Agglomerative Forces and Cluster Shapes

Agglomerative Forces and Cluster Shapes*

William R. Kerr
Harvard University and NBER

Scott Duke Kominers
University of Chicago

May 2013

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 patent technology clusters.

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

Key Words: Agglomeration, Clusters, Networks, Industrial Organization, Silicon Valley, Entrepreneurship, Technology Flows, Patents.

*Comments are appreciated and can be sent to wkerr@hbs.edu and skominers@gmail.com. An earlier version of this paper was titled "Tipping Points and Agglomeration Bubbles." This research is supported by Harvard Business School and the Kauffman Foundation. Alexis Brownell and Kristen Garner provided excellent research assistance. We thank David Albouy, Jerry Carlino, Jim Davis, Scott Dempwolf, Gilles Duranton, Ed Glaeser, Vernon Henderson, Bob Hunt, Guido Imbens, Yannis Ioannides, Sonia Jaffe, Ramana Nanda, Stuart Rosenthal, Steve Ross, Scott Stern, William Strange, and seminar participants at Aalto University School of Economics, American Economic Association, Bank of Finland, Boston University, Boston Urban and Real Estate Seminar, European Regional Science Association, Georgia Institute of Technology, Harvard University, NBER Productivity, NBER Urban Studies, Philadelphia Federal Reserve Board, Rimini Conference in Economics and Finance, University of California San Diego, University of Pennsylvania Wharton School, and University of Virginia for their insights. Kerr thanks the Bank of Finland for hosting him during a portion of this project. The research in this paper was conducted while Kerr was a Special Sworn Status researcher of the US Census Bureau at the Boston Census Research Data Center (BRDC). Support for this research from NSF grant ITR-0427889 [BRDC] and an NSF Graduate Research Fellowship [Kominers] is gratefully acknowledged. Research results and conclusions expressed are the authors' and do not necessarily reflect the views of the Census Bureau or NSF. This paper has been screened to ensure that no confidential data are revealed.

1 Introduction

Agglomeration—industrial clustering—is a key feature of economic geography. A vast body of research now documents the prevalence of agglomeration in many industries and countries, and a number of studies have established its particular importance for firm and worker productivity. Duranton and Puga (2004) and Rosenthal and Strange (2004) provide theoretical and empirical reviews, respectively. Moving from these measurements, researchers have recently sought to identify the economic rationales for firm collocation and thereby the sources of the associated productivity gains. While the list of potential suspects dates back to Marshall (1920)—most notably labor market pooling, customer-supplier interactions, and knowledge flows—we are just beginning to separate the relative importance of these forces.

Research on the spatial horizons over which different agglomerative forces act often takes one of two approaches. A first approach considers "regional" evidence. Examples include Rosenthal and Strange (2001, 2004), Duranton and Overman (2005, 2008), and Ellison et al. (2010). This approach begins by measuring the degree to which each industry is agglomerated across a chosen spatial horizon (e.g., counties, cities, states). A second step then correlates differences in observed agglomeration to the traits of the industries. For example, we might observe that industries intensive in R&D efforts are more agglomerated when using counties or cities as a spatial unit than industries that do not depend upon R&D. Similarly, we might observe that industries with strong customer-supplier linkages are agglomerated at the regional level. A common inference from these patterns, as one example, is that knowledge flows act over a shorter spatial distance than input-output interactions, as the knowledge-intensive industries are more heavily grouped together at the county level.

In parallel, a second strand of work considers "local" evidence on agglomerative interactions. Rather than discerning agglomerative forces from region-industry data, this line of research attempts to measure productivity gains directly at the establishment level. Prominent examples include Rosenthal and Strange (2003, 2008) and Arzaghi and Henderson (2008). A common approach is to estimate a plant-level production function that includes as explanatory variables the count of plants in the same industry observed within five miles of the focal plant, within ten miles, and so on (or within the same county, city, and state). These studies often conclude that spillover effects decay sharply with distance, with the forces being orders of magnitude stronger over the first few city blocks than they are when firms are 2-5 miles apart. These productivity studies are just the tip of the iceberg, however, with many related research strands measuring directly the distances over which humans or firms interact (e.g., patent citations, commuting, etc.). As agglomerative forces depend upon these interactions, these studies also describe the local interactions that give rise to the clusters that we observe.

There is a substantial gap between these two approaches. Despite their individual progress, we have very little understanding of how the "local" interactions aggregate up into "regional"

shapes and sizes of industries we observe in the data. The easiest way to observe this gap is to consider the spatial distances discussed by the two approaches. The regional literature often concludes that technology spillovers have a shorter spatial horizon than labor market pooling by comparing county- and city-level data. But counties have over 75,000 people in population on average, and the spatial size of counties is much greater than what "local" interactions suggest is the relevant range. If studies find that knowledge flows decrease sharply within a single building (e.g., Olson and Olson 2003), why would we believe that we can infer useful comparisons of knowledge spillovers and labor pooling from regional data when the spatial scales of our data swamp the micro-interactions by orders of magnitude?

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 of similar firms. Agglomerative forces in our model are localized because firms face interaction costs. Spillover benefits exceed these interaction costs at short distances, and thus firms choose to interact. Beyond some distance, however, interaction becomes unprofitable and firms no longer engage with each other. 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. By building the cluster up from the micro-interactions, we obtain additional insights into the structure of clusters and the regional data we observe.

Silicon Valley is the world's most famous cluster, and many observers credit its success to technology spillovers. Figure 1 illustrates the foundations of our theoretical framework using technology flows in 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 in the San Francisco area that are 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. While technology sourcing for individual firms is localized, the resulting cluster extends over a larger expanse of land.[1]

[1]The empirical appendix of our NBER working paper contains additional maps that show these small, over-

Our model replicates these features and 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. This connection helps bring together the diverse literature strands described earlier.

After deriving our theoretical predictions, we empirically illustrate the model using 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. This empirical work primarily employs agglomeration metrics that are continuous, like in the metric 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. Early "island" models of agglomeration—in which agglomerative forces act only within sites—implicitly feature maximal radius of interaction 0 (e.g., Krugman 1991, Fujita and Thisse 1996, Ellison and Glaeser 1997). More recently, maximal radii also have been observed in more continuous models (e.g., Arzaghi and Henderson 2008, Duranton and Overman 2005). However, to our knowledge, our framework is the first to identify how variations in the maximal radius govern the shapes and sizes of clusters. At the core of this contribution is the simple mechanism of 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.[2]

Our central empirical contribution is a framework, motivated by our theoretical model, for meaningful analysis of agglomerative forces with continuous distance horizons. Previous work

lapping regions are also evident in the core itself and in other parts of the San Francisco region. These properties are also evident in labor commuting patterns in the region. Arzaghi and Henderson (2008) and Carlino et al. (2012) provide related visual displays.

[2]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.

Studies of agglomeration metrics include Ellison and Glaeser (1997), Maurel and Sédillot (1999), Marcon and Puech (2003), Mori et al. (2005), Barlet et al. (2012), Ellison et al. (2010), Billings and Johnson (2011), and Carlino et al. (2012). Recent related work on cities includes Rozenfeld et al. (2012) and Helsley and Strange (2012).

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, Ellison et al. 2010). Our framework is an important step towards jointly considering agglomeration at different distances (25, 75, and 250 miles) simultaneously. We hope that future research can similarly analyze other factors that govern clusters' shapes and sizes.

In addition to the related work already mentioned, our empirical work with patents relates to two other recent studies also considering continuous density measurements. Carlino et al. (2012) develop a multiscale core-cluster approach to measure the agglomeration of R&D laboratories across continuous space. In many respects, their metric's nesting approach parallels our theoretical focus on overlapping radii of interaction that build to a larger cluster. Likewise, some of their empirical results (e.g., clustering at local scales and at about 40 miles of distance) are also evident under our measures. Similarly, Murata et al. (2012) use continuous density estimations with patent citations to address the question of how localized are knowledge flows. Their careful metric design allows them to bridge the well-known debate between Jaffe et al. (1993, 2005) and Thompson and Fox-Kean (2005) and parse the underlying assumptions embedded in each study. Our work differs from these studies in several ways, but the most important difference is the theoretical focus and hypothesis testing about how different forms of interaction produce observable changes in cluster shapes and sizes.[3]

Section 2 presents our theoretical model. Section 3 describes our empirical strategy and data and provides initial evidence for our model's building blocks using first- and later-generations of patent citations. Section 4 then undertakes specific measurements of technology-level spillover radii and tests our model's predictions one at a time. Section 5 then introduces our continuous density measurements and tests the model predictions. The last section concludes.

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. We keep this initial exposition as simple as possible, and we conclude this section with a discussion of richer frameworks and extensions.

[3]Other related studies not previously mentioned include Audretsch and Feldman (1996), Head and Mayer (2004), Hanson (2005), Greenstone et al. (2010), Delgado et al. (2009), Holmes and Lee (2012), Fallick et al. (2006), Glaeser and Kerr (2009), Menon (2009), Bleakley and Lin (2012), Alcacer and Chung (2007), Pe'er and Vertinsky (2009), Alfaro and Chen (2010), Dauth (2010), Marx and Singh (2012), and Dempwolf (2012). 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, Behrens et al. 2010, Sarvimäki 2010, Fu and Ross 2012). Jackson (2008) outlines a complementary literature on economic networks.

2.1 Basic Framework

There are N firms indexed by i. These firms i sequentially select their locations, denoted $j(i)$, from a fixed set $\mathbf{Z} \subset \mathbb{R}^2$ of potential sites, each of which can hold at most one firm.[4] Sites are drawn at random according to a uniform distribution in advance of any firms' location decision and are fixed. 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 are driven by intra-industry Marshallian forces representing productivity spillovers that firms generate by being in proximity to each other. Three common examples are customer-supplier interactions (e.g., reducing transportation costs for intermediate goods), labor pooling, and knowledge exchanges.

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 deterministic benefit of site $j \in \mathbf{Z}$ to a firm i is given by

$$g_j(i) \equiv \sum_{i' \neq i} G(d_{j,j(i')}),$$

for some continuous, decreasing function G. The value g_j represents the degree to which spillovers from other sites make site j specifically attractive to firms. We assume the standard comparative static that G is decreasing, so that agglomerative forces decline over space. Additionally, for simplicity, we assume that agglomerative forces act across all distances. That is, $G(d) > 0$ for all $d \geq 0$.

We assume that a firm chooses randomly among sites $j_1, \ldots, j_\ell$ over which that firm would be indifferent if forced to choose purely on the basis of spatial attraction.[5] We also assume that firms are not forward-looking, so that the n-th firm to enter, i_n $(1 \leq n \leq N)$, chooses its location $j = j(i_n) \in \mathbf{Z}$ to maximize $g_j(i_n)$ 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. *A firm must pay cost c to interact with another firm.*

These fixed costs c relate to the costs of transporting goods, people, or ideas across firms. Opportunity costs and search costs are the simplest examples, and these costs can be specific

[4] In Section 2.4, we discuss the possibility that multiple firms may occupy (and congest) the same site.

[5] The exact specification of the distribution of random site choice does not matter for our theoretical results and may be conditioned upon the set of sites already occupied, but we do require that it is identical across firms.

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.[6]

Firms invest in establishing contacts when the benefits of doing so equal or exceed the associated costs of interaction. Specifically, firm i only invests in contact with a firm i' if $G(d_{j(i),j(i')}) \geq c$. This defines a strict distance over which firm i finds interactions profitable,

$$d_{j(i),j(i')} \leq \rho \equiv \max\{d : G(d) \geq c\}.$$

Therefore, we immediately observe the following result.

Proposition 1. *Firms at sites further than distance ρ from a firm i cannot profitably interact with i. That is, firm i derives no direct benefits from the presence of firms at locations j with $d_{j,j(i)} > \rho$.*

Proof: Immediate from text.

The key consequence of Proposition 1 is that agglomerative forces in practice act only over finite distances. We call ρ the *maximal radius of interaction* (or just the *maximal radius*). The maximal radius is (weakly) decreasing in the cost c and increasing in the levels of the decay function G. In other words, lower costs or weaker attenuation of benefits lead to larger maximal radii.

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 required is that interaction costs exceed interaction benefits at some distance with a single crossing.

2.3 A Cluster-Based Theory of Agglomeration

We next examine how clusters form in our model and illustrate clusters' properties. Figures 2a-2d provide a graphical presentation of the theory to build intuition. In these graphs, lightly colored circles are potential firm locations, while filled-in circles represent sites populated by firms. 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 firms located in sites interconnected by bilateral interactions. Each firm does not necessarily interact with every other firm in its cluster,

[6] Arzaghi and Henderson (2008) utilize a similar foundation in their model of location choice for ad agencies in Manhattan.

but all firms in a cluster are interconnected. Our measure of agglomeration counts the number of these clusters that are expected to arise; we say that firms exhibit agglomeration if they typically occupy few distinct clusters. (The use of an expectation is necessary because firms choose randomly when indifferent among sites.)

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

$$\mathcal{B}^0_\rho(j) \equiv B_\rho(j) \cap \mathbf{Z}.$$

This formula has a simple interpretation: $\mathcal{B}^0_\rho(j)$ is the set of potential firm locations that can profitably interact with j under maximal radius ρ.

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, $\mathcal{B}^0_\rho$ for site B includes sites A and C. Sites A and C are the only locations within the maximal radius for Marshallian conditions ρ.

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. We define $\mathcal{B}^1_\rho(j)$ to be the set of sites that can profitably interact with the sites in $\mathcal{B}^0_\rho(j)$ through one additional step. In Figure 2a, $\mathcal{B}^1_\rho$(B) further includes the four additional sites within distance ρ from site C that are outside of the spillover range of site B. 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 ($\mathcal{B}^\iota_\rho(j)$ for $\iota = 1, 2, \ldots$). Formally, for any $j \in \mathbf{Z}$,

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

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

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

The ρ-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 $j' \in \mathbf{Z}$ which can profitably interact. The complete set of filled locations in Figure 2a constitute the ρ-cluster for site B in our example. The use of an infinite index in the union defining $\mathcal{B}_\rho(j)$ is ultimately unnecessary, as the finitude of the set of sites implies that $\mathcal{B}^\iota_\rho(j) = \mathcal{B}^{\iota+1}_\rho(j) = \cdots$ for some finite ι.

When the maximal radius ρ is small, clusters are generally small. For two precise examples, define the lower and upper bounds on distances between sites as $\underline{d} \equiv \min_{j_1 \neq j_2 \in \mathbf{Z}} d_{j_1,j_2}$ and $\bar{d} \equiv \max_{j_1 \neq j_2 \in \mathbf{Z}} d_{j_1,j_2}$. When $\rho < \underline{d}$, $\mathcal{B}_\rho(j) = \{j\}$ 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 $\rho > \overline{d}$, $\mathcal{B}_\rho(j) = \mathbf{Z}$ for all $j \in \mathbf{Z}$. A maximal radius longer than the maximal distance between sites results in a single cluster for an industry.

If the maximal radius is ρ and the first firm locates at site j, and the cluster around j contains available locations (i.e., $\mathcal{B}_\rho(j) \neq \{j\}$), then there is some site $j' \in \mathcal{B}_\rho(j)$ which delivers positive deterministic utility flows to the next entrant. It follows that if Marshallian forces are sufficiently strong, then firms select sites in the cluster $\mathcal{B}_\rho(j)$ until $\mathcal{B}_\rho(j)$ is filled. Iterating this analysis shows that when Marshallian forces are strong, firms fill clusters sequentially.

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 at random or based upon idiosyncratic preferences.

These observations suggest a natural notion of agglomeration. 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 fixed, fewer clusters of firms form when the maximal radius is ρ_1 than when it is ρ_2. Formally:

Definition 1. The *level of agglomeration for maximal radius* ρ is $|\mathbf{Z}| - \Xi_\rho$, where Ξ_ρ is the expected number of distinct ρ-clusters $\mathcal{B}_\rho(j)$ about sites $j \in \mathbf{Z}$ occupied by firms.

Note that under this definition, agglomeration increases as the expected number of clusters decreases. Holding industry size constant, increased agglomeration therefore also corresponds to increased cluster size. The additive term $|\mathbf{Z}|$ is a normalization that guarantees that the level of agglomeration is always a positive number. We could equally well define the level of agglomeration for maximal radius ρ to just be $-\Xi_\rho$.

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 due to the decay function G. 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, each industry always forms a single cluster, and entrants generically select sites in a fixed order. This is equivalent to the case in which $\rho \to \infty$ in our model. Once a first entrant picks a location, the set of sites filled by the remaining $N-1$ firms is exactly determined.

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.

