Reversed Citations and the Localization of Knowledge Spillovers

Abstract

Spillover of knowledge is considered to be an important cause of agglomeration of inventive activity. Many studies argue that knowledge spillovers are localized based on the observation that patents tend to cite nearby patents disproportionately. Specifically, patent citations are typically interpreted as marking the transmission of knowledge from the cited invention to the citing invention. The localization of patent citations is therefore taken as evidence that such knowledge transmission is also localized. Localization of knowledge transmission, however, may not be the only reason that patent citations are localized. Using a set of citations that are unlikely to be associated with knowledge transmission from the cited to the citing invention, we present evidence that challenges the view that localization of citations is driven by localized knowledge transmission. While we are silent on the question of whether knowledge transmission is localized, to the extent that such localization exists, we argue that it is unlikely to be captured by patent citations.

Keywords: Patent citations, localization, distance, knowledge spillovers

JEL Classification: O12, O32, O34

1 Introduction

Regional clusters of economic activity are important sources of innovation and economic growth (Martin and Sunley, 2003; Porter, 1990). In the United States, there are over forty clusters, accounting for more than fifty percent of traded employment in U.S. Economic Areas (Porter, 2007). At the same time, clusters vary in their economic and inventive output (Agrawal et al., 2014; Delgado et al., 2008; Porter, 2003; Suire and Vincente, 2009). These differences have attracted scholars to explore potential mechanisms that contribute to their success including labor market pooling (Audretsch and Feldman, 1996; Delgado et al., 2014; Krugman, 1991; Rosenthal and Strange, 2001), specialized resources (Porter, 1990, 2003; Swann, 1998) and knowledge spillovers (Audretsch, 1998; Giuliani, 2007; Jaffe et al., 1993; Saxenian, 1994).

The growing importance of innovation and improved access to patent data have led scholars to use patent citations to study the localization of knowledge spillovers. In a seminal paper, Jaffe, Trajtenberg, and Henderson (1993) find that a citing patent is five to ten times more likely to be in the same SMSA as the cited patent than a matched non-citing patent. The authors and most subsequent studies (a) view patent citation as a marker for knowledge spillovers and (b) interpret the finding that citing and cited

patents are co-located as evidence that knowledge spillovers are localized (Alcácer and Gittelman, 2006; Almeida and Kogut, 1999; Belenzon and Schankerman, 2013; Murata et al., 2014; Singh and Marx, 2013; Thompson, 2006).

The current paper presents evidence inconsistent with the interpretation that localization of citations is driven by localized transmission of knowledge from the cited invention to the citing invention. While our paper is silent on the question of whether knowledge transmission across inventions is localized, to the extent that such localization exists, we argue that it is unlikely to be captured by patent citations. Our empirical test is as follows. We identify a set of citations that are unlikely to reflect knowledge transmission from the cited invention to the citing invention, which we label as "citation reversals", and use these citations as a benchmark to compare the localization of citations for which knowledge transmission is possible, "citation non-reversals." Citation reversals typically occur during the patent examination process when an inventor or an examiner cites a relevant patent. Because ascertaining priority dates (the earliest filing date of a patent on a specific invention) is not always straightforward, citation reversals might occur. If localization of citations is driven by localized knowledge transmission, we expect citation non-reversals to be more localized than citation reversals.

There are at least two broad interpretations for the relationship between patent citations and knowledge spillovers. The first is that a citation from citing patent A to cited patent B indicates that A builds on B (e.g., Belenzon, 2012; Galasso and Schankerman, 2014; Williams, 2013). Patent citations are localized if inventors are more likely to learn from the inventions of nearby inventors. In other words, a citation indicates that inventions A and B are sequentially linked, and localization in this context means that sequential innovation is localized. By definition, the citing invention must come after the cited invention, thus the temporal structure of the citation sequence is key. According to this view, our empirical test should shed new light on the interpretation of localized citations.

