and col. 2 does so for $M = 48$. Column 3 restricts the analysis to firms with at least 100 plants, and col. 4 restricts the analysis to industries where there is a firm with at least 100 plants. For population density, we use the data from the 2010 decennial census, taken from Manson et al. (2021). Robust standard errors are in parentheses.

*** $p < .01$.

and column 4 when we restrict the sample to industries that have at least one firm with 100 or more plants. In all cases, the slope is positive and highly significant.

Table VI, in appendix D, presents a large set of robustness checks, including different thresholds for sample selection and results for major industry groupings, as well as additional spatial resolutions and an exercise with only nonimputed data. In the analysis presented in table 3, there are locations in which multiple firms tie for the highest number of plants. In those cases, we use the average national firm size among these firms. In table VI, we also show that our finding is robust to dropping cases with ties or to using the national size of the largest firm among those tied. Finally, one may worry that the largest firm in a location could be large enough to mechanically and significantly affect the local employment density. In table VI, we show that the results are only marginally affected when excluding the firm's own contribution when calculating local employment density.

C. *The Role of Span-of-Control Costs*

In this section, we present evidence on the particular mechanism driving firm sorting across locations in our model. Lemma 4 shows that higher-productivity firms have a higher cost of increasing their span of control by an additional plant, λ_j. As is evident from equation (4), in choosing the number of plants in a given location, firms trade off these firm-specific fixed costs against profits per plant, which are increasing in a firm's productivity. This trade-off implies that two firms present in the same location, but with different productivity, might decide to have the same number of plants. However, the firm with higher productivity will always have larger plants. Hence, a testable prediction of this mechanism is that, among firms with the same number of plants in a given location, the plants operated by the more productive, and therefore nationally larger, firm should be larger.

Formally, we can write the average plant size of firm j in location s as

$$\bar{l}_{js} = (\varepsilon - 1) z_j^{\varepsilon - 1} \frac{x_s}{W_s} \frac{\kappa(n_{js})}{n_{js}}. \tag{6}$$

Then, it is straightforward to see that, if $z_j > z_{\tilde{j}}$, then $\bar{l}_{js} > \bar{l}_{\tilde{j}s}$ in locations where $n_{js} = n_{\tilde{j}s}$. Note that, in most models of multiplant production, a firm's effective productivity in a given location (e.g., its productivity adjusted by the location's distance to the firm's headquarters) determines both the number of plants and the size of those plants. In contrast to our prediction, this implies that there should be no systematic relationship between firm productivity and plant size after controlling for the number of plants in a location.

Table 4 presents our estimates of the relationship between log average plant employment, $\bar{l}_{js}$, and the log of total firm employment in alternative

TABLE 4
SPAN OF CONTROL

	$\ln \bar{l}_{js}$			
	Baseline		Firms with ≥100 Plants	Industries in Which Largest Firm Has ≥100 Plants
	(1)	(2)	(3)	(4)
$\ln L_{\{j,-s\}}$	.114***	.131***	.275***	.104***
	(.000897)	(.000962)	(.00296)	(.000911)
$\ln n_{js}$	.137***	.172***	−.168***	.0720***
	(.00897)	(.00682)	(.0111)	(.00883)
$(\ln n_{js})^2$	−.0813***	−.0811***	−.00158	−.0564***
	(.00447)	(.00251)	(.00528)	(.00431)
Observations	409,364	386,094	126,999	336,424
R^2	.573	.511	.746	.588
SIC8-location FE	Yes	Yes	Yes	Yes
M	12	48	12	12

NOTE.—The table presents the results of regressing the log of the average plant employment of a firm within a location on the log national employment of the firm (excluding the own-firm contribution of employment in a location from that firm's total employment), industry fixed effects (FE), and controls for the number of plants that the firm has in the location. Column 1 presents the baseline results when $M = 12$, and col. 2 does so for $M = 48$. Column 3 restricts the analysis to firms with at least 100 plants, and col. 4 restricts the analysis to industries where there is a firm with at least 100 plants. Robust standard errors are in parentheses.

*** $p < .01$.

locations, $L_{\{j,-s\}}$, after controlling for the number of plants and industry-location fixed effects. To control for the term $\kappa(n_{js})/n_{js}$, we use a second-order polynomial in $\ln n_{js}$. We exclude the location's employment from a firm's total employment in order to avoid a mechanical relationship between national firm size and local average plant size. In table VIII, in appendix E, we show that the results are similar if we use a higher-order polynomial to approximate $\kappa(n_{js})/n_{js}$ or if we calculate a firm's national size including the location's employment. As before, in table 4 we present estimates for different spatial resolutions, as well as for samples where we restrict attention to firms with more than 100 plants or to industries that have such firms. In all cases, the relationship between average plant size and national firm size is positive and significant, after controlling for the number of establishments.[39]

[39] These results are consistent with those in Fernandes et al. (2018), which finds that a large fraction of the variation of exports in bilateral trade is through the intensive margin of trade. In our model, this variation maps into variation in average plant employment within a location. Moreover, our empirical findings are inconsistent with the application of trade models relying on Pareto distributions to explain the way firms locate their plants across space within the United States, e.g., the ones that rely on the distributional assumptions discussed in Lind and Ramondo (2018).