2.3.2 Agglomeration due to Marshallian Advantages

Figure 2b illustrates the consequences of a longer maximal radius for cluster formation. 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 ρ is the maximal radius for intra-industry spillovers, $\rho \equiv \max\{d : G(d) \geq c\}$. Since $\mathbf{Z}$ is finite, small changes in ρ do not affect location choices. Larger increases in ρ, 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: firms become (weakly) more agglomerated when ρ increases.

Proposition 2. *A longer maximal radius of intra-industry spillovers ρ leads to a (weak) increase in agglomeration.*

Proof: See Appendix.

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 $\mathcal{B}_\rho(j(i))$ are filled, they are more attractive to firms than are unfilled sites outside of $\mathcal{B}_\rho(j(i))$. If ρ grows to $\hat{\rho}$, a radius large enough to cause some cluster $\mathcal{B}_\rho(j)$ to merge with another cluster (i.e. such that $\mathcal{B}_{\hat{\rho}}(j) = \mathcal{B}_\rho(j) \cup \mathcal{B}_\rho(j')$ for some site $j' \notin \mathcal{B}_\rho(j)$), then the expected number of clusters occupied by firms shrinks. Indeed, whenever a firm locates in either $\mathcal{B}_\rho(j)$ or $\mathcal{B}_\rho(j')$, subsequent firms fill all of $\mathcal{B}_{\hat{\rho}}(j)$ before locating in or starting another cluster.

Three empirical implications of this analysis are evident in Figure 2b. 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 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 2.3.4.

2.3.3 Ordering and Characterizing Agglomerative Forces

The theory 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 radii associated with different spillovers. In empirical analyses below, for example, we provide suggestive evidence for the model by plotting an estimate of the maximal radii for different technologies against a measure of technology cluster density. This section introduces terminology and conditions required to jointly test these predictions using the continuous density estimation techniques employed in Section 5.

It is impossible to measure directly the G functions that determine the value of firm clustering. However, observed spatial location patterns allow us to partially model the behavior of the unobserved functions G in a continuous manner.

Proposition 3. *Holding ρ fixed, and assuming that the G is differentiable, an increase in $|G'|$ leads to a (weak) increase in the number of firms clustered at small distances.*

Proof: Immediate from text.

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 at different distances. For an extreme example, if localization of firms is constant across space, then we must have $|G'| = 0$. If localization gradients are very sharp at short distances, then Proposition 3 implies that the underlying G function sharply attenuates. Note that intercept value $G(0)$ is *not* held fixed in Proposition 3. An implication of our framework is that, holding ρ fixed, $G(0)$ impacts the gradient $|G'|$, but does not affect the overall level of agglomeration.

Proposition 3 allows us to use the Duranton and Overman (2005) density estimations in Section 5 to characterize distributions continuously. Adding this more continuous structure to our model, we can compare the full distributions of industries to assess how longer maximal radii affect the shapes of clusters. The predictions that clusters become larger and less dense become jointly visible. Moreover, we can observe this effect's 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 localization nor dispersion occurs when

the maximal radius is ρ, i.e. $g_j = 0$—firms locate randomly—when the maximal radius is ρ. 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, 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 is a substantial departure from counterfactuals having (the same number of) firms occupying sites randomly sampled from $\mathbf{S}$.[7]

2.3.4 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 stronger implication for how spillover length influences the distribution of bilateral distances within clusters.

Proposition 4. *There is some $\bar{\rho} > \underline{d}$ so that whenever ρ and ρ' are such that $\underline{d} < \rho < \rho' < \bar{\rho}$, then the mean intra-cluster firm distance is (weakly) smaller when the maximal radius is ρ than when it is ρ'.*

Proof: See Appendix.

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 very near to each other). This higher frequency should then (weakly) decline as one considers bilateral distances further from the shortest possible connections.[8]

To provide intuition, first consider the impact of the marginal entrant on the bilateral distances in Figure 2c. As site X becomes part of the cluster, the set of bilateral distances grows to incorporate 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 A and D. 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 most at the shortest possible connections and proceeds (weakly) outwards for some distance.

[7] As discussed in the empirical appendix of our NBER working paper and in the paper of Barlet et al. (2012), this approach is slightly strained for the largest industries but is a reasonable baseline for most industries.

[8] The conditions of Proposition 4 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 large as to no longer influence cluster formation.

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 4 further indicates that 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 city 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 that of Duranton and Overman (2005) in assessing 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 and (weakly) declines 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 present 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 clusters and the newly activated sites surrounding them.

2.4 Discussion

We now discuss potential enrichments of the model. We first note that this model is a simplified version of the one contained in our NBER working paper. The present version assumes that all firms belong to the same industry. We also abstract away from the possibility of clustering due to fixed, location-specific natural advantages (e.g., coal mines, universities). The extended theoretical framework relaxes both of these simplifications, and shows that they do not materially affect the predictions for Marshallian clusters that we develop and test here. Our NBER working paper also outlines some basic spatial dynamics for clusters.

For simplicity, our base model only allows at most one firm per site. Our results are unchanged if we allow multiple firms to locate at each site, and assume that collocated firms "congest" each other. Specifically, we may extend our model by assuming that each site $j \in \mathbf{Z}$ has a *maximum capacity* $\kappa_j \geq 1$, and that the set $I(j)$ of firms located at j must always have $|I(j)| \leq \kappa_j$. Congestion is modeled by assuming that firms $i \in I(j) \neq \emptyset$ receive spillover benefits of $\hat{g}_j(i) \equiv (1/|I(j)|) \cdot g_j(i)$. That is, spillovers to location j are divided equally among firms collocated at j. With these notations, our base model corresponds to the case that $\kappa_j = 1$ for all sites $j \in \mathbf{Z}$. Even with congestion, the sequential location choice model is justified: A firm i entering site $j(i)$ has the potential to "crowd" its closest neighbors. But that firm i can never crowd out another firm $i' \in I(j(i))$. Indeed, if firm $i' \in I(j(i))$ were to exit $j(i') = j(i)$ and relocate after the entrance of firm i, then upon relocation, i' would face the same location choice problem previously faced by firm i. The ex ante optimality of $j(i)$ for i would then show that $j(i)$ is the ex post optimal choice for i'.

Second, the micro-interactions across sites that are built into the model are readily generalized. 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 structure of benefit transmission through the cluster so long as ρ is constant, as the spillover radius at the cluster's edge is what determines the marginal entrant's decision. We might also assume that with some probability $p(d) < 1$ firms invest in contact with firms of distance d away, with p declining in d (i.e. $p'(d) < 0$). With our model's structure, we can handle this case by simply replacing the function $G(d)$ with $G(d) \cdot p(d)$. Alternatively, if there is always some (possibly small) fixed probability that a firm chooses its location randomly, as in the model of Ellison and Glaeser (1997), then our qualitative conclusions are maintained: an upwards shift in p leads to a fewer, larger, and less dense clusters.

Finally, the model does not include property prices. One way to introduce property prices to the model is to consider them as the consequence of wanting to be near a fixed feature (e.g., the city center). The version of the model in our NBER working paper shows that this extension does not materially affect our predictions for Marshallian agglomeration so long as feature attraction effects are not too strong.

3 Patent Technology Clusters

3.1 Overview of Empirical Strategy

We illustrate the model's predictions empirically in this section using variation across patent clusters. We proceed in three steps that closely follow the model's structure. We first use patent citation data to illustrate how knowledge flows within US technology clusters resemble the model's maximal radius construct. Patent citations provide a rare window into the distances

over which knowledge interactions and technology flows are occurring within clusters. In a generalization of the Silicon Valley case study in the paper's introduction, we demonstrate how these knowledge flows are limited in distance even within a single cluster. We also show how bilateral interactions form overlapping regions of interaction that cover a larger spatial area than the individual interactions of firms do.

After establishing these properties generally, we use the patent data to calculate differences in the lengths of maximal radii across technology groups. Some technology areas like semiconductors have very localized citation patterns where knowledge flows decay rapidly with distance. Knowledge flows in other technology areas operate effectively over longer distances. After measuring these differences across technologies, we turn to our basic model predictions that a longer radius of interaction generates larger and less-dense clusters (Propositions 1 and 2), showing that each of these basic predictions holds when considered independently. We do not investigate the number of clusters prediction as it is substantially more sensitive to empirical choices than the properties of clusters are.

Our final exercises present a unified empirical framework for analyzing how technology cluster shapes and sizes differ across technologies in relation to their maximal radii of interaction. This framework brings to bear the joint nature of our three main predictions and the more subtle predictions of Propositions 3 and 4 with respect to rates of relative decay. These tests require that we depict the whole distribution of distances within a cluster and analyze the differences in these shapes across technologies. We conduct these tests using a mixture of non-structured plots and the continuous spatial density metrics developed by Duranton and Overman (2005). These depictions provide greater insights into how observable cluster shapes provide information about the underlying agglomeration force.[9]

3.2 Patent Citations and Knowledge Flows

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

[9]This section's investigation most closely relates to knowledge flows as a rationale for agglomeration and cluster formation. Section 4 of our NBER working paper provides additional empirical evidence for the model's structure when comparing the distances over which knowledge flows occur to distances over which agglomeration is driven by labor pooling or natural advantages. These supplementary exercises have the advantage of covering many industries and sectors in the US economy, but the broader approach means that we no longer identify the micro-interactions among firms as we do in patent data. What we show is that the ordering of industries by these various agglomeration rationales produces patterns in line with our model.

number of studies examine the importance of local proximity for scientific exchanges, generally finding that spatial proximity is an important determinant of knowledge flows.[10] 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) 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).[11]

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 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.

3.3 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 or 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 extract 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.

[10] See Jaffe et al. (1993, 2000), Thompson and Fox-Kean (2005), Thompson (2006), and Lychagin et al. (2010). Murata et al. (2012) measure the continuous density of patent citations.

[11] 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. 1998, 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.

Our analyses below consider how distances between zip codes influence patent citation rates. Several 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, listed addresses can represent either home or work addresses. 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.

To ensure that our results are not overly dependent upon this approach—especially with respect to the maximum radii that we calculate by technology—we also calculate a parallel set of distances using a match of USPTO patents to firms in the Census Bureau (Balasubramanian and Sivadasan 2011, Akcigit and Kerr 2010). The Census Bureau data records identify the zip codes of each firm's establishments in a city. We thus take the patents identified to be in Chicago for a particular firm, for example, and assign them the zip codes of the firm's records. Unreported analyses confirm the spillover radii that we identify with our primary dataset.[12]

3.4 Knowledge Flows within Clusters

Our first analysis characterizes how knowledge flows within technology clusters. To do so, we examine patent citation patterns, specifically 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 and clarify terminology, consider a sequence of patents where patent A cites patent B, patent B cites patent C, and patent C cites patent D. Using an arrow to indicate a citation, our sequence is A→B→C→D. Note that in this example the citations are moving from patent A to patent D, while knowledge moves in the opposite direction. That is, patent A is building on patent B, and that is why patent A cites patent B.

We term a first-generation citation as a direct citation of prior work. In our example, these

[12]The primary advantage of the work using the Census Bureau's data is to verify robustness with a second data source. There are two disadvantages. First, we must disclose any results that we wish to report using the Census Bureau's data. Basing our primary estimations on inventor address data allows us much more flexibility for generating graphs of continuous density estimates. Second, the Silicon Valley case study in the introduction (where we manually identified zip codes for work locations) was attractive in that single firm 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 using establishment-based identifiers versus simply using known inventor addresses.

would be A→B, B→C, and C→D. When we discuss the distances over which first-generation citations occur, we are measuring the bilateral distances between these three pairs. We next define a second-generation citation as the culmination of two steps in the citation chain: A→→C and B→→D. When we discuss the distances over which second-generation citations occur, we are measuring the bilateral distances between these end points, removing the intermediate step (i.e., A and C, B and D). Our simple example also has a single third-generation citation, A→→→D, and we would measure this distance as the bilateral distance between patents A and D.

Figure 2d continues with Figure 2a's example to describe our empirical strategy. We place into this graph the A→B→C→D citation sequence just described. We contrived this example to show a pattern where patent A would never have cited patent D directly according to our model. The distance from A to D is too great for the indicated maximal radius, but the distance can be bridged with the intermediate hops through patents B and C.

In reality, some measure of citations occur at distances that stretch across the full cluster (just as academics cite others at distances that span the globe). In fact, even if knowledge travels as in our model from patent D to patent A via sites B and C, we might still observe a patent A citing directly patent D (just as academics cite papers directly that they learned about through other papers). So, the model's structure cannot be taken so strictly as to say that we should never observe citations at distances of the length between A and D. Nevertheless, we can learn a lot about relative distance of knowledge flows by estimating the relative frequencies of citations by distance. Our model suggests that we should observe a higher frequency of first-generation citations when evaluating the shorter distances within clusters, as direct contact can occur at close proximities. Across longer spans, we should observe both fewer first-generation citations and more later-generation citations—indicative of knowledge transmission through a sequence of overlapping interactions.[13]

We demonstrate this pattern through some simple estimations illustrated in Figures 3a and 3b. For Figure 3a, we prepare a dataset that contains bilateral pairs of all zip codes that patent during the post-1975 period. To focus on local exchanges, we restrict these zip code pairs to those that are within 150 miles of each other. For each zip code z_1, we then identify the number of citations that it makes to the other paired zip code z_2. To be conservative in our approach, we do not examine interactions within the same zip code, and we exclude citations that firms make internally among their inventions across zip codes.

With this dataset, we empirically model the count of citations that patents in zip code z_1 make of patents in a second zip code z_2 using the general form:

$$\text{Citations}_{z_1 \to z_2} = \exp^{\beta \cdot d_{z_1,z_2}} (\text{Patents}_{z_1} \cdot \text{Patents}_{z_2})^{\gamma},$$

where as before d_{z_1,z_2} denotes the distance from z_1 to z_2. This expression suggests that citations

[13] The one exception to this would be if knowledge flows are fully transmissible through the cluster such that any site connected to the cluster receives complete effortless access to the knowledge housed at any site in the cluster. The evidence below suggests that this potential exception is not empirically relevant in this setting.

depend upon the interacted stock of patents in the two zip codes and upon the distance between the two zip codes, d_{z_1,z_2}. We would anticipate $\beta < 0$ if knowledge flows are declining with distance, and $\gamma > 0$ if a greater number of patents in the two zip codes provides more opportunities for citations. Rearranging this expression gives

$$\ln(\text{Citations}_{z_1 \to z_2}) = \beta \cdot d_{z_1,z_2} + \gamma \cdot \ln(\text{Patents}_{z_1} \cdot \text{Patents}_{z_2}),$$

which is the starting point for our first estimating equation. We make three further modifications. First, beginning with the citations outcome variable, zero citations may be observed even where patents exist (and this lack of exchange is important information). Our base estimation thus takes $\ln(1 + \text{Citations}_{z_1 \to z_2})$ to be the citations outcome variable, and we model other variations below to test the sensitivity of the $\ln(X) \mapsto \ln(1 + X)$ transformation.

Second, there are multiple ways that one might define distance to potentially allow for non-linear effects that our model emphasizes. Our first approach is to estimate distance's role in a non-parametric format using a series of indicator variables $I(\cdot)$ for distance bands between zip codes. We define a vector of distance bands as within one mile (but not the same zip code), (1,3] miles, (3,5] miles, (5,10] miles, (10,15] miles, ..., (95,100] miles, and (100,150] miles. We denote the set of distance rings as DR, and we include separate indicator variables for each distance band up to zip codes being (95,100] miles apart. Our β_{dr} coefficients will thus measure the difference in citation rates observed for a distance interval compared to the reference category of being more than 100 miles apart in the technology cluster.

Finally, the $\ln(\text{Patents}_{z_1} \cdot \text{Patents}_{z_2})$ control is important, but it is also weak. Our initial tests include all patents, and the patents in the two zip codes may be from very different fields. Thus, while raw citation counts display excessive localization, they may appear localized simply because different types of patenting firms are clustered together. To model this underlying landscape in the most flexible way possible, we generate random citation pairs comparable to our observed sample. For every patent that is actually cited, we randomly draw a counterfactual patent from the pool of all patents with the same technology class and application year as the true citation. This method has been used extensively in the literature, and we make two modifications that reflect our sample design. First, we exclude other patents of the citing firm from the pool of potential draws, just as we exclude within-firm citations in the primary sample. Second, we build the pool of potential patents using only patents within a 250-mile radius of the citing patent. We do not exclude the original cited patent from the random 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. Relative to simple patent counts, this counterfactual distribution has the advantage of very closely matching the underlying properties of local inventions and their technological foundations; it is a much stronger control, for example, than using simple patent counts.