The second interpretation builds on the view that knowledge is "in the air." Accordingly, a citation from patent A to patent B indicates that some background knowledge embodied in invention B is relevant for invention A. Roughly speaking, a citation from A to B implies that the inventors of A are working on a similar problem as the inventors of B. Localization of citations means that inventors in the same locality are working on related problems, not that knowledge spillovers *among inventions* are necessarily

localized.[1] A variant of this interpretation is that a patent citation is a *noisy signal* of local unobserved interactions among inventors. These interactions can manifest as patent citations regardless of whether the citing and cited inventions are actually linked. Clearly, the temporal structure of citations is irrelevant in this case, rendering our empirical test not meaningful.[2] However, this interpretation does not explain examiner-added citations, in contrast to where both A and B draw upon the same background knowledge. Moreover, this interpretation of citations as markers of interactions among inventors is at odds with the legal interpretation of citations as limiting the scope of the invention claimed in the patent.

The distinction between drawing upon shared background knowledge and learning from specific inventions is important for several streams of research. An inventor whose invention is built upon by subsequent inventions can potentially extract licensing revenues from the latter. Indeed, the key questions analyzed in models of cumulative innovation are how to divide rents between sequential inventors (Green and Scotchmer, 1996), choosing patentability criteria or determining patent length and breadth (Bessen and Maskin, 2009; O'Donoghue et al., 1998). If, instead, the invention draws upon a common pool of knowledge, although the innovation may be cumulative it would be very difficult for the inventor to identify subsequent inventions that build upon her invention (Laitner and Stolyarov, 2013). Other studies within cumulative innovation examine whether intellectual property rights hinder sequential innovation (Galasso and Schankerman, 2014; Williams, 2013). Their key assumption is that inventors build on the knowledge contained in specific inventions rather than on background knowledge. These studies interpret citations as sequential links between inventions.

The distinction is also relevant for the study of entrepreneurial spinoffs and regional clusters. Inventors leaving existing employers to start their own firms could build upon either specific inventions (Anton and Yao, 1995) or more general background knowledge (Chatterji, 2009). If spinoffs, which are often co-located with parents, capitalize on discoveries employees make in the course of employment, firms have the ability to design contracts to protect themselves. Such contrivances are less useful if spinoffs exploit more general

---

[1]This explanation leaves open the question of why people working on related problems are located close to each other. One reason could be that such problem solving requires specialized skills (i.e., labor pooling), or access to specialized knowledge that is available only in that region, such as from a local university. If the latter, there may well be knowledge spillovers at work, but not across inventors.

[2]In fact, proponents of the "noisy signal" interpretation of citations might even argue that finding a high degree of localization among reversed citations is consistent with the noisy signal view. If citations are just a noisy signal of inventive activity and do not mark actual sequential links between inventions, reversed links are likely and might be even expected.

knowledge learned in the course of their employment.

The distinction being drawn is vital for innovation management and strategy scholarship. Firms concerned about protecting inventions from spillover to outsiders may try to disperse inventors (Alcácer and Zhao, 2012; Zhao, 2006) or locate them away from competitors (Alcácer and Chung, 2007; Shaver and Flyer, 2000). This strategy is likely to be more effective if inventors are in fact more likely to build upon co-located inventors. If, on the other hand, co-location of citation reflects a reliance upon common knowledge, dispersing inventive activity will merely hurt the firm.

Using 1,356,738 USPTO citations and an equal number of control citations, we confirm the stylized fact that patent citations are localized. Our estimates imply that a 50-mile increase in distance between inventors is associated with a 19% reduction in the probability of citation (close to 40% of the sample's average citation probability). However, when we compare non-reversals to reversals, we find that the two citation groups are equally as localized. That is, the probability of citations falls with distance at the same rate for both groups of citations. This finding is robust to a battery of tests and is inconsistent with the view that citations are localized because inventions are more likely to build on other close by inventions.

The rest of the paper is organized as follows. Section 2 discusses related literature, Section 3 describes the conceptual framework, Section 4 discusses the data, Section 5 presents non-parametric results, Section 6 presents the estimation results, and Section 7 concludes.

