Agglomerative Forces and Cluster Shapes[*]

William R. Kerr
Harvard University and NBER

Scott Duke Kominers
Harvard University

December 2010

Abstract

We model spatial clusters of similar firms. Our model highlights how agglomerative forces lead to localized, individual connections among firms, while interaction costs generate a defined distance over which attraction forces operate. Overlapping firm interactions yield agglomeration clusters that are much larger than the underlying agglomerative forces themselves. Empirically, we demonstrate that our model's assumptions are present in the structure of technology and labor flows within Silicon Valley and its surrounding areas. Our model further identifies how the lengths over which agglomerative forces operate influence the shapes and sizes of industrial clusters; we confirm these predictions using variations across both technology clusters and industry agglomeration.

JEL Classification: J2, J6, L1, L2, L6, O3, R1, R3.

Key Words: Agglomeration, Clusters, Industrial Organization, Silicon Valley, Entrepreneurship, Labor Markets, Technology Flows, Patents, Natural Advantages.

1 Introduction

Agglomeration—industrial clustering—is a key feature of economic geography. Starting with Marshall (1920), a large and diverse literature studies why similar firms often collocate. One body of work provides formal theoretical models of different agglomerative forces like technology sharing or labor pooling among firms. A second literature measures interactions among firms and/or workers directly (e.g., worker moves across firms, patent citation distances). A third literature characterizes differences in clustering behavior across industries (e.g., relating R&D intensity to greater spatial concentration). Despite this progress, we have very little understanding of the actual connection between micro-level interactions and observed cluster shapes and sizes for industries.

This is very worrisome because studies frequently use observed spatial concentrations for industries to infer the properties of the underlying forces that cause them. Economists often infer that technology spillovers are shorter than labor pooling spillovers because patenting firms are relatively more concentrated at the county level than are firms that use similar types of workers. Such reasoning, however, clashes with studies which have concluded that spillover effects are orders of magnitude stronger over the first few city blocks than they are when firms are between two and five miles apart (e.g., Rosenthal and Strange 2003, 2008, Arzaghi and Henderson 2008). Many studies find that knowledge flows decrease sharply across floors within a single building. Why would we believe that we can infer useful information about knowledge spillovers from the number of patents in a county, when a county averages over 75,000 people in population?

This project examines these issues theoretically and empirically. The core of our work is a location choice model that connects limited, localized agglomerative forces with the formation of spatial clusters for similar firms. Agglomerative forces in our model are localized because firms face interaction costs. Spillover benefits exceed these interaction costs at short distances, but beyond some distance interaction becomes unprofitable. For example: while a firm could learn useful technologies from another firm 20 miles away, the costs of doing so may be too great to justify the effort. Clusters are then the product of many small, overlapping regions of interaction.

Silicon Valley is the world's most famous cluster, and many observers credit its success to either technology spillovers or labor pooling. Figures 1a and 1b identify the foundations of our theoretical framework using technology and labor flows, respectively, in the San Francisco Bay Area. Figure 1a focuses on technology sourcing in an area surrounding Silicon Valley. Downtown San Francisco and Oakland are to the north and off of the map. The triangle in the bottom right corner of the map is the core of Silicon Valley. This core contains 76% of industrial patents filed from the San Francisco Bay Area and 18 of the top 25 zip codes in terms of patenting.

To introduce our model, we describe the primary technology sourcing zones for three of the

four largest zip codes for patenting outside of the core. Each focal zip code is marked with a star, and the other points of the shape are the three zip codes that firms in the focal zip code cite most in their work. The orange zone (also labelled with a "1") for Menlo Park extends deepest into the core. The green zone for Redwood City ("2") shifts up and encompasses Menlo Park and Palo Alto but less of the core. The black zone for South San Francisco ("3") further shifts out and brushes the core.

These technology zones are characterized by small, overlapping regions. None of the technology sourcing zones transverse the whole core, much less the whole cluster, and only the closest zip code (Menlo Park) even reaches far enough into the core to include the area of Silicon Valley where the greatest number of patents occur. The empirical appendix contains additional maps that show these small, overlapping regions are also evident in the core itself and in other parts of the San Francisco region. While technology sourcing for individual firms is localized, the resulting cluster extends over a larger expanse of land.

These properties are also evident in labor flows. Figure 1b describes commuting zones for scientists and engineers from the 2000 Census of Populations. Zones on the map are approximations, computed from Public Use Micro Area (PUMA) designations. The colored asterisks indicate locations of work. The zones surrounding the asterisks of similar color indicate the primary PUMAs of residence for the workers. Labor zones are more uniform in direction, with the star for the work PUMA typically nearer to the center of its labor zone. These labor commuting regions again display small, overlapping zones that in turn encompass a large land area.[1]

The central construct of our model replicates these features. Our model thus makes explicit that empirical observation of cluster size in the data does not indicate the length of the micro-interactions that produce the cluster. We show, however, that cluster shape and size does depend systematically on whether the localized interactions for firms in an industry are longer or shorter in length. We demonstrate that a longer effective spillover region, either due to weaker decay in benefits or lower interaction costs, yields a macro-structure with fewer, larger, and less-dense clusters. These regularities allow researchers to use cluster dimensions to rank order spillover lengths even though micro-interactions are not observed.

After deriving these predictions, we empirically test the model in two ways. We first use US patent data to describe differences across technology clusters. Patent citations allow us to measure effective spillover regions by technology. Differences in these spillover regions relate to cluster shapes and sizes as predicted by the model. Technologies with very short distances over which firms interact exhibit clusters that are smaller and denser than technologies that allow for longer distances.

We then use US Census Bureau data on spatial concentration patterns throughout all sectors of the economy. We show that industries with concentrated labor requirements, for whom the benefits of labor pooling are very strong, also exhibit smaller and denser clusters relative to

---

[1]See also the Arzaghi and Henderson (2008) mapping of advertising agencies in Manhattan.

industries without this need. A comparison of differences between agglomeration due to Marshallian forces and those due to natural advantages (e.g., access to waterways) further conform to our model's predictions. This empirical work primarily employs continuous agglomeration metrics like that of Duranton and Overman (2005), and we use traits of industries in the United Kingdom to confirm the causal direction of these relationships (e.g., Ellison et al. 2010).

Our work makes several contributions to the literature on industrial agglomeration. Most importantly, we provide a theoretical connection between observable cluster shapes and the underlying agglomerative forces that cause them. At the core of this contribution is the simple mechanism of including interaction costs among firms. The resulting framework provides a theoretical foundation for inferring properties of agglomerative forces through observed spatial concentrations of industries. We identify settings in which such inference is appropriate, as well as key properties of agglomeration in such settings. An additional contribution of our work, discussed in greater detail later, is to provide a micro-foundation for using continuous spatial density measurements that center on bilateral distances between firms. This class of metrics includes the popular Duranton and Overman (2005) metric.[2,3]

Our central empirical contribution is a framework, motivated by our theoretical model, for meaningful analysis of agglomerative forces with continuous distance horizons. Previous work considers how agglomerative forces affect spatial concentration over different distance horizons, for example up to 75 or 250 miles (e.g., Rosenthal and Strange 2001, 2004; Duranton and Overman 2005, 2008; Ellison et al. 2010). Our framework is an important step towards jointly considering agglomeration at different distances (25, 75, and 250 miles) simultaneously. This paper provides a first rationale for this type of analysis, and we hope that future research can similarly analyze other factors that govern cluster shapes and sizes.[4]

Section 2 presents our theoretical model and connects it to empirical measurement of agglomeration. Section 3 tests the model using variations across patent citation clusters. Section 4 presents the second empirical analysis that uses variations in industrial agglomeration for many sectors of the economy. The last section concludes.

---

[2]Studies of agglomeration metrics include Ellison and Glaeser (1997), Maurel and Sédillot (1999), Marcon and Puech (2003), Mori et al. (2005), Barlet et al. (2009), Ellison et al. (2010), and Billings and Johnson (2010).

[3]A further technical contribution of our work centers on establishing conditions to distinguish agglomeration due to Marshallian spillovers among firms versus proximity to natural advantages. These have been observationally equivalent in prior work (e.g., Ellison and Glaeser 1997).

[4]Related studies include Rosenthal and Strange (2001, 2003, 2004, 2008), Duranton and Overman (2005, 2008), Ellison and Glaeser (1997, 1999), Audretsch and Feldman (1996), Head and Mayer (2004), Arzaghi and Henderson (2008), Ellison et al. (2010), Greenstone et al. (2010), Delgado et al. (2009), Holmes and Lee (2009), Fallick et al. (2006), Glaeser and Kerr (2009), Menon (2009), Bleakley and Lin (2010), Alcacer and Chung (2007), Pe'er and Vertinsky (2009), Alfaro and Chen (2010), and Dauth (2010). Our work also connects to studies of the shapes of cities (e.g., Lucas and Rossi-Hansberg 2002, Baum-Snow 2007, 2010, Glaeser 2008, and Saiz 2010) and of agglomeration and productivity differences across cities and regions (e.g., Ciccone and Hall 1996, Partridge et al. 2009, Fu and Ross 2010, and Behrens et al. 2010, Sarvimäki 2010). Jackson (2008) outlines a complementary literature on economic networks.

2 Theoretical Framework

We now introduce a model of firm location choice that generates large agglomeration clusters from smaller, overlapping spillover zones. To maintain consistency with previous work, we use the notation of Duranton and Overman (2005) whenever possible. The theoretical appendix contains proofs of propositions and additional materials noted below.

2.1 Basic Framework

There are N firms indexed by i, and each firm i belongs to an industry $A(i)$. These firms i sequentially select their locations, denoted $j(i)$, from a fixed set $\mathbf{Z} \subset \mathbb{R}^2$ of potential sites. Sites are drawn at random according to a uniform distribution in advance of any firms' location decision. There are many more possible sites than firms, i.e. $|\mathbf{Z}| \gg N$. To focus on agglomeration economies, we assume that firms compete in broad product markets. Location choice thus affects the productivity of a firm, but not its competitive environment.

The specific benefits of location j to a firm in industry A are driven by two factors that generate spatial concentration. First, intra-industry Marshallian forces represent productivity spillovers that firms generate by being in proximity to each other, independent of geographic positioning relative to attraction sites defined shortly. Three common examples are customer-supplier interactions (e.g., reducing transportation costs for intermediate goods), labor pooling, and knowledge exchanges.

Second, natural advantages make some sites more attractive for firms independent of these Marshallian forces. Shipbuilders are attracted to harbors, and pulp and paper mills are not well-suited for desert areas. Natural advantages can also be man made. State capitols, for example, are attraction sites for law firms, and universities can provide technology transfers to high-tech industries. We accordingly model industry-specific local advantages that represent the value of locating near a fixed set of industry-specific attraction sites, $\mathbf{K}(A)$.[5]

Trade-offs between these factors affect firms' patterns of clustering, as does the distribution of the attraction sites. We denote by d_{j_1,j_2} the spatial distance between $j_1 \in \mathbb{R}^2$ and $j_2 \in \mathbb{R}^2$. We assume that the benefit of site $j \in \mathbf{Z}$ to a firm i in industry $A(i)$ is given by

$$V_j(i) = g_{j,A(i)} + \epsilon_{i,j},$$

where $g_{j,A(i)}$ takes the form

$$g_{j,A(i)} \equiv \sum_{k \in \mathbf{K}(A(i))} G_{\mathbf{K}}^{A(i)}(d_{j,k}) + \sum_{\substack{i' \neq i, \\ A(i')=A(i)}} G_{\mathbf{A}}^{A(i)}(d_{j,j(i')}). \tag{1}$$

---

[5] For technical reasons, we assume that locations of attraction sites $\mathbf{K}(A)$ are realized before potential sites $\mathbf{Z}$ are drawn. All attraction sites $k \in \mathbf{K}(A)$ are equally attractive to firms in industry A. This assumption is made without loss of generality because, to capture the effect of attraction site k that is twice as attractive as site k', it suffices to add a second attraction site k'' infinitesimally close to k. Any rational ratio between attraction site values may be approximated in this way using finitely many sites.

The function $g_{j,A(i)}$ represents the degree to which site j is generally attractive for firms in industry $A(i)$. The two components of (1) respectively encapsulate natural advantages and intra-industry Marshallian forces. We assume the standard basic comparative static of the value function $g_{j,A(i)}$: each function G_Ξ^A in (1) is decreasing, so that agglomerative forces decline over space. Additionally, for simplicity, we assume that agglomerative forces of type Ξ either act across all distances or do not act at all. That is, for a given A, if $G_\Xi^A(0) > 0$, then $G_\Xi^A(d) > 0$ for all $d \geq 0$.

The term $\epsilon_{i,j} \ll g_{j,A(i)}$ is a positive i.i.d. random term representing the degree to which site j is a specific match for firm i. For expositional clarity, we will think of the deterministic "value" of site j to firm i as being given by $g_{j,A(i)}$. We treat the random term $\epsilon_{i,j}$ as representing a process by which firm i chooses randomly among sites $j_1, \ldots, j_\ell$ for which $g_{j_1,A(i)} = \cdots = g_{j_\ell,A(i)}$; that is, among sites over which firm i would be indifferent if forced to choose purely on the basis of spatial attraction. Firms are not forward-looking, so that the n-th firm to enter, i_n $(1 \leq n \leq N)$, chooses its location $j(i_n) \in \mathbf{Z}$ to maximize $V_j(i_n)$ conditional upon the location choices of the first $n-1$ firms.

2.2 Maximal Radius of Interaction

So far, our model has more or less followed a standard structure: proximity to resources and other firms generates benefits, and these benefits decay continuously over distance. However, we now depart from this standard approach via a simple and natural additional assumption.

Assumption. *Firms must pay a fixed cost to interact with other firms or natural advantage sites. Firms in industry A must pay $c_{\mathbf{K}}^A$ to interact with a natural advantage site and $c_{\mathbf{A}}^A$ to interact with another firm in industry A.*

These fixed costs c_Ξ^A relate to the costs of transporting goods, people, or ideas across firms. Opportunity costs and search costs are the simplest examples, but we further model that these costs are specific to industries and spillover types. For example, accessing and understanding codified technologies likely requires a lower fixed cost of establishing interactions than that required for tacit technologies. It is also easier to transport some natural advantages (e.g., lumber) than others (e.g., waterways).

Firms only invest in establishing contacts when the benefits of doing so equal or exceed the associated costs of interaction. Specifically, firm i only invests in type-Ξ contact with site j if $G_\Xi^A(d_{j(i),j}) \geq c_\Xi^A$. This defines a strict distance over which the firm finds type-Ξ interactions profitable,

$$d_{j(i),j} \leq \rho_\Xi^A \equiv \max\{d : G_\Xi^A(d) \geq c_\Xi^A\}.$$

Therefore, we immediately observe the following result.

Proposition 1. *For each industry A, sites further than distance*

$$\max\left\{\rho_{\mathbf{K}}^{A}, \rho_{\mathbf{A}}^{A}\right\} \tag{2}$$

from an industry-A firm cannot profitably interact with that firm. That is, an industry-A firm i derives no direct benefits from the presence of firms or attraction sites at locations j with $d_{j,j(i)} > \max\left\{\rho_{\mathbf{K}}^{A}, \rho_{\mathbf{A}}^{A}\right\}$.

In practice, the key consequence of Proposition 1 is that agglomerative forces act only over finite distances. We call ρ_{Ξ}^{A} the *maximal radius of Ξ-type interaction in industry A*, and abusing terminology slightly, call the vector

$$\rho^{A} \equiv (\rho_{\mathbf{K}}^{A}, \rho_{\mathbf{A}}^{A})$$

the *maximal radius of interaction in industry A* (or just the *maximal radius* when the industry is clear from context). The maximal radius is (weakly) decreasing in the costs c_{Ξ}^{A} and increasing in the levels of the decay functions G_{Ξ}^{A}. In other words, lower costs or weaker attentuation of benefits lead to larger maximal radii.[6]

Our assumption that interaction costs are fixed is only to simplify the discussion below. One might naturally assume that interaction costs rise with distance; such an assumption would also generate the maximal radius described in Proposition 1. The ultimate technical condition is that interaction costs exceed interaction benefits at some distance with a single crossing.

2.3 Agglomeration Clusters

We next examine how agglomeration clusters form in our model and their properties. Figures 2a and 2b provide a graphical presentation of the theory to build intuition. To focus on the most essential parts of the model, Figure 2a only considers Marshallian spillovers (i.e., $G_{\mathbf{K}}^{A}(0) = 0$), while Figure 2b only considers attraction to natural advantage sites (i.e., $G_{\mathbf{A}}^{A}(0) = 0$). In these graphs, lightly colored circles are potential firm locations, while filled-in circles represent sites populated by firms. Stars indicate attraction sites in Figure 2b. Throughout this paper, we use these graphs to explain the model's structure and depict the behavior of marginal entrants.

2.3.1 Basic Definitions and Structure

We define an *agglomeration cluster* to be a group of industry-A firms located in sites interconnected by bilateral interactions. Each firm does not necessarily interact with every other firm in its cluster, but all firms in a cluster are interconnected. Our measure of agglomeration counts the number of these clusters that are expected to arise for an industry. The use of an expectation is necessary because firms choose randomly when indifferent between sites. Firms in industry A exhibit agglomeration if they typically occupy few distinct clusters.

---

[6]Here, by saying that the maximal radius gets "smaller" or "larger," we refer to the standard vector ordering: $(\rho_{\mathbf{K}}, \rho_{\mathbf{A}}) = \rho \leq \rho' = (\rho_{\mathbf{K}}', \rho_{\mathbf{A}}')$, if and only if $\rho_{\mathbf{K}} \leq \rho_{\mathbf{K}}'$ and $\rho_{\mathbf{A}} \leq \rho_{\mathbf{A}}'$.

More formally: For $j \in \mathbb{R}^2$, we denote by $B_d(j) \equiv \{j' \in \mathbb{R}^2 : d_{j,j'} \le d\}$ the closed ball of radius d about j. For $j \in \mathbf{Z} \cup \mathbf{K}(A)$, we set

$$\mathcal{B}^0_{\rho^A}(j) \equiv \begin{cases} (B_{\rho^A_{\mathbf{A}}}(j) \cap \mathbf{Z}) \cup (B_{\rho^A_{\mathbf{K}}}(j) \cap \mathbf{K}(A)) & j \in \mathbf{Z}, \\ \{j\} \cup (B_{\rho^A_{\mathbf{K}}}(j) \cap \mathbf{Z}) & j \in \mathbf{K}(A). \end{cases}$$

This formula, while complex in its notation, has a simple interpretation. $\mathcal{B}^0_{\rho^A}(j)$ is the union of potential firm locations and attraction sites which can profitably interact with j under maximal radius ρ^A.

In Figure 2a, we draw for each populated site a representative maximal radius within which the benefits of interaction exceed the costs for firms. For this example, $\mathcal{B}^0_{\rho^A}$ for site A includes sites B and C. Sites B and C are the only locations within the maximal radius for Marshallian conditions $\rho^A_{\mathbf{A}}$.

We next expand our focus to consider sites that are outside of the profitable spillover range of site j, but can be connected to j via a single interconnection. Define $\mathcal{B}^1_{\rho^A}(j)$ to be all sites which can profitably interact with the sites in $\mathcal{B}^0_{\rho^A}(j)$ through one additional step. In Figure 2a, $\mathcal{B}^1_{\rho^A}(A)$ further includes the four additional sites within distance $\rho^A_{\mathbf{A}}$ from site C that are outside of the spillover range of site A. We continue to iterate this process, successively adding additional sites that are more spatially distant to site j but still connected to site j by increasing numbers of interconnections ($\mathcal{B}^\iota_{\rho^A}(j)$ for $\iota = 1, 2, \ldots$). Formally, for any $j \in \mathbf{Z} \cup \mathbf{K}(A)$,

$$\mathcal{B}^\iota_{\rho^A}(j) \equiv \bigcup_{j' \in \mathcal{B}^{\iota-1}_{\rho^A}(j)} \mathcal{B}^0_{\rho^A}(j')$$

Iterating this construction of clusters to its conclusion, $\mathcal{B}_{\rho^A}(j)$ is the ρ^A-cluster containing $j \in \mathbf{Z} \cup \mathbf{K}(A)$, defined by

$$\mathcal{B}_{\rho^A}(j) \equiv \bigcup_{\iota=0}^{\infty} \mathcal{B}^\iota_{\rho^A}(j).$$

The ρ^A-cluster containing site j is the largest cluster of sites that 1) contains j and 2) is connected by a chain of "hops" between sites $j' \in \mathbf{Z} \cup \mathbf{K}(A)$ which can profitably interact. The complete set of filled locations in Figure 2a constitute the ρ^A-cluster for site A in our example.[7]

When the maximal radius ρ^A is small, clusters are generally small. For two precise examples, define the lower and upper bounds on distances between sites as $\underline{d} \equiv \min_{j_1 \ne j_2 \in \mathbf{Z} \cup \mathbf{K}} d_{j_1,j_2}$ and $\bar{d} \equiv \max_{j_1 \ne j_2 \in \mathbf{Z} \cup \mathbf{K}} d_{j_1,j_2}$. When $\rho^A < (\underline{d}, \underline{d})$, $\mathcal{B}_{\rho^A}(j) = \{j\}$ for each $j \in \mathbf{Z}$. In other words, a maximal radius that is shorter than the shortest distance between two sites results in each cluster only containing a single firm. By contrast, when $\rho^A > (\bar{d}, \bar{d})$, $\mathcal{B}_{\rho^A}(j) = \mathbf{Z} \cup \mathbf{K}(A)$ for all $j \in \mathbf{Z}$. A maximal radius longer than the maximal distance between sites results in a single cluster for an industry.

---

[7]The use of an infinite index in the union defining $\mathcal{B}_{\rho^A}(j)$ is ultimately unnecessary, as the finitude of the set of sites implies that $\mathcal{B}^\iota_{\rho^A}(j) = \mathcal{B}^{\iota+1}_{\rho^A}(j) = \cdots$ for some finite ι.

Suppose that natural advantages do not act in an industry A, i.e. that $G_{\mathbf{K}}^A(0) = 0$, so that $\rho_{\mathbf{K}}^A = 0$. (In this case, here and hereafter, we simplify our notation slightly by writing $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j) \equiv \mathcal{B}_{\rho^A}(j)$.) If the first firm in industry A with maximal Marshallian radius $\rho_{\mathbf{A}}^A$ locates at site j, and the cluster around j contains available locations (i.e., $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j) \neq \{j\}$), then there is some site $j' \in \mathcal{B}_{\rho_{\mathbf{A}}^A}(j)$ which delivers positive utility flows to the next entrant in industry A. It follows that if Marshallian forces are sufficiently strong, then firms in industry A select sites in the cluster $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j)$ until $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j)$ is filled. Iterating this analysis shows that when Marshallian forces are strong, firms occupy clusters in sequence.

The sequential filling of clusters explains how large-area clustering may arise in an industry even if agglomerative forces act only over short distances. Cluster sizes associated with a given maximal radius can be much larger than the underlying radius itself. Clusters may span large regions even if each firm derives benefits only from its immediate neighbors.

A consequence of the maximal radius, however, is that clusters can reach their capacity, at which point the next entrant for the industry will locate elsewhere. In Figure 2a, the closest remaining site to the existing cluster is site X, but this location is beyond the spillover ranges of any of the populated sites in the cluster. As the marginal entrant cannot profitably interact with the cluster, it is indifferent among sites X, Y, Z, and any other unoccupied site. It will choose its location based upon its idiosyncratic preferences $\epsilon_{i,j}$.

Likewise, when natural advantages are important to industry A, firms fill clusters near attraction sites $K(A)$ whenever possible. Figure 2b depicts the simple case where an industry does not have any Marshallian spillovers, but a series of attraction sites are present and important for location choice. The maximal radius in this figure shows the range around each attraction site where firms can profitably interact with the attraction site. A boundary or edge of the cluster again emerges.

The preceding observations suggest a natural notion of agglomeration in our model:

Definition 1. The *level of agglomeration in industry A for maximal radius ρ^A* is the expected number of distinct ρ^A-clusters $\mathcal{B}_{\rho^A}(j)$ about sites $j \in \mathbf{Z} \cup \mathbf{K}(A)$ occupied by industry-A firms.

We say that firms are *(weakly) more agglomerated with respect to maximal radius ρ_1 than they are with respect to radius ρ_2* if, holding N^A fixed, fewer clusters of industry-A firms form when the maximal radius is ρ_1 than when it is ρ_2. Note that agglomeration increases as the expected number of clusters decreases. Holding industry size constant, increased agglomeration therefore also corresponds to increased cluster size.

Our discussion of the marginal entry decision also highlights the core difference between our structure and prior work. Without considering interaction costs, strictly positive spillover benefits exist at all distances. Industries may differ in how fast or slow Marshallian benefits decay, but these differences in Marshallian forces do not impact the number of clusters. Regardless of whether the potential spillover benefit is large or miniscule, marginal entrants always select sites

closest to the developing cluster regardless of distance (i.e., site X next in Figure 2a). As a consequence, one cluster always forms for an industry, and entrants select sites in an exact order in all industries.[8]

Thus, the simplest framework does not provide a foundation for relating differences in spatial concentration for industries to their underlying agglomerative forces. Yet, our intuitive addition of interaction costs provides additional traction by establishing a spatial range over which interactions are relevant. This localization in turn provides meaningful differences in cluster formation. We now turn to these comparative statics.[9]

2.3.2 Agglomeration due to Marshallian Advantages

Figure 2a illustrates the consequences of a longer maximal radius for cluster formation. We enlarge the effective spillover range for site B to the wider circle. While we only enlarge site B's circle for visual clarity, the comparative static we have in mind is one in which all firms in the industry have a longer spillover range equivalent to site B's. Under the larger maximal radius, the marginal entrant is no longer indifferent over sites, but would instead choose site X. Thus, a longer maximal radius is (weakly) associated with greater industry agglomeration as fewer clusters form in expectation.