With these three adjustments, our core estimating equation for Figure 3a becomes,

$$\ln(1+\text{Citations}_{z_1\to z_2}) = \alpha + \sum_{dr\in DR} \beta_{dr}\cdot I(d_{z_1,z_2}=dr) + \gamma\cdot\ln(\text{Patents}_{z_1}\cdot\text{Patents}_{z_2}) + \eta\cdot\ln(1+\text{Expected Citations}_{z_1\to z_2}) + \varepsilon_{z_1\to z_2}. \quad (1)$$

The solid line in Figure 3a plots the β_{dr} coefficients for first-generation citations like the A→B example discussed in Figure 2d. First-generation citations are quite concentrated at short distances and decline almost monotonically with increasing distance. The citation premium loses half of its strength by (15,20] miles, and zip codes that are 40 miles or more apart are very similar to those in the reference category of being 100-150 miles apart. This substantial decay echoes very closely the localized networking results of Arzaghi and Henderson (2008) and the spillover estimations of Rosenthal and Strange (2003, 2008). It is important to recall that we have excluded interactions within the same zip code, in order to be conservative. The within-zip code citation premium is larger in magnitude than that observed for neighboring zip codes within one mile of each other.

The dashed line in Figure 3a graphs the spatial patterns of second and third generations of patent citations, equivalent to the A→→C and A→→→D example. To construct the second-generation citation profile for zip code z_1, we start with the patents that were cited directly by firms in zip code z_1 within a 250-mile radius around zip code z_1. We then collect the citations that those patents made to other patents within 250 miles of their zip code. We then calculate the distances from the focal zip code z_1 to these citations, and we will focus again on the second-generation citations that fall within 150 miles of zip code z_1. We take this approach to provide very flexible local distances. Note, for example, that the distance from zip code z_1 to a given second-generation citation can be closer than the first-generation citation that links it. We repeat the same process for third-generation citations. The specification (1) is again used to compare rates in local distances to the rates that exist over 100-150 miles.

The results are intuitive and agree with the developed model. At very small distances, later-generation citations are substantially less frequent than first-generation citations. This gap quickly closes, and from distances of 10-25 miles, the relative frequencies are very similar. After 25 miles or so, there follows a distance interval where later-generation citations have a greater relative frequency than first-generation citations. These relative differences slowly decay thereafter, and at longer distances, the spatial overlaps become very similar across generations.

Figure 3b plots comparable evidence from a second approach. Rather than utilize bilateral zip code pairs, we sum the activity of zip codes that falls into the distance rings utilized above. This approach renders our analysis less sensitive to vagaries of zip code mappings and the issues that one encounters with zero citations; the corresponding disadvantage is that we sacrifice some of the granularity that the bilateral estimations allow. The consolidated empirical framework also allows us to include in the estimations a vector of fixed effects ϕ_{z_1} for citing zip codes. These

fixed effects remove persistent differences that exist across zip codes in citation counts such that we are exploiting only variation in how much zip code z_1 cites other zip codes in its technology cluster more or less than typical for zip code z_1. The second estimating equation takes the form

$$\ln(1+\text{Citations}_{z_1\to dr}^{Ring}) = \sum_{dr\in DR}\beta_{dr}\cdot I(d_{z_1,z_2}=dr)+\gamma\cdot\ln(\text{Patents}_{dr}^{Ring}) + \eta\cdot\ln(1+\text{Expected Citations}_{z_1\to dr}^{Ring})+\phi_{z_1}+\varepsilon_{z_1\to dr}. \quad (2)$$

These regressions measure the β_{dr} coefficients relative to the activity observed in excluded distance ring of 100-150 miles apart. The solid and dashed lines in Figure 3b again plot the first- and later-generation citations, respectively. At very short distances, first-generation citations show greater relative frequency compared to later-generation citations. The differences reverse at moderate distance ranges.

Appendix Tables 1, 2a, and 2b provide complete details on these estimations and descriptive statistics. Appendix Tables 2a and 2b report very similar results to Figures 3a and 3b, respectively, when zero-citation cells are excluded, when we drop the expected citations controls, and when we include own-zip code citations.

These differences across citation generations suggest that knowledge flows 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 2d, the chain of interconnected hops A→B→C→D aids site A's access to knowledge from sites around sites C and D. Moreover, the extra strength for first-generation citations over very short distances offers an approach to identifying maximal radii of interactions—we investigate this next. While it is important to note that other models may be able to generate these patterns, this framework does provide suggestive evidence on how knowledge movements through clusters conform to our model's structure.

4 Maximal Radii and Spatial Cluster Patterns

Our theory connects the maximal radius of firm interactions with cluster structure. We illustrate these predictions by looking at differences across 36 technologies using the sub-category level of the USPTO system. Hall et al. (2001) describe these technology groups, and examples include Semiconductors, Optics, and Resins. Similar to the analysis conducted in Figures 3a and 3b using all patents, we exploit patent citations separately within these individual technology fields so as to measure their radii of interaction. We then examine whether patterns across technologies' cluster shapes and sizes and our measured radii conform to our model's predictions. This section analyzes predictions individually, and the next section models the predictions jointly using continuous density measurement techniques.

4.1 US-Based Maximal Radii

We proxy the maximal radius of interaction for each technology through the citation localization patterns evident among patents within that technology. 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 0-70 miles. In the context of Figures 3a and 3b, this second technology 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.

Our sample preparation for these estimations is similar to that used for the above graphs. The sample is again restricted to zip codes that are observed to patent in a technology. We again consider citations that are outside of the same zip code to be conservative, excluding self-citations for firms. We also exclude cases where we believe that an inventor has moved and is self-citing his or her prior work. There are several ways that one can attempt to measure these spillover radii from the data, and we consider three different formats below. These approaches are all simpler than the flexible estimations undertaken in (1) and (2), but similar in spirit. These simpler formats are necessary given the substantial reduction in data points when estimating citation patterns on a technology-by-technology basis (especially when extended to the United Kingdom as noted below). Table 1 lists by technology the radii measured.

Our first technique considers each technology j in isolation, measuring its citation decay with distance in a log-linear form,

$$\ln(1+\text{Citations}_{j,z_1\rightarrow z_2}) = \beta_j \ln(d_{z_1,z_2}) + \gamma_j \cdot \ln(\text{Patents}_{j,z_2}) + \phi_{j,z_1} + \varepsilon_{j,z_1\rightarrow z_2} \text{ for all } j. \quad (3)$$

Thus, we estimate a single β_j parameter for how the rate of citations declines with distance. By estimating only one parameter for distance's role, we greatly increase our empirical power for these technology-level estimations. As we are only looking at patents and patent citations within a single technology, we no longer calculate the random citation counterfactual as the patents themselves capture the underlying technology landscape. These estimations are weighted by an interaction of patent counts in the two zip codes. With this technique, Semiconductor and Electrical Devices show the greatest citation localization (most negative β), while Heating and Apparel & Textiles show the weakest role for distance (β in the neighborhood of 0).

Our second technique makes several changes to (3) to ensure robustness of technique. We estimate

$$\begin{aligned}\ln(1+\text{Citations}_{j,z_1\rightarrow z_2}) &= \sum_j \beta_{j,0-10} \cdot I(d_{z_1,z_2} \leq 10) + \sum_j \beta_{j,10-30} \cdot I(10 < d_{z_1,z_2} \leq 30) \quad (4)\\ &\quad +\gamma \cdot \ln(\text{Patents}_{j,z_1} \cdot \text{Patents}_{j,z_2}) + \phi_j + \varepsilon_{j,z_1\rightarrow z_2}.\end{aligned}$$

The core differences between this approach and (3) are: 1) we estimate all of the citation declines

jointly so that γ is restricted to be the same across technologies, 2) we return to our indicator variable approach for estimating distances role in a more flexible manner, and 3) we include a vector of fixed effects for technologies instead of zip codes. We do not have the data to estimate distance rings as finely grained as those considered in the preceding exercises, so 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 second measure of technology spillover horizons is the observed premium $\beta_{j,0-10}$ over the first ten miles compared to the reference group. With this technique, Information Storage and Semiconductors show the greatest citation localization (most positive β), while Furniture and Receptacles show the weakest role for distance (β in the neighborhood of zero). This measure has a 0.7 correlation with that calculated through the (3) measure.

Finally, our third approach is completely non-parametric and relies on the relative prevalence of first- versus later-generation citations by distance for technologies (using up to six generations). All technologies start with first-generation citations having the highest relative prevalence, and all technologies eventually at some distance have later-generation citations more prevalent. For each technology, we identify the distance at which this crossing point occurs in two-mile increments. The series can be jumpy, especially for smaller technologies, so we make the specific requirement that one of two conditions be met: 1) the relative frequency of later-generation citations exceeds first-generation citations by 2% or more, or 2) that the relative frequency of later-generation citations exceeds first-generation citations for three consecutive distances. Many technologies show crossing points at 10 miles or less, while Receptacles and Pipes & Joints show the longest crossings at more than 20 miles. Overall, this measure is less correlated with the first two metrics at 0.2-0.3.

4.2 UK-Based Maximal Radii

We find evidence of a strong correlation between lengths of micro-interactions among firms (within technologies) and their associated cluster shapes and sizes. It is natural to worry in this setting 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 citation premia similar to our first two metrics using patent data from the United Kingdom. Ellison et al. (2010) introduce this technique and discuss its strengths and limitations. The central idea behind this identification strategy can be illustrated with the semiconductors technology. Many semiconductor firms are located in Silicon Valley, and as the map in Figure 1 illustrates, Silicon Valley is circled by water, mountains, and protected land. It could be that the cluster density and short citation ranges that we observed are due to this industry having developed in a location with natural features that pushed it towards

density and tight connections. Perhaps if the semiconductors industry had instead grown up in Houston, the industry would not display citation localization. If so, the data would describe features like our model's predictions but the connection would be spurious.

We can provide a safeguard against these concerns by measuring citation premia in the United Kingdom, which are not influenced by the local terrain of the United States or similar factors. This test does not solve every potential endogeneity concern, but it certain provides traction against some of the most worrisome endogeneity. To implement this strategy, we geocode all city names and postal codes associated with UK inventors. To provide more accurate city assignments, we also manually search for addresses of firms in the United Kingdom with more than fifty USPTO patents. 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 importantly, there are significantly fewer data points to estimate these citation premia (the UK sample is less than a tenth of the US sample size). Second, the geocoding has greater measurement error, perhaps most concentrated around London, and is coarser than in the United States. As a consequence, we do not attempt to exclude same-region citations as we do for the United States data. We also do not attempt to implement our third approach of measuring the crossing point of citation generations, as the data are too sparse with respect to later-generations citations. While these limitations 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 some reverse causality concerns.

Table 1 lists the UK metrics. The correlation between the US and UK metrics using our first specification (3) is 0.4. The correlation between the US and UK metrics using our second specification (4) is 0.2. The two UK metrics have a 0.5 correlation among themselves.

4.3 Analyses of Single Predictions

Figure 4a provides a cross-sectional plot of cluster density and our first proxy for maximal radius by technology that uses log-linear decay rates. 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 visible association between longer spillover horizons (weaker decay rates that approach zero) and less patent density. On the other hand, technologies that display very rapid decay rates and short technology spillover horizons are very tightly clustered. Recall that citation decay rates are calculated controlling for the underlying spatial patent distribution, so this relationship is not mechanical. The slope of the trend line is -1.336 (0.226). Very clearly, some of the industries in information technology show exceptional densities. The slope of the

trend line is -0.612 (0.082) when capping the density ratio at 50%.

Figure 4b provides a cross-sectional plot of US cluster density against UK citation decay rates. The vertical axis is the same as Figure 4a, but we substitute the UK citation decay rates for the horizontal axis's measure of technology spillover ranges. The UK has an outlier, raw decay rate of -0.507 (Earth Working & Wells); we cap this rate at the second-highest decay rate. There are some material adjustments among some information technology industries in Figure 4b compared to Figure 4a, with most noticeably semiconductors' decay rate not being as steep as we measured in the US. Nevertheless, a close connection exists between the decay rates for technologies in the UK and associated cluster density in the United States. The slope of the trend line is -0.979 (0.293); it is even sharper at -0.466 (0.115) when capping the US density ratio at 50%. The slope of the trend line is -0.745 (0.158) without the cap for Earth Working & Wells. These patterns provide confidence that these relationships are not being solely determined by unmodeled factors.

Table 2 continues these analyses of single predictions regarding the size and shape of clusters. Each entry in the table is from a separate regression where the outcome variable is indicated in the column header, and the five panel headers indicate the metric used to model the maximal radius of interaction. Panels A and D consider the log-linear citation decay rates estimated through technique (3) measured in the US and UK, respectively. Panels B and E similarly consider the US' and UK's citation premium observed over 10 miles from estimation (4). Finally, Panel C models spillover lengths through the crossing points observed for technologies between first- and later-generation citation frequencies.

To make our estimates easily comparable to each other, we transform variables to have unit standard deviation. We also multiply the raw $\beta_{j,0-10}$ coefficient for Panels B and E by -1 so that the predicted signs for Table 2's regressions are aligned in the same direction. These regressions exploit variation across the 36 technologies, and we control for the size of the technology using its patent count during the 1975-2009 period. Regressions are unweighted and report robust standard errors. We find very similar patterns when weighting technologies by size.

The first three columns examine the size of clusters, where we have the prediction that a longer maximal spillover radius produces a larger cluster. We take metropolitan statistical areas (MSAs) as the unit of observation, measuring the patenting that occurs within the zip codes of each MSA. In the next section we consider more flexible techniques that do not depend upon MSA definitions, as technology clusters may extend past MSA boundaries or across MSAs. This simple starting point is attractive, however, as it does not depend upon the structure of the continuous density techniques.

We identify the leading or dominant zip code per MSA in terms of patent counts by technology. Columns 1 and 2 describe the mean and median distance, respectively, from the dominant zip code to other patents in the MSA by technology. These distances are calculated as the weighted averages of the distances from the dominant zip code using zip code centroids. There

is a positive relationship in Columns 1 and 2, such that a one standard-deviation increase in the estimated maximal radius of a technology is associated with a 0.24-0.58 standard-deviation increase in these mean and median distances when using US-based radii in Panels A-C. The estimated elasticity is 0.11-0.43 when using UK-based radii in Panels D and E. Overall, with the exception of weaker performance in Panel E, these results highlight that a greater spillover range for a technology is associated with longer mean and median distances within MSAs for the technology's patents.

Column C evaluates an alternative metric where we calculate the normalized Herfindahl index of patents over the zip codes in a given MSA by technology. A second way that we might observe a greater size of technology clusters within a given MSA is if the patents for the technology are spread out over more zip codes (a weaker Herfindahl index). This prediction connects with our radii as measured in Panels A and B, with very strong elasticities of about -0.63 to -0.73, and with the UK log-linear decay function in Panel D, with a strong elasticity of -0.36. On the other hand, the support in Panels C and E is weak. The coefficient elasticity retains the predicted sign, but the results are not statistically significant.

Columns 4-6 shift the focus towards the prediction that clusters with longer spillover radii will be less dense. Column 4 continues with the density metric examined in Figures 4a and 4b, where we measure the fraction of bilateral distances between patents that are 150 miles or less apart that are in fact 50 miles or less apart (i.e., count of patents with bilateral distances of 50 miles or less / count of patents with bilateral distances of 150 miles or less). This prediction finds support with all of our metrics. After controlling for the size of technology, the estimated elasticity using US-based radii is 0.25-0.85; the UK-based elasticities are 0.30-0.48. These elasticities are precisely measured. Column 5 shows comparable results when capping density at 50%, and Column 6 shows similar patterns when we instead consider the density among patents that are 50 miles or less apart by looking at the fraction of these patents that are 25 miles or less apart.

5 Continuous Density Estimations

Overall, these regressions in Table 2 suggest that a longer spillover radius for a technology is associated with larger and less-dense clusters. To some degree, of course, the different outcome measures that we model in Table 2 are variations on a similar theme. Our six outcomes are also ad hoc in their design, in that we do not have any particular reason to examine, for example, the density over 50 miles compared to the density over 43 or 72 miles. This section provides a joint test of our model's predictions in a more rigorous manner using continuous density estimations. We first introduce the Duranton and Overman (2005) methodology that we utilize, and then we show how the shapes of local technology clusters relate to technology spillover horizons.

5.1 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 in our NBER working paper 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 an industry A through a continuous function:

$$\hat{K}_A(d) = \frac{1}{hN^A(N^A-1)} \sum_{i=1}^{N^A-1} \sum_{i'=i+1}^{N^A} f\left(\frac{d - d_{j(i),j(i')}}{h}\right). \tag{5}$$

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 1. 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 working paper's 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:

$$\begin{aligned} \gamma_A(d) &\equiv \max\left[\hat{K}_A(d) - K_A^{LCI-U}(d), 0\right] \\ \psi_A(d) &\equiv \max\left[K_A^{LCI-L}(d) - \hat{K}_A(d), 0\right] \text{ if } \gamma_A(d) = 0 \\ &\quad \text{and 0 otherwise.} \end{aligned} \tag{6}$$

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:

$$\gamma_A^C(d) \equiv \gamma_A(d) - \psi_A(d). \tag{7}$$

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.