2   Related literature

The study of knowledge spillovers using patent citations can be traced back to Jaffe et al. (1993) who argue that knowledge spillovers are localized because, as they demonstrate, citations are significantly more likely than controls to be in the same geographic area. However, the relationship between citations and spillovers remains unclear in the Jaffe et al. paper and in most follow up research. On the one hand, (i) localized citations might indicate that a citing patent is more likely to build on a cited patent if the two inventions are local (a citation from A to B means that A builds on B). On the other hand, (ii) a citation might be a *noisy signal* of local inventive activities either because knowledge is "in the air," or regional specialization causes local inventors to work on similar problems and draw from common background

knowledge. These inventors might cite even if the citing invention does not build on the cited invention. The present paper proposes an empirical test to help clarify the interpretation of localized citations. At a minimum, it pushes authors to take a clearer stand on whether they view citations as a sequential link between patented inventions or as a noisy measure of related inventive activities whereby the citing patent may not necessarily build on the cited patent. Our discussion in the introduction explains why this distinction is important.

There are numerous studies on regional clusters that utilize patent citations to examine whether knowledge spillovers are localized (e.g. Jaffe et al., 1993; Belenzon and Schankerman, 2013; Murata et al., 2014; Nomaler and Verspagen, 2016; Verspagen and Schoenmakers, 2004) and whether co-located firms benefit from localized knowledge spillovers (e.g. Alcácer and Zhao, 2012; Menon, 2015). While these studies demonstrate that citations are localized, they are usually unclear on whether this localization is due to local transmission of knowledge from the citing to the cited patent (localized sequential innovation), or due to a noisy signal of some underlying inventive activity whereby local inventors might learn from each other or just be working on related problems. For example, Verspagen and Schoenmakers (2004) examine citation frequencies within and across regional clusters and find that citations tend to occur more frequently within clusters than across clusters. In another study, Nomaler and Verspagen (2016) use direct citations and long-run citation chains (citation sequences linking multiple patents both directly and indirectly) to confirm that direct citations tend to be concentrated within clusters, but citation chains can span across clusters. In exploring whether firms benefit from localized knowledge spillovers, Menon (2015) shows that firms tend to patent more when they are located near other more inventive firms and argues that this increase in patenting is due to benefits of localized knowledge spillovers.

Other studies try to uncover potential mechanisms underlying the localization of citations. An influential stream of research examines inventor mobility. This line of work identifies inventors' intra-regional mobility based on inventor addresses and patent assignees and examine the relationship between mobility and localized citations. These studies show that citations are local partly because inventor mobility is local. Almeida and Kogut (1999) identify top twenty inventors in terms of number of forward citations and estimate the effect of their mobility on intra-regional citations. The authors find a positive relationship between mobility of top inventors within a region and the probability that a patent cites another

patent in the same region. This finding is interpreted as evidence that inventors' tendency to stay in a specific region drives the localization of knowledge spillovers in that region. In a related study, Breschi and Lissoni (2009) use patent citations data from the European Patent Office to examine the contribution of inventor mobility to localized knowledge spillovers. They compare the probability that citing and cited patents are in the same MSA with and without self citations (inventors citing their own patents). They find that the localization effect drops substantially when self citations are excluded. A different reason for why citations might be localized is that they can be added by local intermediaries. Wagner et al. (2014) argue that patent attorneys build their knowledge repositories while interacting with their clients and reference patents from those repositories when they prepare new patent applications. If patent attorneys are engaged primarily with local inventors and knowledge repositories are built using the inventions of those local inventors, then patent citations added by patent attorneys would likely be localized.

In recent years, several studies cast doubt on the interpretation of patent citations and the notion of localized knowledge spillovers. In their critical review of the localized knowledge spillovers literature, Breschi and Lissoni (2001) argue that prior studies have not been clear about the notion of localized knowledge spillovers and that the conceptual framework of localized knowledge spillovers needs to be reassessed and improved. They also argue that more studies should examine the mechanisms underlying localized knowledge spillovers, such as labor mobility and role of university and research institutions. Criticizing their method, Thompson and Fox-Kean (2005) argue that the matching approach in Jaffe et al. (1993) is too coarse and does not adequately control for existing regional specialization. Imposing stricter matching criteria, they find that the localization effect almost disappears at the SMSA and state levels. In a response to this criticism, Henderson et al. (2005) argue that the lack of localization in Thompson and Fox-Kean (2005) is likely to be due to selection bias arising from inability to match a substantial portion of patents at the subclass level, which in turn drastically reduces sample size. In a subsequent study, Thompson (2006) utilizes examiner citations, which arguably are less likely to reflect knowledge spillovers than inventor citations, and shows that inventor citations are more localized than examiner citations, a finding consistent with the interpretation in Jaffe et al. (1993), yet without relying on their citation matching methodology. In their re-examination of the findings from Jaffe et al. (1993) and Thompson and Fox-Kean (2005), Murata et al. (2014) use distance-based K-density tests to show

that citation pairs are more localized than control pairs in about thirty percent of technology areas even when citing and control patents are matched on six-digit technology classification codes.