More formally, recall that $\rho_{\mathbf{A}}^A$ is the maximal radius for intra-industry spillovers, $\rho_{\mathbf{A}}^A \equiv \max\{d : G_{\mathbf{A}}^A(d) \geq c_{\mathbf{A}}^A\}$. Since $\mathbf{Z} \cup \mathbf{K}(A)$ is finite, sufficiently small changes do not affect location choices. Larger increases in $\rho_{\mathbf{A}}^A$, either due to weaker attenuation in spillover benefits or lower interaction costs, can lead firms to organize into fewer clusters. In fact, we may sign this change: so long as Marshallian forces are sufficiently strong, industry-A firms become more agglomerated when $\rho_{\mathbf{A}}^A$ increases. To simplify our demonstration of this fact in the next proposition, we "turn off" natural advantage effects by temporarily assuming $G_{\mathbf{K}}^A(0) = 0$.[10]

Proposition 2. *If $G_{\mathbf{K}}^A(0) = 0$, then a longer maximal radius for intra-industry spillovers $\rho_{\mathbf{A}}^A$ leads to a (weak) increase in the agglomeration level of industry A.*

The idea behind the proof of Proposition 2 is intuitive. A firm i that is indifferent across sites chooses its location $j(i) \in \mathbf{Z}$ randomly. But until the sites in cluster $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j(i))$ are filled, they are more attractive to industry-A firms than are unfilled sites outside of $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j(i))$. If $\rho_{\mathbf{A}}^A$ grows to $\hat{\rho}_{\mathbf{A}}^A$, a radius large enough that $\mathcal{B}_{\hat{\rho}_{\mathbf{A}}^A}(j) = \mathcal{B}_{\rho_{\mathbf{A}}^A}(j) \cup \mathcal{B}_{\rho_{\mathbf{A}}^A}(j')$ for some $j, j' \in \mathbf{Z}$, then the expected

---

[8]This is equivalent to $\rho^A \to \infty$ in our model. Once a first entrant for an industry picks a location, the set of sites filled by the remaining $N^A - 1$ firms (where $N^A \equiv |\{i : A(i) = A\}|$) is exactly determined.

[9]Our discussion and proofs focus on the simple case where spillover benefits do not transfer through the cluster. Interaction costs are incurred on a bilateral basis, and firms at the periphery of a cluster only receive benefits from their immediate neighbors. More generally, our predictions hold for any transmission structures of benefits through the cluster so long as $\rho_{\mathbf{A}}^A$ is constant.

The theoretical appendix discusses other extensions and attributes of this framework. Most notably, we outline some basic spatial dynamics for clusters.

[10]The conclusion of Proposition 2 holds whenever natural advantage benefits are sufficiently small. However, the assumption that $G_{\mathbf{K}}^A(0) = 0$ substantially simplifies the demonstration of the result.

number of clusters occupied by industry-A firms shrinks. Indeed, whenever a firm locates in either $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j)$ or $\mathcal{B}_{\rho_{\mathbf{A}}^A}(j')$, industry-$A$ firms fill all of $\mathcal{B}_{\hat{\rho}_{\mathbf{A}}^A}(j)$ before locating in or starting another cluster.

Three empirical implications of the proof of Proposition 2 are evident in Figure 2a. First, industries with a longer maximal radius have larger clusters in the sense of having more firms and covering a greater spatial area. Intuitively, a longer spillover radius makes sites at the edges of clusters attractive that are not attractive with a shorter radius. This induces marginal entrants into choosing these sites rather than starting new clusters. A longer radius can be due to weaker decay of spillover benefits or lower interaction costs.

The second and third predictions are closely related. A longer spillover radius yields fewer clusters for a given industry size. As clusters grow in size, fewer clusters are needed to house the N^A firms in the industry. Finally, clusters are less dense. The longer radius activates sites at the edges of a cluster that are too spatially distant to profitably interact with previous entrants if the radius is shorter. Thus, growth in cluster size is simultaneous with reduction in cluster density.

Our result that clusters due to a longer maximal radius are less dense is the same as saying that average bilateral distances among firms within the clusters increase. The model's structure, however, contains a much more powerful implication regarding spillover lengths and the complete distribution of bilateral distances within clusters. We draw out this implication in Section 2.4.3 after introducing empirical measurement more formally.

2.3.3 Agglomeration due to Natural Advantages

To this point, our discussion has focused mostly on Marshallian agglomeration where firms derive benefits from being near to each other. In partial analogy with Proposition 2, industries sufficiently dependent upon natural advantages also exhibit increased agglomeration as $\rho_{\mathbf{K}}^A$ grows. Moreover, this effect is mediated by the density of the set of potential firm locations.

Proposition 3. *If natural advantages are sufficiently more important to industry-A firms than Marshallian benefits are, and $\rho_{\mathbf{K}}^A$ is such that*

$$(\mathcal{B}_{(0,\rho_{\mathbf{K}}^A)}(k)) \cap \mathbf{Z} \neq \emptyset \quad \text{for each } k \in \mathbf{K}(A), \tag{3}$$

then a longer maximal radius for natural advantages spillovers $\rho_{\mathbf{K}}^A$ leads to a (weak) increase in the agglomeration level of industry A.

The first condition is a requirement that natural advantages be sufficiently important to the industry, and one sufficient condition is given in equation (10) of the theoretical appendix. The second condition, (3) in Proposition 3, requires that all attraction sites have been discovered. This second condition prevents situations where $\rho_{\mathbf{K}}^A$ is so small that some attraction sites are out

of range of potential firm locations. If (3) does not hold, increasing $\rho^A_{\mathbf{K}}$ can lead to discovery of new resources, causing formation of new clusters.

If these two conditions hold, expansion of the radius $\rho^A_{\mathbf{K}}$ both 1) expands the number of sites with access to the resources at attraction sites $k \in \mathbf{K}(A)$ and 2) grows the clusters $\mathcal{B}_{\rho^A_{\mathbf{K}}}(k)$ ($k \in K(A)$) at the peripheries where Marshallian forces act. So long as Marshallian forces outside of the clusters $\{\mathcal{B}_{(0,\rho^A_{\mathbf{K}})}(k)\}_{k \in \mathbf{K}(A)}$ do not dominate the draw of natural advantages, the result holds. Figure 2b illustrates this case, and the same empirical predictions regarding the relationships of cluster size, density, and number to the effective spillover length for an industry exist.

2.4 Empirical Measurement of Agglomeration

As a final step, we connect our theoretical framework to empirical metrics of agglomeration that utilize continuous bilateral distances among firms within an industry. This connection is important given that micro-interactions among firms are generally unobserved and must instead be inferred from realized firm locations. We thus need to build an apparatus to distinguish the agglomerative forces of interest from idiosyncrasies of economic geography.

Establishing this connection also has the nice corollary of showing how our model micro-founds the use of continuous agglomeration metrics like that of Duranton and Overman (2005).[11]

2.4.1 Duranton and Overman (2005)

Our empirical work in large part uses a slight variant of the Duranton and Overman (2005, hereafter DO) metric or its underlying smoothed kernel density. This discussion summarizes the DO methodology to show the connection to our theory. The empirical appendix further describes the DO metric and the empirical modifications required for our specific datasets.

The DO metric considers bilateral distances among establishments in an industry. The central calculation is the spatial density of industry A through a continuous function

$$\hat{K}_A(d) = \frac{1}{hN^A(N^A - 1)} \sum_{i=1}^{N^A-1} \sum_{i'=i+1}^{N^A} f\left(\frac{d - d_{j(i),j(i')}}{h}\right). \tag{4}$$

Here, as in our basic model set-up, $d_{j(i),j(i')}$ is the Euclidean distance between the spatial locations of establishments $j(i)$ and $j(i')$ within industry A. The double summation considers every pairwise bilateral distance within the industry analyzed (i.e., $N^A(N^A - 1)/2$ distances). Establishments receive equal weight, and the function f is a Gaussian kernel density function with bandwidth h that smooths the series.

The resulting density function provides a distribution of bilateral distances for establishments within an industry. Across all potential distances—ranging from firms being next door to each

---

other to being across the country from each other—this distribution sums to one. Smoothed density functions are calculated separately for each technology or industry analyzed. Industries where establishments tend to pack together tightly in cities, for example, are measured to have higher densities $\hat{K}_A(d)$ at short distance ranges.

While the density function is of direct interest, it is also important to compare the observed distributions of bilateral distances to general activity in the underlying economy. This comparison provides a basis for saying whether an industry's spatial concentration at a given distance is abnormal or not. Because the density functions for small industries with fewer plants are naturally more lumpy, these comparisons are specific to industry size. Operationally, comparisons are calculated through 1000 random draws of hypothetical industries of equivalent size to the focal industry A and repeating the density estimation. This procedure, which is further discussed in the empirical appendix, provides 5%/95% confidence bands for each industry and distance that we designate as $K_A^{LCI-U}(d)$ and $K_A^{LCI-L}(d)$.

Industry localization γ_A and dispersion ψ_A at distance d are defined using the DO formulae:

$$
\begin{aligned}
\gamma_A(d) &\equiv \max\left[\hat{K}_A(d) - K_A^{LCI-U}(d), 0\right] \\
\psi_A(d) &\equiv \max\left[K_A^{LCI-L}(d) - \hat{K}_A(d), 0\right] \text{ if } \gamma_A(d) = 0 \\
&\quad \text{and } 0 \text{ otherwise.}
\end{aligned}
\tag{5}
$$

Positive localization is observed when the kernel density exceeds the upper confidence band; similarly, positive dispersion occurs when the kernel density is below the lower confidence band. In between, an industry is said to be neither localized nor dispersed, and both metrics have a zero value. To allow for consistent and simple graphical presentation, we present a combined measure of localization and dispersion:

$$
\gamma_A^C(d) \equiv \gamma_A(d) - \psi_A(d).
\tag{6}
$$

An industry is neither localized nor dispersed at a given distance if its density is within the 5%/95% confidence bands. In such cases, $\gamma_A^C(d)$ has a value of zero. Excess density at distance d has a positive value, while abnormally low density carries a negative value. Our estimations analyze these local departures in a systematic manner across industries.[12]

2.4.2 Ordering and Characterizing Agglomerative Forces

Having introduced the DO metric, we now connect our model's focus on spillover lengths and cluster shapes to this empirical measurement. Before proceeding, however, it is helpful to recall that the structure required for Propositions 2 and 3 is minimal. These propositions build directly

---

[12] The empirical appendix further discusses the summation of localized agglomeration levels into aggregate statistics similar to DO's global indices. Our technique does not require this step, as we instead interpret local deviations directly (although one can use global metrics and their gradients as well). We thank Gilles Duranton for pointing this feature out to us.

on the agglomeration cluster theory to demonstrate that adjustments of the unobserved maximal radius affect cluster shapes and sizes in predictable ways.

The theory specifically suggests that longer maximal spillover radii are associated with fewer, larger, and less-dense clusters. It is feasible to use these observed traits in different industries to rank order the lengths of different spillovers. In empirical analyses below, for example, we provide suggestive evidence for the model by plotting an estimate of maximal radius for different technologies against a measure of the density of technology clusters over 50 miles. These tests, however, are partial representations of the complete predictions and in some sense ad hoc. They also do not provide a comparison to expected behavior similar to the DO confidence intervals.

With additional structure, we can jointly quantify these predictions within the DO framework. It is impossible to measure directly the G_{Ξ}^A functions that determine the value of firm clustering. However, observed spatial location patterns within industry A allow us to model the behavior of the unobserved functions G_{Ξ}^A via an observation that is immediate from the functional form (1) of $g_{j,A}$.

Proposition 4. *Holding $G_{\mathbf{K}}^A$ ($G_{\mathbf{A}}^A$) and ρ^A fixed, and assuming that the G_{Ξ}^A are differentiable functions, an increase in the magnitude $|(G_{\mathbf{A}}^A)'|$ ($|(G_{\mathbf{K}}^A)'|$) of the derivative $(G_{\mathbf{A}}^A)'$ ($(G_{\mathbf{K}}^A)'$) leads to a (weak) increase in the number of firms clustered at low distances.*

The decay of agglomerative forces across space correlates with observed distances between clustered firms. Thus, we may understand the speed at which the benefits of localization decay by measuring the degree of localization within industry A at different distances. For an extreme example, if localization of firms in industry A is constant across space, then we must have $|(G_{\mathbf{K}}^A)'| = |(G_{\mathbf{A}}^A)'| = 0$. If localization gradients are very sharp at short distances, then Proposition 4 implies that the underlying G_{Ξ}^A function sharply attenuates.

Proposition 4 provides the additional conditions—differentiability—which allow us to use DO density estimations to characterize distributions continuously. Adding this more continuous structure to our model, we can compare the full distributions of industries to assess how a longer maximal radius affects the shapes of clusters. The predictions that clusters become larger and less dense become jointly visible. Moreover, we can observe the influence using regular step sizes in distance.

Let $\mathbf{S}$ denote the set of sites occupied by firms in equilibrium, with many industries present in the economy. The null hypothesis is that neither agglomeration nor dispersion occurs in industry A at a given bilateral distance d when the maximal radius is ρ^A. This hypothesis is equivalent to the observation that firms in industry A locate randomly when a maximal radius ρ^A is assumed. Said differently, the null hypothesis is that $g_{j,A} = 0$ at a given bilateral distance d conditional upon the maximal radius being ρ^A.

We empirically proxy the set of potential sites $\mathbf{Z}$ with the observed set of actual sites $\mathbf{S}$ for all businesses. With this assumption, the DO density measures can quantify localization by

comparing observed localization levels to counterfactuals representing the underlying distribution of economic activity typical for a bilateral distance. The null hypothesis is rejected if the localized density of firms in industry A is a substantial departure from counterfactuals having the same number of firms as A and occupying sites randomly sampled from $\mathbf{S}$.[13]

2.4.3 Bilateral Distance Gradients in Agglomeration Clusters

There is an additional benefit to connecting our model to these continuous structures. We earlier noted that our empirical implication of smaller, denser clusters for a shorter maximal radius is equivalent to saying that the mean bilateral density for clusters declines. The model, however, has a much stronger implication for how spillover length influences the distribution of bilateral distances within clusters.

Proposition 5. *There is some $\bar{\rho} > (\underline{d}, \underline{d})$ so that whenever ρ and ρ' are such that $(\underline{d}, \underline{d}) < \rho < \rho' < \bar{\rho}$, then the mean intra-cluster firm distance is (weakly) smaller when the maximal radius is ρ than when it is ρ'.*

This result describes a key comparative static across spillover lengths. When comparing two industries, we earlier established that the industry with the shorter maximal radius should exhibit denser clusters such that very close bilateral distances are common. This proposition further identifies that this greater representation should be at its highest at the shortest bilateral distances possible (i.e., among locations that are located very near to each other). This higher frequency should then (weakly) decline as one considers bilateral distances further from the shortest possible connections.[14]

To provide intuition, first consider the impact of the marginal entrant on the bilateral distances in Figure 2a. As site X becomes part of the cluster, the kernel density (4) incorporates the bilateral distance from site X to every other populated site in the cluster into the spatial description. Some of the added bilateral distances are shorter than those that already existed in the cluster, with the distance between sites X and B, for example, being less than the distance between sites B and C. Yet, all of the additional bilateral distances are longer than the closest connections possible (e.g., those surrounding site C). Thus, as the cluster expands and becomes less dense, the relative impact on densities is highest at the shortest possible connections and proceeds (weakly) outwards for some distance.

An empirical example can also help. Assume that the premium for proximity is higher for investment bankers than it is for accountants. We predict that clusters of investment bankers should exhibit shorter mean bilateral distances among firms than clusters of accountants do. When comparing the spatial distributions of their clusters, Proposition 5 further indicates that

---

[13] As discussed in the empirical appendix and Barlet et al. (2009), this approach is slightly strained for the largest industries but is a reasonable baseline for most industries.

[14] The conditions of Proposition 5 indicate that this effect may disappear when the maximal radius is very large. This is a natural consequence of approaching a limit where the maximal radius is so long as to no longer influence cluster formation.

the greater density for investment banking should be at its highest at the spatial level of being in the same building or on the same block. When looking at firms being five blocks away from each other, the spatial density for investment bankers can still exceed that of accountants, but the difference should not be higher than it is when looking at being next door to each other.

This requirement micro-founds use of continuous density metrics like (4) to assess whether differences in agglomerative forces across industries yield meaningful deviations in agglomeration behavior. To summarize, we should empirically see that the greater density associated with a shorter maximal radius is at its maximum at the closest possible distance on the spatial scale (the intercept of DO) and (weakly) decline thereafter for some distance. Eventually, a distance is reached where the bilateral densities are the same even with the differences in maximal radius. Continuing with our earlier example, the offices of investment bankers and accountants may be equally represented when looking at firms that are ten city blocks apart.

After this point, a distance interval follows with relative under-representation for the cluster associated with the shorter maximal radius. Finally, once spatial distances are reached that represent distances between agglomeration clusters for Marshallian industries, the relative densities again converge. In our example, accounting firms should be more represented than investment bankers when looking at businesses 15-20 blocks from each other. This higher representation of accountants should then decline as we consider progressively longer distances that start to exceed the sizes of cities.

By contrast, our model generally does not make predictions for bilateral distances across Marshallian industries beyond the spatial horizons of individual clusters. The behavior of longer horizons depends upon the underlying distribution of cluster sites and it is thus ambiguous in our general framework. The median bilateral distance for all firms within an industry, for example, can increase or decrease with a longer maximal radius depending upon the spatial distances among the multiple, growing agglomeration clusters and the newly activated sites surrounding them. Our predictions at this scale instead relate to the distinctions between Marshallian spillovers and natural advantages.

2.4.4 Distinguishing between Marshallian and Natural Advantage Forces

Finally, we determine conditions under which we can empirically separate agglomeration due to Marshallian spillovers from that due to natural advantages. We generally do not observe the locations of natural attraction sites, hence it is not possible to distinguish these forces in every setting. Put another way, there is always some inherent limit to our observational capacity: we may not be able to determine whether firms are clustering around an unobserved natural advantage or each other.[15]

Analogous results have arisen in the work of Ellison and Glaeser (1997, 1999), DO, and

---

[15]It may be possible to elucidate this with a dynamic approach; some intuitions along these lines are given in the theoretical appendix.

others. Historically, this feature of agglomeration measurement has been viewed as an advantage, rather than a deficiency, since it means that localization levels are determined in a fashion unencumbered by distinctions between sources of agglomerative forces. The solution to the Ellison and Glaeser (1997) location choice model, which is the micro-foundation of their widely used metric, is the same in both scenarios. However, this indeterminacy is not desirable when analyzing and contrasting the determinants of agglomeration (e.g., Ellison and Glaeser 1999).

Unlike in previous work, our framework allows us to partially infer the trade-off between Marshallian forces and natural advantage dependency when attraction sites are sufficiently far apart.

Proposition 6. *Suppose that N^A is sufficiently small. Then, there is a constant Δ such that if the minimal distance between attraction sites for A is bounded below by Δ, and $\rho_{\mathbf{K}}^A$ and $G_{\mathbf{K}}^A$ are held fixed, scaling $G_{\mathbf{A}}^A$ by a constant $\alpha > 1$ will (weakly) increase agglomeration.*

Proposition 6 indicates that when natural advantages are sufficiently far away from each other, dense agglomeration in small industries is likely driven by Marshallian factors, while natural advantage dependency is likely to drive agglomeration observed at longer horizons. If coal mines or state colleges have sufficient distance between them, and an industry's location choices are heavily influenced by being near to one of these attraction sites, the cluster sites for this industry will generally be more dispersed than industries that are not influenced by these attraction sites. This observation is empirically confirmed below for a set of natural advantages.[16]

2.5   Our Empirical Approach

Our model connects the maximal radius of interaction for an industry with the shape, density, and number of clusters it forms. We test these predictions empirically in the next two sections. For clarity, we first describe how our empirical analyses fit together.

The graphical analysis in Figure 2a highlights that one approach for testing the model would be to identify marginal entrants into clusters by industry and infer from their relative spatial location the maximal radius. This would be particularly powerful as the technique could be applied to any setting. While quite interesting, multiple data constraints, measurement issues, and interpretation questions suggest this be left for future research. We thus take two simpler, more transparent approaches that separately test aspects of the model.

Our first test, presented in Section 3, uses patent citation data to measure the maximal radii that exist in different technology fields. This work closely links with technology or knowledge spillovers as a rationale for agglomeration. Returning to Figure 2a, our strategy is to use citation rates across bilateral distances *within* clusters to determine whether (and how) distant firms interact. By estimating the relative frequencies of B→A and B→C citations among the

---

[16]It is important to note that, while our general framework can encompass many varieties of natural advantages, this prediction focuses on more structured settings with minimum distances across attraction sites.

sites in Figure 2a, we can develop a reasonable proxy of the maximal radius. We can then investigate whether differences in maximal radius by technology produce the predicted variations in clustering traits.

Our second test in Section 4 is much broader in scope. We examine differences in agglomeration across many industries and sectors in the US economy. This broader approach means that we can no longer identify the micro-interactions among firms. We instead pursue a more reduced-form approach where we take from the agglomeration literature traits of industries that increase their desire to agglomerate for labor pooling or around natural advantage sites. We show how ordering industries by these traits produces the patterns predicted in our model.

3   Patent Technology Clusters

3.1   Patent Citations and Knowledge Spillovers

We employ individual records of patents granted by the United States Patent and Trademark Office (USPTO) from January 1975 to May 2009. Each patent record provides information about the invention (e.g., technology classification, firm or institution) and the inventors submitting the application (e.g., name, address). Hall et al. (2001) provide extensive details about these data, and Griliches (1990) surveys the use of patents as economic indicators of technology advancement. The data are extensive, with over eight million inventors and four million granted patents during this period.

A long literature exploits patent citations to measure knowledge diffusion or spillovers. A number of studies examine the importance of local proximity for scientific exchanges, generally finding that spatial proximity is an important determinant of knowledge flows.[17] Additional work links these local exchanges and economic clusters. Carlino et al. (2007) find that higher urban employment density is correlated with greater patenting per capita within cities. Rosenthal and Strange (2003) and Ellison et al. (2010) both find that intellectual spillovers are strongest at the very local levels of proximity. These empirical patterns closely link to ethnographic accounts of economic activity within clusters (e.g., Saxenian 1994).[18]

Patent citations thus offer us a unique opportunity to quantify differences in spillover radii and cluster shapes. It is important, however, to recall several boundaries of this approach. First, patent citations can reasonably proxy for technology exchanges, but there are many other forms of knowledge spillovers that may behave differently (e.g., Glaeser and Kahn 2001, Arzaghi and Henderson 2008). Second, several studies find that patent citations reflect Marshallian spillovers among firms other than pure knowledge exchange. Breschi and Lissoni (2009) closely

---

[17]See Jaffe et al. (1993, 2000), Thompson and Fox-Kean (2005), Thompson (2006), and Lychagin et al. (2010).

[18]Recent theoretical and empirical work further ties innovation breakthroughs to the clustering of activity around the discovery location, suggestive of very short spillover ranges (e.g., Zucker et al. 1996, Duranton 2007, Kerr 2010). These concepts are central to endogenous growth theory (e.g., Romer 1986), and Desmet and Rossi-Hansberg (2010) presents a recent model of spatial endogenous growth.

link citations to inventor mobility across neighboring firms in their sample, and Porter (1990) emphasizes how technologies embodied in products and machinery can be transferred directly through customer-supplier exchanges. Our measurements below may encompass these effects to the extent that they operate.[19]

3.2 Patent Data Construction

Inventors are required to cite the prior work on which their current patent builds. The total count of citations made by USPTO domestic and foreign patents granted after 1975 is about 41 million citations. We first restrict this sample to citations where the citing and cited patents are both applied for after 1975. This restriction is necessary for collecting inventor addresses. Our second restriction is that both patents have inventors resident in the United States at the time of the invention with identifiable cities and zip codes. About 15 million citations remain after these restrictions. Our primary dataset further focuses on the 4.3 million citations that are made in a geographical radius of 250 miles or shorter from the citing patent.

To identify these distances, we employ zip codes from addresses given for inventors. This dataset combines both zip codes listed directly on patents and representative zip codes taken from city addresses where zip codes are not listed. Where multiple inventors exist for a patent, we take the most frequent zip code; ties are further broken using the order of inventors listed on the patent. The spatial radius is defined using geographic centroids of zip codes and the Haversine flat earth formula. We assign a distance of less than one mile to cases where the citing and cited patents are in the same zip code.

Our analyses below consider how distances between zip codes influence patent citation rates. Two issues with using inventor zip codes should be noted. A small concern is that our approach does not consider all of the zip codes associated with inventors for some patents, and this may lead to mismeasurement in our distance measure over short spatial scales (specifically, an upward bias on the minimum distance). As a check against this concern, we find very similar results when instead employing only patents with single inventors.

More substantively, addresses listed for inventors can be either their home or work address. It would be nice to model both distances between work locations and distances between inventor home locations. Both of these distances can influence technology diffusion, and it is not clear which is more important. The patent data do not let us separate these two, however, and this measurement error biases us against finding shorter spillover effects.[20]

---

[19]As an additional note, the literature on intellectual spillovers is divided on whether the development of new innovations is most aided by having a large concentration of one's own industry or by industrial diversity. The view stressing industrial concentration is most often associated with Marshall, Arrow, and Romer (MAR). The MAR model emphasizes the benefits of concentrated industrial centers, particularly citing the gains in increasing returns and learning-by-doing that occur within industries. The second view, often associated with Jacobs (1970), argues that major innovations come when the ideas of one industry are brought into a new industrial sector. This perspective stresses that a wealth of industrial diversity is needed to create the cross-fertilization that leads to new ideas and entrepreneurial success. Duranton and Puga (2001) formalize theoretical foundations for this model. Our empirical analysis focuses solely on the MAR determinant.

[20]While we manually identified work locations for the description of the San Francisco Bay Area in the intro-

3.3   Generations of Citations within Clusters

Our first analysis examines differences in spatial scope within clusters for first generation citations compared to later generations of citations. This analysis is useful because it provides evidence for the interconnections among firms built into our model's structure. It also introduces the empirical framework that we use to calculate the maximal radius for each technology.

To introduce the terminology, a first generation citation is simply patent A (our focal patent) citing patent B. A second generation analysis analyzes the distances from patent A to a patent C cited by patent B. Third generation citations are similarly calculated through the citations made by patent C, and so on. If spillovers are not fully transmissible through a cluster,[21] the maximal radius model predicts that first generation citations will be spatially nearer than later generation citations within clusters. This framework thus offers evidence about whether technology movements through clusters conform to our model's structure, although other models may generate similar predictions.

We undertake this test using patent citations among bilateral pairs of zip codes that are within 150 miles of each other. To allow for maximum flexibility in the estimations, we construct a series of distance rings indexed by dr. Our distance rings start with a ring of one mile or less, which includes all citations that occur in the same zip code or between close neighbors. The next ring is one mile to five miles, and subsequent rings are every five miles to 50 miles. Denote the set of distance rings as DR.