5.2 Descriptive Statistics

For each technology, we estimate the continuous DO spatial density metric described above using patent data from 1990-1999.[14] Distances are calculated using zip code centroids. Figures 5a and 5b provide 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).

Figure 5a simply provides the average kernel density (5) by distance for the technologies that are contained within each grouping. The technologies within the Computers/Electronics grouping show high spatial concentration over the first 30 miles, but then exhibit very low density at moderate to long distances. The Chemicals/Medical grouping has lower average density levels at short ranges, but then exhibits the highest average spatial densities over medium distances. By contrast, the Mechanical/Miscellaneous grouping does not exhibit very strong patterns.

Figure 5b uses the localization metric (6), plotting the fraction of technologies within each group that are localized. 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. 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 is little material variation by distance in the number of technologies localized for the Mechanical/Miscellaneous grouping.

These patterns roughly conform with our predictions, as our measures of technology spillover radii in Table 1 tend to be smaller for Computers/Electronics than for Chemicals/Medical or Mechanical/Miscellaneous. Greater requirements for very close knowledge exchange are visibly associated with shorter, denser spatial clusters across these broad groups. This description,

[14] 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.[15]

5.3 Complete Density Plots

While transparent, Table 2's analyses are incomplete in that they do not describe the full distribution of firm localization behavior. They also do not account for differences in technology 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 (5) for a technology A. The process of assigning localization (6) 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 6a and 6b present coefficients from empirical specifications of the form

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

These estimations provide a continuous description of how technology cluster shapes vary with technology horizons. SpilloverRadius_A is the technology spillover radius for industry A calculated through through our five techniques and listed in Table 1. Greater values of SpilloverRadius_A correspond to longer maximal radii in our model, and we thus anticipate finding larger and less-dense clusters for these technologies. We transform $\hat{K}_{A,d}$ and SpilloverRadius_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 SpilloverRadius_A. Higher values of β_{dr} indicate that technologies with longer spillover radii show greater spatial density at that distance. The cross of 51 distances and 36 technologies yields 1836 observations per estimation.

Figures 6a and 6b present these density estimations using the US and UK measures of SpilloverRadius_A, respectively, estimated with the log-linear decay rates. Triangles report β_{dr} coefficients. The dashed lines provide 90% confidence bands with standard errors clustered by technology.

[15] 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.

Technologies with greater SpilloverRadius$_A$ (i.e. longer maximal radii) are substantially less agglomerated at very short distance horizons. A standard deviation increase in SpilloverRadius$_A$ is associated with a 1.5 standard deviation decrease in the density of activity conducted at five miles or closer using the US measure; the UK-based estimate is 0.9 standard deviations. By 60-75 miles, the abnormal spatial concentration is no longer statistically different from zero.

Looking further, technologies with longer SpilloverRadius$_A$ are over-represented after 75 miles or thereabouts. Using the US estimate of citation density, these clusters show an abnormal 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 statistically different from zero for a shorter distance range. 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 match our model and the predictions given in Section 2.3.4 regarding maximal radii and cluster shapes. Note that the patterns of under-representation followed by over-representation are not mechanical. Other attributes, for example, could predict higher spatial concentration for a technology at all spatial distances to 250 miles.[16]

Figures 7a and 7b take the next step of calculating localized deviations from technology-specific confidence intervals using (7). The patterns are very similar to Figures 6a and 6b. The lack of density at very short spatial horizons is robustly different from the random counterfactuals and very similar to the kernel plots. The abnormally high spatial concentration at moderate spatial horizons is weaker than in the raw kernel density plots, with the US and UK estimators both exhibiting a narrower range where they are statistically different from zero.

Overall, these figures jointly illustrate our central model predictions. A longer maximal radius, or weaker spillover density, is very strongly associated with reduced agglomeration at very short spatial horizons (i.e., the cluster is less dense). These same technologies tend to be over-represented at moderate spatial horizons (i.e., the clusters are larger). The latter result is very strong in the raw US data, and it is mostly 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 4 is robustly supported.

5.4 Robustness Checks and Extensions

Tables 3 and 4 provide robustness checks on these results. Panel A provides estimates using the kernel densities of technologies, and Panel B provides estimates using patent localization. To facilitate reporting, we estimate a single parameter per 25-mile distance interval, with spatial densities at 225-250 miles serving as the reference group. Column 1 in Table 3 repeats Figures 6a and 7a under this approach.

[16] The kernel density functions (5) 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.

Column 2 shows very similar results if weighting technologies by their size, with somewhat greater persistence evident for the abnormal densities observed at moderate distances. Column 3 shows slightly stronger patterns when excluding the five technology groups that are defined as residuals (e.g., Miscellaneous Drugs), where consistent clustering concepts may not apply. Finally, Column 4 reports bootstrapped standard errors, showing them to be smaller than the clustered standard errors that we otherwise report.

Columns 5-8 show the results with our four maximal radius metrics. While the patterns and levels can be different, we discern three key features from this work. First, all five approaches exhibit the basic joint patterns predicted by the model of a longer technology spillover radius being associated with larger and less-dense clusters. Second, the reduced-density prediction is robustly confirmed with results holding and precisely measured over the first 50 miles or thereabouts. Finally, the longer prediction finds more moderate support. It is evident in the patterns of all five measures, but it is not statistically different from zero across any distance range in Column 8. In addition, the exact distance intervals at which the increased density is evident varies somewhat by measure. Thus, we find good confirmation of the longer-cluster prediction, but it is generally just directional in nature.

Similar results are found using three additional specification variants. The first employs the density function (5) 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.

Table 4 next provides some sample splits that consider features not emphasized in our model and baseline empirics. We seek to establish the robustness of our results by looking at variations within each subsample to see if similar results hold. The first two columns split the sample by the degree to which patents in the technology cite other patents in the same zip code. We have excluded these own-zip code citations in our maximal radius calculations, and so this sample split utilizes independent data. Our model's structure does not emphasize the intensity of very local interactions (i.e., the $G(0)$ intercepts) but instead the maximum radii. This sample split tests this feature. The empirical patterns that we emphasize are present in both samples, confirming robustness, with the interesting finding that these patterns are more accentuated in industries with very intense local interactions.

Second, our model and baseline empirics only consider technology flows within the same industry, while the development of new patents often draws from several technology areas.[17] To test the robustness to cross-fertilization of technologies, we split technologies by the share of

[17] 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.

their patents that go to other technology areas. The relationships that we emphasize in this paper look quite similar in the two halves.

Third, our baseline estimations do not restrict patent citations to be within a specified time interval, but diffusion occurs with time that makes knowledge widely available in a local area and beyond. We anticipate our model's predictions to be more important in industries where access to very recent knowledge is critical. To test this feature, we calculate the share of citations nationally by technology area that occur to patents within the prior five years. The patterns are substantially stronger in the sample of technologies that rely on very recent knowledge, with only the less dense part of the prediction holding in the lower half of the distribution.

Fourth, our model does not include input prices that can generate further sorting across locations by firms. We test this feature by calculating from the 1990 Census of Populations a weighted average of expected scientists and engineering wages using the top 10 cities for each technology in terms of patent counts. Science and engineering wages are reflective of the wages to be paid to inventors. The patterns are present in both parts of the distribution, with some emphasis towards the technologies developed in areas with above median input costs.[18]

Finally, Tables 5 and 6 provide broader robustness checks on the technique of using continuous density estimation. Rather than undertaking the DO transformations, we simply group observed bilateral distances between patents in technologies that are within 250 miles of each other into a set of distance bins. We then calculate for each technology the fraction of the bilateral distances that fall within each bin. The top of Table 5 provides the mean and standard deviation of these shares. The first four columns provide break-outs for the first 20 miles at five-mile intervals, while Columns 5-12 consider 20-mile increments across the full range to 140 miles.

Table 5 conducts a set of point-by-point regressions on the shares of patents by technology that fall within each distance bin. Similar to the structure of Table 2, the five panels of Table 5 provide a simple set of regressions with each of our techniques for measuring spillover radii. Radius measures are normalized to have unit standard deviation, and we control for the number of patents in the technology. Unlike Table 2, however, we leave the outcome variables in their raw shares since these shares are easy to interpret.

The results in Table 5 show that our conclusions are not being driven by the construction of continuous density metrics.[19] With all five radius measures, we again see evidence that a longer spillover radius is associated with larger and less-dense clusters. For example, Column 1 finds that a one standard-deviation increase in the spillover radius lowers the share of patenting within [0,5) miles by 1.7%-4.1% compared to a base of 4.4%. Similarly, Column 5 shows that this same radius increase lowers the [0,20) share by 4%-12% compared to a base of 16.5%. On

[18] In addition to these four sample splits, we find very similar results to our baseline estimations when include four single control variables for these dimensions.

[19] There are several key differences of the point-by-point regressions compared to the DO estimations. These raw shares are not smoothed, and they are not being measured relative to confidence intervals. The shares are also constrained to sum to 100% over the 250-mile range, which is not imposed upon the continuous density estimates.

the other hand, the later columns show an increase in shares at longer ranges. There are several advantages of employing the DO technique, but these estimates show that our conclusions are robust to variations on this approach.

On a related note, Table 6 reports similar point-by-point regressions where we consider each patent assignee as a single observation. Unassigned patents, which represent about a quarter of all patents, are also retained. Our baseline estimations consider bilateral distances between patents, similar to the employment-weighted estimations of DO. Table 6 shows quite similar patterns when instead considering bilateral distances among unweighted assignees and individual inventors.

6 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 variation across patent technology clusters.

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, 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 describing patent technology clusters, but we believe that many more applications in industrial agglomeration are possible. For example, future work could look to price the marginal sites of clusters or identify spillover lengths by examining the location decisions of marginal entrants. Our framework highlights the important information that is contained in those agents' 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 maximal radii.

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* 93:1 (2011), 126-146.

Barlet, Muriel, Anthony Briant, and Lauire Crusson, "Location Patterns of Services in France: A Distance-Based Approach", *Regional Science and Urban Economics* (2012), forthcoming.

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 Nonparametric Test for Industrial Specialization", Working Paper (2011).

Bleakley, Hoyt, and Jeffrey Lin, "Portage: Path Dependence and Increasing Returns in U.S. History", *Quarterly Journal of Economics* 127 (2012), 587-644.

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, Jake Carr, Robert Hunt, and Tony Smith, "The Agglomeration of R&D Labs", Working Paper (2012).

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).

Dempwolf, C. Scott, "A Network Model of Regional Innovation Clusters and their Influence on Economic Growth", Working Paper (2012).

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?", *Journal of Labor Economics* (2012), forthcoming.

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.

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

Hanson, Gordon, "Market Potential, Increasing Returns and Geographic Concentration", *Journal of International Economics* 67:1 (2005), 1-24.

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, "Coagglomeration", Working Paper (2012).

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", *Review of Economics and Statistics* 94:1 (2012), 1-19.

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).

Marx, Matt, and Jasjit Singh, "More than Just Distance? Borders and Institutions in Knowledge Spillovers", Working Paper (2012).

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.

Murata, Yasusada, Ryo Nakajima, Ryosuke Okamoto, and Ryuichi Tamura, "Localized Knowledge Spillovers and Patent Citations: A Distance-Based Approach", *Review of Economics and Statistics* (2012), forthcoming.

Olson, Gary, and Judith Olson, "Mitigating the Effects of Distance on Collaborative Intellectual Work", *Economics of Innovation and New Technology* 12:1 (2003), 27-42.

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.

Rozenfeld, Hernán, Diego Rybski, Xavier Gabaix, and Hernán Makse, "The Area and Population of Cities: New Insights form a Different Perspective on Cities", *American Economic Review* 101 (2011), 2205-2225.

Saiz, Albert, "The Geographic Determinants of Housing Supply", *Quarterly Journal of Economics* 125:3 (2010), 1253-1296.

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.

Theoretical Appendix

Proof of Proposition 2

We assume a generic distribution of sites.

We order the firms by entry period $i = 1, \ldots, N$, 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 $g_{j(i)}$, and $\mathbf{Z}_{i+1} = \mathbf{Z}_i \cup \{j(i)\}$.

The preceding discussion implies that the set of locations occupied by firms following all firms' entries is completely determined by an instantiation of the random cluster selection which occurs whenever firms are forced to choose sites 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.

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

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

Denoting the probability of draw $\phi \in \Phi$ by $\text{Prob}(\phi)$,[20] we compute that $E_\phi[\#_\rho(\phi)]$, the expected level of agglomeration, is given by

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

We suppose that ρ increases to $\rho' > \rho$. 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 $\Phi \ni \phi = \phi^1 \cdots \phi^{B(\rho)}$, 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 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

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

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

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

which proves the desired result.

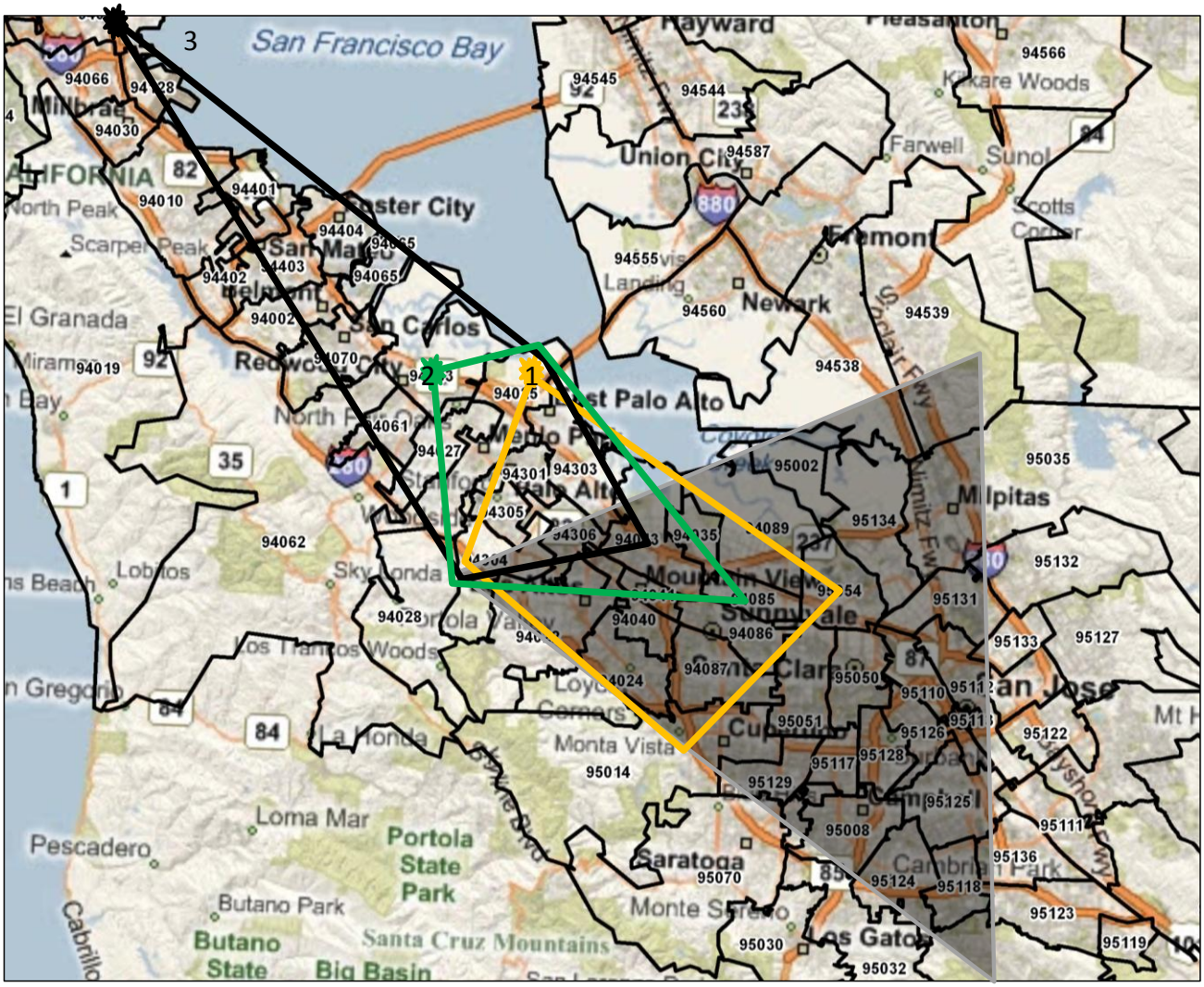
Proof of Proposition 4

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.[21] Thus, the existence of the desired $\bar{\rho}$ is immediate.