The present paper contributes to a broader debate on the meaning of patent citations (Duguet and MacGarvie, 2005; Jaffe et al., 2000; Moser et al., 2017; Roach and Cohen, 2013). Using survey data, Roach and Cohen (2013) show that patent citations are likely to underestimate knowledge transmission from certain types of inventive activities, such as basic research, and overestimate knowledge transmission when firms engage in strategic behaviors of citing prior art to mitigate patent invalidation risk. In another survey, Duguet and MacGarvie (2005) show that citations are associated with technology flows through some channels (e.g. equipment sales), but not through others (e.g. R&D outsourcing). Additional survey evidence by Jaffe et al. (2000) shows that one third of inventors did not learn about the inventions they cite until the citing invention was completed and that close to one third of inventors did not learn about the cited inventions at all. Using data from field trials for hybrid corn, Moser et al. (2017) show that self citations are likely to capture follow-on inventions, but not citations added by examiners.

As noted above, we are agnostic about whether knowledge spillovers are localized. However, we argue that the evidence presented in this paper is inconsistent with the view that patent citations are localized because inventions are more likely to build on other local inventions. Put differently, our results imply that either patent citation is a poor measure of knowledge spillover or that knowledge spillover is not localized.

3 Preliminary concerns

3.1 Reversed citations and local interactions

To test whether the localization of citations is driven by localized knowledge transmission, we identify a set of citations that are not likely to be driven by knowledge transmission, "citation reversals". If localization of citations is driven by localized knowledge transmission, we expect citation non-reversals to be more localized than citation reversals.

An important concern about our methodology is that citation reversals might reflect highly localized knowledge transmission among local inventors who share knowledge about new ideas and inventions that have not been disclosed to the public. If this were true, reversed citations would be localized due to highly

localized knowledge transmission and our key assumption that reversed citations are not associated with knowledge transmission would be violated.

Table 1 summarizes different mechanisms through which both knowledge about inventions and background knowledge could be transmitted, including those that might lead to citation reversals. There are three main learning mechanisms indicated by columns 1–3. The first is learning from patent documents—learning about inventions by reading published patent documents. There is no reason to expect this type of learning to be localized. Thus, we are not concerned that learning from patent documents would generate localized citation reversals. The second mechanism is learning about inventions from inventors. In this case, inventors might share among themselves knowledge about new inventions. Such learning might be localized if inventors tend to discuss their inventions more frequently with nearby inventors than faraway inventors. The third source of learning is patent intermediaries, such as patent attorneys. By using the same patent attorneys, inventors might learn about new inventions before they are disclosed to the public. Because inventors are more likely to use the same patent attorneys if they are near one another, engaging with patent attorneys might generate localized citation reversals.

Section 6 presents several tests aiming at mitigating the concern that reversals are generated through highly localized learning.

Insert Table 1 here

3.2   Comparing the localization effect of reversed and non-reversed citations

We propose the following motivation for why localized knowledge transmission might lead to a stronger negative effect of distance on citation probability for citation non-reversals than for reversals.

We distinguish between citations that reflect knowledge transmission, i.e., where one invention builds upon another, improving or extending it, and citations that reflect other relationships, where no such knowledge transmission takes place. A patent may cite another patent because both draw upon a common pool of knowledge. Patents may be related in other ways as well. For instance, the citing patent may accomplish the same outcome as the cited patent using a different method, or use similar methods to accomplish different goals.