Our empirical specification takes the form

$$\ln(\text{Citations}_{z_1 \to z_2}) = \sum_{dr \in DR} \beta_{dr} \cdot I(d_{z_1,z_2} = dr) + \gamma \cdot \ln(\text{Patents}_{z_2}) + \phi_{z_1} + \varepsilon_{z_1 \to z_2}, \qquad (7)$$

where z_1 indexes source zip codes and z_2 indexes destination zip codes. The regressors of interest are the indicator variables $I(\cdot)$ for each distance ring. ϕ_{z_1} is a vector of citing zip code fixed effects that controls for the overall citation rates by the source zip code. $\ln(\text{Patents}_{z_2})$ controls for the level of patenting in the destination zip code, which naturally influences the likelihood of citations occurring. We drop zip codes for which patents and citations are never observed locally. This regression thus describes the deviations in citation rates at different distance bands from what the spatial distribution of patents would predict. Zip code pairs with distances of 50

---

duction, we are unable to replicate this approach for the full patent dataset. This is partly due to time constraints, but more importantly it becomes increasingly difficult to assign inventions to facilities in multi-unit firms. The San Francisco Bay Area is a nice case in that single locations typically house both corporate headquarters and innovation facilities. This collocation is much less prevalent in the New York City region, for example, where major firms frequently have offices in Manhattan and in surrounding areas. These multiple offices even within 250 mile circles limit the gain from this approach versus simply using known inventor addresses. Aarland et al. (2007) study the separation of firm activity into multiple locations.

[21]Marshallian spillovers are *fully transmissible* through a cluster if all sites in a cluster receive identical spillover benefits from one another, independent of location. In Figure 2a, this would result in firms at sites A and B receiving equal benefits from a firm at site C, even though site C is only within the maximal radius of interaction of site A. The intuition is that the spillovers of site C are "transmitted" to the firm at site B through the firm at site A.

to 150 miles serve as the reference group.[22]

Figure 3a plots the β_{dr} coefficients. First generation citations are remarkably concentrated at short distances. There is a particular spike at less than one mile, which includes citations within the same zip code. Over half of this excess effect is gone when looking at zip codes one to five miles away, which connects to the very localized networking of Arzaghi and Henderson (2008) and the spillover estimations of Rosenthal and Strange (2003, 2008). The additional decline to zip codes five to ten miles away is very modest, after which a steeper decline ensues. A little more than half of the excess effect evident at one to five miles is gone once 20-25 miles is reached. Coefficients for first generation citations are statistically different from zero at all distances.

Figure 3a also graphs the spatial patterns of subsequent generations of patent citations. To construct these later generations, we randomly select one subsequent citation at each generation per cited patent. We require that newly selected citations be within 250 miles of the citing patent to maintain geographic proximity, but not necessarily within 250 miles of the original citing patent. The specification (7) is again used to compare rates in local distances to the rates that exist over 50-150 miles.

The results are intuitive and agree with the developed model. Earlier generations, or more immediate interactions, are significantly stronger at less than one mile than are later generations. Differences at one to five miles, however, are much more modest. These relative differences slowly decay from five miles onward. At longer distances, the spatial overlaps become very similar across generations.

To provide a second baseline of comparison, we generate random citation pairs comparable to our observed sample. For every patent cited, we randomly draw a patent with a similar patent class and application year. This method has been used extensively in the literature; we add the innovation that the new counterfactual citation must be within a 250 mile radius of the original citing patent. We do not exclude the original cited patent from the draws, and thus we use the original citation if there are no other patents with the same technology and application year in the defined spatial radius. This approach allows us to focus on local geographic areas and provide a rough comparison of underlying patent densities with similar technology traits.

The random citation counterfactual series is significantly different from zero at short distances. This suggests that firms with closely related technologies locate near each other even within spatial distances of 150 miles or less. Thus, the citation premiums evident at short distances combine both a higher likelihood to cite proximate neighbors and a higher likelihood to cite generally firms in closely related technologies. While our basic model is interested in both of these mechanisms, we can derive further insights into the structure of interconnections within clusters by comparing our citation premiums by generation with these counterfactuals.

---

[22]The total sample count for these estimations is 429,400 zip code pairs with non-zero citation counts. We find very similar results when collapsing destination zip codes into consolidated rings around citing zip codes. Our core specification drops zero citation cells due to the log transformation. We find similar results when instead recoding zero citations to be one citation for cells where patenting exists but a citation does not occur.

The citation premiums for first and second generation citations at distances of one mile or less are substantially stronger than those of the random counterfactuals. These deviations gradually decay as distances between locations within clusters lengthen, with the initial decay for first generation interactions again being much more rapid. These results confirm that observed citations are not simply following from underlying spatial distributions of technologies. Moreover, even against the random baseline, initial interactions among firms are most important at very localized distances and decline thereafter.

The contrast with third and fourth generation citations is also very informative. Third generation citations are identical to the random counterfactual at distances of one mile or less. They then achieve relative importance between distances of five and 20 miles, before converging again to the random counterfactuals. Thus, extended chains of citations only mirror localized technology distributions at very short spatial scales, but the interconnected hops do aid accessing technologies that are farther away in the cluster. Finally, fourth generation citations are under-represented at distances of less than one mile, whereas they mirror the random baseline thereafter. This suggests that more than three interconnections do not provide additional access in technology clusters for the first 50 miles compared to 50-150 miles.[23]

These differences across citation generations suggest that knowledge spillovers are not fully transmissible through a cluster, but instead follow a pattern indicated by the Silicon Valley example and our model's structure. In Figure 2a, the chain of interconnected hops B→A→C aids site B's access to spillovers from sites around site C. Moreover, the extra strength for first generation citations over very short distances offers a path for identifying maximal radii of firm interactions that we investigate next.

3.4 Maximal Radius and Spatial Cluster Patterns

Our theory connects the maximal radius of firm interactions with cluster structure. We test these predictions by looking at differences across patent technologies in their citation behaviors similar to Figure 3a. We analyze 36 technologies using the sub-category level of the USPTO system and Hall et al. (2001). Examples include Semiconductors, Optics, and Resins.

3.4.1 Basic Calculations and Descriptive Evidence

We proxy the maximal radius of interaction for each technology through the citation patterns evident among patents within that technology. We again measure how close citing patents are within individual clusters. One technology, for example, may show that most of the citations that exist within local areas occur across firms with a bilateral distance of ten miles or less. On the other hand, a second technology's local citations could occur more evenly over distances

---

[23]The patterns in Figure 3a allow for citations within firms so that later generations of citations and the random counterfactual are simple to define. Excluding citations within firms, the citation premium for first generation citations at one mile or less is 1.8. The subsequent patterns look very similar. We exclude citations within firms in all subsequent citation analyses.

0-70 miles. In the context of Figure 3a, this second industry would have a much flatter citation premium for short distances. While we cannot put an exact distance on each technology's maximal radius, we can use the differences across technologies in these observable citation patterns to proxy relative differences in their maximal radii.

Operationally, we estimate for each technology a first generation citation regression similar to that undertaken for all citations in Figure 3a. To develop a single measure of short technology exchanges, however, we only include indicator variables for bilateral distances of [0,10] miles and (10,30] miles. Thus, the reference group is bilateral zip code pairs of distances between 30-150 miles. Our measure of technology spillover horizons is the observed premium over the first ten miles compared to this reference group. We exclude within-firm citations from these calculations.

A shorter maximal radius is associated with a greater citation premium for the first ten miles. The average premium is 0.43. Semiconductors has the greatest value of 0.74 followed by Surgery & Medical Instruments at 0.53. Values of less than 0.3 are found in Gas and Pipes & Joints.[24]

For each technology, we estimate the continuous DO spatial density metric described above. Figure 3b provides descriptive evidence on patent cluster shapes. We group our 36 technologies into three broad buckets based upon the categories of the USPTO system following Hall et al. (2001): Chemicals, Pharmaceuticals, and Medical (categories 1 and 3), Computers, Communications, Electrical, and Electronics (2 and 4), and Mechanical and Miscellaneous (5 and 6). The chart simply provides the fraction of technologies considered localized by distance within each grouping using the metric (5).[25]

Every technology within the Computers/Electronics grouping shows abnormally high spatial concentration over the first 30 miles. After 35 miles, however, localization within this group decays rapidly and is mostly gone by 70 miles or thereabouts. On the other hand, the Chemicals/Medical grouping shows abnormally high spatial concentration over 30-60 miles, with a much slower decay rate thereafter. Finally, there are no material variations by distance in the number of technologies localized for the Mechanical/Miscellaneous grouping.

These patterns roughly conform with our predictions, as the unweighted average citation premium for technologies in Computers/Electronics (0.48) is greater than in Chemicals/Medical (0.43) and then Mechanical/Miscellaneous (0.39). These differences indicate that the maximal radius for technology spillovers tends to be shortest in Computers/Electronics, followed by Chemicals/Medical and Mechanical/Miscellaneous. Greater requirements for very close knowledge exchange are visibly associated with shorter, denser spatial clusters. This description,

---

[24]We find similar results to those described below when using other techniques to calculate these spillover horizons. Perhaps most interestingly, the patterns are the same when estimating the distance interval required to halve the citation premium evident at one to five miles compared to 70-75 miles or 145-150 miles. These half-life calculations mirror typical decay function estimations. The reported technique with the localized citation premium over 0-10 miles has the advantage of requiring much less data to compute consistently. This efficiency becomes critical when we estimate these patterns at the technology level in the United Kingdom, where we have only about a tenth of the patents available in the United States.

[25]Distances are calculated using zip code centroids. Computational limitations, primarily around constructing the counterfactuals, require that we calculate these densities using patents from 1990-1999. We calculate very similar densities for a few smaller technologies when instead considering 1975-2009.

however, does not take advantage of the heterogeneity within groups or the intensity of agglomeration, to which we turn next.[26]

3.4.2 Cross-Sectional Density Plots

Figure 3c provides a cross-sectional plot of cluster density and our proxy for maximal radius by technology. Density is measured by the share of bilateral distances among patents for a technology over 0-50 miles divided by the share of 0-150 miles. Shares range from 30% to over 80%, with a very high share indicating that patents in the technology are very densely packed in one cluster and then mostly absent until the next cluster.

There is a strong upward relationship between greater citation premiums (shorter maximal radii) and greater patent density. Recall that citation premiums are calculated controlling for the underlying spatial patent distribution, so this relationship is not mechanical. The slope of the trend line is 0.875 (0.163). Very clearly, some of the IT industries show exceptional densities. The slope of the trend line is 0.336 (0.075) when capping the density ratio at 50%.[27]

While this plot suggests a strong correlation between lengths of micro-interactions among firms and cluster density at the technology level, it is natural to worry about reverse causality. Existing cluster shapes and economic geography likely influence citation behavior. Moreover, technology clusters may have their spatial locations for unmodeled reasons (e.g., historical accidents, fixed university locations). The length of patent citations could then be determined by the geographical features of these locations.

To address this, we calculate similar citation premiums using patent data from the United Kingdom. Ellison et al. (2010) introduce this technique and discuss its strengths and limitations. We geocode all city names and postal codes associated with UK inventors. We also manually search for addresses of firms in the United Kingdom with more than fifty USPTO patents to provide more accurate city assignments. Calculating bilateral distances among pairwise city combinations, we then estimate a second set of technology-level citation regressions that parallels our US estimations.

The UK calculations face several important limitations relative to the US calculations. First, and most important, there are significantly fewer data points to estimate these citation premiums (the UK sample is less than a tenth of the US sample size). Second, the geocoding has greater measurement error, perhaps most centrally dealing with London. Finally, there is some selection bias in the set of UK inventors applying for US patent protection. While these limitations

---

[26] At first it may appear odd that a majority of technologies are deemed localized when the confidence bands are selected such that only 5% of the counterfactuals reach them. This is to be expected if agglomerative forces exist, however, as the counterfactuals build upon all patent locations. The counterfactuals are not selected such that only 5% of technologies will be deemed agglomerated. This levels effect for localization, along with its overall decline with distance, is predicted by our model if sites are distributed uniformly but agglomerative forces exist in nearly all technologies. The same pattern is shown in the industrial data in the next section.

[27] Five technology sub-categories are residual groups (e.g., Miscellaneous Drugs) where consistent clustering concepts may not apply. Our results generally strengthen when excluding these groups, but we keep them in the presented results.

restrict our analysis somewhat, the UK results in this section and the next provide important confirmation of our model's predictions in a manner that addresses reverse causality concerns.

Figure 3d provides a cross-sectional plot of US cluster density against UK citation premiums. The vertical axis is the same as Figure 3c, but we substitute the UK citation premiums for the horizontal axis. Interestingly, the UK citation premiums are weaker than in the United States on average (mean of 0.24 versus 0.43). This decline may reflect measurement error in distances or more interesting economic differences. We do not further investigate, however, as we are primarily interested in relative differences across technologies as a proxy for differences in maximal radii.

There are some substantial changes among some IT industries in Figure 3d compared to Figure 3c. Most noticeably, semiconductors declines substantially (i.e., its maximal radius is not as short as we measure in the US data). Nevertheless, the positive relationship persists with cluster density in the United States. A greater citation premium for a technology in the United Kingdom, corresponding to a shorter maximal radius, links to a denser cluster in the United States. The slope of the trend line is 0.230 (0.075). The slope of the trend line is 0.159 (0.031) when capping the US density ratio at 50%. These patterns provide confidence that these relationships are not being solely determined by unmodeled factors.

3.4.3 Complete Density Plots

While transparent, the cross-sectional plots are incomplete in that they do not describe the full distribution of location behavior. They also do not account for differences in industry size, which can have a mechanical effect on density estimates. We now use the DO methodology to describe these patterns more completely.

We begin with the kernel density $\hat{K}_A(d)$ defined in (4) for a technology A. The process of assigning localization (5) involves non-monotonic transformations of the data, and it is thus useful to view the simpler density functions first. With some abuse of notation, we define $\hat{K}_{A,d}$ as the sum of the kernel density over five mile increments starting from zero to five miles and extending to 245-250 miles. We again index distance rings with dr and denote the set of distance rings as DR, although the distance rings are different from the citations analysis.

Figures 3e and 3f present coefficients from empirical specifications of the form

$$\hat{K}_{A,d} = \sum_{dr \in DR} \beta_{dr} \cdot I(d = dr) \cdot SpilloverDensity_A + \phi_d + \varepsilon_{A,d}. \tag{8}$$

These estimations provide a continuous description of how technology cluster shapes vary with technology horizons. $SpilloverDensity_A$ is the technology spillover density for industry A calculated through patent citations (i.e., the same as the horizontal axis values in Figures 3c and 3d). Greater values of $SpilloverDensity_A$ correspond to a shorter maximal radius ρ^A in our model, and we thus anticipate finding denser and smaller clusters for these technologies. We

24

transform $\hat{K}_{A,d}$ and $SpilloverDensity_A$ to have unit standard deviation to aid interpretation, and we evaluate β_{dr} at each distance ring.

A vector of distance fixed effects ϕ_d controls for typical agglomeration densities by distance. They thus directly account for the overall spatial density of patenting so that our estimations consider differences across technologies. As the vector of distances fully contains the support of distances, we do not include a main effect for $SpilloverDensity_A$. Higher values of β_{dr} indicate that technologies with high spillover densities show greater spatial density at that distance. The cross of 51 distances and 36 technologies yields 1836 observations per estimation.

Figures 3e and 3f present these density estimations using the US and UK measures of $SpilloverDensity_A$, respectively. Triangles report β_{dr} coefficients. The dashed lines provide 90% confidence bands with standard errors clustered by technology. We find similar precision when bootstrapping standard errors.

Technologies with greater $SpilloverDensity_A$, or a shorter maximal radius, are substantially more agglomerated at very short distance horizons. A standard deviation increase in $SpilloverDensity_A$ is associated with a 1.3 standard deviation increase in agglomeration at five miles using the US measure; the UK-based estimate is 0.6 standard deviations. By 60-75 miles, the abnormal spatial concentration is no longer statistically different from zero.

Looking further, technologies with greater $SpilloverDensity_A$ are under-represented after 75 miles or thereabouts. Using the US estimate of citation density, these clusters show an abnormal lack of density from 80 to 185 miles that is statistically different from zero at every five mile increment. The UK estimation shows a similar pattern, although its point estimates are not statistically different from zero. In both cases, the point estimates converge to zero as distances approach 250 miles. At the edge of this spatial scale, differences in maximal radius are not systematically associated with different agglomeration intensities.

These patterns closely follow our theory model and the predictions given in Section 2.4.3 regarding maximal radii and cluster shapes. Note that the patterns of over-representation followed by under-representation are not mechanical. Other attributes, for example, could predict higher spatial concentration for a technology at all spatial distances to 250 miles.[28]

Figures 3g and 3h take the next step of calculating localized deviations from industry specific confidence intervals using (6). The patterns are very similar to Figures 3e and 3f. The excess density at very short spatial horizons is robustly different from the random counterfactuals and very similar to the kernel plots. The abnormally low spatial concentration at moderate spatial horizons is weaker than in the raw kernel density plots, with the US estimator significantly different from zero over a narrower range of 120-155 miles. The UK estimator does not show material deviations at moderate horizons once the confidence intervals are considered.

Overall, these figures jointly illustrate our central model predictions. A shorter maximal

---

[28] The kernel density functions (4) sum to one over the support of all bilateral distances in US, stretching from next door to several thousand miles. This does not materially influence the cluster descriptions we develop here over the first 250 miles.

radius, or greater spillover density, is very strongly associated with greater agglomeration at very short spatial horizons (i.e., the cluster is denser). These same technologies tend to be under-represented at moderate spatial horizons (i.e., the clusters are smaller). The latter result is very strong in the raw US data, and it is partially confirmed with the UK estimator. Moreover, in all cases the initial decline in bilateral densities from the closest feasible values that is predicted by Proposition 5 is robustly supported.[29]

4   Marshallian Spillovers and Natural Advantages

This section uses data from the US Census Bureau to more broadly examine industry agglomeration in the United States. We first outline the US Census Bureau data employed. We then compute spatial densities using bilateral distances among establishments by industry. We finally link differences in spatial agglomeration patterns to differences across industries in Marshallian spillovers and dependency on natural advantages.

4.1   Spatial Densities for Establishments

Our estimates of industrial agglomeration patterns are developed through confidential data housed by the US Census Bureau. The Longitudinal Business Database (LBD) provides annual observations for every private-sector establishment with payroll from 1976 onward. The Census Bureau data are an unparalleled laboratory for studying the industrial structure of US firms. Sourced from US tax records and Census Bureau surveys, the micro-records document the universe of establishments and firms rather than a stratified random sample or published aggregate tabulations. In addition, the LBD lists physical locations of establishments rather than locations of incorporation, circumventing issues related to higher legal incorporations in states like Delaware. We define 362 industries through the three-digit level of the 1987 Standard Industrial Classification (SIC3). Industry assignment is at the establishment level for multi-unit firms. The data include 108 million workers and 5.8 million establishments in 1997.

We use bilateral distances among establishments to create continuous density metrics similar to those described for our patent data. Our primary metric weights each establishment equally, and we have confirmed that an employment-weighted index yields similar results. Localization is typically higher in employment-weighted distributions than in those using plant counts (e.g., Holmes and Stevens 2002). We measure bilateral distances using distances between county centroids. Establishments in the same county are given a uniform distance of one mile.[30]

---

[29]Similar results are found using three additional specification variants. The first employs the density function (4) and introduces the confidence bands $K_A^{LCI-U}(d)$ and $K_A^{LCI-L}(d)$ as precision controls. The second calculates a global index similar to DO's main metric and then evaluates the gradient of this concentration measure across distances. Finally, the DO confidence bands can be adjusted to a 1%/99% significance level.

[30]Multiple factors yield conflicting recommendations regarding this latter choice—for example, accounting for larger land areas versus congestion effects with higher urban density. As shown below, the major movements we analyze in this section happen at distances of 50 miles or greater, making this choice of secondary importance.

Figure 4a graphs the fraction of industries that are localized at each distance level, akin to Figure 3b. We analyze 25 mile increments to 1000 miles. Approximately 60% of industries display positive localization at each increment to 125 miles. Of the remainder, about 30% display dispersion and 10% show neither pattern. Localization attenuates over moderate distances, and only 30% of industries are localized at distances of 900-1000 miles. Manufacturing industries are more agglomerated at short spatial scales than non-manufacturing industries, a pattern that reverses at longer horizons. This result is to be expected, given the transportable and tradeable nature of manufactured products. The empirical appendix provides additional descriptive statistics.

4.2 Agglomerative Forces

We develop two traits of industries to model the Marshallian agglomeration and natural advantages forces analyzed in our theoretical model. For the former, we focus on labor pooling as a rationale for firm clustering. Multiple models specifically describe the benefits of firm clustering due to specialized workforces.[31] We focus on this feature, recognizing that other forms of Marshallian interactions exist (e.g., Rosenthal and Strange 2004, Ellison et al. 2010).

We measure how concentrated or specialized the occupational needs of an industry are through the 1999 National Industry-Occupation Employment Matrix (NIOEM) published by the Bureau of Labor Statistics. The NIOEM provides industry-level employments in over 700 occupations. We calculate the occupational concentration for each industry as a Herfindahl-Hirschman Index (HHI) of employments across occupations. Higher values of this index suggest greater concentration of workers in a small number of occupations, which increases the spillover benefits of industrial agglomeration.

Figure 4b provides some simple evidence that this labor concentration relates to spatial scales. We divide industries into terciles based upon their occupational concentration. Highly concentrated industries are more likely to be localized at shorter spatial scales; they are less likely to exhibit localization at longer spatial horizons. Figure 4b's trends are for all industries, and the same patterns hold both within and outside of manufacturing.

We next model natural advantages for manufacturing industries using data from the Census of Manufacturers. There are many natural advantages that one could model, and we pick five for this exercise: coal inputs, natural gas inputs, lumber inputs, agricultural and livestock inputs, and dependency upon water transportation. We normalize each of these dependencies

---

[31]Marshall (1920) described how an agglomeration of workers and firms shields workers from firm-specific shocks. Workers can be more productive and better insured by moving from firms that are hit with negative shocks to better opportunities (e.g., Diamond and Simon 1990, Krugman 1991, Overman and Puga 2010). Larger labor pools further promote more efficient matches (e.g., Helsley and Strange 1990), and multiple firms protect workers against ex post appropriation of investments in human capital (e.g., Rotemberg and Saloner 2000). All of these mechanisms suggest that firms that employ similar types of workers will tend to locate near one another and that firms will benefit from thick local markets for their specific labor needs. Glaeser and Kerr (2009) provide a further discussion of this pooling concept. Duranton and Puga (2004) and Glaeser (2008) provide a general theoretical taxonomy of agglomeration determinants.

by industry shipments, and we define a single natural advantages metric for each industry as its maximum dependency across these five traits. Some of the largest dependencies on natural advantages are found in food processing, pulp and paper, and steel industries.

Figure 4c again provides simple evidence that natural advantages are systematically related to spatial scales. We divide manufacturing industries into terciles. In contrast to Figure 4b, the most dependent industries on natural advantages are the least likely to be abnormally concentrated at small spatial distances. This dependency instead exhibits itself over horizons of 250 to 700 miles. The industries most concentrated to 200 miles are those with the least dependency on natural advantages. The rates of decay at the edges for natural advantages also appear more gradual.

4.3   Labor Pooling and Spatial Cluster Patterns

To more formally test these patterns, we run a series of regressions similar to the patent regression model (8). We organize the data by industries A and distances d. We use 40 distance intervals, ranging from 0-25 miles to 975-1000 miles, and 353 industries for a total of 14,120 observations. We exclude nine industries from our general sample of 362 industries for which we cannot calculate all of the explanatory variables.

We present the patterns using the basic kernel estimator (4); the results are very similar with the other variants studied in Section 4. We again transform variables to have unit standard deviation to aid in interpretation. The size variation of these industries is substantially larger than among patent technologies.[32] To provide the most representative depictions, we conservatively cap extreme densities at their 1% and 99% values by distance and weight by log employment of the industry. These specification steps are not very important, however, and the patterns we describe are very robust across specification variants.

Figure 4d first tests occupational concentration across all industries. Higher occupational concentration promotes localization over the first 100 miles, with substantial and monotonic attenuation thereafter. By 250 miles, higher occupational concentration is no longer associated with greater localization at a 90% confidence level. The magnitudes are important in economic size. A standard deviation increase in occupational concentration is associated with a 0.6 standard deviation increase in localization to 100 miles. It is also clear that the total length of the agglomeration cluster associated with labor pooling is substantially longer than typical commuting distances (which are frequently less than 40 minutes), akin to the San Francisco example in Figure 1b.

We have confirmed this pattern using several other techniques for measuring occupational concentration or specialization. We find similar results when calculating the concentration ratio of employment in the most frequent occupations or in the top four occupations. Similarly, we find

---

[32]Counts of patents used in our estimations range from 2811 to over 66,000. Among the industrial data, 18 industries have fewer than 100 establishments, while the maximum is over 400,000 establishments.

the opposite pattern of results when modeling the number of occupations listed. Industries with more listed occupations, regardless of employment levels, are associated with greater dispersion. We also find that the results are invariant to including other distance-based interactions for average wage, college share, labor intensity, and the number of firms in the industry.

To check that reverse causality is not the case for our labor measures, we calculate the corresponding occupational concentration of manufacturing industries in the United Kingdom. We use data from the UK Labour Force Survey, which is similar to the US's Current Population Survey. The patterns in Figure 4e are very similar. This suggests that our results are not due to endogenous selections of production techniques by spatially isolated industries to match local worker environments.[33] A second advantage of the UK test is that the occupational concentrations are calculated from a second data source, suggesting that our outcomes are not dependent upon idiosyncratic features of the BLS's NIOEM.

The shapes of these distributions mirror Figures 3e-3h. It is important to note, however, the more subtle differences in the estimation. A central strength of the patent data is the capacity to model directly the micro-level interactions of firms to calculate a measure of the maximum radii. While the Census Bureau data allows us to consider many more industries and sectors, this broader dataset requires a reduced-form approach that models traits of industries.

4.4 Natural Advantages and Spatial Cluster Patterns

Figure 4f considers natural advantages within the manufacturing sector. These estimations control for occupational concentration, with the unreported coefficients being very similar to Figure 4d. The contrast between natural advantages and labor agglomeration is very interesting and supports the crossing property of our model described in Proposition 6. High dependency upon natural advantages is associated with dispersion (a lack of density) at short spatial scales. This follows from coal mines, coastal ports, and similar features naturally having some distance between them. The effective radius for transporting goods and material inputs is also longer than for labor pooling.

Over the next 500 miles, however, the importance of natural advantages grows. By distances of 400 miles, the agglomerative force of natural advantages exceeds that of labor, and remains strong until about 800 miles. Very clearly, the spatial dimension of natural advantages is longer, and it appears that the attenuation rate is weaker, than for the Marshallian agglomeration proxied by our labor features. This is most evident in the modest gradients and long spatial durations of the slope in Figure 4f. The shapes of these densities conform to our model's predictions.

---

[33]Several theoretical models describe spatial location choice and production technology as jointly determined (e.g., Duranton and Puga 2001, Glaeser et al. 2010).

5   Conclusion

This paper introduces a new model of location choice and agglomeration behavior. From a simple and general framework, we show that agglomeration clusters generally cover a substantially larger area than the micro-interactions upon which they build. In turn, agglomerative forces with longer micro-interactions are associated with fewer, larger, and less-dense clusters. The theory thereby provides a basis for the use of continuous agglomeration metrics that build upon bilateral distances among firms. The theory also rationalizes the use of observable cluster shapes and sizes to rank order the lengths of underlying agglomerative forces. We find confirmation of our theoretical predictions using both variations across patent technology clusters and variations across industry concentrations.