[21] 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 $\mathcal{B}_\rho(j)$.

Fig. 1: 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.

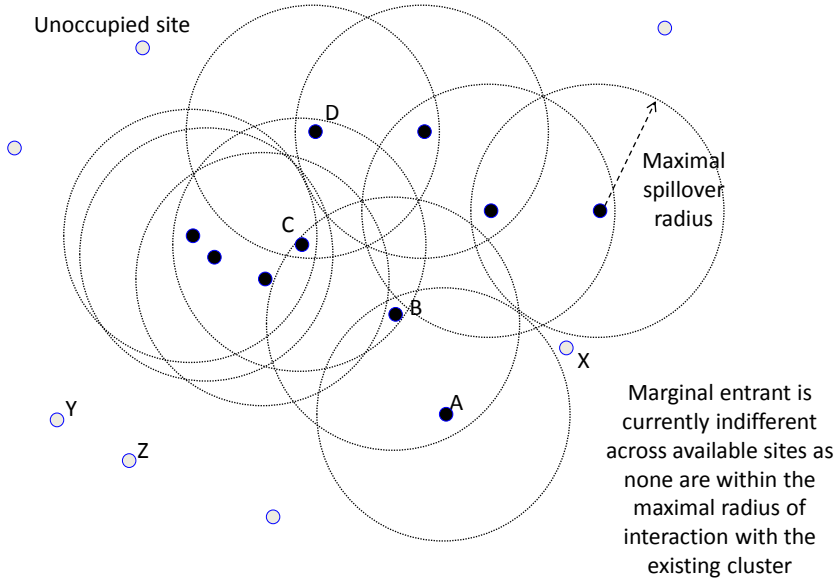
The stars indicate the focal zip codes, and the shape of each technology sourcing zone is determined by the three zip codes that firms in the focal zip code cite most in their work. The orange zone (1) for Menlo Park extends deepest into the core. The green zone (2) for Redwood City shifts up and encompasses 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. 2a: Marshallian Clusters

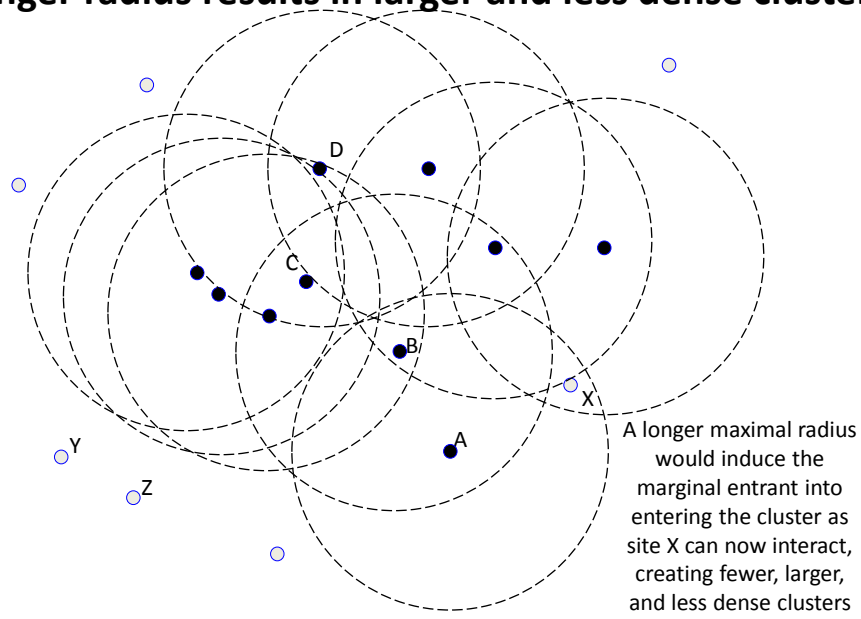
Agglomeration due to interactions among firms



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. Spillover radii are limited due to fixed costs of interaction. Large area clustering is due to small, contained interaction effects that overlap each other. The next entrant is indifferent among available sites, including sites X, Y, and Z.

Fig. 2b: Clusters with Longer Spillover Radius

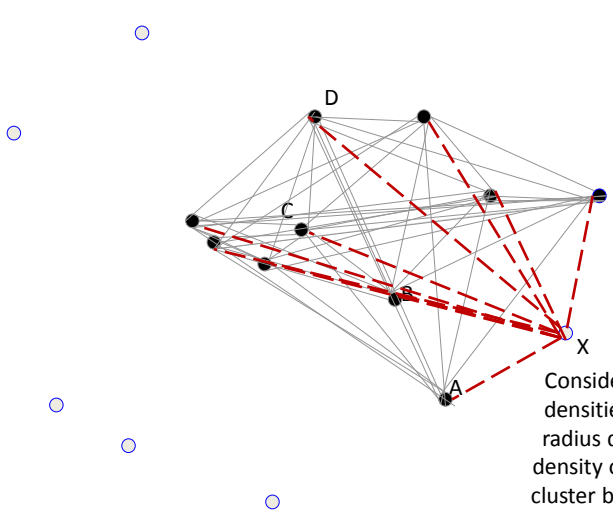
Longer radius results in larger and less dense clusters



Notes: Image illustrates a cluster for an agglomeration force with a longer maximal radius. The larger dashed circles show that a longer maximal radius would induce the marginal entrant into the cluster at site X over the other sites, resulting in (weakly) larger and less dense clusters. An additional prediction is that there should be fewer clusters for a technology given a fixed number of firms.

Fig. 2c: Bilateral Distances and Radius Length

A longer radius raises density the most among longer distances

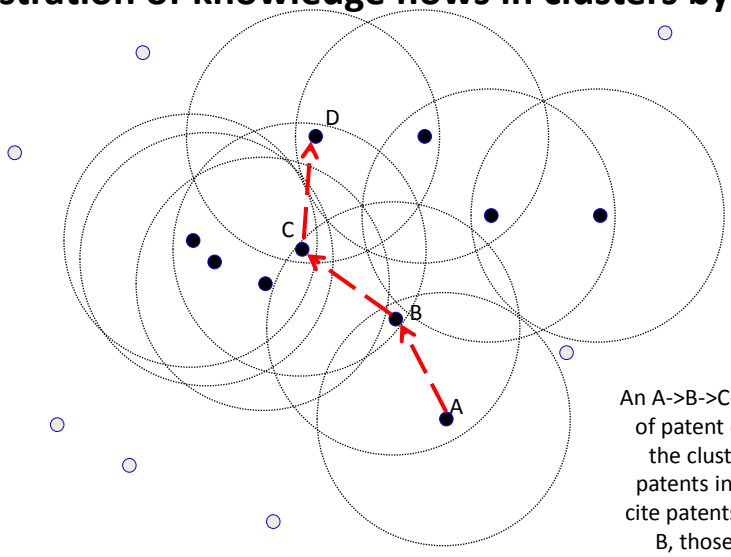


Considering bilateral distance densities, raising the maximal radius does not impact on the density of short distances in the cluster but increases the density at moderate to long distances

Notes: Image illustrates bilateral distances within clusters. The light grey lines show bilateral connections for the cluster formed under the shorter maximal radius in Figure 2a. The large, red, dashed lines show the additional bilateral connections formed when the cluster grows due to the longer maximal radius of interaction in Figure 2b. The length of the bilateral connections from the induced entry at site X will be longer than the shortest existing bilateral connections.

Fig. 2d: Patent Citation Example

Illustration of knowledge flows in clusters by distance

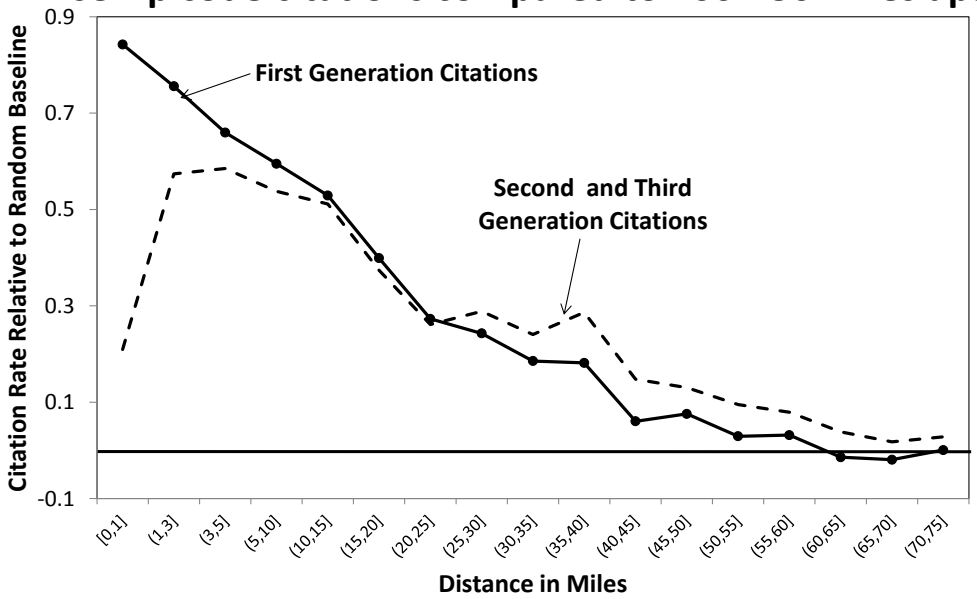


An A->B->C->D example of patent citations in the cluster, where patents in location A cite patents in location B, those in B cite location C, and those in C cite location D

Notes: Image provides an example of knowledge flows within a cluster discussed in the text. Knowledge is flowing from site D to site C to site B to site A, and we thus observe patent citations in the reverse order. This example is contrived so that site A will only interact with site D under the given maximal radius via interconnections provided by other sites.

Fig. 3a: Local Patent Technology Horizons

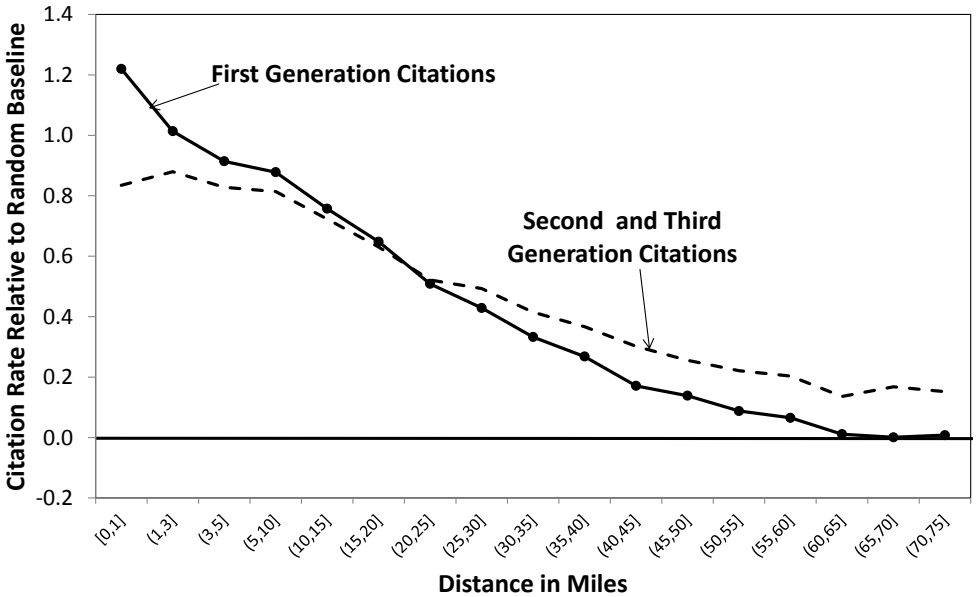
Pairwise zip code citations compared to 100-150 miles apart



Notes: Figure plots coefficients from regressions of log patent citation counts between pairwise citing and cited zip codes within 150 miles of each other. Explanatory variables are indicator variables for distance bands with effects measured relative to zip codes 100-150 miles apart (unreported bands for 75-100 miles resemble 70-75 miles). Regressions control for an interaction of log patenting in the pairwise zip codes and log expected citations based upon random counterfactuals that have the same technologies and years as true citations. Citations within the same zip code are excluded.

Fig. 3b: Local Patent Technology Horizons

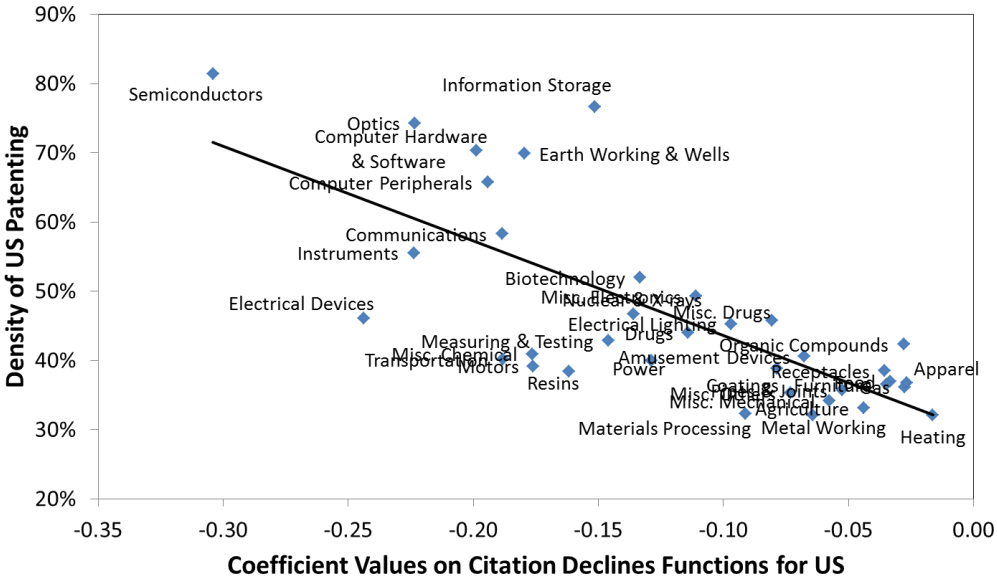
Citations by consolidated rings compared to 100-150 miles apart



Notes: See Figure 3a. This figure plots coefficients from regressions of log patent citation counts that employ consolidated distance rings around citing zip codes rather than pairwise combinations of zip codes. Explanatory variables are indicator variables for distance rings with effects measured relative to zip codes 100-150 miles (unreported bands for 75-100 miles resemble 70-75 miles). Regressions control log patenting in the distance ring, log expected citations based upon random counterfactuals that have the same technologies and years as true citations, and citing zip code fixed effects. Citations within the same zip code are excluded.

Fig. 4a: Patent Cluster Density & Spillover Radius

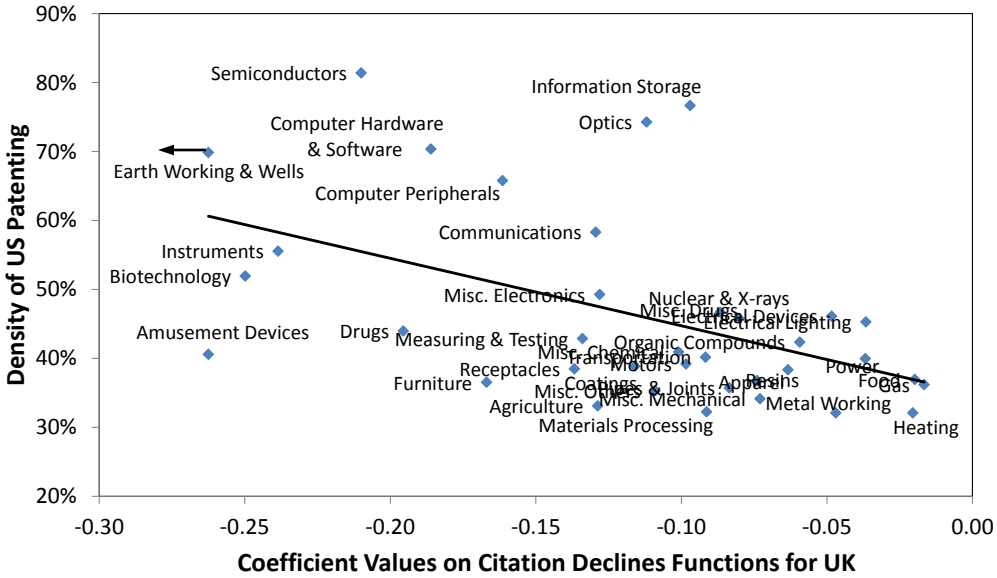
Cross-section of invention density and tech. spillover lengths



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. The horizontal axis measures by technology the log rate of citation decay by distance, controlling for underlying patenting and citing zip codes fixed effects. Longer spillover horizons (i.e., weak decay rates) are associated with less dense clusters. The slope of the trend line is -1.336 (0.226).