V. Conclusions

In this paper, we propose a novel methodology to analyze the problem of how to serve customers distributed across heterogeneous locations when firms face transport costs, fixed costs of setting up new plants, and span-of-control costs of managing multiple plants. Although the basic trade-off between transport costs and cannibalization is clear, characterizing the solution to this core problem in economics has proven elusive, given its complexity. In order to make progress, we propose a limit problem in which firms choose a density of plants in space. A large combinatorial problem is therefore reduced to a much simpler calculus-of-variations problem. The solution can be easily characterized, and the problem can be readily incorporated into a general equilibrium spatial setup with labor mobility.

The solution to the firm's problem has a number of unique predictions. First, and most important, is that span-of-control considerations imply that firms sort in space. Specifically, more productive firms operate relatively more plants in locations with higher rents. Less productive firms, in turn, operate more plants in low-rent locations. Furthermore, conditional on the number of plants, more productive firms operate larger plants. These and other predictions of the theory are empirically verified using NETS establishment-level data for 2014 in the United States, both when we look at all industries and when we restrict the sample to industries with small catchment areas that might be better approximated by the limit economy.

The methodology proposed in this paper can readily be used to understand the role of changes in transport infrastructure on plant locations. We illustrate numerically how firms in a "small" industry—one that does not affect local rents or wages—adjust by opening fewer plants but in more locations. We also carry out a similar quantitative exercise to illustrate the effects of improvements in the span-of-control technology, where we see large firms expanding into low-rent markets. Studying general equilibrium counterfactuals for "large" industries that affect local factor prices, or for the whole economy, is left for future research. A quantitative general equilibrium analysis of such changes could be used to study the implications of secular technological changes for the spatial distribution of economic activity, as well as local competition and concentration. These are exciting avenues that our methodology now makes feasible.

Data Availability

Codes to produce all tables and figures in this article, and information about the proprietary data used can be found in Oberfield et al. (2023), in the Harvard Dataverse, https://doi.org/10.7910/DVN/FZB46K.

References

Aghion, Philippe, Antonin Bergeaud, Timo Boppart, Peter J. Klenow, and Huiyu Li. 2019. "A Theory of Falling Growth and Rising Rents." Working Paper no. 2019–11, Fed. Reserve Bank, San Francisco.

Arkolakis, Costas, Fabian Eckert, and Rowan Shi. 2023. "Combinatorial Discrete Choice: A Quantitative Model of Multinational Location Decisions." Working Paper no. 31877 (November), NBER, Cambridge, MA.

Arkolakis, Costas, Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple. 2018. "Innovation and Production in the Global Economy." *A.E.R.* 108 (8): 2128–73.

Balinski, M. L. 1965. "Integer Programming: Methods, Uses, Computations." *Management Sci.* 12 (3): 253–313.

Behrens, Kristian, Gilles Duranton, and Frédéric Robert-Nicoud. 2014. "Productive Cities: Sorting, Selection, and Agglomeration." *J.P.E.* 122 (3): 507–53.

Bilal, Adrien, and Esteban Rossi-Hansberg. 2021. "Location as an Asset." *Econometrica* 89 (5): 2459–95.

Bollobás, Béla. 1973. "The Optimal Arrangement of Producers." *J. London Math. Soc.*, 2nd series, 6 (4): 605–13.

Byrka, Jaroslaw, and Karen Aardal. 2010. "An Optimal Bifactor Approximation Algorithm for the Metric Uncapacitated Facility Location Problem." *SIAM J. Computing* 39 (6): 2212–31.

Cao, Dan, Erick Sager, Henry Hyatt, and Toshihiko Mukoyama. 2019. "Firm Growth through New Establishments." 2019 Meeting Paper 1484, Soc. Econ. Dynamics, Stony Brook, NY.

Christaller, Walter. 1933. *Die zentralen Orte in Süddeutschland*. Jena: Fischer.

Crane, Leland, and Ryan Decker. 2019. "Business Dynamics in the National Establishment Time Series (NETS)." Finance and Econ. Discussion Series 2019-034, Board of Governors, Fed. Reserve System, Washington, DC. https://doi.org/10.17016/FEDS.2019.034.

Davis, Donald R., and Jonathan I. Dingel. 2019. "A Spatial Knowledge Economy." *A.E.R.* 109 (1): 153–70.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *A.E.R.* 106 (3): 479–524.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5): 1741–79.

Eeckhout, Jan, Roberto Pinheiro, and Kurt Schmidheiny. 2014. "Spatial Sorting." *J.P.E.* 122 (3): 554–620.

Fejes Tóth, László. 1953. *Lagerungen in der Ebene, auf der Kugel und im Raum*. Berlin: Springer.