We hope that our theoretical framework proves an attractive model for incorporating additional factors that influence firm location and agglomeration behavior. Important extensions include: modelling the dynamics of industry life-cycles, incorporating interactions across firms in different industries, allowing for pricing of locations, and incorporating the development of new sites. We likewise believe our setting is an attractive laboratory for structural modelling that would allow recovery of the underlying lengths of micro-interactions. These parameters could in turn be useful for understanding spillover transmissions in networks and for studying spatial propagation of economic shocks.

We have applied our framework to two empirical settings, but we believe that many more applications in industrial agglomeration are possible. For example, future work can look to price clusters or identify spillover lengths by examining the location decisions of marginal entrants. Our framework highlights the important information that is contained in these indifference conditions if properly identified. As important, we believe our framework describes interactions in many other contexts as well. For example, studies find that knowledge flows within firms or universities are substantially shaped by the physical layout of facilities (e.g., Liu 2010). We hope that future work similarly analyzes parallel situations where costs of interaction generate a maximal radius.

References

Alcacer, Juan, and Wilbur Chung, "Location Strategies and Knowledge Spillovers", *Management Science* 53:5 (2007), 760-776.

Alfaro, Laura, and Maggie Chen, "The Global Agglomeration of Multinational Firms", HBS Working Paper 10-043 (2010).

Aarland, Kristin, James Davis, J. Vernon Henderson, and Yukako Ono, "Spatial Organization of Firms: The Decision to Split Production and Administration", *RAND Journal of Economics* 38:2 (2007), 480-494.

Akcigit, Ufuk, and William Kerr, "Growth through Heterogeneous Innovations", NBER Working Paper 16443 (2010).

Arzaghi, Mohammad, and J. Vernon Henderson, "Networking off Madison Avenue", *Review of Economic Studies* 75:4 (2008), 1011-1038.

Audretsch, David, and Maryann Feldman, "R&D Spillovers and the Geography of Innovation and Production", *American Economic Review* 86 (1996), 630-640.

Balasubramanian, Natarajan, and Jagadeesh Sivadasan, "What Happens When Firms Patent? New Evidence from US Economic Census Data", *Review of Economics and Statistics* (2010), forthcoming.

Barlet, Muriel, Anthony Briant, and Lauire Crusson, "Location Patterns of Services in France: A Distance-Based Approach", Working Paper (2009).

Baum-Snow, Nathaniel, "Did Highways Cause Suburbanization", *Quarterly Journal of Economics* 122:2 (2007), 775-805.

Baum-Snow, Nathaniel, "Changes in Transportation Infrastructure and Commuting Patterns in U.S. Metropolitan Areas, 1960-2000", *American Economic Review Papers and Proceedings* 100 (2010), 378-382.

Behrens, Kristian, Gilles Duranton, and Frederic Robert-Nicoud, "Productive Cities: Sorting, Selection, and Agglomeration", Working Paper (2010).

Billings, Stephen, and Erik Johnson, "A Test for Agglomeration with Location Assignment", Working Paper (2010).

Bleakley, Hoyt, and Jeffrey Lin, "Portage: Path Dependence and Increasing Returns in U.S. History", Working Paper (2010).

Breschi, Stefano, and Francesco Lissoni, "Mobility of Skilled Workers and Co-invention Networks: An Anatomy of Localized Knowledge Flows", *Journal of Economic Geography* 9 (2009), 439-468.

Carlino, Gerald, Satyajit Chatterjee, and Robert Hunt, "Urban Density and the Rate of Invention", *Journal of Urban Economics* 61 (2007), 389-419.

Ciccone, Antonio, and Robert Hall, "Productivity and the Density of Economic Activity", *American Economic Review* 86:1 (1996), 54-70.

Dauth, Wolfgang, "The Mysteries of the Trade: Interindustry Spillovers in Cities", Working Paper (2010).

Delgado, Mercedes, Michael Porter, and Scott Stern, "Convergence, Clusters and Economic Performance", Working Paper (2009).

Desmet, Klaus, and Esteban Rossi-Hansberg, "Spatial Development", Working Paper (2010).

Diamond, Charles, and Chris Simon, "Industrial Specialization and the Returns to Labor", *Journal of Labor Economics* 8 (1990), 175-201.

Duranton, Gilles, "Urban Evolutions: The Fast, the Slow, and the Still", *American Economic Review* 97:1 (2007), 197-221.

Duranton, Gilles, and Henry Overman, "Testing for Localization Using Micro-Geographic Data", *Review of Economic Studies* 72 (2005), 1077-1106.

Duranton, Gilles, and Henry Overman, "Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data", *Journal of Regional Science* 48:1 (2008), 313-343.

Duranton, Gilles, and Diego Puga, "Nursery Cities: Urban Diversity, Process Innovation, and the Life Cycle of Products", *American Economic Review* 91:5 (2001), 1454-1477.

Duranton, Gilles, and Diego Puga, "Micro-foundations of Urban Agglomeration Economies", in J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4* (Amsterdam: North-Holland, 2004), 2063-2117.

Ellison, Glenn, and Edward Glaeser, "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy* 105 (1997), 889-927.

Ellison, Glenn, and Edward Glaeser, "The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?", *American Economic Review Papers and Proceedings* 89 (1999), 311-316.

Ellison, Glenn, Edward Glaeser, and William Kerr, "What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns", *American Economic Review* 100 (2010), 1195-1213.

Fallick, Bruce, Charles Fleischman, and James Rebitzer, "Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster", *Review of Economics and Statistics* 88:3 (2006), 472-481.

Fu, Shihe, and Stephen Ross, "Wage Premia in Employment Clusters: Does Worker Sorting Bias Estimates?", Working Paper (2010).

Glaeser, Edward, *Cities, Agglomeration and Spatial Equilibrium* (Oxford: Oxford University Press, 2008).

Glaeser, Edward, and Matthew Kahn, "Decentralized Employment and the Transformation of the American City", NBER Working Paper 8117 (2001).

Glaeser, Edward, and William Kerr, "Local Industrial Conditions and Entrepreneurship: How Much of the Spatial Distribution Can We Explain?", *Journal of Economics and Management Strategy* 18:3 (2009), 623-663.

Glaeser, Edward, William Kerr, and Giacomo Ponzetto, "Clusters of Entrepreneurship", *Journal of Urban Economics* 67:1 (2010), 150-168.

Greenstone, Michael, Richard Hornbeck, and Enrico Moretti, "Identifying Agglomeration Spillovers: Evidence from Winners and Losers of Large Plant Openings", *Journal of Political Economy* 118:3 (2010), 536-598.

Griliches, Zvi, "Patent Statistics as Economic Indicators: A Survey", *Journal of Economic Literature* 28:4 (1990), 1661-1707.

Harrison, Brett, and Scott D. Kominers, "A Concentration-Based Index of Industrial Agglomeration", Working Paper (2008).

Hall, Bronwyn, Adam Jaffe, and Manuel Trajtenberg, "The NBER Patent Citation Data File: Lessons, Insights and Methodological Tools", NBER Working Paper 8498 (2001).

Head, Keith, and Thierry Mayer, "The Empirics of Agglomeration and Trade", in J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4* (Amsterdam: North-Holland, 2004).

Helsley, Robert, and William Strange, "Matching and Agglomeration Economics in a System of Cities", *Regional Science and Urban Economics* 20 (1990), 189-212.

Holmes, Thomas, and Sanghoon Lee, "Economies of Density versus Natural Advantage: Crop Choice on the Back Forty", Working Paper 14704 (2009).

Holmes, Thomas, and John Stevens, "Geographic Concentration and Establishment Scale", *Review of Economics and Statistics* 84:4 (2002), 682-690.

Jackson, Matthew, *Social and Economic Networks* (Princeton: Princeton University Press, 2008).

Jacobs, Jane, *The Economy of Cities* (New York, NY: Vintage Books, 1970).

Jaffe, Adam, Manuel Trajtenberg, and Michael Fogarty, "Knowledge Spillovers and Patent Citations: Evidence from a Survey of Inventors", *American Economic Review Papers and Proceedings* 90 (2000), 215-218.

Jaffe, Adam, Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations", *Quarterly Journal of Economics* 108:3 (1993), 577-598.

Jarmin, Ron, and Javier Miranda, "The Longitudinal Business Database", Working Paper (2002).

Kerr, William, "Ethnic Scientific Communities and International Technology Diffusion", *Review of Economics and Statistics* 90 (2008), 518-537.

Kerr, William, "Breakthrough Inventions and Migrating Clusters of Innovation", *Journal of Urban Economics* 67 (2010) 46-60.

Kerr, William, and Shihe Fu "The Survey of Industrial R&D—Patent Database Link Project", *Journal of Technology Transfer* 33:2 (2008), 173-186.

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

Liu, Christopher, "A Spatial Ecology of Structure Holes: Scientists and Communication at a Biotechnology Firm", Working Paper (2010).

Lucas, Robert, and Esteban Rossi-Hansberg, "On the Internal Structure of Cities", *Econometrica* 70:4 (2002), 1445-1476.

Lychagin, Sergey, Joris Pinkse, Margaret Slade, and John Van Reenen, "Spillovers in Space: Does Geography Matter?", Working Paper 2010.

Marcon, Eric, and Florence Puech, "Evaluating the Geographic Concentration of Industries using Distance-based Methods", *Journal of Economic Geography* 3:4 (2003), 409-428.

Marshall, Alfred, *Principles of Economics* (London, UK: MacMillan and Co., 1920).

Maurel, Francoise, and Béatrice Sédillot, "A Measure of the Geographic Concentration in French Manufacturing Industries", *Regional Science and Urban Economics* 29 (1999), 575-604.

Menon, Carlo, "The Bright Side of Gerrymandering: An Enquiry on the Determinants of Industrial Agglomeration in the United States", Working Paper (2009).

Mori, Tomoya, Koji Nishikimi, and Tony Smith, "A Divergence Statistic for Industrial Localization", *Review of Economics and Statistics* 87:4 (2005), 635-651.

Overman, Henry, and Diego Puga, "Labor Pooling as a Source of Agglomeration: An Empirical Investigation", in Edward Glaeser (ed.) *Agglomeration Economics* (Chicago: University of Chicago Press, 2010).

Partridge, Mark, Dan Rickman, Kamar Ali, and Rose Olfert, "Agglomeration Spillovers and Wage and Housing Cost Gradients Across the Urban Hierarchy", *Journal of International Economics* 78:1 (2009), 126-140.

Pe'er, Aviad, and Ilan Vertinsky, "Survival-Enhancing Strategies of De Novo Entrants in Clusters and Dispersal", Working Paper (2009).

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

Rauch, James, "Business and Social Networks in International Trade", *Journal of Economic Literature* 39 (2001), 1177-1203.

Romer, Paul, "Increasing Returns and Long-Run Growth", *Journal of Political Economy* 94:5 (1986), 1002-1037.

Rosenthal, Stuart, and William Strange, "The Determinants of Agglomeration", *Journal of Urban Economics* 50 (2001), 191-229.

Rosenthal, Stuart, and William Strange, "Geography, Industrial Organization, and Agglomeration", *Review of Economics and Statistics* 85:2 (2003), 377-393.

Rosenthal, Stuart, and William Strange, "Evidence on the Nature and Sources of Agglomeration Economies", in J. Vernon Henderson and Jacques-François Thisse (eds.) *Handbook of Regional and Urban Economics, Volume 4* (Amsterdam: North-Holland, 2004), 2119-2171.

Rosenthal, Stuart, and William Strange, "The Attenuation of Human Capital Spillovers", *Journal of Urban Economics* 64 (2008), 373-389.

Rotemberg, Julio, and Garth Saloner, "Competition and Human Capital Accumulation: A Theory of Interregional Specialization and Trade", *Regional Science and Urban Economics* 30 (2000), 373-404.

Saiz, Albert, "The Geographic Determinants of Housing Supply", *Quarterly Journal of Economics* (2010), forthcoming.

Sarvimäki, Matti, "Agglomeration in the Periphery", Working Paper (2010).

Saxenian, AnnaLee, *Regional Advantage: Culture and Competition in Silicon Valley and Route 128* (Cambridge, MA: Harvard University Press, 1994).

Thompson, Peter, "Patent Citations and the Geography of Knowledge Spillovers: Evidence from Inventor- and Examiner-Added Citations", *Review of Economics and Statistics* 88:2 (2006), 383-388.

Thompson, Peter, and Melanie Fox-Kean, "Patent Citations and the Geography of Knowledge Spillovers: A Reassessment", *American Economic Review* 95:1 (2005), 450-460.

Zucker, Lynne, Michael Darby, and Marilynn Brewer, "Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises", *American Economic Review* 88 (1998), 290-306.

Fig. 1a: Technology Sourcing from Silicon Valley
Top patenting zip codes outside of core and their sourcing zones
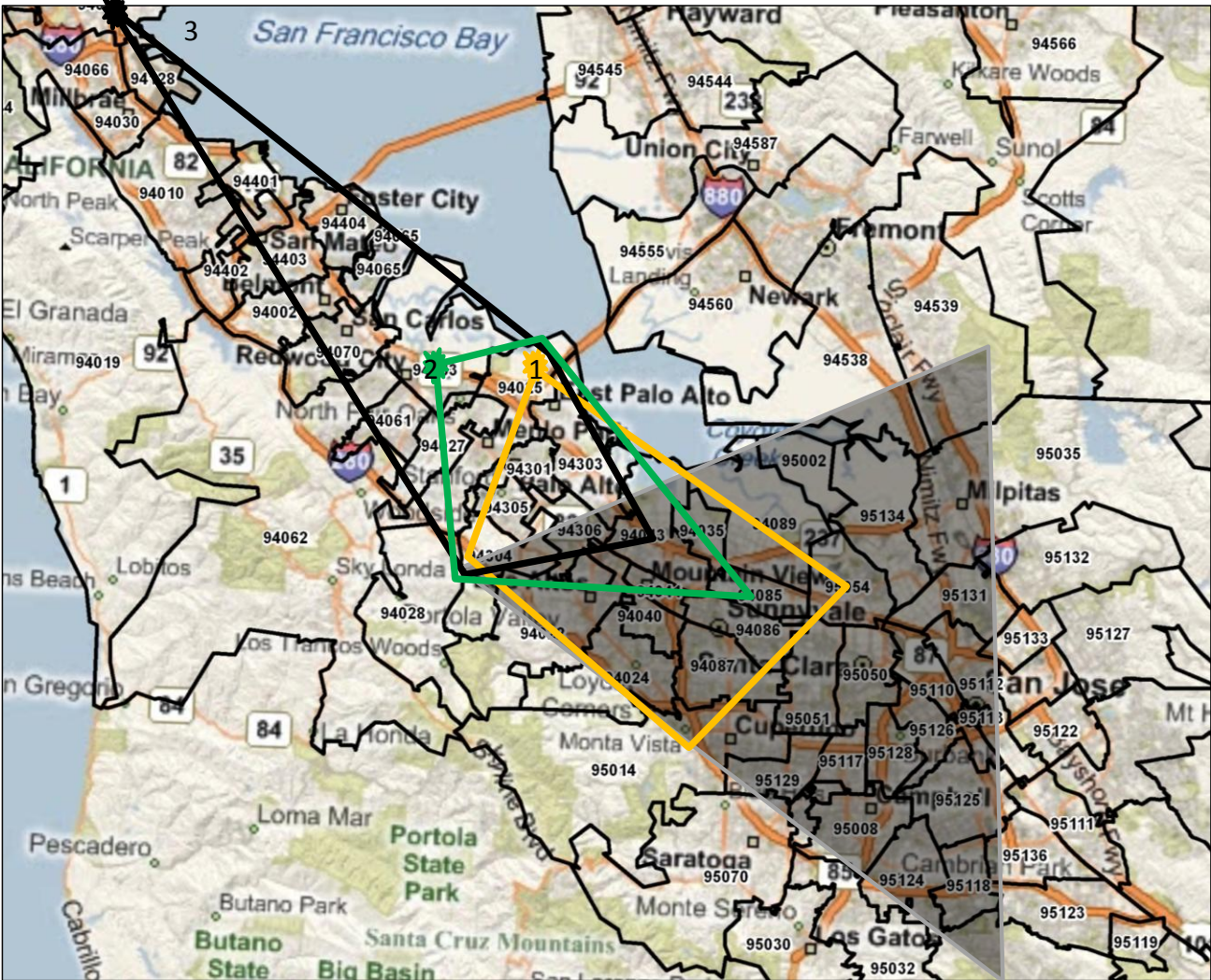
Notes: Figure characterizes technology flows for the San Francisco area. The core of Silicon Valley is depicted with the shaded triangle. The Silicon Valley core contains 76% of the patenting for the San Francisco region. This map describes the technology sourcing for three of the four largest zip codes for patenting not included in the core itself. Technology sourcing zones are determined through patent citations.

The stars indicate the focal zip codes, and the shape of each technology sourcing zone is determined by the three zip codes that firms in the focal zip code cite most in their work. The orange zone (1) for Menlo Park extends deepest into the core. The green zone (2) for Redwood City shifts up and encompasses Palo Alto but less of the core. The black zone (3) for South San Francisco further shifts out and brushes the core.

These technology zones are characterized by small, overlapping regions. None of the technology sourcing zones transverse the whole core, and only the technology zone of the closest zip code (Menlo Park) reaches far enough into the core to include the area of the core where the greatest number of patents occur. Transportation routes and geographic features influence the shapes and lengths of these sourcing zones.

The empirical appendix contains additional maps that show these small, overlapping regions are also evident in the core itself and in other areas outside of the core.

Fig. 1b: Labor Commutes in San Francisco Area
Work commutes for scientists and engineers across PUMAs

Notes: Figure characterizes labor commutes in the San Francisco area for scientists and engineers. Data are taken from the 2000 Census and organized by PUMA of residence and work. The colored asterisks indicate locations of work. The zones surrounding the asterisks of similar color indicate the primary PUMAs of residence for the workers. Zone boundaries are only representative given the coarse nature of PUMA designations. The labor commuting patterns also exhibit multiple, overlapping sourcing zones, none of which transverses the whole region.

Fig. 2a: Marshallian Clusters
Agglomeration due to interactions among firms

Unoccupied site

Existing
spillover
radius

C

A
X

B
Y
Z

A larger spillover radius
for site B would induce
the marginal entrant—
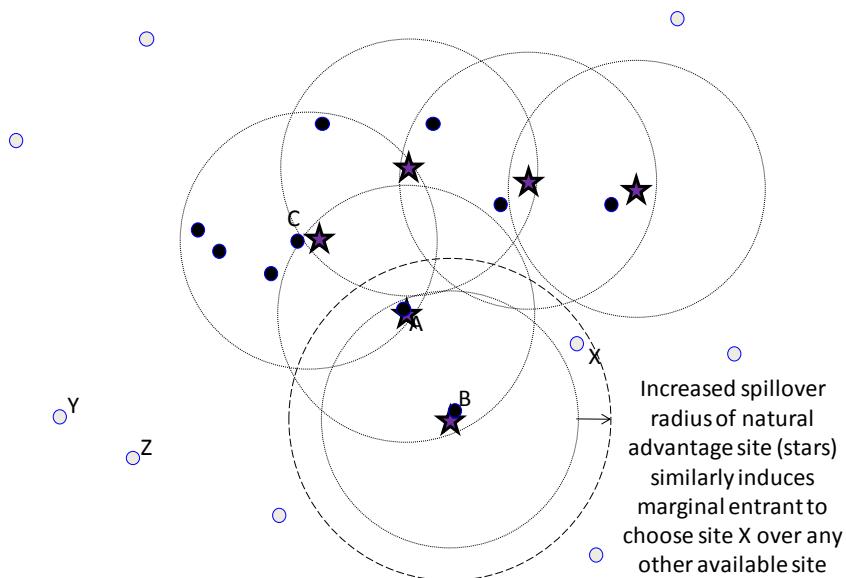currently indifferent
across available sites–
into choosing site X

On the other hand, all
sites chosen at random
if radius is very small
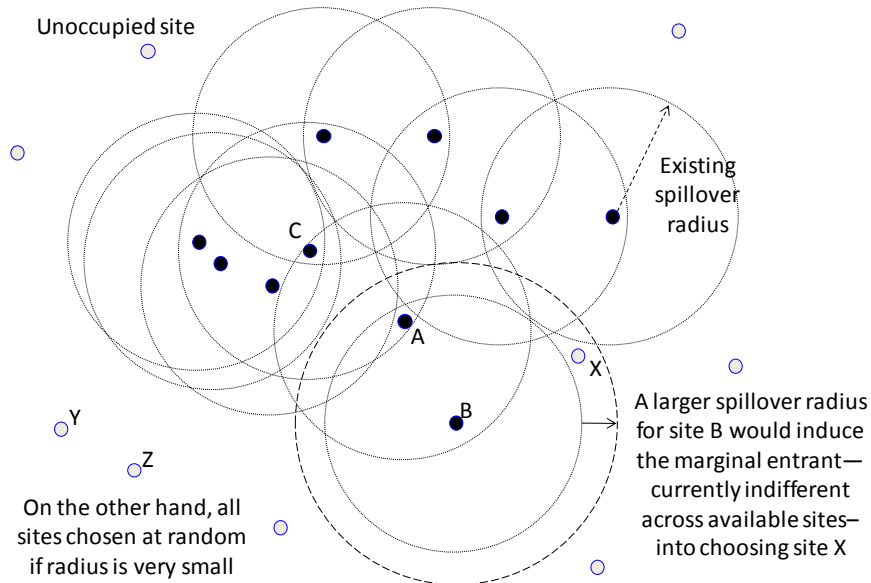
Notes: Image illustrates a Marshallian cluster. Entry is sequential, without foresight, and potential sites are fixed. Black dots are chosen sites, and circles represent maximal spillover radii. Large area clustering is due to small, contained interaction effects that overlap each other. The next entrant is indifferent among available sites, including X, Y, and Z. The larger dashed circle shows that a larger maximal radius would induce the marginal entrant into the cluster over other sites, resulting in (weakly) fewer, larger, and less dense clusters.

Fig. 2b: Natural Advantage Clusters
Agglomeration due to proximity to fixed resources

C

A
X

B
Y
Z

Increased spillover
radius of natural
advantage site (stars)
similarly induces
marginal entrant to
choose site X over any
other available site

Notes: Image illustrates a natural advantage cluster. The stars represent fixed natural advantage sites, with circles now representing the maximal spillover radii of those advantages. Large area clustering follows from firm attraction to fixed resources and an uneven spatial distribution of natural advantage sites. The larger dashed circle again shows that a longer maximal radius induces marginal entrants into the cluster. Moreover, this tipping behavior is weakly increasing in Marshallian complementarity if these latter spillovers exist among firms.

Fig. 3a: Local Patent Technology Horizons
Pairwise citations over 50 miles compared to 50-150 miles
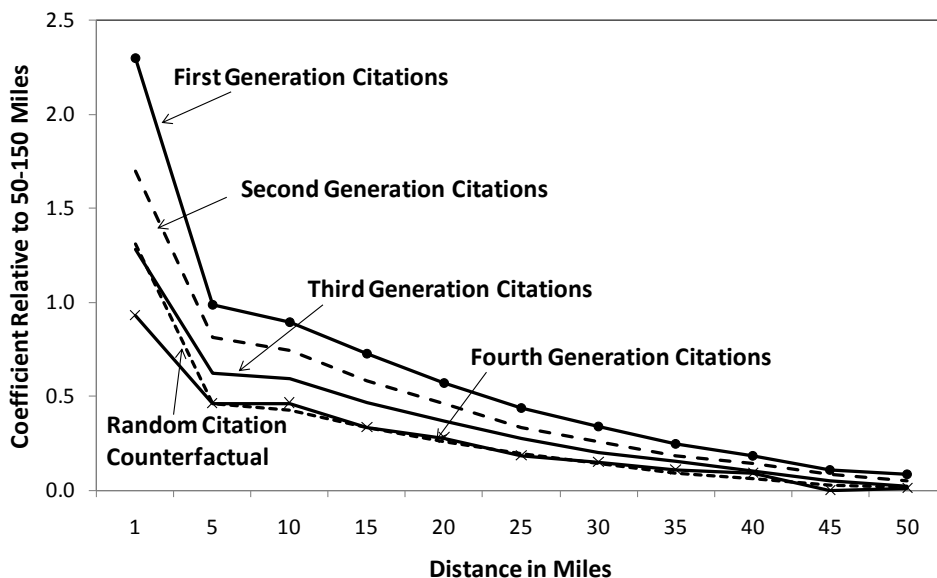
Notes: Figure plots coefficients from regressions of log patent citation counts on distances. Observations are pairwise citing and cited zip codes within 150 miles of each other. Distance bands are <1 mile, 1-5 miles, and each subsequent five mile increment to 50 miles. Explanatory variables are indicator variables for each distance band, with 50-150 miles apart serving as the reference category. Regressions control for log patenting and citing zip code fixed effects. The random counterfactual considers same technologies and years as true citations.

Fig. 3b: Patent Spatial Densities
Share of sub-technologies in indicated group localized by distance
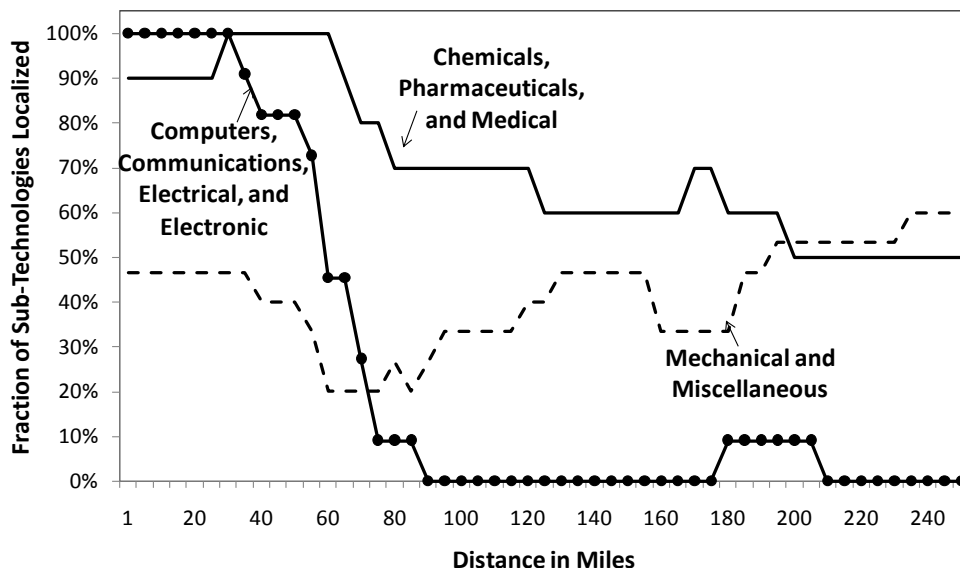
Notes: Figure plots the fraction of technologies showing localization by each distance. The sample includes 36 sub-categories of the USPTO system organized into three simple divisions. Localization is calculated using pairwise distances among inventors in a technology with the Duranton and Overman (2005) metric. Technologies are considered localized at a distance if they exhibit abnormal density compared to 1000 random draws of US inventors of a similar size to the technology. Local confidence bands are set at 5%/95% for this determination.