Fig. 4b: Patent Cluster Density & Spillover Radius

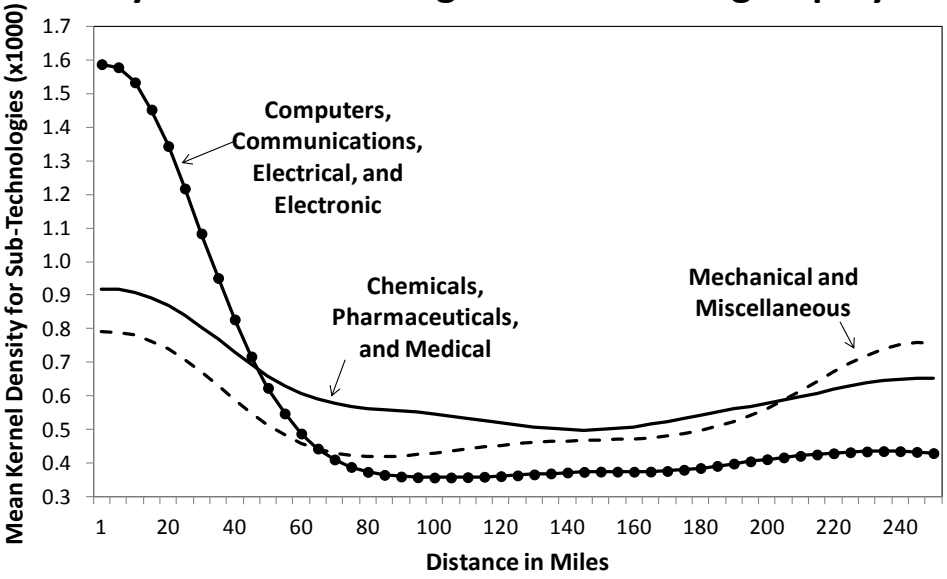
Using UK tech. spillover lengths to predict US density levels



Notes: See Figure 4a. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths. The outlier, raw decay rate of -0.507 for Earth Working & Wells is capped at the second-highest decay rate. The slope of the trend line is -0.979 (0.293). The slope of the trend line is -0.745 (0.158) without the cap.

Fig. 5a: Patent Kernel Densities

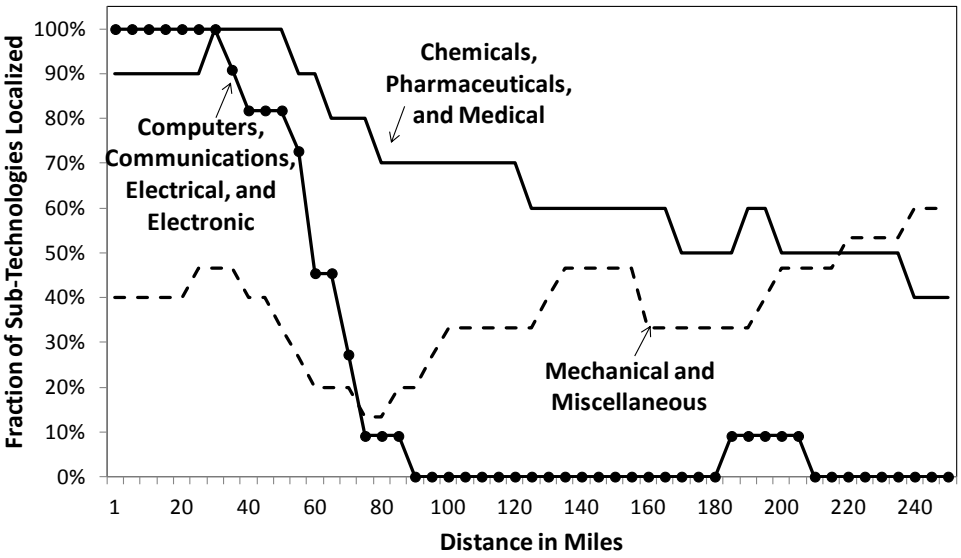
Mean density of sub-technologies in indicated group by distance



Notes: Figure plots the mean kernel density of technologies by each distance (x1000 for scale). The sample includes 36 sub-categories of the USPTO system organized into three simple divisions. Kernel density is calculated using pairwise distances among inventors in a technology.

Fig. 5b: Patent Localization Measures

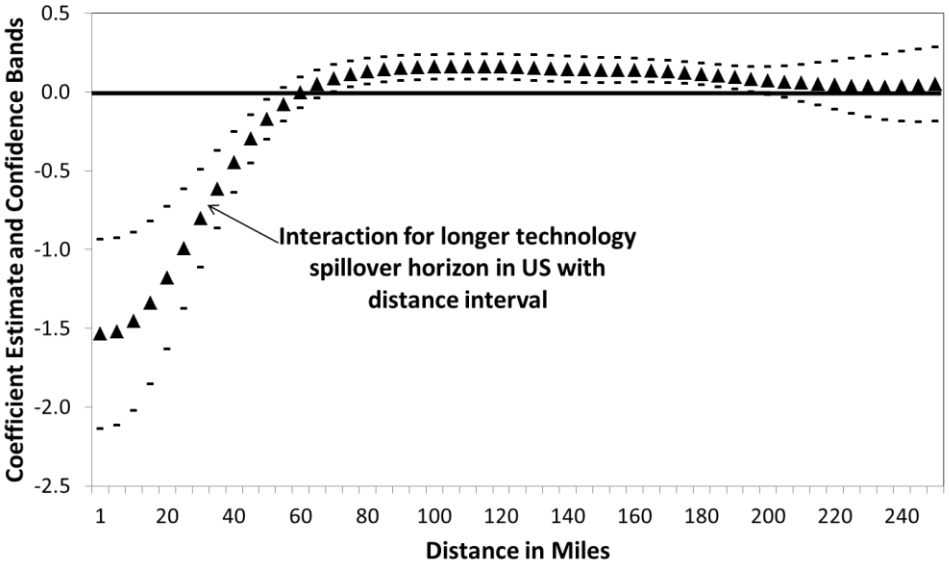
Share of sub-technologies in indicated group localized by distance



Notes: See Figure 5a. Localization is calculated through a comparison of the kernel density estimations for technologies with Monte Carlo confidence bands under the Duranton and Overman (2005) technique. 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. Localization looks very similar with 1%/99% confidence bands.

Fig. 6a: Patent Cluster Shape & Spillover Radius

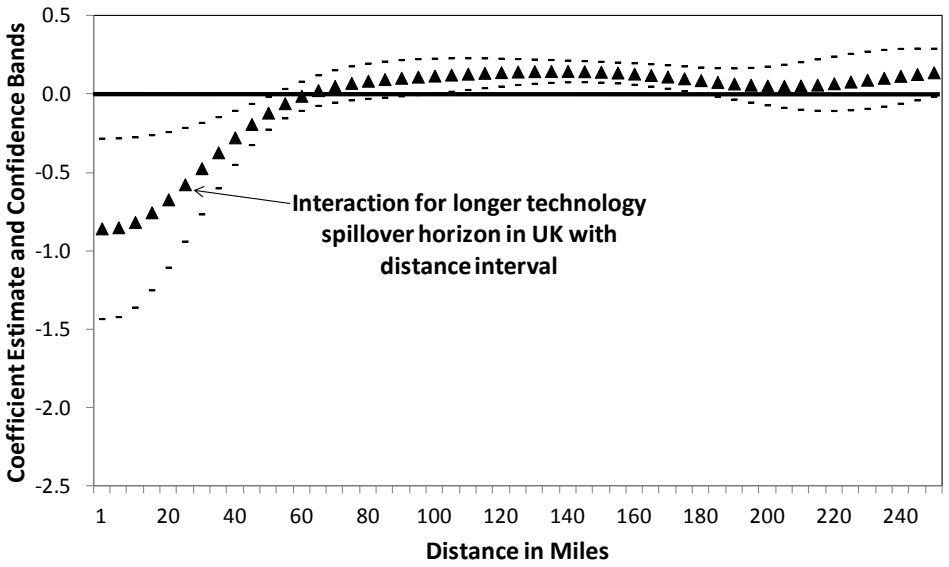
Kernel estimations of cluster shape and tech. spillover lengths



Notes: Figure plots coefficients from regressions of kernel densities by distance for 36 technologies. Technology decay rates are measured as in Figure 4a by the log rate of citation decay by distance, controlling for underlying patenting and citing zip codes fixed effects. Regressions include fixed effects for each distance. Dashed lines are 90% confidence bands. Technologies with longer spillover ranges (i.e., weaker decay functions) show lower density at short distances and increased activity over medium distances (i.e., larger and less dense clusters).

Fig. 6b: Patent Cluster Shape & Spillover Radius

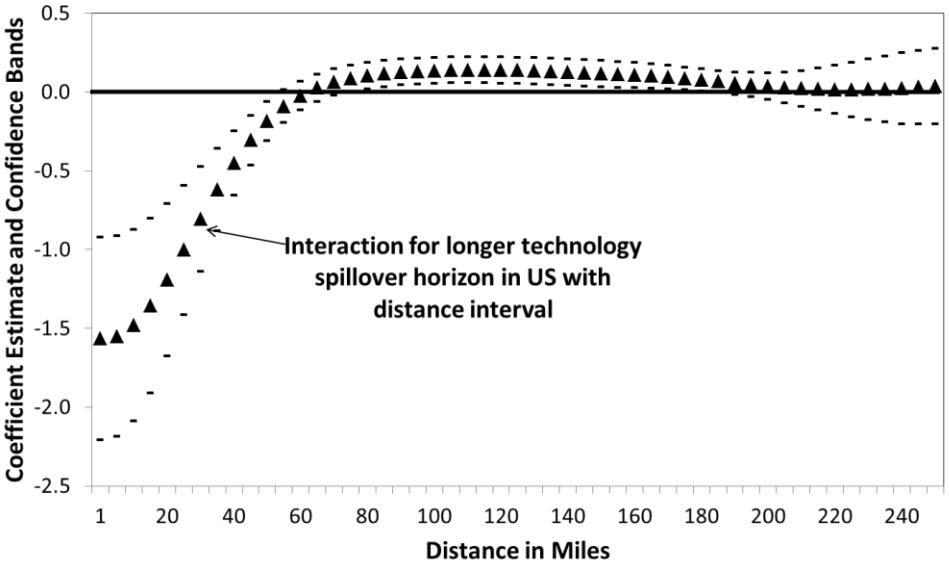
Using UK tech. spillover lengths to predict US cluster shapes



Notes: See Figure 6a. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths. Technology decay rates are measured as in Figure 4b by the log rate of citation decay by distance in the UK, controlling for underlying patenting and citing zip codes fixed effects.

Fig. 7a: Patent Localization & Spillover Radius

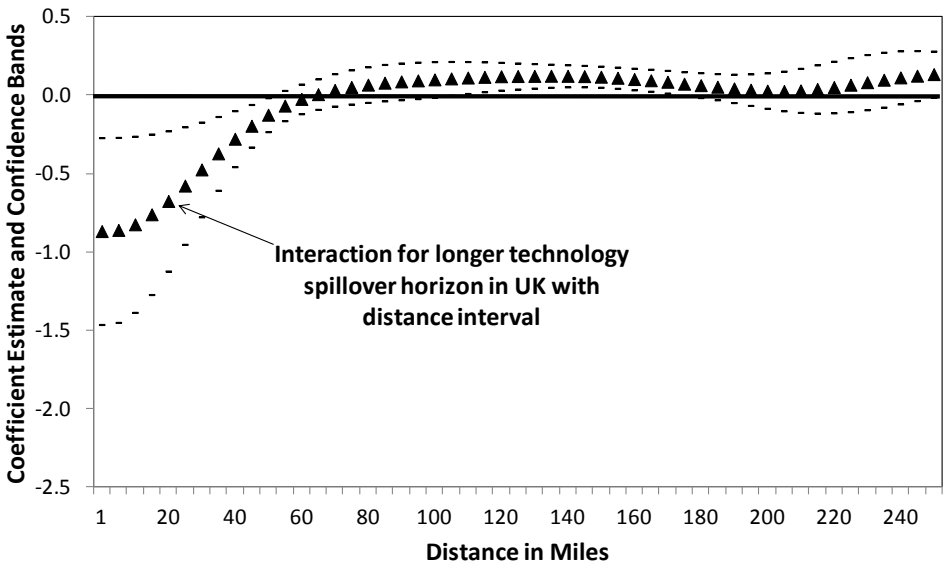
Localization estimations and technology spillover lengths



Notes: See Figure 6a. The dependent variable is updated from the kernel density in Figure 6a 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 longer spillover ranges again exhibit larger and less dense clusters with this technique.

Fig. 7b: Patent Localization & Spillover Radius

Using UK tech. spillover lengths to predict US localization



Notes: See Figure 7a. Estimates use technology spillover lengths in the UK to address potential reverse causality where US cluster shapes determine spillover lengths.

Table 1: Estimations of maximal spillover radius by technology

Technology	USPTO code	Raw value from technique (more negative values in columns 3 and 6 are steeper decays, more positive values in columns 4 and 7 are steeper decays): US parametric citation decays	US non-parametric citation decays	US first- vs. later-gen. citations	UK parametric citation decays	UK non-param. citation decays
(1)	(2)	(3)	(4)	(5)	(6)	(7)
Category 1: Chemicals		**-0.090**	**0.544**	**8**	**-0.084**	**0.510**
Agriculture, Food, Textiles	11	-0.035	0.373	13	-0.033	0.365
Coating	12	-0.081	0.771	5	-0.079	0.757
Gas	13	-0.031	0.303	1	-0.028	0.288
Organic Compounds	14	-0.034	0.291	7	-0.028	0.258
Resins	15	-0.171	0.784	19	-0.162	0.731
Miscellaneous Chemical	19	-0.185	0.742	5	-0.177	0.659
Category 2: Computers & Communications		**-0.183**	**1.944**	**4**	**-0.143**	**0.102**
Communications	21	-0.188	1.127	5	-0.129	0.107
Computer Hardware & Software	22	-0.199	1.784	3	-0.186	0.132
Computer Peripherals	23	-0.194	1.420	3	-0.161	0.082
Information Storage	24	-0.152	3.447	3	-0.097	0.088
Category 3: Drugs and Medical		**-0.138**	**0.668**	**7**	**-0.191**	**0.140**
Drugs	31	-0.114	0.387	15	-0.196	0.243
Surgery & Medical Instruments	32	-0.224	1.385	3	-0.239	0.121
Biotechnology	33	-0.133	0.470	3	-0.250	0.117
Miscellaneous Drug & Medical	39	-0.081	0.431	5	-0.080	0.079
Category 4: Electrical and Electronic		**-0.167**	**1.080**	**6**	**-0.097**	**0.120**
Electrical Devices	41	-0.244	1.272	3	-0.048	0.032
Electrical Lighting	42	-0.097	0.731	3	-0.037	0.129
Measuring & Testing	43	-0.146	0.936	15	-0.134	0.113
Nuclear & X-rays	44	-0.136	0.876	5	-0.087	0.160
Power Systems	45	-0.129	0.798	3	-0.037	0.077
Semiconductor Devices	46	-0.304	2.262	3	-0.210	0.230
Miscellaneous Electrical	49	-0.111	0.688	9	-0.128	0.096
Category 5: Mechanical		**-0.133**	**0.567**	**6**	**-0.086**	**0.121**
Materials Processing & Handling	51	-0.091	0.372	3	-0.091	0.072
Metal Working	52	-0.064	0.212	5	-0.047	0.053
Motors, Engines & Parts	53	-0.176	0.590	7	-0.098	0.158
Optics	54	-0.224	1.455	5	-0.112	0.168
Transportation	55	-0.188	0.570	11	-0.092	0.145
Miscellaneous Mechanical	59	-0.058	0.204	3	-0.073	0.129
Category 6: Other		**-0.059**	**0.325**	**9**	**-0.138**	**0.129**
Agriculture, Husbandry, Food	61	-0.044	0.198	1	-0.129	0.232
Amusement Devices	62	-0.068	0.502	3	-0.262	0.116
Apparel & Textile	63	-0.027	0.148	7	-0.074	0.081
Earth Working & Wells	64	-0.179	1.001	1	-0.262	0.252
Furniture, House Fixtures	65	-0.035	0.061	9	-0.167	0.083
Heating	66	-0.016	0.125	1	-0.020	0.031
Pipes & Joints	67	-0.053	0.515	21	-0.084	0.193
Receptacles	68	-0.036	0.139	31	-0.137	0.125
Miscellaneous Others	69	-0.073	0.235	5	-0.109	0.052

Table 2: Basic cluster traits and maximal radius estimations

	Size of clusters			Density of clusters		
	Mean distance to other patents in MSA from the dominant zip code per tech.	Median distance to other patents in MSA from the dominant zip code per tech.	Herfindahl index of patent distribution over zip codes within MSA for tech.	Share of patents that occur within 150 miles of each other that occur within 50 miles	Column 4's measure with the maximum density capped at 50%	Share of patents that occur within 50 miles of each other that occur within 25 miles
	Prediction: Longer distance	Prediction: Longer distance	Prediction: Weaker HHI	Prediction: Less dense	Prediction: Less dense	Prediction: Less dense
	(1)	(2)	(3)	(4)	(5)	(6)
	Variables are in unit standard deviations, regressions control for patent count by technology					
	A. Measuring radius through log-linear patent citation decay functions					
Maximal radius of interaction	0.416	0.392	-0.631	-0.776	-0.758	-0.818
	(0.151)	(0.172)	(0.176)	(0.134)	(0.109)	(0.098)
	B. Measuring radius through non-parametric patent citation decay functions					
Maximal radius of interaction	0.579	0.566	-0.721	-0.851	-0.689	-0.858
	(0.107)	(0.111)	(0.148)	(0.154)	(0.185)	(0.171)
	C. Measuring radius through comparing first- vs. later-generation citation distributions					
Maximal radius of interaction	0.241	0.257	-0.151	-0.253	-0.177	-0.238
	(0.093)	(0.104)	(0.132)	(0.113)	(0.112)	(0.099)
	D. Measuring radius through log-linear patent citation decay functions in United Kingdom					
Maximal radius of interaction	0.427	0.363	-0.361	-0.484	-0.505	-0.390
	(0.103)	(0.121)	(0.179)	(0.160)	(0.137)	(0.162)
	E. Measuring radius through non-parametric patent citation decay functions in United Kingdom					
Maximal radius of interaction	0.163	0.110	-0.229	-0.301	-0.229	-0.254
	(0.148)	(0.158)	(0.181)	(0.193)	(0.173)	(0.187)

Notes: Table quantifies the relationship between traits of technology clusters and the maximal radius of interactions for technologies. Dependent variables are indicated by column headers, and panel titles indicate the technique employed to measure the maximal radius. A longer maximal radius is predicted to have larger and less dense clusters. Cluster traits are measured during the 1990-1999 period. Sample includes 36 technologies at the sub-category level of the USPTO classification system. Variables are transformed to have unit standard deviation for interpretation. Estimations are unweighted, control for the number of US patents in the technology, and report robust standard errors.