Fernandes, Ana M., Peter J. Klenow, Sergii Meleshchuk, Denisse Pierola, and Andrés Rodríguez-Clare. 2018. "The Intensive Margin in Trade." Working Paper no. 25195 (October), NBER, Cambridge, MA.

Fowler, Robert J., Michael S. Paterson, and Steven L. Tanimoto. 1981. "Optimal Packing and Covering in the Plane Are NP-Complete." *Information Processing Letters* 12 (3): 133–37.

Gaubert, Cecile. 2018. "Firm Sorting and Agglomeration." *A.E.R.* 108 (11): 3117–53.

Gersho, Allen. 1979. "Asymptotically Optimal Block Quantization." *IEEE Transactions Information Theory* 25 (4): 373–80.

Guha, Sudipto, and Samir Khuller. 1999. "Greedy Strikes Back: Improved Facility Location Algorithms." *J. Algorithms* 31 (1): 228–48.

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

Holmes, Thomas J., and S. Lee. 2010. "Cities as Six-by-Six-Mile Squares: Zipf's Law?" In *Agglomeration Economics*, edited by Edward L. Glaeser, 105–31. Chicago: Univ. Chicago Press (for NBER).

Hopenhayn, Hugo A. 1992. "Entry, Exit, and Firm Dynamics in Long Run Equilibrium." *Econometrica* 60 (5): 1127–50.

Hsieh, Chang-Tai, and Esteban Rossi-Hansberg. 2023. "The Industrial Revolution in Services." *J.P.E. Macroeconomics* 1 (1): 3–42.

Hu, Kathleen, and Rowan Shi. 2019. "Solving Combinatorial Discrete Choice Problems in Heterogeneous Agents Models." Working paper, Princeton Univ.

Jia, Panle. 2008. "What Happens When Wal-Mart Comes to Town: An Empirical Analysis of the Discount Retailing Industry." *Econometrica* 76 (6): 1263–316.

Jovanovic, Boyan. 1982. "Selection and the Evolution of Industry." *Econometrica* 50 (3): 649–70.

Li, Shi. 2013. "A 1.488 Approximation Algorithm for the Uncapacitated Facility Location Problem." *Information and Computation* 222:45–58.

Lind, Nelson, and Natalia Ramondo. 2018. "Trade with Correlation." Working Paper no. 24380 (March), NBER, Cambridge, MA.

Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2021. "IPUMS [Integrated Public Use Microdata Series], National Historical Geographic Information System [dataset]." Version 16.0. Minneapolis: IPUMS. http://doi.org/10.18128/D050.V16.0.

Melitz, Marc J. 2003. "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity." *Econometrica* 71 (6): 1695–725.

Nocke, Volker. 2006. "A Gap for Me: Entrepreneurs and Entry." *J. European Econ. Assoc.* 4 (5): 929–56.

Oberfield, Ezra, Esteban Rossi-Hansberg, Pierre-Daniel Sarte, and Nicholas Trachter. 2020. "Plants in Space." Working Paper no. 27303 (June), NBER, Cambridge, MA.

———. 2023. "Replication data for: 'Plants in Space.'" Harvard Dataverse, https://doi.org/10.7910/DVN/FZB46K.

Ramondo, Natalia. 2014. "A Quantitative Approach to Multinational Production." *J. Internat. Econ.* 93 (1): 108–22.

Ramondo, Natalia, and Andrés Rodríguez-Clare. 2013. "Trade, Multinational Production, and the Gains from Openness." *J.P.E.* 121 (2): 273–322.

Rossi-Hansberg, Esteban, Pierre-Daniel Sarte, and Nicholas Trachter. 2021. "Diverging Trends in National and Local Concentration." *NBER Macroeconomics Ann.* 35:115–50.

Rossi-Hansberg, Esteban, and Mark L. J. Wright. 2007. "Establishment Size Dynamics in the Aggregate Economy." *A.E.R.* 97 (5): 1639–66.

Stollsteimer, John Fred. 1961. "The Effect of Technical Change and Output Expansion on the Optimum Number, Size, and Location of Pear Marketing Facilities in a California Pear Producing Region." PhD dissertation, Univ. California, Berkeley.

Sviridenko, Maxim. 2002. "An Improved Approximation Algorithm for the Metric Uncapacitated Facility Location Problem." In *Integer Programming and Combinatorial Optimization: 9th International IPCO Conference*, Proc., edited by William J. Cook and Andreas S. Schulz, 240–57. Lecture Notes in Computer Science, vol. 2337. Berlin: Springer.

Tintelnot, Felix. 2017. "Global Production with Export Platforms." *Q.J.E.* 132 (1): 157–209.

Walsh, Conor. 2019. "Firm Creation and Local Growth." Working paper. Available at SSRN: https://doi.org/10.2139/ssrn.3496782.
Weber, Alfred. 1909. *Über den Standort der Industrien* [*Theory of the Location of Industries*]. Tübingen: Mohr.
Ziv, Oren. 2019. "The Micro-geography of Productivity, Density, and Sorting." Working paper.