Fig. 3c: Patent Cluster Density & Spillover Radius

Cross-section of invention density and tech. spillover lengths
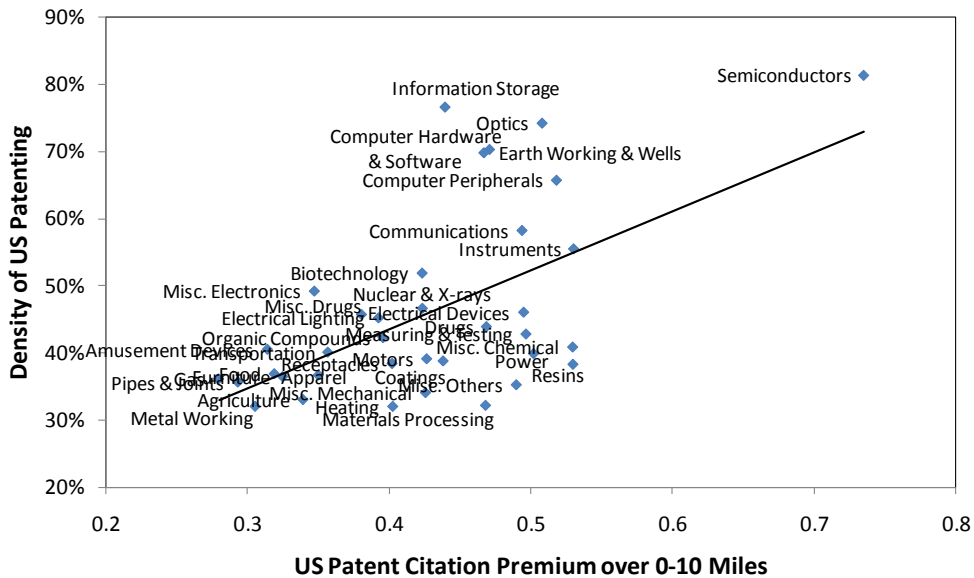
Notes: Figure provides a cross-sectional plot of cluster density and technology spillover lengths. Cluster density is measured through bilateral patent distances in each technology. It is the share of patenting that occurs within 50 miles relative to the share within 150 miles. Technology spillover horizon measures by patent class the citation premium among patents 0-10 miles apart relative to 30-150 miles apart. Technologies with shorter spillover ranges (i.e., greater citation premiums) have denser clusters. The slope of the trend line is 0.875 (0.163).

Fig. 3d: Patent Cluster Density & Spillover Radius

Using UK tech. spillover lengths to predict US density levels
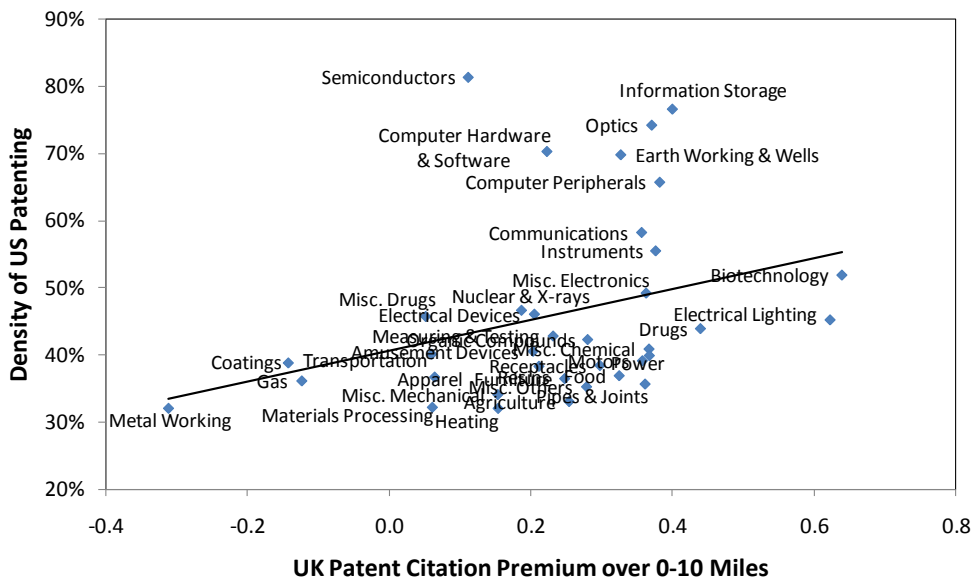
Notes: See Figure 3c. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths. Using patents filed from the UK to the USPTO, the technology spillover horizon measures by patent class the citation premium among patents 0-10 miles apart relative to 30-150 miles apart. The slope of the trend line is 0.230 (0.075).

Fig. 3e: Patent Cluster Shape & Spillover Radius
Kernel estimations of cluster shape and tech. spillover lengths

Notes: Figure plots coefficients from regressions of kernel densities by distance for 36 technologies. Technology spillover horizon measures by technology the citation premium for 0-10 miles relative to 30-150 miles. This metric is interacted with indicator variables for each five mile increment in distances to 250 miles. Regressions include fixed effects for each distance. Dashed lines are 90% confidence bands. Technologies with shorter spillover ranges show excess density at short distances and increase dispersion over medium distances (i.e., denser, smaller clusters).

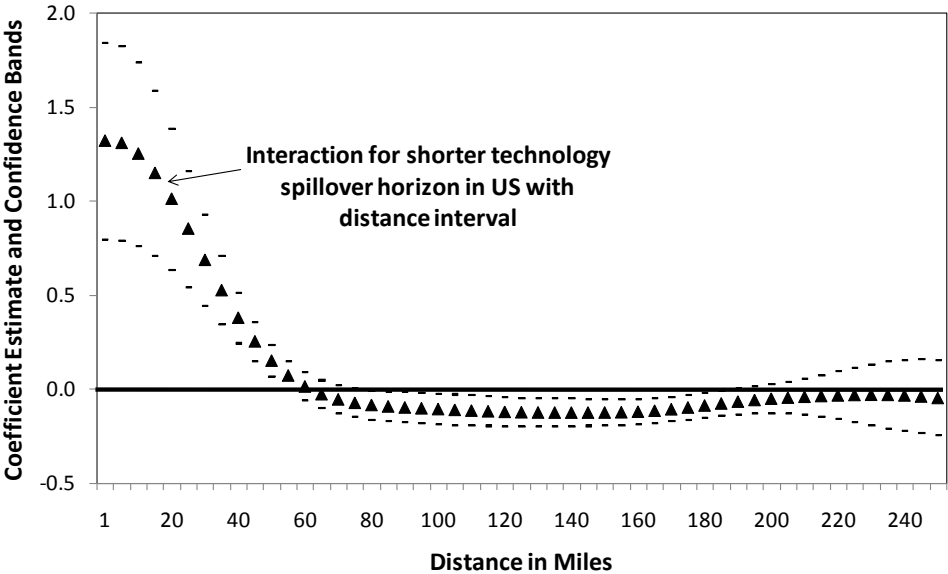# Fig. 3f: Patent Cluster Shape & Spillover Radius
Using UK tech. spillover lengths to predict US cluster shapes
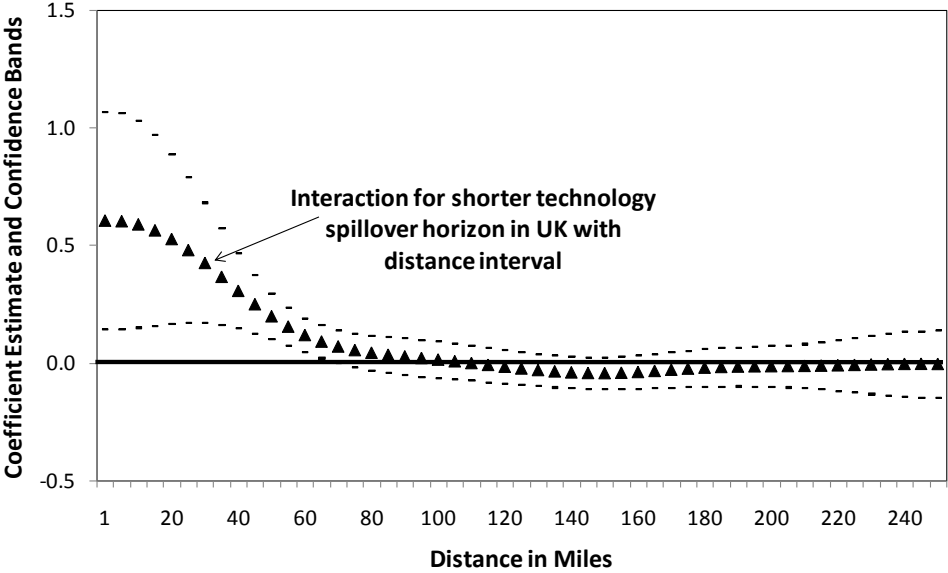
Notes: See Figure 3e. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths.

Fig. 3g: Patent Localization & Spillover Radius
Localization estimations and technology spillover lengths

Notes: See Figure 3e. The dependent variable is updated from the kernel density in Figure 3e to be the measurement of localization developed by Duranton and Overman (2005). Technologies are considered localized at a distance if they exhibit abnormal density compared to 1000 random draws of US inventors of a similar size to the technology. Local confidence bands are set at 5%/95% for this determination. Technologies with shorter spillover ranges again exhibit denser, smaller clusters with this technique.

Fig. 3h: Patent Localization & Spillover Radius
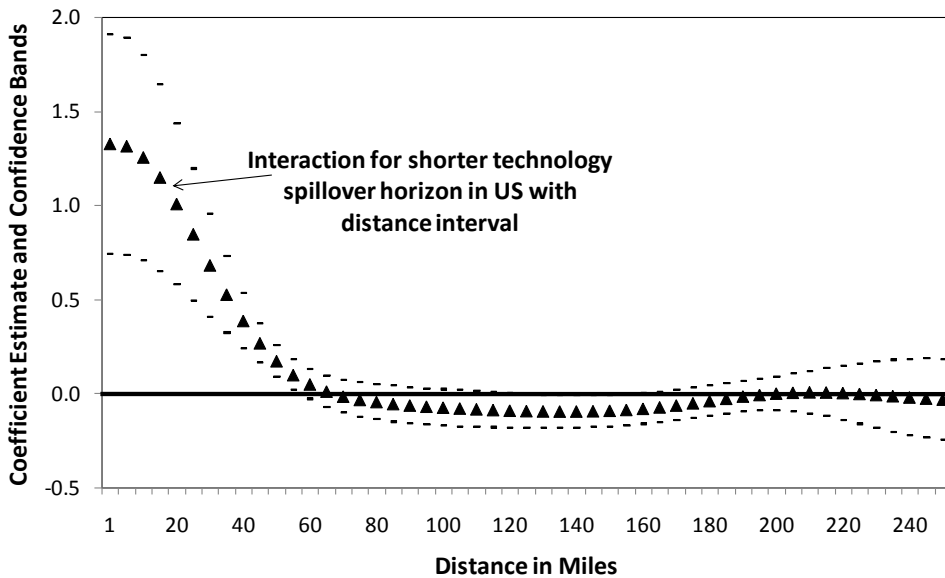## Using UK tech. spillover lengths to predict US localization

Notes: See Figures 3e-3g. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths.
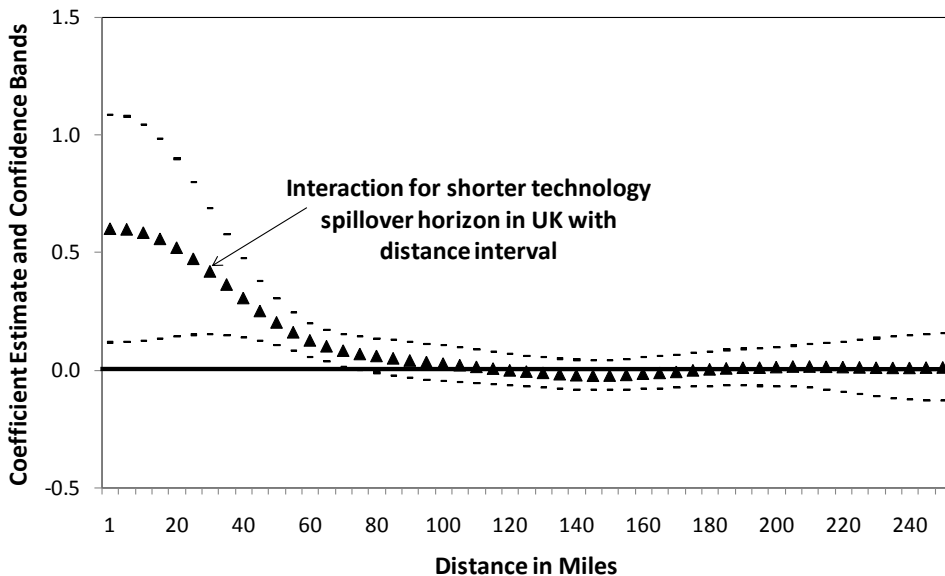
Fig. 4a: Industry Spatial Densities
Share of SIC3 industries in sector localized by distance
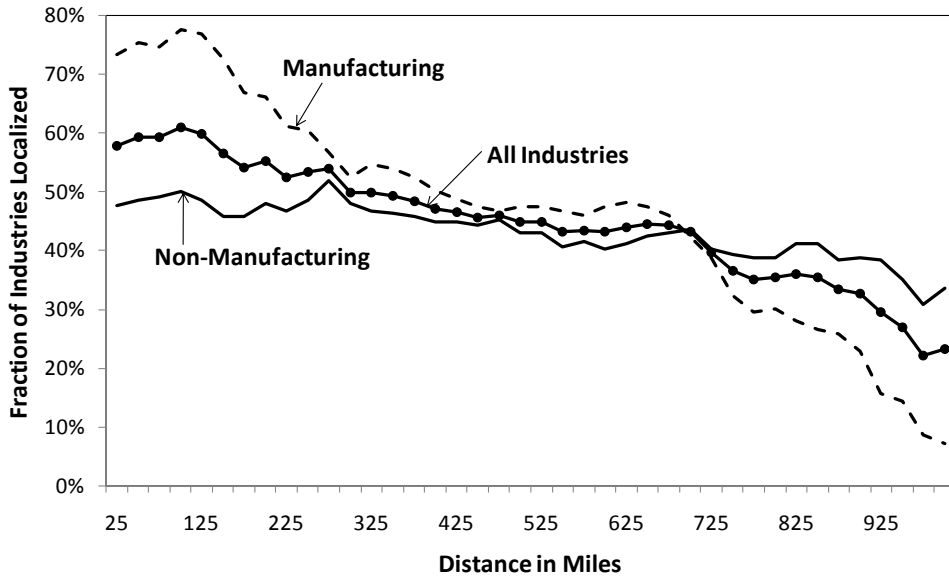
Notes: Figure plots the fraction of industries showing localization by each distance. The total sample includes 362 SIC3 industries; 140 industries are in manufacturing. Localization is calculated using pairwise distances among plants in an industry. Industries are considered localized at a distance if they exhibit abnormal density compared to 1000 random draws of US establishments of a similar size to the industry. Local confidence bands are set at 5%/95% for this determination.

Fig. 4b: Densities & Occupational Concentration
Share of SIC3 industries in tercile localized by distance
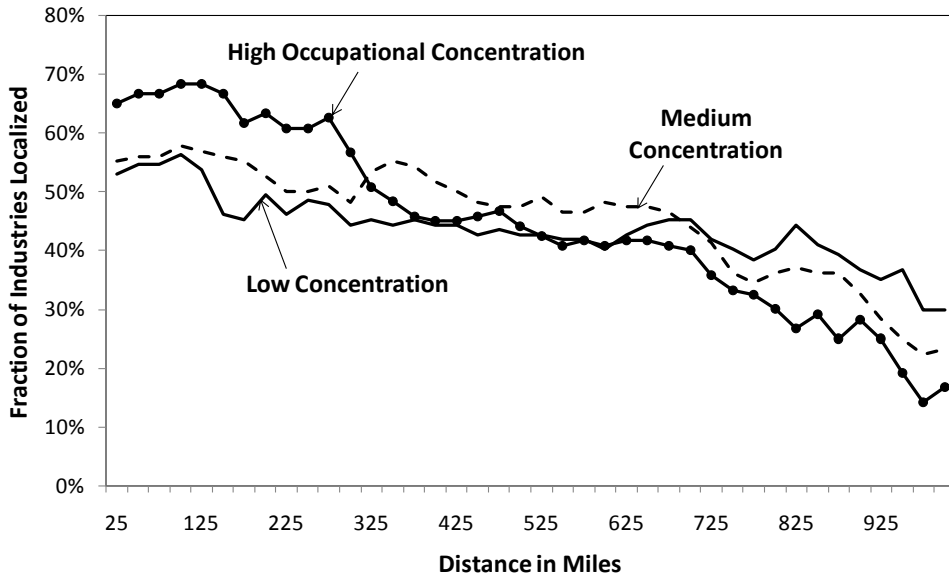
Notes: See Figure 4a. This figure separates the 362 industries into terciles based upon how concentrated their occupational employments are. Occupation concentration is taken from the National Industry-Occupational Employment Matrix of the Bureau of Labor Statistics. Greater occupational concentration increases the benefits of labor pooling. Greater occupational concentration is associated with higher spatial concentration at short distances and lower spatial concentration at longer distances.

Fig. 4c: Densities & Natural Advantages

Share of SIC3 mfg. industries in tercile localized by distance

Fraction of Industries Localized (y-axis: 0% to 100%)

High Natural Advantages Dependency

Medium Dependency

Low Dependency

Distance in Miles (x-axis: 25 to 925)
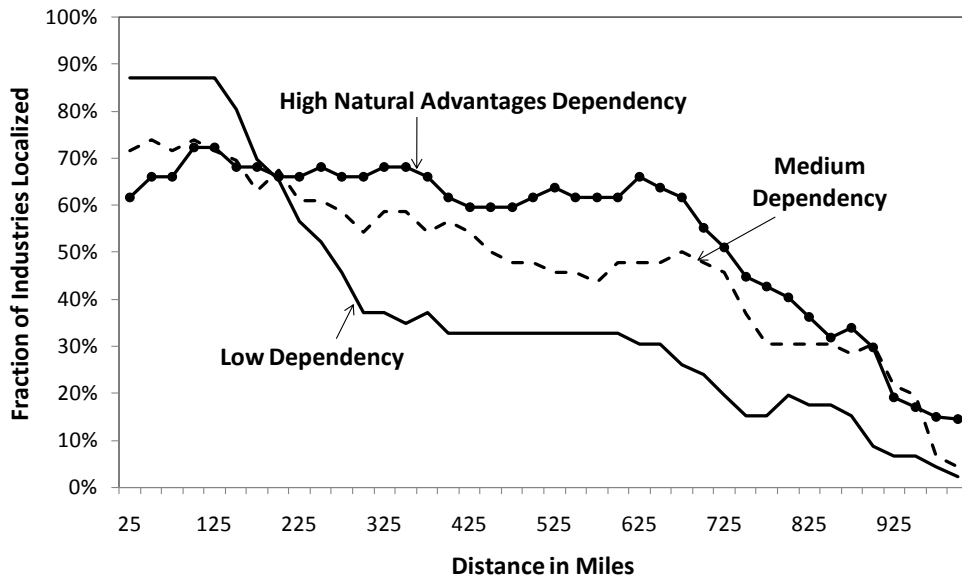
Notes: See Figure 4a. This figure separates the 140 manufacturing industries into terciles based upon their dependency for natural advantages. Dependency is taken as the maximum of dependency for water transportation, coal, natural gas, lumber, and agricultural products. Greater dependency for natural advantages is associated with lower spatial densities at short distances and higher spatial densities at longer distances. While unobserved, the spillover range for natural advantages related to these factors is often longer than for labor pooling.

Fig. 4d: Labor Pooling Cluster Shape

Estimations of cluster shape and occupation concentration
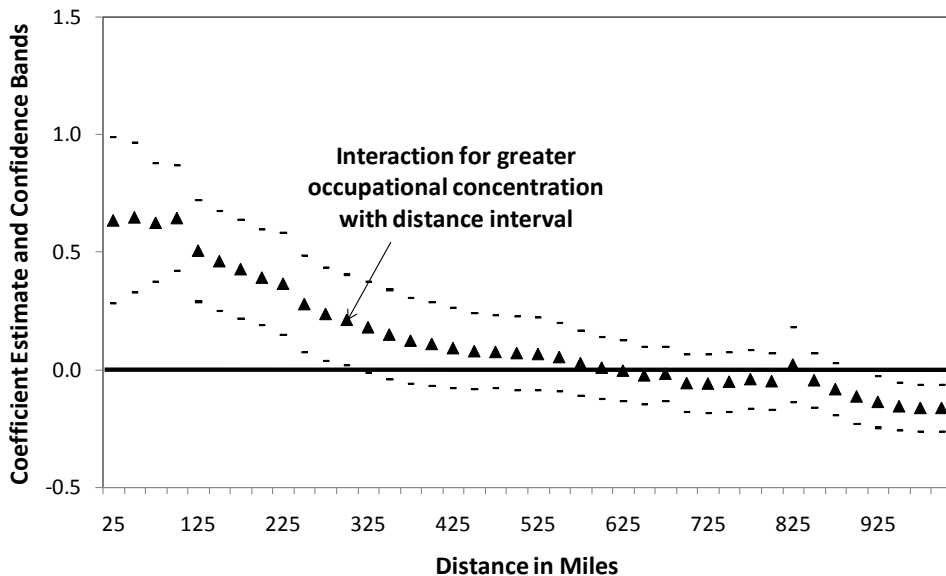
Coefficient Estimate and Confidence Bands (y-axis: -0.5 to 1.5)

Interaction for greater occupational concentration with distance interval

Distance in Miles (x-axis: 25 to 925)

Notes: Figure plots coefficients from regressions of agglomeration densities by distance for 362 SIC3 industries. Occupational concentration measures the HHI index of occupational employment within an industry. This metric is interacted with indicator variables for each 25 mile increment in distances to 1000 miles. Regressions include fixed effects for each distance. Dashed lines are 90% confidence bands. Occupational concentration promotes excess concentration over short to moderate distances. The shape of these functions mirrors the technology regressions in Figure 3g.

Fig. 4e: Labor Pooling Cluster Shape
Estimations of US cluster shape and UK occupation concentration
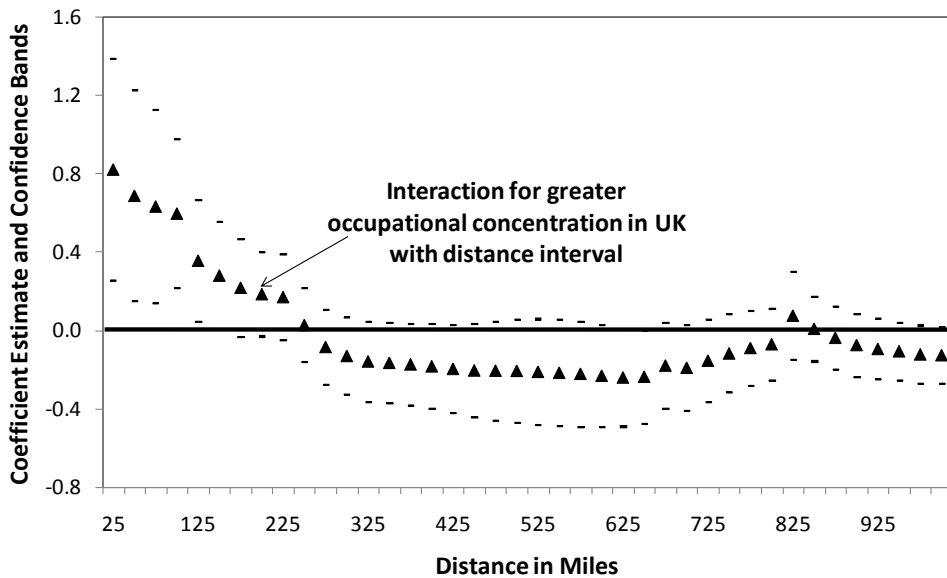
Notes: See Figure 4d. Estimates use UK occupational concentration measures to confirm reverse causality is not determining the results. Occupational concentration is defined as the HHI index of occupational employment within an industry using data from the UK.

Fig. 4f: Natural Advantage Cluster Shape
Estimations of cluster shape and natural advantage dependency
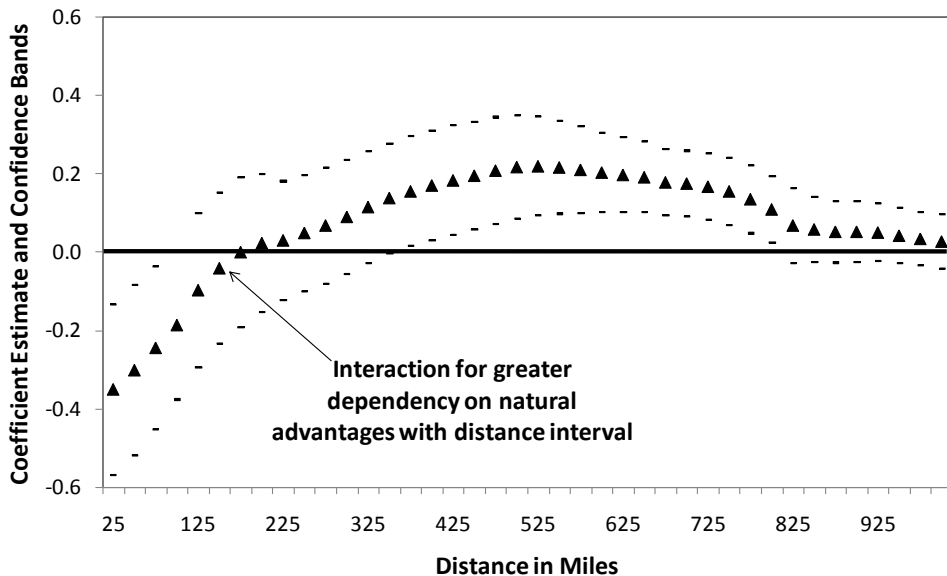
Notes: Figure plots coefficients from regressions of agglomeration densities by distance for 140 SIC3 manufacturing industries. Natural advantages measure maximum dependency of the industry on one of five natural advantages (water transportation, coal, natural gas, lumber, and agricultural products). These metrics are interacted with indicator variables for each 25 mile increment in distances to 1000 miles. Regressions include fixed effects for each distance and control for occupational concentration. Compared to occupational concentration, industries dependent upon natural advantages show larger and less dense agglomeration patterns.