Table 3: Extensions on continuous density estimations

Distance interval	Measuring radius using US parametric citation decays		Column 1 with weights for technology size		Column 1 excluding miscellaneous categories		Column 1 with bootstrapped standard errors		Measuring radius using US non-parametric citation decays		Measuring radius using first- vs. later-gen. citations		Measuring radius using UK parametric citation decays		Measuring radius using UK non-param. citation decays	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
A. Dependent variable is kernel density by distance in unit standard deviations																
[0,25]	-1.338	(0.310)	-1.364	(0.353)	-1.354	(0.313)	-1.338	(0.139)	-1.773	(0.206)	-0.430	(0.209)	-0.754	(0.297)	-0.599	(0.370)
(25,50]	-0.469	(0.120)	-0.423	(0.119)	-0.470	(0.123)	-0.469	(0.080)	-0.623	(0.074)	-0.107	(0.095)	-0.286	(0.107)	-0.244	(0.145)
(50,75]	0.032	(0.053)	0.062	(0.062)	0.043	(0.057)	0.032	(0.030)	0.023	(0.050)	0.073	(0.063)	0.016	(0.057)	-0.017	(0.078)
(75,100]	0.146	(0.047)	0.169	(0.057)	0.159	(0.050)	0.146	(0.022)	0.163	(0.060)	0.128	(0.076)	0.103	(0.068)	0.036	(0.091)
(100,125]	0.159	(0.048)	0.184	(0.049)	0.171	(0.050)	0.159	(0.020)	0.193	(0.058)	0.133	(0.073)	0.135	(0.057)	0.044	(0.079)
(125,150]	0.144	(0.048)	0.147	(0.044)	0.156	(0.050)	0.144	(0.021)	0.209	(0.038)	0.116	(0.063)	0.146	(0.042)	0.045	(0.064)
(150, 175]	0.130	(0.043)	0.128	(0.042)	0.144	(0.044)	0.130	(0.019)	0.217	(0.027)	0.097	(0.061)	0.121	(0.043)	0.018	(0.062)
(175,200]	0.091	(0.045)	0.110	(0.047)	0.107	(0.046)	0.091	(0.022)	0.202	(0.031)	0.080	(0.067)	0.070	(0.060)	-0.051	(0.067)
(200,225]	0.050	(0.081)	0.119	(0.071)	0.071	(0.082)	0.050	(0.037)	0.214	(0.059)	0.076	(0.072)	0.064	(0.096)	-0.113	(0.090)
B. Dependent variable is localization metric using 5%/95% confidence bands by distance in unit standard deviations																
[0,25]	-1.359	(0.332)	-1.382	(0.376)	-1.375	(0.333)	-1.359	(0.149)	-1.838	(0.213)	-0.444	(0.216)	-0.761	(0.307)	-0.631	(0.387)
(25,50]	-0.476	(0.126)	-0.426	(0.127)	-0.478	(0.128)	-0.476	(0.084)	-0.642	(0.078)	-0.116	(0.098)	-0.289	(0.111)	-0.250	(0.150)
(50,75]	0.010	(0.052)	0.044	(0.057)	0.020	(0.055)	0.010	(0.029)	0.016	(0.045)	0.068	(0.066)	0.000	(0.058)	-0.013	(0.082)
(75,100]	0.121	(0.049)	0.147	(0.055)	0.133	(0.051)	0.121	(0.022)	0.153	(0.055)	0.122	(0.077)	0.085	(0.069)	0.041	(0.094)
(100,125]	0.137	(0.049)	0.164	(0.047)	0.148	(0.050)	0.137	(0.021)	0.184	(0.054)	0.126	(0.072)	0.115	(0.057)	0.044	(0.081)
(125,150]	0.123	(0.050)	0.128	(0.047)	0.133	(0.051)	0.123	(0.022)	0.201	(0.036)	0.108	(0.060)	0.121	(0.043)	0.043	(0.064)
(150, 175]	0.098	(0.046)	0.102	(0.045)	0.109	(0.046)	0.098	(0.019)	0.205	(0.028)	0.088	(0.058)	0.093	(0.042)	0.015	(0.062)
(175,200]	0.054	(0.044)	0.079	(0.044)	0.067	(0.044)	0.054	(0.021)	0.189	(0.036)	0.073	(0.066)	0.045	(0.055)	-0.050	(0.067)
(200,225]	0.019	(0.079)	0.086	(0.071)	0.039	(0.079)	0.019	(0.036)	0.206	(0.063)	0.070	(0.071)	0.043	(0.090)	-0.105	(0.088)

Notes: Table quantifies the relationship between the density of technology clusters by distance intervals and the technology's maximal radius of interaction. Panel A considers the kernel density estimates for technologies, and Panel B considers the localization metric of Duranton and Overman (2005). The explanatory variables are interactions of indicator variables for distance bands with technology-level spillover lengths. Technologies with longer spillover ranges are predicted to show lower density at short distances and increased activity over medium distances (i.e., larger and less dense clusters). For columns 1-4, technology spillover lengths are measured as in Figure 4a by the log rate of citation decay by distance, controlling for underlying patenting and citing zip codes fixed effects. Variations on technology spillover lengths are employed in Columns 5-8 similar to Table 2. Sample includes 36 technologies at the sub-category level of the USPTO classification system. Variables are transformed to have unit standard deviation for interpretation. Except where noted, estimations are unweighted, control for fixed effects by distance, and report robust standard errors.

Table 4: Extensions on continuous density estimations, continued

Distance interval	Splitting technologies by rates of citations to patents within the same zip code				Splitting technologies by rates of citations to patents in other technology areas				Splitting technologies by rates of citations to very recent patents (last five years)				Splitting technologies by average wage costs of MSAs where patenting conducted			
	Above median		Below median		Above median		Below median		Above median		Below median		Above median		Below median	
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)	
A. Dependent variable is kernel density by distance in unit standard deviations																
[0,25]	-1.825	(0.289)	-0.837	(0.277)	-1.483	(0.570)	-1.254	(0.369)	-1.542	(0.420)	-0.695	(0.340)	-1.829	(0.436)	-1.081	(0.453)
(25,50]	-0.630	(0.139)	-0.315	(0.109)	-0.615	(0.276)	-0.413	(0.118)	-0.509	(0.149)	-0.312	(0.174)	-0.712	(0.182)	-0.346	(0.156)
(50,75]	0.049	(0.089)	-0.002	(0.041)	-0.028	(0.125)	0.038	(0.041)	0.046	(0.052)	-0.027	(0.115)	-0.058	(0.068)	0.070	(0.069)
(75,100]	0.185	(0.081)	0.088	(0.035)	0.165	(0.102)	0.122	(0.050)	0.158	(0.039)	0.008	(0.145)	0.104	(0.062)	0.157	(0.067)
(100,125]	0.190	(0.081)	0.113	(0.047)	0.227	(0.100)	0.119	(0.053)	0.173	(0.044)	-0.029	(0.141)	0.138	(0.078)	0.162	(0.063)
(125,150]	0.181	(0.080)	0.095	(0.048)	0.238	(0.089)	0.097	(0.057)	0.163	(0.051)	-0.056	(0.127)	0.114	(0.089)	0.155	(0.060)
(150, 175]	0.155	(0.074)	0.096	(0.045)	0.218	(0.074)	0.088	(0.055)	0.153	(0.047)	-0.077	(0.108)	0.076	(0.080)	0.152	(0.052)
(175,200]	0.065	(0.082)	0.111	(0.049)	0.146	(0.096)	0.062	(0.053)	0.118	(0.050)	-0.173	(0.098)	-0.031	(0.085)	0.147	(0.045)
(200,225]	-0.039	(0.156)	0.134	(0.063)	0.004	(0.190)	0.051	(0.078)	0.091	(0.096)	-0.362	(0.131)	-0.212	(0.160)	0.180	(0.051)
B. Dependent variable is localization metric using 5%/95% confidence bands by distance in unit standard deviations																
[0,25]	-1.898	(0.306)	-0.804	(0.279)	-1.554	(0.609)	-1.257	(0.398)	-1.565	(0.452)	-0.663	(0.383)	-1.858	(0.480)	-1.101	(0.482)
(25,50]	-0.655	(0.144)	-0.303	(0.113)	-0.642	(0.288)	-0.412	(0.125)	-0.519	(0.158)	-0.306	(0.183)	-0.724	(0.197)	-0.352	(0.164)
(50,75]	0.033	(0.090)	-0.028	(0.033)	-0.034	(0.123)	0.014	(0.043)	0.012	(0.046)	-0.057	(0.118)	-0.076	(0.052)	0.047	(0.070)
(75,100]	0.167	(0.084)	0.057	(0.030)	0.144	(0.106)	0.101	(0.054)	0.129	(0.041)	-0.026	(0.149)	0.079	(0.060)	0.134	(0.071)
(100,125]	0.178	(0.083)	0.082	(0.042)	0.213	(0.105)	0.099	(0.057)	0.148	(0.046)	-0.061	(0.141)	0.114	(0.079)	0.143	(0.066)
(125,150]	0.167	(0.082)	0.068	(0.043)	0.218	(0.095)	0.078	(0.060)	0.139	(0.052)	-0.089	(0.124)	0.088	(0.093)	0.137	(0.061)
(150, 175]	0.132	(0.076)	0.055	(0.043)	0.191	(0.081)	0.058	(0.058)	0.115	(0.049)	-0.120	(0.104)	0.044	(0.083)	0.122	(0.056)
(175,200]	0.042	(0.081)	0.061	(0.048)	0.127	(0.098)	0.023	(0.053)	0.067	(0.045)	-0.216	(0.086)	-0.060	(0.076)	0.108	(0.047)
(200,225]	-0.050	(0.153)	0.085	(0.066)	0.003	(0.185)	0.013	(0.081)	0.048	(0.090)	-0.402	(0.113)	-0.247	(0.141)	0.154	(0.051)

Notes: See Table 3. Estimations measure radius using US parametric citation decays.

Table 5: Point-by-point regressions of density functions for technologies

	Raw share of patent counts by bilateral distance in miles between patents											
Distance range	[0,5)	[5,10)	[10,15)	[15,20)	[0,20)	[20,40)	[40,60)	[60,80)	[80,100)	[100,120)	[120,140)	[140,250]
Mean	0.044	0.041	0.044	0.036	0.165	0.106	0.062	0.059	0.061	0.064	0.064	0.420
Standard deviation	(0.047)	(0.037)	(0.035)	(0.016)	(0.130)	(0.027)	(0.016)	(0.014)	(0.016)	(0.016)	(0.018)	(0.090)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Radius measures are in unit standard deviations, regressions control for patent count by technology											
	A. Measuring radius through log-linear patent citation decay functions											
Maximal radius of interaction	-0.032	-0.026	-0.024	-0.012	-0.093	-0.013	0.008	0.009	0.011	0.012	0.011	0.056
	(0.010)	(0.005)	(0.005)	(0.002)	(0.021)	(0.005)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.012)
	B. Measuring radius through non-parametric patent citation decay functions											
Maximal radius of interaction	-0.041	-0.032	-0.033	-0.014	-0.120	-0.017	0.010	0.011	0.011	0.011	0.013	0.080
	(0.006)	(0.004)	(0.003)	(0.001)	(0.014)	(0.004)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)	(0.007)
	C. Measuring radius through comparing first- vs. later-generation citation distributions											
Maximal radius of interaction	-0.014	-0.011	-0.010	-0.005	-0.040	-0.003	0.006	0.005	0.006	0.004	0.005	0.017
	(0.007)	(0.004)	(0.005)	(0.002)	(0.017)	(0.004)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)	(0.012)
	D. Measuring radius through log-linear patent citation decay functions in United Kingdom											
Maximal radius of interaction	-0.021	-0.015	-0.014	-0.010	-0.059	-0.018	0.002	0.005	0.005	0.008	0.009	0.050
	(0.010)	(0.007)	(0.007)	(0.002)	(0.025)	(0.004)	(0.004)	(0.003)	(0.004)	(0.003)	(0.002)	(0.015)
	E. Measuring radius through non-parametric patent citation decay functions in United Kingdom											
Maximal radius of interaction	-0.017	-0.009	-0.008	-0.006	-0.040	-0.008	0.001	0.004	0.004	0.006	0.006	0.026
	(0.013)	(0.008)	(0.008)	(0.003)	(0.032)	(0.005)	(0.004)	(0.004)	(0.005)	(0.004)	(0.003)	(0.018)

Notes: Table quantifies the point-by-point relationship between spatial distributions for technology clusters and the maximal radius of interactions for technologies. Dependent variables are shares of bilateral distances between patents for each technology that fall within the indicated distance band. Shares are calculated relative to all bilateral distances observed for the technology over the distances of zero miles to 250 miles. The table header provides the raw descriptive statistics. Panel titles indicate the technique employed to measure the maximal radius. A longer maximal radius is predicted to have larger and less dense clusters. Cluster traits are measured during the 1990-1999 period. Sample includes 36 technologies at the sub-category level of the USPTO classification system. Radius measures are transformed to have unit standard deviation for interpretation. Estimations are unweighted, control for the number of US patents in the technology, and report robust standard errors.

Table 6: Point-by-point regressions of density functions for technologies using unweighted assignees

	Raw share of patent assignee counts by bilateral distance in miles between patent assignees											
Distance range	[0,5)	[5,10)	[10,15)	[15,20)	[0,20)	[20,40)	[40,60)	[60,80)	[80,100)	[100,120)	[120,140)	[140,250]
Mean	0.024	0.032	0.038	0.035	0.128	0.117	0.074	0.064	0.066	0.067	0.066	0.419
Standard deviation	(0.018)	(0.026)	(0.029)	(0.016)	(0.086)	(0.032)	(0.011)	(0.011)	(0.011)	(0.012)	(0.013)	(0.080)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Radius measures are in unit standard deviations, regressions control for patent count by technology											
	A. Measuring radius through log-linear patent citation decay functions											
Maximal radius of interaction	-0.010	-0.015	-0.016	-0.009	-0.050	-0.020	0.001	0.005	0.007	0.008	0.007	0.042
	(0.003)	(0.004)	(0.004)	(0.002)	(0.012)	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)	(0.010)
	B. Measuring radius through non-parametric patent citation decay functions											
Maximal radius of interaction	-0.015	-0.022	-0.025	-0.014	-0.076	-0.028	0.002	0.008	0.009	0.011	0.010	0.063
	(0.001)	(0.002)	(0.003)	(0.001)	(0.007)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)
	C. Measuring radius through comparing first- vs. later-generation citation distributions											
Maximal radius of interaction	-0.005	-0.006	-0.007	-0.004	-0.022	-0.005	0.004	0.004	0.004	0.004	0.004	0.008
	(0.002)	(0.003)	(0.004)	(0.002)	(0.011)	(0.004)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.010)
	D. Measuring radius through log-linear patent citation decay functions in United Kingdom											
Maximal radius of interaction	-0.008	-0.009	-0.009	-0.006	-0.033	-0.015	0.000	0.004	0.004	0.006	0.007	0.028
	(0.003)	(0.004)	(0.005)	(0.002)	(0.014)	(0.005)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.012)
	E. Measuring radius through non-parametric patent citation decay functions in United Kingdom											
Maximal radius of interaction	-0.004	-0.005	-0.004	-0.003	-0.016	-0.007	0.000	0.002	0.003	0.003	0.003	0.012
	(0.004)	(0.006)	(0.006)	(0.003)	(0.018)	(0.006)	(0.002)	(0.003)	(0.003)	(0.003)	(0.002)	(0.015)

Notes: See Table 5. Estimations treat each assignee or individual location as a single observation independent of patent levels.

App. Table 1: Descriptive tabulations on citation distribution

Distance interval between zip code centroids	Raw citation counts			Citation share distribution over 0-150 mile			Shares relative to random	
	First-generation citations	Later-generation citations	Random citation distribution	First-generation citations	Later-generation citations	Random citation distribution	First-generation citations	Later-generation citations
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[0,1] miles	314,381	214,049	368,653	15.4%	10.3%	11.6%	1.331	0.887
(1,3] miles	22,911	21,846	23,329	1.1%	1.1%	0.7%	1.533	1.430
(3,5] miles	69,206	72,998	103,219	3.4%	3.5%	3.2%	1.046	1.080
(5,10] miles	229,207	255,767	338,860	11.3%	12.3%	10.7%	1.056	1.153
(10,15] miles	222,620	254,700	328,497	10.9%	12.2%	10.3%	1.058	1.184
(15,20] miles	167,916	181,016	233,941	8.3%	8.7%	7.4%	1.120	1.182
(20,25] miles	132,951	148,024	195,444	6.5%	7.1%	6.2%	1.062	1.157
(25,30] miles	105,913	124,996	165,891	5.2%	6.0%	5.2%	0.996	1.151
(30,35] miles	76,287	85,113	116,435	3.7%	4.1%	3.7%	1.023	1.117
(35,40] miles	68,753	82,777	108,620	3.4%	4.0%	3.4%	0.988	1.164
(40,45] miles	42,987	48,086	70,882	2.1%	2.3%	2.2%	0.947	1.036
(45,50] miles	37,897	42,108	61,927	1.9%	2.0%	1.9%	0.955	1.039
(50,55] miles	29,463	32,373	54,897	1.4%	1.6%	1.7%	0.838	0.901
(55,60] miles	28,818	30,639	49,035	1.4%	1.5%	1.5%	0.917	0.954
(60,65] miles	26,359	27,749	52,351	1.3%	1.3%	1.6%	0.786	0.810
(65,70] miles	27,160	30,065	56,252	1.3%	1.4%	1.8%	0.754	0.816
(70,75] miles	25,380	25,991	49,427	1.2%	1.2%	1.6%	0.801	0.803
(75,80] miles	25,105	25,230	49,988	1.2%	1.2%	1.6%	0.784	0.771
(80,85] miles	26,738	28,086	49,023	1.3%	1.4%	1.5%	0.851	0.875
(85,90] miles	26,633	27,292	48,432	1.3%	1.3%	1.5%	0.858	0.861
(90,95] miles	25,599	25,739	47,810	1.3%	1.2%	1.5%	0.836	0.822
(95,100] miles	25,554	25,950	50,480	1.3%	1.2%	1.6%	0.790	0.785
(100,150] miles	277,289	268,946	552,915	13.6%	12.9%	17.4%	0.783	0.743

Notes: Table provides descriptive statistics on citations sample. The sample builds off of pairwise citing and cited zip codes within 150 miles of each other. First-generation citations are citations made directly between patents. Later-generation citations are second- and third-generation citations (patent A cites patent B that cites patent C). Random citation counterfactuals are drawn with the same technologies and years as true citations within a 250-mile distance interval to model the underlying technology landscape. Columns 1-3 count citations by distance interval between zip codes, with the levels driven in part by the spatial distribution of zip codes in the US. The specific levels of random citations in Column 3 does not have meaning. Columns 4-6 provide these distributions as shares that sum to 100% over 0-150 miles. Columns 7-8 divide the shares for first- and later-generation citations by the random baseline distribution. These columns show the higher concentration of direct citations between inventors at local distances compared to the random counterfactuals. They also show the more localized nature of first-generation citations compared to later-generation citations.

App. Table 2a: Coefficients for Figure 3a's pairwise approach

	First-generation citations	Later-generation citations	Column 1 with non-zero citations	Column 2 with non-zero citations	Column 1 with-out expected citations	Column 2 with-out expected citations	Column 1 including own zip code	Column 2 including own zip code
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[0,1] miles	0.843 (0.008)	0.209 (0.019)	1.054 (0.026)	0.455 (0.036)	1.217 (0.021)	0.603 (0.026)	1.040 (0.089)	0.555 (0.182)
(1,3] miles	0.756 (0.037)	0.574 (0.074)	0.946 (0.016)	0.836 (0.034)	1.570 (0.116)	1.428 (0.143)	0.760 (0.036)	0.572 (0.072)
(3,5] miles	0.660 (0.064)	0.585 (0.125)	0.830 (0.040)	0.824 (0.087)	1.850 (0.206)	1.836 (0.254)	0.667 (0.063)	0.585 (0.123)
(5,10] miles	0.595 (0.062)	0.538 (0.122)	0.759 (0.040)	0.771 (0.086)	1.724 (0.201)	1.724 (0.248)	0.601 (0.061)	0.537 (0.119)
(10,15] miles	0.529 (0.054)	0.511 (0.105)	0.680 (0.035)	0.721 (0.075)	1.441 (0.176)	1.470 (0.217)	0.534 (0.053)	0.510 (0.103)
(15,20] miles	0.399 (0.034)	0.374 (0.066)	0.545 (0.016)	0.584 (0.035)	0.994 (0.111)	0.999 (0.136)	0.402 (0.034)	0.373 (0.065)
(20,25] miles	0.273 (0.024)	0.261 (0.047)	0.392 (0.009)	0.441 (0.018)	0.669 (0.080)	0.677 (0.099)	0.275 (0.024)	0.261 (0.047)
(25,30] miles	0.243 (0.024)	0.289 (0.045)	0.359 (0.010)	0.459 (0.021)	0.560 (0.078)	0.622 (0.097)	0.245 (0.023)	0.288 (0.044)
(30,35] miles	0.185 (0.017)	0.241 (0.032)	0.289 (0.006)	0.400 (0.012)	0.360 (0.058)	0.424 (0.071)	0.186 (0.017)	0.240 (0.031)
(35,40] miles	0.182 (0.023)	0.287 (0.042)	0.285 (0.016)	0.442 (0.033)	0.403 (0.078)	0.519 (0.096)	0.183 (0.023)	0.285 (0.042)
(40,45] miles	0.060 (0.004)	0.148 (0.007)	0.127 (0.006)	0.280 (0.012)	0.060 (0.015)	0.148 (0.019)	0.060 (0.004)	0.147 (0.007)
(45,50] miles	0.076 (0.003)	0.130 (0.005)	0.138 (0.005)	0.231 (0.009)	0.049 (0.013)	0.103 (0.016)	0.076 (0.003)	0.130 (0.006)
(50,55] miles	0.029 (0.002)	0.095 (0.004)	0.062 (0.010)	0.184 (0.021)	-0.046 (0.005)	0.016 (0.006)	0.029 (0.002)	0.095 (0.004)
(55,60] miles	0.032 (0.001)	0.079 (0.003)	0.075 (0.007)	0.133 (0.016)	-0.056 (0.002)	-0.013 (0.003)	0.031 (0.001)	0.079 (0.003)
(60,65] miles	-0.014 (0.001)	0.039 (0.003)	-0.011 (0.006)	0.074 (0.013)	-0.091 (0.002)	-0.042 (0.002)	-0.015 (0.001)	0.038 (0.003)
(65,70] miles	-0.019 (0.002)	0.018 (0.003)	-0.027 (0.002)	0.037 (0.005)	-0.049 (0.007)	-0.014 (0.008)	-0.019 (0.002)	0.018 (0.003)
(70,75] miles	0.001 (0.001)	0.028 (0.003)	0.032 (0.005)	0.066 (0.011)	-0.041 (0.004)	-0.015 (0.005)	0.000 (0.001)	0.028 (0.003)
(75,80] miles	0.004 (0.002)	0.037 (0.003)	0.006 (0.002)	0.041 (0.003)	-0.018 (0.007)	0.014 (0.009)	0.004 (0.002)	0.037 (0.003)
(80,85] miles	0.022 (0.002)	0.077 (0.002)	0.023 (0.001)	0.093 (0.001)	-0.017 (0.007)	0.036 (0.008)	0.022 (0.002)	0.077 (0.002)
(85,90] miles	0.035 (0.002)	0.066 (0.004)	0.039 (0.003)	0.062 (0.005)	0.013 (0.009)	0.044 (0.011)	0.035 (0.002)	0.066 (0.004)
(90,95] miles	0.002 (0.001)	0.018 (0.002)	-0.005 (0.001)	0.017 (0.004)	-0.024 (0.003)	-0.009 (0.004)	0.002 (0.001)	0.018 (0.002)
(95,100] miles	-0.023 (0.000)	-0.011 (0.000)	-0.044 (0.000)	-0.031 (0.000)	-0.025 (0.001)	-0.013 (0.001)	-0.023 (0.000)	-0.011 (0.000)
Patent interaction	0.043 (0.016)	0.031 (0.017)	0.209 (0.035)	0.265 (0.055)	0.706 (0.062)	0.728 (0.077)	0.045 (0.017)	0.034 (0.019)
Expected citations	0.842 (0.008)	0.885 (0.021)	0.672 (0.016)	0.667 (0.019)			0.838 (0.009)	0.882 (0.020)
Observations	12,779,214	12,779,214	268,417	228,000	12,779,214	12,779,214	12,803,952	12,803,952

Notes: Table provides coefficient estimates and standard errors for Figure 3a. The sample builds off of pairwise citing and cited zip codes within 150 miles of each other. Explanatory variables are indicator variables for distance bands with effects measured relative to zip codes 100-150 miles apart. Regressions control for an interaction of log patenting in the pairwise zip codes and log expected citations based upon random counterfactuals that have the same technologies and years as true citations. Citations within the same zip code are excluded. Regressions are weighted by the log interaction of patenting in the two zip codes and report standard errors clustered by distance interval.

App. Table 2b: Coefficients for Figure 3b's distance ring with fixed effects approach

	First-generation citations	Later-generation citations	Column 1 with non-zero citations	Column 2 with non-zero citations	Column 1 without expected citations	Column 2 without expected citations	Column 1 including own zip code	Column 2 including own zip code
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
[0,1] miles	1.220 (0.027)	0.834 (0.052)	1.406 (0.032)	0.993 (0.049)	1.719 (0.044)	1.348 (0.055)	1.240 (0.045)	0.976 (0.064)
(1,3] miles	1.014 (0.033)	0.880 (0.058)	1.145 (0.041)	1.078 (0.075)	1.510 (0.024)	1.390 (0.047)	0.961 (0.055)	0.866 (0.056)
(3,5] miles	0.914 (0.042)	0.828 (0.060)	1.021 (0.051)	0.996 (0.076)	1.419 (0.034)	1.348 (0.047)	0.886 (0.044)	0.821 (0.057)
(5,10] miles	0.878 (0.035)	0.814 (0.048)	0.964 (0.041)	0.944 (0.058)	1.356 (0.029)	1.305 (0.038)	0.862 (0.033)	0.811 (0.047)
(10,15] miles	0.757 (0.029)	0.723 (0.044)	0.832 (0.034)	0.833 (0.053)	1.154 (0.023)	1.131 (0.034)	0.740 (0.028)	0.719 (0.043)
(15,20] miles	0.648 (0.018)	0.630 (0.032)	0.711 (0.022)	0.724 (0.039)	0.962 (0.012)	0.953 (0.023)	0.623 (0.026)	0.625 (0.031)
(20,25] miles	0.509 (0.015)	0.521 (0.029)	0.563 (0.018)	0.604 (0.035)	0.764 (0.012)	0.784 (0.024)	0.481 (0.027)	0.515 (0.028)
(25,30] miles	0.429 (0.015)	0.493 (0.028)	0.478 (0.018)	0.564 (0.035)	0.630 (0.015)	0.701 (0.027)	0.400 (0.027)	0.489 (0.026)
(30,35] miles	0.333 (0.013)	0.415 (0.025)	0.377 (0.017)	0.480 (0.032)	0.466 (0.017)	0.552 (0.029)	0.302 (0.029)	0.410 (0.024)
(35,40] miles	0.268 (0.013)	0.367 (0.025)	0.306 (0.017)	0.422 (0.032)	0.377 (0.017)	0.479 (0.029)	0.241 (0.026)	0.364 (0.022)
(40,45] miles	0.171 (0.012)	0.302 (0.021)	0.202 (0.015)	0.350 (0.026)	0.232 (0.016)	0.364 (0.028)	0.137 (0.032)	0.297 (0.019)
(45,50] miles	0.138 (0.010)	0.256 (0.019)	0.165 (0.014)	0.297 (0.023)	0.162 (0.014)	0.280 (0.025)	0.104 (0.032)	0.250 (0.018)
(50,55] miles	0.088 (0.012)	0.221 (0.022)	0.116 (0.016)	0.260 (0.030)	0.096 (0.018)	0.230 (0.030)	0.052 (0.033)	0.216 (0.020)
(55,60] miles	0.065 (0.012)	0.203 (0.020)	0.090 (0.017)	0.238 (0.026)	0.047 (0.018)	0.185 (0.027)	0.031 (0.033)	0.199 (0.018)
(60,65] miles	0.011 (0.012)	0.136 (0.019)	0.032 (0.017)	0.165 (0.026)	-0.018 (0.018)	0.106 (0.027)	-0.024 (0.033)	0.132 (0.018)
(65,70] miles	0.001 (0.012)	0.168 (0.018)	0.022 (0.018)	0.195 (0.026)	-0.023 (0.018)	0.144 (0.026)	-0.033 (0.031)	0.165 (0.017)
(70,75] miles	0.008 (0.012)	0.152 (0.018)	0.038 (0.016)	0.190 (0.027)	-0.016 (0.016)	0.127 (0.025)	-0.027 (0.033)	0.147 (0.017)
(75,80] miles	-0.011 (0.012)	0.099 (0.019)	0.018 (0.017)	0.127 (0.032)	-0.050 (0.017)	0.060 (0.028)	-0.046 (0.032)	0.095 (0.018)
(80,85] miles	-0.002 (0.011)	0.140 (0.021)	0.018 (0.016)	0.159 (0.033)	-0.045 (0.015)	0.096 (0.029)	-0.038 (0.033)	0.136 (0.019)
(85,90] miles	0.019 (0.012)	0.131 (0.021)	0.041 (0.018)	0.150 (0.036)	-0.033 (0.016)	0.078 (0.030)	-0.016 (0.033)	0.127 (0.020)
(90,95] miles	-0.027 (0.011)	0.105 (0.021)	-0.005 (0.016)	0.136 (0.035)	-0.073 (0.014)	0.057 (0.030)	-0.062 (0.034)	0.100 (0.019)
(95,100] miles	-0.038 (0.009)	0.106 (0.020)	-0.016 (0.016)	0.125 (0.035)	-0.087 (0.012)	0.056 (0.028)	-0.074 (0.033)	0.100 (0.019)
Patent interaction	0.335 (0.027)	0.348 (0.036)	0.423 (0.028)	0.497 (0.040)	0.893 (0.005)	0.922 (0.009)	0.315 (0.030)	0.340 (0.036)
Expected citations	0.610 (0.030)	0.628 (0.036)	0.525 (0.030)	0.502 (0.036)			0.615 (0.029)	0.633 (0.036)
Observations	522,459	522,459	87,280	67,559	522,459	522,459	547,197	547,197
Citing zip code FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: See App. Table 2a. Table provides coefficient estimates and standard errors for Figure 3b.