Theoretical Appendix

Proof of Proposition 1

Immediate from text.

Proof of Proposition 2

Throughout, we fix the industry A and only consider Marshallian forces, so that $\rho_{\mathbf{K}}^A = 0$. Thus, for simplicity, we henceforth suppress the indices A and $\mathbf{A}$ whenever doing so does not introduce confusion. We also assume for notational convenience that the firms in industry A enter in uninterrupted sequence. The results are unchanged if we instead assume that firms in other industries enter between firms i and i', so long as the entry order of all firms is fixed at the beginning. Finally, we assume a generic distribution of sites.

We order the firms in industry A by entry period $i = 1, \ldots, N^A$, and let $\mathbf{Z}_i$ be the set of sites occupied at the stage in which firm i chooses its location. Since sites are generically distributed and only Marshallian forces affect the location decision of firm i, we see that i is indifferent between sites if and only if $\mathcal{B}_\rho(j(i-1)) \setminus \mathbf{Z}_i = \emptyset$. In that case, i chooses locations randomly. Otherwise, i chooses the unique site $j(i) \in \mathcal{B}_\rho(j(i-1)) \setminus \mathbf{Z}_i$ which maximizes $g_{j(i),A}$, and $\mathbf{Z}_{i+1} = \mathbf{Z}_i \cup \{j(i)\}$.

The preceding discussion implies that the set of locations occupied by industry-A firms following all firms' entries is completely determined by an instantiation of the random cluster selection which occurs whenever a firm i is forced to choose a site randomly because of indifference. We may represent such a sequence of random draws by an ordering $\phi \equiv \phi^1 \cdots \phi^{B(\rho)}$ of the $B(\rho)$ disjoint ρ-clusters $\mathcal{B}_\rho(j_{\phi^b})$. (Here, $\mathbf{Z} \ni j_{\phi^b} \in \mathcal{B}_\rho(j_{\phi^b})$ is a representative site in the ρ-cluster assigned index b in the ordering ϕ.) We denote the set of possible orderings of ρ-clusters by Φ.[34]

The number of ρ-clusters actually *occupied* by industry-A firms if the random entry sequence is drawn as $\phi \in \Phi$ is given by

$$\#_\rho(\phi) \equiv \min \left\{ b \in \mathbb{N} : N^A \leq \sum_{b'=1}^{b} \left| \mathcal{B}_\rho(j_{\phi^{b'}}) \right| \right\}.$$

Denoting the probability of draw $\phi \in \Phi$ by $\text{Prob}(\phi)$,[35] we compute that $E_\phi[\#_\rho(\phi)]$, the expected level of agglomeration level in industry A, is given by

$$E_\phi[\#_\rho(\phi)] = \sum_{\phi \in \Phi} \#_\rho(\phi) \cdot \text{Prob}(\phi).$$

We suppose that ρ increases to $\rho' > \rho$ (with $\rho'_{\mathbf{K}} = 0$). If $\mathcal{B}_\rho(j) = \mathcal{B}_{\rho'}(j)$ for all $j \in \mathbf{Z}$, then clearly agglomeration behavior is unchanged. Thus, we may assume that

$$\mathcal{B}_\rho(j) \neq \mathcal{B}_{\rho'}(j) \tag{9}$$

for at least one $j \in \mathbf{Z}$. Moreover, since the distribution of sites is generic, we may assume without loss of generality that $B(\rho') = B(\rho) - 1$, so that there is some $j' \in \mathbf{Z}$ such that (9) holds for *exactly* the sites $j \in \mathcal{B}_{\rho'}(j')$. Iterating our arguments for that case (over successive expansions of ρ) show the proposition in general.

Now, we let β and β' denote the two ρ-clusters which are merged when the maximal radius expands to ρ', so that $\mathcal{B}_{\rho'}(j_\beta) = \mathcal{B}_{\rho'}(j_{\beta'})$ but $\mathcal{B}_\rho(j_\beta) \neq \mathcal{B}_\rho(j_{\beta'})$. Abusing notation slightly, for

---

[34]We maintain the subscript ρ to remind the reader which maximal radius is being considered.

[35]Here, we do not specify the actual distribution of draws, since it is not needed for the proposition.

$\Phi \ni \phi = \phi^1 \cdots \phi^{B(\rho)}$, we let $\beta(\phi)$ denote the index b such that $\phi^b = \beta$. We define $\beta'(\phi)$ analogously.

We may again associate the possible firm location patterns to the orders of possible ρ-cluster selection $\phi \in \Phi$, with the understanding that the ρ'-clusters are occupied in the order

$$\mathcal{B}_{\rho'}(j_{\phi^1}), \ldots, \mathcal{B}_{\rho'}(j_{\phi^{\beta(\phi)}}), \ldots, \mathcal{B}_{\rho'}(j_{\phi^{\beta'(\phi)-1}}), \mathcal{B}_{\rho'}(j_{\phi^{\beta'(\phi)+1}}), \ldots, \mathcal{B}_{\rho'}(j_{\phi^{B(\rho)}})$$

until all N^A firms have entered. The actual number of clusters occupied by firms, denoted $\#_{\rho'}(\phi)$, will not in general be equal to $\#_\rho(\phi)$. There are two cases to consider: $\beta'(\phi) \leq \#_\rho(\phi)$ and $\beta'(\phi) > \#_\rho(\phi)$. In each case, we have that $\#_{\rho'}(\phi) \leq \#_\rho(\phi)$.

As the probability that cluster $\mathcal{B}_{\rho'}(j_{\phi^\beta})$ is selected by a firm choosing randomly among available sites $\mathbf{Z} \setminus \mathbf{Z}_i$ is equal to the sum of the probability of choosing $\mathcal{B}_\rho(j_{\phi^\beta})$ and that of choosing $\mathcal{B}_\rho(j_{\phi^{\beta'}})$, direct computation shows that the level of agglomeration expected when the maximal radius is ρ' is equal to $E_\phi[\#_{\rho'}(\phi)]$. Since we have shown that $\#_{\rho'}(\phi) \leq \#_\rho(\phi)$ for all $\phi \in \Phi$, we have

$$E_\phi[\#_{\rho'}(\phi)] = \sum_{\phi \in \Phi} \#_{\rho'}(\phi) \cdot \mathrm{Prob}(\phi) \leq \sum_{\phi \in \Phi} \#_\rho(\phi) \cdot \mathrm{Prob}(\phi) = E_\phi[\#_\rho(\phi)],$$

which proves the desired result.

Proof of Proposition 3

As in the proof of Proposition 2, we fix the industry A and suppress the index A whenever doing so does not introduce confusion. (In particular, we write $\mathbf{K} \equiv \mathbf{K}(A)$ throughout this proof.) We also assume for notational convenience that the firms in industry A enter in uninterrupted sequence, and that the distribution of sites is generic.

We suppose that the cost of extracting natural advantages falls from $c_{\mathbf{K}}$ to $c'_{\mathbf{K}} < c_{\mathbf{K}}$, so that $\rho_{\mathbf{K}}$ rises to $\rho'_{\mathbf{K}}$. We suppose for simplicity that $\rho_{\mathbf{K}} > \rho_{\mathbf{A}}$.[36]

Under the condition

$$G_{\mathbf{A}}(0) - c_{\mathbf{A}} < \min \left\{ G_{\mathbf{K}} \left(\max_{\{(j,k) \in \mathbf{Z} \times \mathbf{K}: d_{j,k} < \rho\}} d_{j,k} \right) - c_{\mathbf{K}}, G_{\mathbf{K}} \left(\max_{\{(j,k) \in \mathbf{Z} \times \mathbf{K}: d_{j,k} < \rho'\}} d_{j,k} \right) - c'_{\mathbf{K}} \right\} \quad (10)$$

(so that natural advantages are always more valuable than Marshallian advantages) and the cost $c_{\mathbf{K}}$, if there exists an unoccupied site $j \in \mathcal{B}_\rho(k)$ for some $k \in \mathbf{K}$, then any entering industry-A firm i will choose a location $j(i) \in \mathcal{B}_\rho(k')$ for some $k' \in \mathbf{K}$.[37] We write

$$\mathcal{B}_\rho(\mathbf{K}) \equiv \bigcup_{k \in \mathbf{K}} \mathcal{B}_\rho(k). \quad (11)$$

If $|\mathcal{B}_\rho(\mathbf{K})| \geq N^A$, then all firms in industry A will locate in the clusters $\{\mathcal{B}_\rho(k)\}_{k \in \mathbf{K}}$ when facing the cost $c_{\mathbf{K}}$. When the cost falls to $c'_{\mathbf{K}}$, firms' location decisions will be unchanged, hence there is nothing to show. This follows from the fact that the cost term enters linearly into firm utility from contact with a site. Likewise, if $N^A > |\mathcal{B}_\rho(\mathbf{K})|$ but $|\mathcal{B}_\rho(\mathbf{K})| = |\mathcal{B}_{\rho'}(\mathbf{K})|$, that is, if the cost decrease leaves the set of sites in the maximal-radius clusters about attraction sites $k \in \mathbf{K}$

---

[36] There are two cases in which $\rho_{\mathbf{A}} > \rho_{\mathbf{K}}$, according as $\rho_{\mathbf{A}}$ is larger or smaller than $\rho'_{\mathbf{K}}$. In each of these cases, the result follows via an argument analogous to that presented here for the case $\rho_{\mathbf{A}} < \rho_{\mathbf{K}}$.

[37] Note that this is true even if no sites within $\rho_{\mathbf{K}}$ of an attraction site $k \in \mathbf{K}$ are available, as in this case if there is some $k \in \mathbf{K}$ for which there exists an unoccupied site $j \in \mathcal{B}_\rho(k)$, then there exists some $k' \in \mathbf{K}$ with a site $j' \in \mathcal{B}_\rho(k')$ which delivers Marshallian agglomeration benefits.

unchanged, then there is nothing to show since the clusters in $\mathbf{Z} \setminus \mathcal{B}_\rho(\mathbf{K})$ are left unchanged by the change in $c_\mathbf{K}$.[38]

Thus, we need only consider the case in which $|\mathcal{B}_{\rho'}(\mathbf{K})| > |\mathcal{B}_\rho(\mathbf{K})|$. Because the site distribution is generic, it suffices to prove the result in the case that there is only one cluster $\mathcal{B}_{\rho'}(j) \subset \mathcal{B}_{\rho'}(\mathbf{K}) \setminus \mathcal{B}_\rho(\mathbf{K})$. The full result follows upon induction.

The expected number of ρ-clusters occupied by firms under maximal radius ρ, conditional upon some firm choosing a site $j' \in \mathcal{B}_{\rho'}(j)$, is left unchanged upon expansion of the maximal radius to ρ'. To complete the proof, we consider the case in which, under maximal radius ρ, the first $|\mathcal{B}_\rho(\mathbf{K})|$ firms choose locations in $\mathcal{B}_\rho(\mathbf{K})$ and the remaining $N^A - |\mathcal{B}_\rho(\mathbf{K})|$ firms choose locations in $\mathbf{Z} \setminus (\mathcal{B}_\rho(\mathbf{K}) \cup \mathcal{B}_\rho(j)) = \mathbf{Z} \setminus \mathcal{B}_{\rho'}(\mathbf{K})$.[39] A decrease in $c_\mathbf{K}$ leaves the number of distinct clusters in $\mathbf{Z} \setminus (\mathcal{B}_\rho(\mathbf{K}) \cup \mathcal{B}_\rho(j))$ fixed, hence to bound the number of clusters occupied under maximal radius ρ', it suffices to consider, for each collection of clusters $\{\mathcal{B}_\rho(j^b) \subset \mathbf{Z} \setminus (\mathcal{B}_\rho(\mathbf{K}) \cup \mathcal{B}_\rho(j))\}_b$ occupied, the effect on the number of occupied ρ-clusters of moving $|\mathcal{B}_\rho(j)|$ firms to $\mathcal{B}_\rho(j)$ from the last clusters $\mathcal{B}_\rho(j^b)$ to be occupied. Clearly, this process can at most increase the number of occupied ρ-clusters by 1 (in the case in which no cluster $\mathcal{B}_\rho(j^b)$ is emptied). It follows that for any instantiation of random site selection, the number of ρ-clusters outside $\mathcal{B}_{\rho'}(\mathbf{K})$ occupied under maximal radius ρ' is at most as large as the number of ρ-clusters outside $\mathcal{B}_{\rho'}(\mathbf{K})$ occupied under maximal radius ρ. The site discovery condition (3) guarantees that the number of distinct ρ'-clusters inside $\mathcal{B}_{\rho'}(\mathbf{K})$ occupied under maximal radius ρ' is always weakly smaller than the number of ρ-clusters inside $\mathcal{B}_{\rho'}(\mathbf{K})$ occupied under maximal radius ρ, hence we have the result.

Proof of Proposition 4

Immediate from text.

Proof of Proposition 5

The result is trivial for the case in which ρ is such that $|\mathcal{B}_\rho(j)| > 1$ for only two $j \in \mathbf{Z}$ (i.e. the case in which only one cluster contains more than one site), as in that case expansion to ρ' either does not change the composition of clusters or increases mean bilateral distances between sites in clusters.[40] Thus, the existence of the desired $\bar{\rho}$ is immediate.

Proof of Proposition 6

As in the proofs of Propositions 2 and 3, we fix the industry A and suppress the index A whenever doing so does not introduce confusion. (In particular, we write $\mathbf{K} \equiv \mathbf{K}(A)$ throughout this proof.) Additionally, we suppose a generic distribution of sites.

The result follows from Proposition 2 if $G_\mathbf{K}(0) = 0$, so we assume that $G_\mathbf{K}(0) > 0$. We define $\mathcal{B}_\rho(\mathbf{K})$ as in (11), and suppose that $N^A \le |\mathcal{B}_\rho(\mathbf{K})|$. With these assumptions, it is clear that all firms (deterministically) choose sites in $\mathcal{B}_\rho(\mathbf{K})$.

Scaling $G_\mathbf{A}$ by α (possibly) expands the maximal radius to $\rho' \ge \rho$. We assume that the attraction sites $k \in \mathbf{K}$ are sufficiently dispersed that $\mathcal{B}_{\rho'}(k) = \mathcal{B}_{\rho'}(k')$ implies that $\mathcal{B}_\rho(k) = \mathcal{B}_\rho(k')$

---

[38]If $\mathcal{B}_\rho(\mathbf{K}) = \mathcal{B}_{\rho'}(\mathbf{K})$, then the total number of clusters occupied by firms in the maximal-radius clusters about attraction sites weakly decreases. Meanwhile, since again $N^A - |\mathcal{B}_\rho(\mathbf{K})|$ firms must locate outside $\mathcal{B}_\rho(\mathbf{K})$, the expected number of clusters occupied by sites choosing sites in $\mathbf{Z} \setminus \mathcal{B}_\rho(\mathbf{K})$ is unchanged.

[39]Note that $\mathcal{B}_{\rho'}(j) \setminus \mathcal{B}_{\rho'}(\mathbf{K}) = \mathcal{B}_\rho(j)$ because $c_\mathbf{A}$ is unchanged.

[40]Note that for $\bar{\rho}$ sufficiently small, if $\mathcal{B}_{\rho'}(j) \supsetneq \mathcal{B}_\rho(j)$, then the mean bilateral distance between sites in $\mathcal{B}_{\rho'}(j)$ is larger than that between sites in $\mathcal{B}_\rho(j)$.

for all $k, k' \in \mathbf{K}$.[41] This guarantees that no ρ-cluster $\mathcal{B}_\rho(k)$ ($k \in \mathbf{K}$) which is not occupied under maximal radius ρ is occupied under maximal radius ρ'. Indeed, if the cluster $\mathcal{B}_\rho(k)$ is not occupied under maximal radius ρ, then (by firm-revealed preference) its natural advantage benefits must be lower than the benefits offered by the other clusters occupied under maximal radius ρ. Since scaling $G_\mathbf{A}$ by α leaves the natural advantage benefits fixed, firms could only want to occupy a site $j \in \mathcal{B}_\rho(k)$ under maximal radius ρ' if the Marshallian advantages afforded by j are sufficiently strong. These benefits must be derived from contact with sites $j' \notin \mathcal{B}_\rho(k)$; the assumption on the dispersion of $\mathbf{K}$ rules out this possibility.

Cluster Dynamics

Although firms' location decisions are deterministic in the model, it is still possible for agglomeration clusters due to Marshallian forces to move over time in a dynamic setting. We describe several limit cases and general observations here, leaving a complete analysis of our model's dynamics for future work.

When a first-mover firm $i_1 \in A$ chooses its location, this decision must be driven by natural advantage concerns or idiosyncratic factors. But then, Marshallian forces drive subsequent entrants $i_2, \ldots, i_m \in A$ to locate around $j(i_1)$, creating a network of localized spillovers. If these forces are sufficiently strong, then even were i_1 to fold and vacate its site $j(i_1)$, a newly entering firm i_{m+1} might not choose to locate at $j(i_1)$. Instead, the new entrant could locate in some location $j \neq j(i_1)$ close to multiple sites of $\{j(i_2), \ldots, j(i_m)\}$ that provides more spillovers in the now-populated economy than the first location does.

For an example of this behavior, we will assume that all relevant sites are located on integer points of a coordinate plane $\mathbb{R}^2$. We have $\mathbf{K} = \{(0,1)\}$ and $\mathbf{Z} = \{(0,1), (0,2), (0,3), (1,3)\}$.[42] We assume that $\rho_\mathbf{A}^A = \rho_\mathbf{K}^A = \sqrt{2} + \delta$, where $\delta > \frac{\epsilon\sqrt{2}}{1-\epsilon}$ for some tiny $\epsilon > 0$ and take

$$G_\mathbf{A}(d) = \begin{cases} \rho_\mathbf{A}^A - d & d \leq \rho_\mathbf{A}^A, \\ 0 & d \geq \rho_\mathbf{A}^A. \end{cases} \qquad G_\mathbf{K}(d) = \begin{cases} \epsilon(\rho_\mathbf{K}^A - d) & d \leq \rho_\mathbf{K}^A, \\ 0 & d \geq \rho_\mathbf{K}^A. \end{cases}$$

The first firm i_1 will choose to locate at site $(0,1)$, and then the second and third firms will locate at sites $(0,2)$ and $(0,3)$. Now, suppose that i_1 folds and vacates $(0,1)$. The value of site $(0,1)$ to a new firm is now

$$\epsilon(\rho_\mathbf{K}^A - d_{(0,1),(0,1)}) + (\rho_\mathbf{A}^A - d_{(0,1),(0,2)}) = \epsilon(\sqrt{2} + \delta) + (\sqrt{2} + \delta - 1) = (1 + \epsilon)(\sqrt{2} + \delta) - 1,$$

while the value of site $(1,3)$ to a new firm is now

$$(\rho_\mathbf{A}^A - d_{(1,3),(0,2)}) + (\rho_\mathbf{A}^A - d_{(1,3),(0,3)}) = (\sqrt{2} + \delta - 1) + (\sqrt{2} + \delta - \sqrt{2}) = 2\delta + \sqrt{2} - 1.$$

Since $\delta > \frac{\epsilon\sqrt{2}}{1-\epsilon}$, we see that the value of site $(1,3)$ exceeds that of site $(1,0)$.[43] Thus, a newly entering firm will locate at site $(1,3)$.

The preceding discussion shows that growth of clusters at their peripheries arises endogenously in our static model and is supported in a simple dynamic setting, too. By extension,

---

[41] This assumption implicitly defines the constant Δ of the theorem statement.

[42] We say that both $(0,1) \in \mathbf{K}$ and $(0,1) \in \mathbf{Z}$ as shorthand denoting the fact that there is a site in $\mathbf{K}$ and a site in $\mathbf{Z}$, both of which are ϵ-close to $(0,1)$ for some small ϵ.

[43] To see this, note that

$$(2\delta + \sqrt{2} - 1) - ((1 + \epsilon)(\sqrt{2} + \delta) - 1) = \delta(1 - \epsilon) - \sqrt{2}\epsilon > 0.$$

complete clusters could be dynamically vacated in favor of more appropriate locations due to entry and exit behavior.

By contrast, when natural advantages dominate Marshallian forces, as under the conditions of Proposition 3, the pattern of location choices for firms is dynamically stable. If a firm exits in this scenario, then the newly vacated site j is the most valuable site available. This is immediate by virtue of the fact that site j was previously selected and all resource attraction sites are known. Since natural advantage spillovers dominate, location j must be (weakly) optimal relative to other available sites. Thus, if the number of firms in industry A is constant, then the pattern of firm locations is constant if natural advantages are of central importance.

The dynamics are determined by Marshallian conditions if a natural advantage site loses its innate importance. If local Marshallian forces are strong, agglomeration can persist absent the original natural advantages. If local Marshallian forces are weak, the cluster may be vacated over time. The former scenario is in line with the related empirical work by Bleakley and Lin (2010) on the persistence of US cities originally centered on portage sites. This natural advantage has lost its economic importance, but Bleakley and Lin (2010) find continued and growing importance of these locations.

Looking ahead to future work, this framework provides a tractable setting for modelling agglomeration patterns and industry evolution. As industries mature, their needs evolve and the concomitant maximal radii shift. A classic example is moving from heavy dependence upon knowledge spillovers at industry birth to more production-centered factors at industry maturity. The associated changes in spillover lengths can be linked to dynamics in cluster patterns through this framework. Extending the model to have interesting interactions across industries could also afford evaluation of the role of diversity in industry birth similar to the work of Jacobs (1970) and Duranton and Puga (2001).

Additional Theoretical Notes

- One can further introduce inter-industry Marshallian forces (coagglomeration) by adding a third term to $g_{j,A(i)}$ of the form $\sum_{\substack{i' \neq i, \\ A(i') \neq A(i)}} G_{\mathbf{C}}^{A(i)}(d_{j,j(i')})$. An earlier version of this paper traces out some of the properties that arise due to coagglomeration (e.g., the spatial scale of coagglomeration can exceed agglomeration even if the interaction effect is weaker). We will pursue these coagglomeration traits further in future work.

- Our framework assumes that interaction costs exceed spillover benefits at some distance, with a single crossing property. While this condition holds in many settings, it is interesting to note an empirical case where it does not. Ethnic networks have long been associated with reducing interaction costs at great distances and thereby facilitating economic exchanges (e.g., Rauch 2001, Kerr 2008). In effect, these networks transmit the benefits of clusters that are typically confined spatially. Future extensions of this model could allow for networks, ethnic or otherwise, to connect clusters.

- In some industries, natural advantages are transmitted through contact with other firms. That is, if an industry-A firm i has access to a natural advantage site $k \in \mathbf{K}(A(i))$, and some other industry-A firm i' is proximate to $j(i)$ but not to k, then i' derives some of the benefit of the attraction site k. Although a full treatment of such a model is outside the scope of this paper, one special case stands out. If $G_{\mathbf{A}}(0) = 0$ and natural advantages are fully transmissible[44], then firms behave as in the case of Marshallian forces with attraction sites dictating the spatial distribution of firm locations. In this case, agglomeration again increases with a longer maximal radius of interaction.

---

[44]We say that natural advantages are *fully transmissible* if when $j(i') \in \mathcal{B}_{\rho^A}(j(i))$ and $d_{j(i),k} < \rho_{\mathbf{K}}^A$, then i' receives utility equal to $G_{\mathbf{K}}^A(d_{j(i),k})$, the full value of the natural advantage to i.

Empirical Appendix

This appendix begins with a deeper presentation of the technology flow maps discussed in the introduction. Color images of these maps are available at http://www.people.hbs.edu/wkerr/. We then introduce the US Census Bureau data employed and our approximation of the continuous Duranton and Overman (2005) metric. We close with additional descriptive statistics on US industrial concentration. The main text summarizes this material.

Technology Flows in Silicon Valley

We start with a characterization of technology flows in Silicon Valley and its surrounding region. The underlying structures of these technology flows exhibit several properties that are central to our location choice model, making this visual description a useful reference point. We employ individual records of all patents granted by the United States Patent and Trademark Office (USPTO) between January 1975 and May 2009 to inventors in the San Francisco Bay Area. We define the San Francisco Bay Area as any zip code that is within 75 miles of Redwood City, CA. Our analysis only considers industrial patents and the citations among them (e.g., we exclude technology transfer from Stanford University). We exclude self-citations.

We discuss in Section 3 the literature that exploits patent citations to measure knowledge diffusion or technology spillovers. We use localized citations among patents to portray technology sourcing zones at the zip code level within the San Francisco Bay Area. We assign firms to zip codes based on business directories and web-based searches. This data development process requires extensive manual effort as the San Francisco Bay Area produced as many industrial patents over the last 30 years as the bottom 29 states combined. We thus map with certainty firms that had two or more patents during the 1975-2009 period and use 20% subsamples for assignees with just a single patent. In total, we obtain physical addresses for 88% of the potential citations for our sample.[1]

We start with the structure of technology flows in the core area of Silicon Valley and work outwards. App. Figure 1a describes the core area that contains three-quarters of industrial patents filed from the San Francisco Bay Area. The core contains the top 10 zip codes in the whole region in terms of patenting (and 18 of the top 25). We first place these 10 zip codes on the map and describe their technology zones.

Panel A of App. Figure 1a shows the primary technology sourcing for zip code 95054, which is marked with a star. This zip code produced the most industrial patents of any zip code in the San Francisco Bay Area (more than 20,000, which is equal to the seventeenth largest state). It is home to firms like Intel, Sun, Advanced Micro Devices, and Applied Materials. The other points of the red shape are the three zip codes that firms in 95054 cited most in their work.

These three zip codes contain about 30% of 95054's external citations, with 95054 itself accounting for about 15%. On average, the top three external zips contain 41% of local external

---

[1]Our approach has several limitations. First, the vast majority of web searches yield current business addresses. In some cases, these searches also yield historical addresses for firms that have been acquired or gone out of business, but addresses are nonetheless more likely to be found for firms currently in operation. Second, even within the San Francisco Bay Area, some firms have multiple facilities. In such cases, our top priority is for locations where research is performed. Our results are robust when multiple assignments are possible to either using the most likely facility or to weighting each facility equally.

A second source of address information comes from parallel work matching firms into the Census Bureau data described in Section 4 (e.g., Kerr and Fu 2008, Balasubramanian and Sivadasan 2010, Akcigit and Kerr 2010). This approach produces a similar rank order of zip codes and description of cluster zones as those discussed below. Disclosure restrictions prohibit further tabulating these results.

citations for a zip code. This dependency grows modestly as one moves from the core (39%) to the periphery (47%). Likewise, the relative importance of external citations within the focal zip code declines as one moves to the periphery. It averages 10% across the region. In total, about 17% of external citations for the firms are contained in the San Francisco Bay Area. This share is greater than the more typical 10% share found at the metropolitan level due to both a larger expanse of land considered and due to this region being the dominate technology cluster.

Panel B further incorporates the second and third largest zip codes and their technology sourcing. The second largest zip code is 95134, home to firms like Altera, Micron, and Cisco. It has reciprocal sourcing from 95054, and it draws extensively from its neighbors 95035 and 95124. The third largest zip code is 94304 (e.g., Hewlett-Packard, Incyte, Palo Alto Research Center). Its technology sourcing extends down to include 95014 (e.g., Apple, Tandem, GTE) and 95120 (e.g., IBM Almaden Research Center).

Panels C and D show the top 6 and 10 zip codes, respectively. Although unrestricted in design, the primary technology sourcing zones for each zip code in the core are also contained in the core. This is true for all of the 18 zip codes in the core that are among the top 25. These zones are characterized by small, overlapping regions that often exhibit directional transmission. No technology sourcing zone spans the entire core. In larger maps, we represent the core as a shaded triangle, and for visual ease we will no longer maintain the micro technology flows. Panel D shows the boundaries of this triangle; its longest side is 25 miles in length.

App. Figure 1b places the Silicon Valley core in a larger map that includes surrounding parts of the San Francisco Bay Area. Downtown San Francisco and Oakland are to the north and still off the map. We map on this figure the seven largest zip codes for patenting that are not contained in the core itself (numbers 12, 13, 17, 19, 22, 24, and 25). The shape of each technology sourcing zone is again determined by the three zip codes that firms in the focal zip code cited most in their work. For visual ease, San Ramon and Santa Clara are indicated on the edge of the map at the location of their primary transportation route.

Similar to the core, these zones are characterized by small, overlapping regions that often exhibit directional transmission. The northwestern portion is a good example of the phenomena our theory describes. The orange zone (also labelled with a "1") for Menlo Park extends deepest into the core. The green zone for Redwood City ("2") shifts up and encompasses Menlo Park and Palo Alto but less of the core. The black zone for South San Francisco ("3") further shifts out and brushes the core. The northern side is similar. Fremont's pink zone extends deepest through the core, while points further out attach more to the core's edge. Finally, Santa Clara's zone connects up the eastern edge of the core.

The background of the map demonstrates the roles of geographical features and transportation costs. The shape of the core and outlying technology zones are dictated by mountains, protected land, and the bay. Likewise, the edges of sourcing zones conform to major transportation routes. Highway networks play a central role (e.g., horizontal flows across the top of the core, vertical flows on the eastern edge of the core). Bridges help cross the bay. While not shown in App. Figure 1a due to visual limitations, these same factors, especially the highway transportation routes, govern the technology flows within the core itself.

The length of each zone is influenced by congestion. The distance along the upper edge of the orange zone from Menlo Park into the core is 14 miles with an estimated travel time of 28 minutes. The distance along the upper edge of the black zone from South San Francisco to Menlo Park is 20 miles but with a comparable travel time of 32 minutes. A full spatial equilibrium

depends on additional factors like commercial and residential real estate prices.[2]

App. Figure 1c further expands the size of the map to include the northern portion of the San Francisco region. The bottom of the triangle for the Silicon Valley core is truncated in this map. The other regions are representations of major patent citation flows across zip codes in the San Francisco region. We use citation data from about 70 zip codes to identify these zones. These zip codes are subject to minimum requirements for counts of patent, citations, and firms.

The tan region represents the northwestern expansion from the core up to South San Francisco discussed in App. Figure 1b. The large blue region represents the northern expansion. Three zip codes from this group (i.e., Fremont, Hayward, and San Ramon) are also depicted in App. Figure 1b. The red zone from the bottom represents the region surrounding Santa Clara and its access. Finally, the green region represents interconnections between downtown San Francisco and Oakland and their access to Silicon Valley. These macro groupings also exhibit multiple, overlapping technology sourcing zones, none of which transverses the whole span. The importance of geography and transportation networks in shaping these zones is again apparent.

This mapping exercise is a quick tour of the structure of technology flows in the San Francisco Bay Area. We find similar outcomes in unreported variants. Examples include using four zip codes to define technology sourcing zones, separating major technology groups, separating entrants from long-lived incumbents, and examining only the most recent citations. More important for the development of our theory, Figure 1b in the main text shows similar patterns using predominant commuting zones for scientists and engineers from the 2000 Census of Populations.

US Census Bureau Data

Our estimates of industrial agglomeration patterns are developed through confidential data housed by the US Census Bureau. The Longitudinal Business Database (LBD) provides annual observations for every private-sector establishment with payroll from 1976 onward. The Census Bureau data are an unparalleled laboratory for studying the industrial structure of US firms. Sourced from US tax records and Census Bureau surveys, the micro-records document the universe of establishments and firms rather than a stratified random sample or published aggregate tabulations. In addition, the LBD lists physical locations of establishments rather than locations of incorporation, circumventing issues related to higher legal incorporations in states like Delaware.

We define industries through the three-digit level of the 1987 Standard Industrial Classification (SIC3). Industry assignment is at the establishment level for multi-unit firms. We focus on 362 SIC3 industries that are consistently measured from 1976 to 2001, after which time the transition to the NAICS industry classifications occurs. Our data contains 4.0 million establishments employing 68 million workers in 1977. In 1997, the data include 108 million workers and 5.8 million establishments. While we calculate agglomeration in all 26 years, our current discussion focuses just on 1997.

---

[2]Glaeser (2008) provides a detailed introduction. Saiz (2010) demonstrates the linkage between geography and city shape, and Baum-Snow (2007, 2010) studies the role of highways and transportation routes for city shapes.

Duranton and Overman (2005)

Duranton and Overman (2005, DO) construct a continuous metric of agglomeration. DO criticize indices like Ellison and Glaeser (1997) that employ discrete spatial units (e.g., states) to measure geographic concentration. This discreteness in effect makes the distance from Detroit, MI, to Chicago, IL, equivalent to that of Detroit, MI, to Miami, FL. The discreteness also fails to account for within-geographic unit distances that arise, for example, due to California's and Texas' larger spatial size relative to states within New England. DO instead propose analyzing agglomeration of industry A through a continuous index

$$
\hat{K}_A^{Ct}(d) = \frac{1}{hN^A(N^A-1)} \sum_{i=1}^{N^A-1} \sum_{i'=i+1}^{N^A} f\left(\frac{d - d_{j(i),j(i')}}{h}\right) \tag{1}
$$

$$
\hat{K}_A^{Emp}(d) = \frac{1}{h\sum_{i=1}^{N^A-1}\sum_{i'=i+1}^{N^A} e(i)e(i')} \sum_{i=1}^{N^A-1} \sum_{i'=i+1}^{N^A} e(i)e(i') f\left(\frac{d - d_{j(i),j(i')}}{h}\right).
$$

where $d_{j(i),j(i')}$ is the Euclidean distance between establishments i and i' within the focal industry A. The function f is a Gaussian kernel density function with bandwidth h. The summations are over every pairwise bilateral distance of establishments within the industry analyzed (i.e., $N^A(N^A-1)/2$ distances). The second formulation weights densities by establishment employments, $e(i)$ and $e(i')$, while the top formulation focuses on plant counts regardless of size, $e = 1$. We exclude the Ct and Emp superscripts in the main text and below for convenience.

These observed agglomeration densities are then compared to an underlying distribution of economic activity. This baseline is calculated through 1000 draws of hypothetical industries of equivalent size to the focal industry A and repeating the density estimation. An industry is said to be localized or dispersed at a given distance d when its observed agglomeration density shows substantial deviation from the random draws. Our estimation procedure analyzes these local departures in a systematic manner. DO likewise propose a global localization and dispersion measure that captures sustained deviations up to a distance threshold specified by the researcher. We discuss this measure after describing the base calculations.

Base Index Calculations

Our approximations to the DO metric employ county-to-county distances to measure bilateral distances between establishments. The county spatial unit is the most detailed spatial unit available for establishments in the LBD. We collapse establishment counts and employments into a county-industry spatial distribution for each year. We then apply the almost ten million potential pairwise distances between the 3141 counties. We calculate distances between county centroids using the Haversine flat-earth formula that abstracts from the earth's curvature. Establishments in the same county are given a uniform distance of one mile. Multiple factors yield conflicting recommendations regarding this latter choice—for example, accounting for larger land areas versus congestion effects with higher urban density. As shown below, the major movements/tipping we analyze happen at distances of 50 miles or greater, making this choice of secondary importance.

The combination of county-industry spatial distributions and county-to-county distances affords a spatial distribution of establishments ranging from [same county] to just over 6,000 miles.

We then apply the Gaussian kernel density function f to smooth the distance series. For this smoothing, the data are reflected around zero and the bandwidth h is chosen to minimize the mean integrated squared error. This smoothed density function is calculated for both establishment counts and employments. This process is repeated for each of the 362 industries annually.[3]

Local Confidence Intervals

These distance distributions are compared to counterfactuals of distances among randomly-drawn establishments from the LBD. These counterfactual baselines control for the overall spatial distribution of economic activity. Industries are deemed to be substantially agglomerated only to the extent they are more spatially clustered than US establishments generally. In preparing this test, it is important that each counterfactual have a similar number of observations (and therefore precision) as the focal distribution. It is also important, given the random nature of the establishment draw, to replicate the counterfactual baseline multiple times. DO converge on 1000 counterfactuals as the comparison for their agglomeration densities.

Computational restrictions require that we simplify how the confidence bands are calculated. Our first simplification is that we only calculate confidence bands using data from 1977 and 1997, rather than annually as in the base index calculation above. These two years are chosen as sample end points when economic censuses are conducted using the SIC system. When calculating our agglomeration measures, we find very similar patterns across all years 1976-2001 when using either the 1977 or 1997 baselines. Our current discussion focuses just on the 1977 and 1997 cross-sections, and the confidence intervals are specific to these years.

Second, we approximate the establishment count in the industry. Industries differ substantially in their establishment counts, ranging in 1997 from 18 industries having fewer than 100 establishments to 12 industries having more than 100,000. The maximum number of establishments was over 400,000 for Eating and Drinking Places (581). It is computationally prohibitive to match exactly the establishment count in all industries and years. We thus randomly draw without replacement 100,000 establishments for a hypothetical industry; this sampling is repeated 1000 times in both 1977 and 1997.

We then calculate sub-draws of these samples using increments of 100 establishments from a smallest industry size of 100 establishments up to 1500 establishments (i.e., 100, 200, ..., 1500). In addition, we calculate sub-draws using increments of 500 establishments from 1500 establishments to 5000 establishments and increments of 10,000 establishments up to 100,000 establishments. This procedure creates 32 comparable industry counts ranging from 100 establishments to 100,000 establishments. As this procedure is repeated in two years and for 1000 samples, 64,000 random baselines are created. Industries are paired with constructed confidence bands that most closely reflect the underlying observation counts (e.g., an industry with 223 establishments is compared to the confidence bands for 200 establishments).

To compute local confidence baselines akin to 5% and 95% confidence intervals, we apply the Gaussian kernel density function f to smooth the 64,000 counterfactuals. For each distance, we then identify the 50th and 950th extreme values in the smoothed series. We designate these confidence intervals as $K_A^{LCI-U}(d)$ and $K_A^{LCI-L}(d)$. These local confidence intervals are specific

---

[3]The largest values for pairwise distance are due to our inclusion of Alaska and Hawaii. Densities are very small at this level, however. We obtain equivalent results if pairwise establishment distances are capped at 3,000 miles, which is roughly the distance from Miami, FL, to Seattle, WA. The correlation for capped and uncapped distributions is 0.99.

to industries due to differences in industry size. While not expressed in the notation, the local baselines are separately calculated for establishment counts and employments.

Industry localization γ and dispersion ψ at distance d are defined using the formulas:

$$
\begin{aligned}
\gamma_A(d) &\equiv \max\left[\hat{K}_A(d) - K_A^{LCI-U}(d), 0\right] \\
\psi_A(d) &\equiv \max\left[K_A^{LCI-L}(d) - \hat{K}_A(d), 0\right] \text{ if } \gamma_A(d) = 0 \\
&\quad \text{and } 0 \text{ otherwise.}
\end{aligned}
\tag{2}
$$

An industry is said to be neither localized nor dispersed at a given distance unless its density falls outside of the 5% confidence bands. We also construct a combined index $\gamma_A^C(d)$:

$$
\gamma_A^C(d) \equiv \gamma_A(d) - \psi_A(d).
$$

This index does not create a lower or upper bound for localization or dispersion but instead views them symmetrically. This combined index is convenient for graphs and empirical analyses and continues to satisfy the necessary properties of an agglomeration estimator.

Global Localization and Dispersion

Most industries exhibit local clustering or dispersion at one distance or another. The DO approach therefore also calculates indices of global localization and dispersion. These measures quantify sustained deviations from the baselines of random placement. The procedure first specifies a distance threshold to analyze, and we denote this threshold as thr.

After defining the threshold thr, the DO approach first looks for evidence of substantial localization up to the threshold. If localization is not evident by the threshold, then evidence for global dispersion is sought up to the threshold. If an industry is spatially distributed similarly to overall economic activity, neither global localization nor dispersion will be found. This multi-step procedure and omission of distances beyond the threshold is defined in this way since density functions sum to one over the support. Thus, localization in one distance range will correspond to dispersion in other distances.

Formally, the global indices of localization Γ and dispersion Ψ are calculated using the formulas:

$$
\begin{aligned}
\Gamma_A(thr) &\equiv \sum_{d=1}^{thr} \max\left[\hat{K}_A(d) - K_{A,thr}^{GCI-U}(d), 0\right] \\
\Psi_A(thr) &\equiv \sum_{d=1}^{thr} \max\left[K_{A,thr}^{GCI-L}(d) - \hat{K}_A(d), 0\right] \text{ if } \Gamma_A(thr) = 0 \\
&\quad \text{and } 0 \text{ otherwise.}
\end{aligned}
\tag{3}
$$

$K_{A,thr}^{GCI-U}(d)$ and $K_{A,thr}^{GCI-L}(d)$ are confidence baselines over distances d for measuring global localization and dispersion, respectively. These global confidence intervals are wider than the local confidence intervals calculated with the 50th and 950th extreme values for each distance. $K_{A,thr}^{GCI-U}(d)$ and $K_{A,thr}^{GCI-L}(d)$ instead represent the local confidence values that when applied uniformly over the support $[1, thr]$ results in 5% of the counterfactual baselines showing deviation at both tails.

Thus, $K_{A,thr}^{GCI-U}(d)$ and $K_{A,thr}^{GCI-L}(d)$ are specific to the industry and threshold studied. The global bands depend upon the industry studied as the counterfactual random draws replicate

the number of establishments by industry. Industries with fewer establishments have wider confidence bands. While not included in DO's original formulation, the global bands also depend on the threshold analyzed. Confidence bands are wider for longer intervals studied; equivalently, the local confidence intervals become tighter with greater thresholds.[4]

When studying the UK, DO employ a distance threshold for identifying localization of 180 kilometers, which is the median distance among UK manufacturing establishments. The US' large land mass and scattered regional industrial centers do not provide as clear of a threshold, and the appropriate threshold is not dictated by theory. There is also a question of edge effects as discussed in Harrison and Kominers (2008). 1000 miles is approximately the median distance between US establishments and is on the order of magnitude for the distance between Detroit, MI, and Dallas, TX. The 25th, 10th, and 3rd percentiles of establishment bilateral distances are approximately 500, 250, and 100 miles, respectively. These are on the order of magnitude of Detroit, MI, to Washington, DC, Cincinnati, OH, and Lansing, MI, respectively.[5]

We therefore calculate global localization and dispersion across increments from 25 miles to 1000 miles. The 100 mile threshold is of comparable distance to the UK hurdle employed by DO (180 kilometers = 112 miles), while 1000 miles maintains the median concept. Finally, we again define a combined index $\Gamma_A^C(thr)$:

$$\Gamma_A^C(thr) \equiv \Gamma_A(thr) - \Psi_A(thr). \tag{4}$$

Descriptive Statistics

Table 1A describes industrial concentration patterns in 1997 using global metrics. Distance thresholds increase from 50 miles in Panel A to 1000 miles in Panel E. Within each panel, the four rows document localization and dispersion using establishment counts. Localization is typically higher in employment-weighted distributions than in those using plant counts (e.g., Holmes and Stevens 2002); the pairwise correlation of our agglomeration statistics using counts and employment is 0.7. Employment-weighted distributions and estimates for 1977 are available upon request.

To a distance threshold of 50 miles, a substantial number of industries show global localization or dispersion. Approximately 60% of industries display non-zero localization, while 30% display non-zero dispersion. 10% of industries show zero global localization and dispersion, and about 25% of industries show minimal deviation from the baseline economic activity of LBD establishments (i.e., localization or dispersion values less than 0.001).

At first it may appear odd that 60% of industries are deemed globally localized when the confidence bands are selected such that only 5% of the counterfactuals reach them. This is

---

[4]Our procedure creates a grid of confidence bands by industry size and distance threshold for 1977 and 1997. The industry sizes are the 32 interval ranges outlined earlier. We calculate exact global bands for eight thresholds and use linear interpolation for distance thresholds between them. These eight thresholds and the extreme values used for them are: 1000 miles (5th/995th), 500 (6/994), 375 (7/993), 250 (9/991), 175 (10/990), 100 (13/987), 75 (14/986), and 50 (18/982). As a comparison, DO also find that the 10/990 extreme values are appropriate for 5% global confidence bands at a threshold of approximately 100 miles in their UK sample. These extreme values are assigned through analyses of the complete distributions for the random draws of 100, 1000, 10,000, and 100,000 establishments and are most representative for 1000 establishments. Appropriate extreme values tend to widen (i.e., move to greater extremes) with large industry size due to the interaction of randomness and greater distance granularity with more observations. Barlet et al. (2009) closely analyze this issue.

[5]More precisely, the percentiles of unweighted establishment distances in 1997 are 1013 miles (50th), 573 (25th), 280 (10th), and 106 (3rd). The 1977 percentiles are 967, 541, 259, and 95, respectively. Employment weighted percentiles are shorter than count percentiles.

to be expected if agglomeration forces exist, however, as the counterfactuals are built upon all industrial establishments. The counterfactuals are not selected such that only 5% of industries will be deemed agglomerated. This result is expected if, as in our model, sites are distributed uniformly but agglomerative forces exist in nearly all industries.

Moving from Panel A to Panel E, the share of industries that are localized and the values of their localization increase as the distance threshold increases. By 1000 miles, 88% of industries exhibit localization. Our combined metric (4) also increases in value four fold. This is again despite the fact that the confidence intervals are being continually widened at longer thresholds to maintain only 5% of the counterfactual series reaching them. This confirms the model's implication that an increase in the distance horizon should typically lead to increased agglomeration.

Unreported analyses find that these patterns also hold within each one-digit SIC sector up to 1000 miles. It is important to note, however, that global localization can decrease across some distance ranges. For example, law offices can be concentrated at short distances within each city but also present in many cities. A short threshold captures the initial spike. As the threshold lengthens, the confidence intervals against which the densities are compared widen. This results in a weaker measured concentration and, in some case, even a determination that the industry is not globally concentrated at all.

Table 1B describes the local deviations that are behind these patterns. We also graph in the main paper the fraction of industries that are localized at each distance level to 1000 miles. Localization is most powerful at short distance horizons, with substantial attenuation commencing after 125 miles or so. Whereas about 60% of industries show localization at short spatial scales, only 30% or thereabouts are localized at distances of 900-1000 miles. This pattern is predicted in our model so long as some agglomerative force exists. Manufacturing industries are more agglomerated at short spatial scales than non-manufacturing industries, a pattern that reverses at longer horizons. This is to be expected given the transportable and tradeable nature of manufactured products.

Tables 2A-2D document extreme localization and dispersions values for 1997. Table 2A begins with the global metric. Some industries are extremely localized regardless of the threshold chosen. Most prominently, many of the tobacco, textiles, and apparel industries are very concentrated in the southeast of the US; moreover, these industries are very concentrated spatially at small distances within this region. Other cases are due to very concentrated natural advantages (e.g., mining industries) or idiosyncrasies in definitions (e.g., freight transport on Great Lakes).

On the other hand, some industries are quite localized at small distances but have clusters throughout the country. Many of the finance industries and professional services, for example, fall into this pattern. They have extreme values at small distances but not at larger distances. A third pattern is where industries are not very locally clustered but tend to congregate in one region of the US. This may again follow from natural advantages (e.g., regional oil and gas deposits) or from Marshallian interactions among firms (e.g., metalworking machinery).

Table 2B similarly documents extreme dispersion values. The dispersed industries are intuitive. Many represent services and trade that require homogeneous distribution like gas stations, hotels, hospitals, and utilities. Other dispersed industries like lumber and concrete are associated with transportation of heavy products, which limits the advantages of even local agglomeration. A third rationale is cheaper available land for dealers of used cars, mobile homes, and similar.

These purchases are substantial enough that consumers are willing to travel to the dealerships.[6]

Tables 2C and 2D repeat this analysis using the localization metrics by distance. These tables provide some simple intuition for the behavior described by our model. The localized concentration and dispersion at 50 miles is naturally closely aligned with the global levels noticed earlier. Most of the localized industries at 50 miles, however, are among the most dispersed industries when looking at concentric rings farther away. This is most easily seen in comparing the list of highly localized industries at 50 miles to highly dispersed industries at 1000 miles. There are also several cases of industries dispersed at short distances that show abnormal concentration at moderate to longer distances horizons (e.g., hospitals).

Industry movements across the panels can identify underlying forces of agglomeration and their spatial scales. This behavior is embedded in global concentrations metrics, like Tables 2A and 2B, but is even more prominent in localized density functions. These local density functions describe the edge of firm location decisions in ways that aggregate agglomeration functions obscure. This gradient or edge is the central piece that we study in the main text.

Table 3A documents the correlation of the combined concentration metric (4) by threshold. The correlation is very high among small thresholds and longer thresholds, but weaker in between. Index values using a 50 mile threshold have a 0.98 correlation with those using a 100 mile threshold. The correlation of the 50 mile threshold to a 375 mile threshold, however, declines to 0.81. From here, the correlation only further declines to 0.73 when examining a 1000 mile threshold.

Table 3B documents the correlation for local deviations. There is a much faster attenuation of the correlation in local deviations across neighboring distances than for the global metric, and a negative correlation exists for very distant values. These features are to be expected as the underlying kernel densities sum to one over the support. Excess density in one distance range yields below average density in another distance range. The non-linear summation behind the global metric in Table 3A, however, will typically continue to identify an industry as agglomerated at longer distance thresholds once it is agglomerated at lower thresholds.

The correlation between industry size and Γ_A^C is on the order of -0.1. Smaller industries are typically slightly more localized. The adjustment of confidence bands insures that this is not due to greater randomness in the series with fewer observations. It is more of an economic feature that descends from why the industry has a small number of plants (e.g., economies of scale in manufacturing versus consumer-oriented services). We have separately analyzed Γ_A^C values at several thresholds for different size industries. Growth in industry size is always associated with declining Γ_A^C values; likewise, longer distance thresholds are associated with greater Γ_A^C values in all size buckets.

---

[6]Tables 2A and 2B show evidence of agglomeration of production-related activities and dispersion of distribution-related activities among some extreme industries. While Motion Picture Production and Services (781) is extremely localized, Motion Picture Theaters (783) is quite dispersed. Other examples include textile production versus clothing stores and gas and oil refinement versus service stations.

App. Fig. 1a: Technology Flows in Silicon Valley Core
Patent citations across zip codes among industrial firms
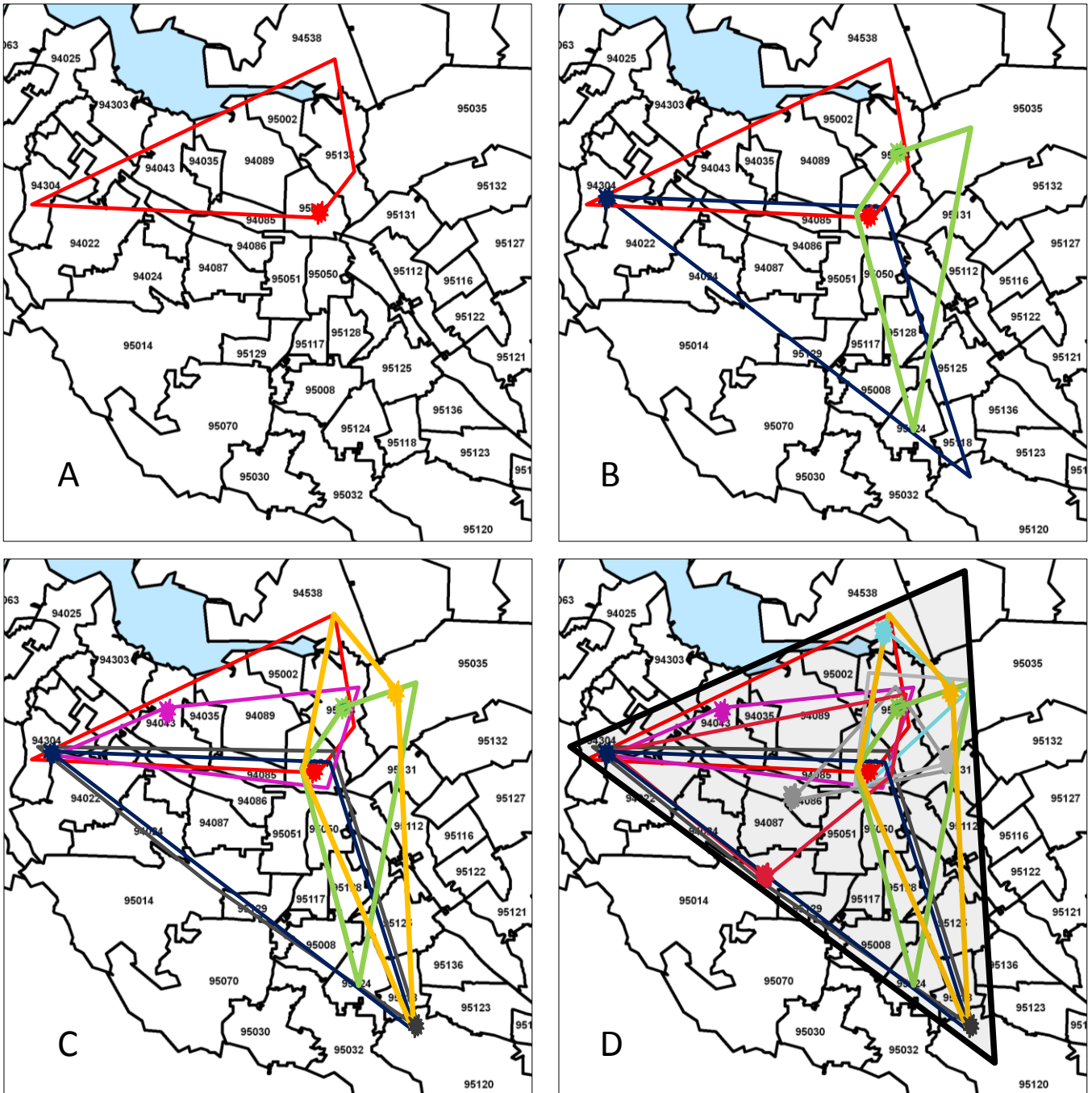
Notes: Figure describes the construction of the technology core for Silicon Valley. The core includes the top 10 zip codes in SF for patenting and 18 of the top 25. Panel A shows the primary technology sourcing for zip code 95054, which is marked with the star. This zip code produced the most industrial patents of any SF zip code (>20,000, equal to 17th largest state). The other points of the shape are the three zip codes that firms in 95054 cited most in their work. On average, the top three external zips contain 41% of local external citations for a zip code. Panel B incorporates the 2nd and 3rd largest zip codes and their technology sourcing. Panels C and D show the top 6 and 10 zip codes, respectively. While unrestricted in design, the primary technology sourcing zones for zip codes in the core are also contained in the core. These zones are small, overlapping regions that often exhibit directional transmission. In larger maps, we represent the core as a shaded triangle whose longest side is 25 miles in length.

App. Fig. 1b: Technology Sourcing around the Core
Seven largest patenting zip codes outside core and their sourcing
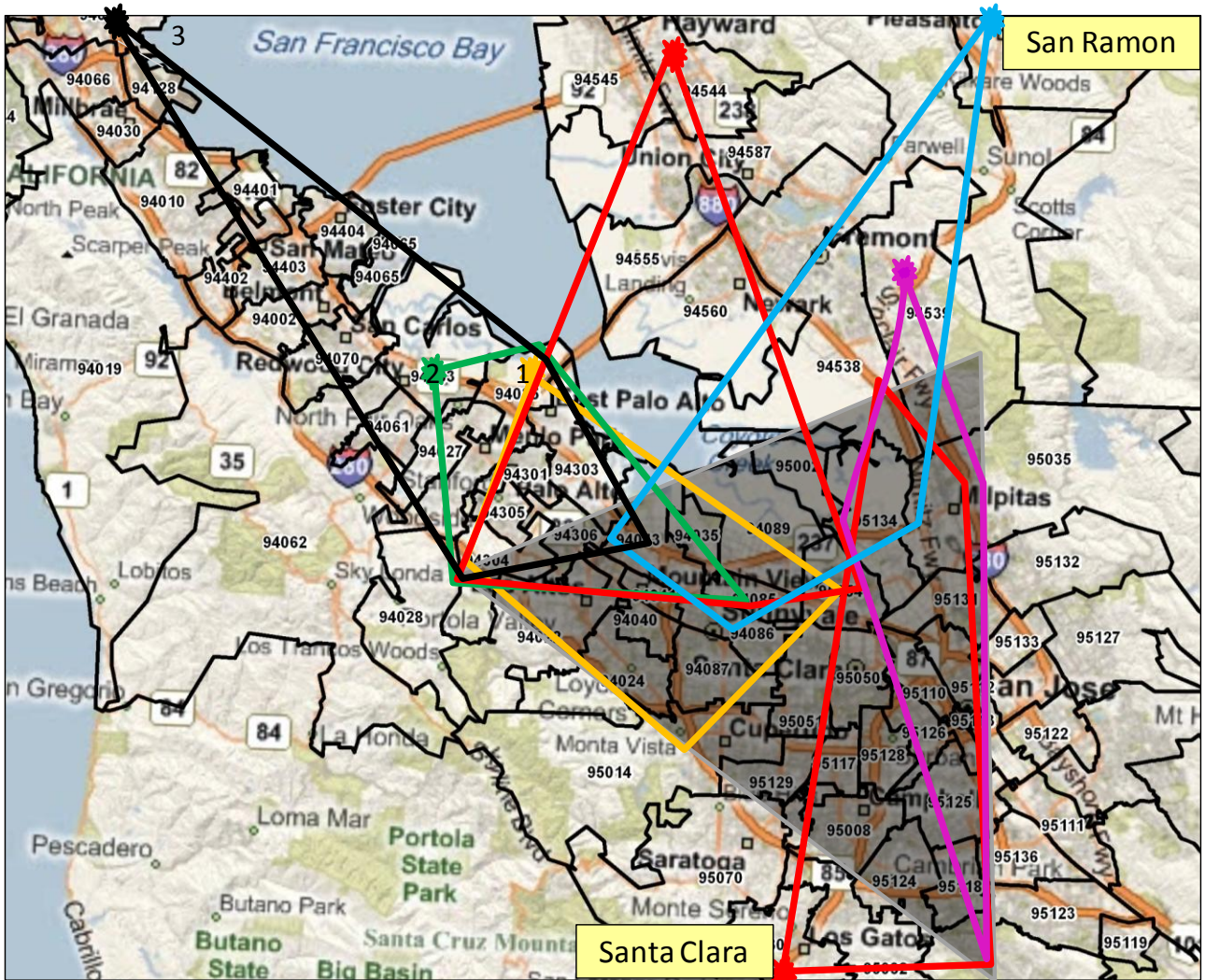
Notes: Figure continues to characterize technology flows for the San Francisco area. The Silicon Valley core depicted in App. Figure 1a is represented on this larger map as the shaded triangle. The Silicon Valley core contains 18 of the top 25 zip codes for patenting in the San Francisco area. This figure includes the seven largest zip codes for patenting that are not contained in the core itself (#12, 13, 17, 19, 22, 24, 25). The top 25 zip codes contain 87% of industrial patenting in the San Francisco area, with 76% in the core.

Similar to App. Figure 1a, the shape of each technology sourcing zone is determined by the three zip codes that firms in the focal zip code cite most in their work. For visual ease, San Ramon and Santa Clara are indicated on the edge of the map at the location of their primary transportation route. The northwestern portion is discussed in Figure 1a. The northeastern portion displays a similar structure.

Similar to the core, these technology zones are characterized by small, overlapping regions that exhibit directional transmission. The map also demonstrates the roles of geographical features (e.g., mountains, protected land) and transportation costs (e.g., highways, bridges). These forces substantially govern the shape of the technology zones and access to the core. These same features also play an important role in the technology flows evident in the core, especially the highway transportation routes.

App. Fig. 1c: Technology Sourcing at SF Region Level
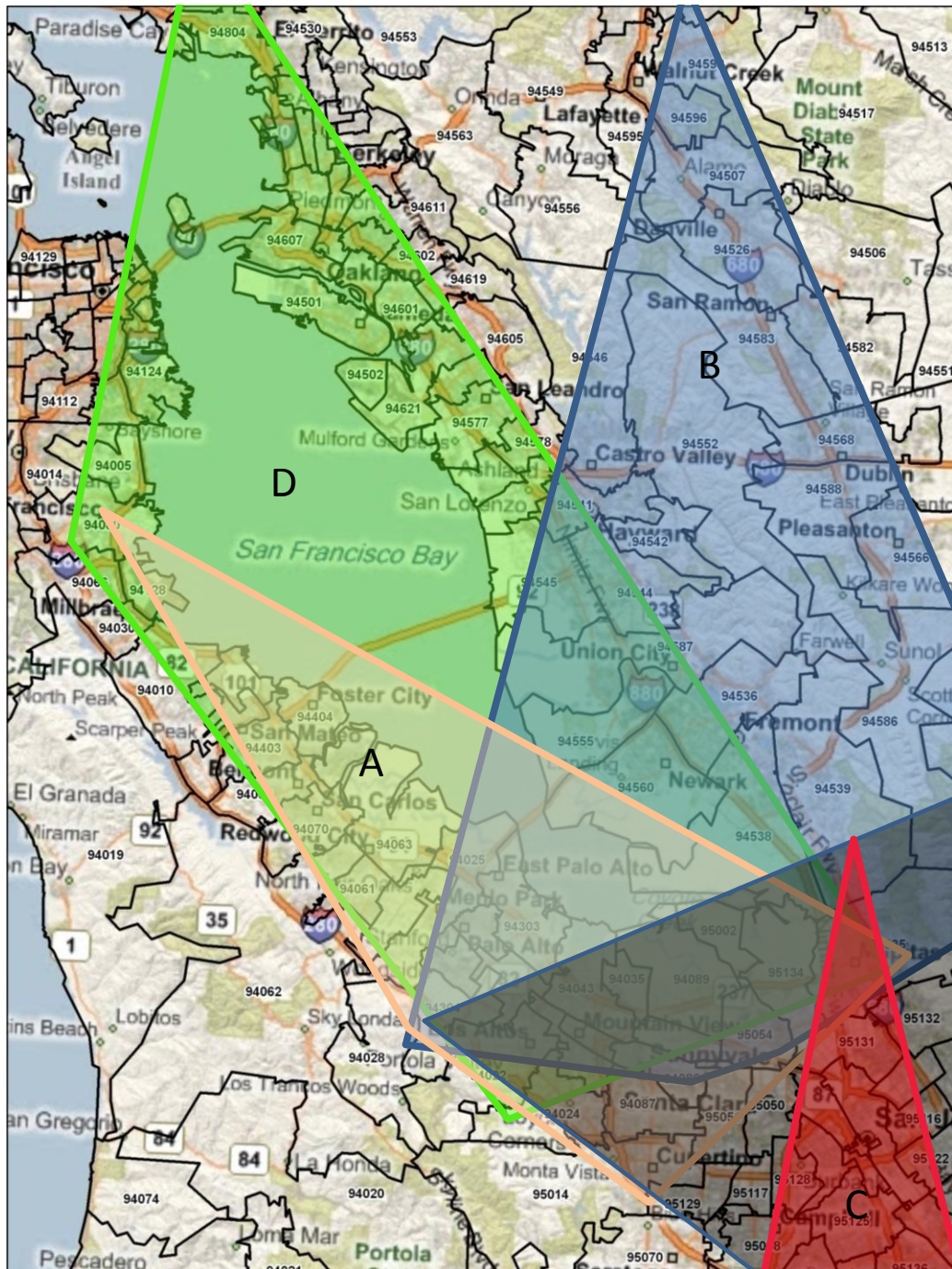## Broad technology sourcing zones across full region

Notes: Figure continues to characterize technology flows for the San Francisco area. The Silicon Valley core depicted in App. Figure 1a is represented as the black shaded triangle whose bottom is truncated at the edge of the map. The other regions are representations of major patent citation flows across zip codes in the San Francisco region. The tan region (A) represents the northwestern expansion from the core discussed in App. Figure 1b. The large blue region (B) represents the northern expansion. Three zip codes from this group are depicted in App. Figure 1b. The red zone (C) from the bottom represents the region surrounding Santa Clara and its access. Finally, the green region (D) represents interconnections between San Francisco and Oakland and their access to Silicon Valley. These macro groupings also exhibit multiple, overlapping technology sourcing zones, none of which transverses the whole region. Geography and transportation networks are important in defining shapes.

Table 1A: Descriptive Statistics for 1997 Global Agglomeration Measures

	Industry Count	Relevant Industries (non-zero)		
		Mean	Stand. Dev.	Maximum
Panel A. Distance Threshold of 50 Miles				
DO Global Localization - Establishment Counts	216	0.019	0.061	0.737
DO Global Dispersion	112	0.003	0.002	0.007
DO Combined Metric	362	0.011	0.049	
DO Localization and Dispersion < 0.001	84			
Panel B. Distance Threshold of 100 Miles				
DO Global Localization - Establishment Counts	236	0.026	0.066	0.751
DO Global Dispersion	93	0.004	0.003	0.012
DO Combined Metric	362	0.016	0.055	
DO Localization and Dispersion < 0.001	76			
Panel C. Distance Threshold of 250 Miles				
DO Global Localization - Establishment Counts	277	0.036	0.078	0.801
DO Global Dispersion	59	0.007	0.005	0.022
DO Combined Metric	362	0.026	0.070	
DO Localization and Dispersion < 0.001	67			
Panel D. Distance Threshold of 500 Miles				
DO Global Localization - Establishment Counts	302	0.046	0.089	0.816
DO Global Dispersion	34	0.011	0.011	0.038
DO Combined Metric	362	0.037	0.084	
DO Localization and Dispersion < 0.001	45			
Panel E. Distance Threshold of 1000 Miles				
DO Global Localization - Establishment Counts	318	0.054	0.092	0.815
DO Global Dispersion	25	0.019	0.020	0.083
DO Combined Metric	362	0.046	0.089	
DO Localization and Dispersion < 0.001	37			

Notes: Table describes Duranton and Overman (2005, DO) agglomeration metrics calculated from the Longitudinal Business Database of the US Census Bureau for 1997. 362 industries are defined through SIC3 classifications; all private, non-farm sectors are included in the sample. The distance threshold for determining global localization or dispersion is adjusted across panels. Within each panel, localization and dispersion statistics are calculated over observations displaying that trait. The combined metric is calculated for all industries as the value of the localization metric minus the value of the dispersion metric. The final row within each panel counts industries with approximately zero localization or dispersion. The text further describes the construction of these metrics.

Table 1B: Descriptive Statistics for 1997 Local Agglomeration Measures

	Industry Count	Relevant Industries (non-zero)		
		Mean	Stand. Dev.	Maximum
Panel A. Distances of 50 Miles				
DO Localization - Establishment Counts	212	0.329	0.957	12.528
DO Dispersion	113	0.056	0.042	0.306
DO Combined Metric	362	0.175	0.755	
DO Localization and Dispersion < 0.001	37			
Panel B. Distances of 100 Miles				
DO Localization - Establishment Counts	219	0.175	0.289	1.911
DO Dispersion	89	0.034	0.031	0.240
DO Combined Metric	362	0.098	0.245	
DO Localization and Dispersion < 0.001	55			
Panel C. Distances of 250 Miles				
DO Localization - Establishment Counts	194	0.118	0.185	1.173
DO Dispersion	91	0.038	0.038	0.260
DO Combined Metric	362	0.053	0.153	
DO Localization and Dispersion < 0.001	80			
Panel D. Distances of 500 Miles				
DO Localization - Establishment Counts	162	0.101	0.101	0.555
DO Dispersion	132	0.086	0.083	0.439
DO Combined Metric	362	0.014	0.119	
DO Localization and Dispersion < 0.001	73			
Panel E. Distances of 1000 Miles				
DO Localization - Establishment Counts	90	0.024	0.024	0.116
DO Dispersion	181	0.057	0.059	0.315
DO Combined Metric	362	-0.022	0.056	
DO Localization and Dispersion < 0.001	94			

Notes: See Table 1A. This table considers local agglomeration characteristics.

Table 2A: Most Globally Localized Industries in 1997 by Distance Threshold

Industry	Local.	Industry	Local.	Industry	Local.	Industry	Local.
50 Mile Threshold		100 Mile Threshold		500 Mile Threshold		1000 Mile Threshold	
Anthracite Mining (123)	0.737	Anthracite Mining (123)	0.751	Anthracite Mining (123)	0.816	Anthracite Mining (123)	0.815
Fur Goods (237)		Fur Goods (237)		Bituminous Coal and Lignite Mining (122)		Fur Goods (237)	
Carpets and Rugs (227)		Carpets and Rugs (227)		Fur Goods (237)		Bituminous Coal and Lignite Mining (122)	
Women's, Misses', & Jrs' Outerwear (233)		Bituminous Coal and Lignite Mining (122)		Freight Transport on Great Lakes (443)		Tobacco Stemming and Redrying (214)	
Apparel, Piece Goods and Notions (513)	d	Women's, Misses', & Jrs' Outerwear (233)	0.201	Tobacco Stemming and Redrying (214)	d	Freight Transport on Great Lakes (443)	d
Tobacco Stemming and Redrying (214)		Tobacco Stemming and Redrying (214)		Coal Mining Services (124)		Cigarettes (211)	
Bituminous Coal and Lignite Mining (122)		Commodity Brokers and Dealers (622)		Cigarettes (211)		Coal Mining Services (124)	
Commodity Brokers and Dealers (622)		Foreign Banks (608)		Yarn and Thread Mills (228)		Yarn and Thread Mills (228)	
Motion Picture Product'n & Services (781)		Cigarettes (211)		Carpets and Rugs (227)		Carpets and Rugs (227)	
Foreign Banks (608)	0.079	Apparel, Piece Goods and Notions (513)	d	Crude Petroleum and Natural Gas (131)	0.215	Knitting Mills (225)	0.269
Costume Jewelry and Notions (396)		Costume Jewelry and Notions (396)		Women's, Misses', & Jrs' Outerwear (233)		Crude Petroleum and Natural Gas (131)	
Knitting Mills (225)		Freight Transport on Great Lakes (443)		Costume Jewelry and Notions (396)		Commodity Brokers and Dealers (622)	
Cigarettes (211)		Coal Mining Services (124)		Knitting Mills (225)		Metalworking Machinery (354)	
Freight Transport on Great Lakes (443)		Knitting Mills (225)		Metalworking Machinery (354)		Natural Gas Liquids (132)	
Taxicabs (412)	0.061	Motion Picture Product'n & Services (781)	0.106	Foreign Banks (608)	0.176	Women's, Misses', & Jrs' Outerwear (233)	0.206

Notes: See Tables 1A and 1B. "d" indicates suppressed value due to disclosure restrictions.

Table 2B: Most Globally Dispersed Industries in 1997 by Distance Threshold

Industry	Disp.	Industry	Disp.	Industry	Disp.	Industry	Disp.
50 Mile Threshold		100 Mile Threshold		500 Mile Threshold		1000 Mile Threshold	
Petroleum and Petroleum Products (517)	0.007	Petroleum and Petroleum Products (517)	0.012	Misc. Foods and Kindred Products (209)	0.038	Misc. Foods and Kindred Products (209)	0.083
Electric Services (491)		Hotels and Motels (701)		Airports, Flying Fields & Services (458)		Camps and Recreat'n Vehicle Parks (703)	
Hotels and Motels (701)		Electric Services (491)		Camps and Recreat'n Vehicle Parks (703)		Airports, Flying Fields & Services (458)	
Concrete, Gypsum and Plaster Prod. (327)		Radio and Television Broadcasting (483)		Hotels and Motels (701)		Preserved Fruits and Vegetables (203)	
Water Supply (494)	0.006	Heavy Construction, Exc Highway (162)	0.010	Automotive Dealers, Nec (559)	0.025	Toys and Sporting Goods (394)	0.038
Heavy Construction, Exc Highway (162)		Water Well Drilling (178)		Toys and Sporting Goods (394)		Recreational Vehicle Dealers (556)	
Water Well Drilling (178)		Concrete, Gypsum and Plaster Prod. (327)		Recreational Vehicle Dealers (556)		Public Warehousing and Storage (422)	
Farm Product Raw Materials (515)		Recreational Vehicle Dealers (556)		Used Merchandise Stores (593)		Guided Missiles, Space Vehicles, Parts (376)	
Radio and Television Broadcasting (483)		Hospitals (806)		Public Warehousing and Storage (422)		Footwear, Exc Rubber (314)	
Fed. & Federally-Sponsored Credit (611)	0.006	Fed. & Federally-Sponsored Credit (611)	0.008	Water Well Drilling (178)	0.013	Family Clothing Stores (565)	0.014
Camps and Recreat'n Vehicle Parks (703)		Newspapers (271)		Family Clothing Stores (565)		Miscellaneous Nonmetallic Minerals (149)	
Used Car Dealers (552)		Water Supply (494)		Miscellaneous Nonmetallic Minerals (149)		Pottery and Related Products (326)	
Lumber and Oth. Building Materials (521)		Camps and Recreat'n Vehicle Parks (703)		Agricultural Chemicals (287)		Millwork, Plywood and Struct. Mbrs. (243)	
Newspapers (271)		Variety Stores (533)		Millwork, Plywood and Struct. Mbrs. (243)		Motion Picture Theaters (783)	
Mobile Home Dealers (527)	0.005	Automotive Dealers, Nec (559)	0.008	Motion Picture Theaters (783)	0.010	Job Training & Related Services (833)	0.010

Notes: See Tables 1A and 1B. "d" indicates suppressed value due to disclosure restrictions.

Industry	Local.	Industry	Local.	Industry	Local.	Industry	Local.
50 Miles		100 Miles		500 Miles		1000 Miles	
Anthracite Mining (123)	12.528	Bituminous Coal and Lignite Mining (122)	1.911	Knitting Mills (225)	0.555	Title Abstract Offices (654)	0.116
Carpets and Rugs (227)		Tobacco Stemming and Redrying (214)		Tobacco Stemming and Redrying (214)		Gas Production and Distribution (492)	
Women's, Misses', & Jrs' Outerwear (233)		Cigarettes (211)		Natural Gas Liquids (132)		Automotive Dealers, Nec (559)	
Bituminous Coal and Lignite Mining (122)	2.059	Coal Mining Services (124)		Crude Petroleum and Natural Gas (131)		Cigars (212)	
Tobacco Stemming and Redrying (214)		Freight Transport on Great Lakes (443)	d	Oil and Gas Field Services (138)	0.402	Offices of Osteopathic Physicians (803)	0.071
Commodity Brokers and Dealers (622)		Yarn and Thread Mills (228)		Freight Transport on Great Lakes (443)		Water Well Drilling (178)	
Apparel, Piece Goods and Notions (513)		Foreign Banks (608)		Cigarettes (211)		Petroleum and Petroleum Products (517)	
Foreign Banks (608)		Costume Jewelry and Notions (396)		Tires and Inner Tubes (301)		Fed. & Federally-Sponsored Credit (611)	
Motion Picture Product'n & Services (781)		Carpets and Rugs (227)		Farm Product Raw Materials (515)		Boat Dealers (555)	
Cigarettes (211)	d	Fur Goods (237)	0.966	Wooden Containers (244)	0.285	Airports, Flying Fields & Services (458)	0.061
Costume Jewelry and Notions (396)		Knitting Mills (225)		Blast Furnace & Basic Steel Products (331)		Water Transportation Services (449)	
Freight Transport on Great Lakes (443)		Jewelry, Silverware and Plated Ware (391)		Railroad Equipment (374)		Oil and Gas Field Services (138)	
Knitting Mills (225)		Broadwoven Fabric Mills, Wool (223)		Crushed and Broken Stone (142)		Combination Utility Services (493)	
Coal Mining Services (124)		Copper Ores (102)		Sawmills and Planing Mills (242)		Hospitals (806)	
Jewelry, Silverware and Plated Ware (391)	1.147	Footware Cut Stock (313)	0.640	Variety Stores (533)	0.249	Miscellaneous Investing (679)	0.048

Notes: See Tables 1A and 1B. "d" indicates suppressed value due to disclosure restrictions.

Table 2D: Most Dispersed Industries in 1997 by Distance

Industry	Disp.	Industry	Disp.	Industry	Disp.	Industry	Disp.
50 Miles		100 Miles		500 Miles		1000 Miles	
Fur Goods (237)	0.306	Anthracite Mining (123)	0.240	Fur Goods (237)	0.439	Anthracite Mining (123)	0.315
Petroleum and Petroleum Products (517)		Hotels and Motels (701)		Foreign Banks (608)		Freight Transport on Great Lakes (443)	
Electric Services (491)		Petroleum and Petroleum Products (517)		Women's, Misses', & Jrs' Outerwear (233)		Tobacco Stemming and Redrying (214)	
Hotels and Motels (701)		Automotive Dealers, Nec (559)		Costume Jewelry and Notions (396)		Bituminous Coal and Lignite Mining (122)	
Radio and Television Broadcasting (483)	0.120	Recreational Vehicle Dealers (556)	0.076	Jewelry, Silverware and Plated Ware (391)	0.322	Knitting Mills (225)	0.236
Concrete, Gypsum and Plaster Prod. (327)		Radio and Television Broadcasting (483)		Motion Picture Product'n & Services (781)		Yarn and Thread Mills (228)	
Water Well Drilling (178)		Airports, Flying Fields & Services (458)		Anthracite Mining (123)		Fur Goods (237)	
Water Supply (494)		Electric Services (491)		Motion Picture Distribution & Services (782)		Coal Mining Services (124)	
Heavy Construction, Exc Highway (162)		Heavy Construction, Exc Highway (162)		Deep Sea Domestic Trans. of Freight (442)		Women's, Misses', & Jrs' Outerwear (233)	
Fed. & Federally-Sponsored Credit (611)	0.110	Miscellaneous Nonmetallic Minerals (149)	0.068	Apparel, Piece Goods and Notions (513)	d	Carpets and Rugs (227)	0.168
Hospitals (806)		Used Merchandise Stores (593)		Aircraft and Parts (372)		Metalworking Machinery (354)	
Camps and Recreat'n Vehicle Parks (703)		Hospitals (806)		Water Transportation of Passengers (448)		Metal Forgings and Stampings (346)	
Newspapers (271)		Agricultural Chemicals (287)		Producers, Orchestras & Entertainers (792)		Motion Picture Product'n & Services (781)	
Used Car Dealers (552)		Water Well Drilling (178)		Handbags & Pers. Leather Goods (317)		Cigarettes (211)	
Recreational Vehicle Dealers (556)	0.101	Auto and Home Supply Stores (553)	0.058	Electronic Components & Accessories (367)	0.169	Screw Machine Products, Bolts, etc. (345)	0.144

Notes: See Tables 1A and 1B. "d" indicates suppressed value due to disclosure restrictions.

Table 3A: Correlation of Combined 1997 Global Agglomeration Index by Distance Threshold

	50 mi.	75 mi.	100 mi.	250 mi.	375 mi.	500 mi.	100 mi.
50 mi.	1.00						
75 mi.	0.99	1.00					
100 mi.	0.98	1.00	1.00				
250 mi.	0.88	0.93	0.96	1.00			
375 mi.	0.81	0.86	0.90	0.98	1.00		
500 mi.	0.77	0.82	0.86	0.96	0.99	1.00	
1000 mi.	0.73	0.78	0.82	0.92	0.96	0.98	1.00

Notes: See Tables 1A and 1B.

Table 3B: Correlation of Combined 1997 Local Agglomeration Index by Distance

	50 mi.	75 mi.	100 mi.	250 mi.	375 mi.	500 mi.	100 mi.
50 mi.	1.00						
75 mi.	0.56	1.00					
100 mi.	0.35	0.92	1.00				
250 mi.	0.18	0.56	0.70	1.00			
375 mi.	0.20	0.33	0.38	0.84	1.00		
500 mi.	-0.18	-0.05	0.05	0.50	0.76	1.00	
1000 mi.	-0.53	-0.64	-0.65	-0.52	-0.40	-0.08	1.00

Notes: See Tables 1A and 1B.