The distinction bears on whether the citing invention benefited from the knowledge created in the

form of the cited invention. The citing inventions that are *related* would be unaffected. The citing invention that builds upon the cited invention would have had to acquire the knowledge created by the cited invention. The extant literature has implicitly or explicitly assumed that, when patent class is controlled for, citations reflect knowledge transmission. We argue that citation reversals, by construction, cannot reflect knowledge transmission from the cited invention to the citing invention.[3] It is conceivable that the inventor may have learned useful knowledge from the inventor of the cited invention, but it is far more likely that a citation reversal represents the recognition of a related invention, albeit one that should not have been cited because it is not prior art.[4]

We do not observe whether a citation represents transmission of knowledge or merely some type of relatedness. Instead, we propose a simple structure that clarifies how one can infer the relative importance of the two types of citations in generating the observed pattern of localization of patent citations. If citation reversals reflect relatedness, and if knowledge transmission falls with distance, then citation reversals should be less localized than non-reversals.

Consider a pair of patents i and j, where i is the focal patent and j is invented after i. Patent j can build upon patent i with probability π_n if it is near and π_f if it is far. With probability θ_n, patent j can be related to patent i if they are near each other and θ_f if they are far apart. For simplicity, assume a patent that builds upon another will cite it with probability α. Similarly, a patent that is related to another will cite with probability β. The probability that j cites i if they are near is $\alpha\pi_n + \beta\theta_n$.[5] Similarly, the probability of a citation if they are far apart is $\alpha\pi_f + \beta\theta_f$. The difference in probability of citation between patents located near each other and patents located far is $\alpha(\pi_n - \pi_f) + \beta(\theta_n - \theta_f)$.[6] This difference has two components: $\pi_n - \pi_f$ representing the extent to which knowledge transmission decreases with

---

[3]A citation implies that the inventor, the patent agent, or the examiner became aware of the cited invention. However, in citation reversals, the citing invention could not have benefited from this knowledge although the citing patent application may have been modified.

[4]The issue is subtle. The inventor of a citing patent may acquire the required knowledge from the inventor of the cited patent. Alternatively, the inventor may learn from the cited patent, or the inventor of the cited patent, knowledge that is not unique to the cited invention. We address many of these issues in greater detail in Section 6, where we analyze the localization patterns in self citations, citations inserted by examiners, and citations made to unpublished patents.

[5]We are assuming, for simplicity, that these are mutually exclusive outcomes. However a citation may reflect both building upon and relatedness. In that case, the probability that j cites i if they are near is $\alpha\pi_n + \beta\theta_n - \alpha\beta\pi_n\theta_n$. Both $\alpha\pi_n$ and $\beta\theta_n$ are very small in magnitude, so that the product term $\alpha\beta\pi_n\theta_n$ can be a neglected.

[6]Existing studies have tried to use patent classes to control for relatedness: It is implicitly or explicitly assumed that two patents in the same class are equally likely to be related independently of whether they are far or near. Formally, for patents in the same patent class, $\theta_n - \theta_f = 0$.

distance, and $\theta_n - \theta_f$ representing the extent to which patents near each other are related relative to those far apart. Each component is weighted by the relevant citation propensity. The consensus is that both components are positive. That is, nearby inventions are more likely to be related, and knowledge transmission is more likely among nearby inventions.

Now consider the probability that i cites j, i.e., a citation reversal. Since i cannot have built upon j, the probability of the patents being related is simply θ_n if they are near each other and θ_f if they are far. If the propensity to cite is β as before, difference in probability of citation between near and far is simply $\beta(\theta_n - \theta_f)$. The effect of proximity on the probability of citation for non-reversals is $\alpha(\pi_n - \pi_f) + \beta(\theta_n - \theta_f)$ so that the difference in the effect of proximity with respect to reversals is $\alpha(\pi_n - \pi_f)$. The extant literature has asserted that knowledge transmission increases with proximity, i.e., $\alpha(\pi_n - \pi_f) > 0$. This implies that the effect of proximity on the probability of a citation reversal should be smaller than the effect of proximity on the probability of a non-reversal citation. In the empirical analysis we use distance, so that we expect that distance should have a smaller absolute effect on reverse citations than on normal or non-reversal citations.

The validity of our test hinges on the assumption that the underlying knowledge described in a patent cannot be changed as the patent moves forward in the examination process. This assumption assures that the priority date is the date when the invention is created and allows us to determine whether knowledge could have been transmitted from one invention to another. This assumption is validated by the two statutory provisions, 35 U.S.C. 132 and 35 U.S.C. 251, which prohibit introduction of new matter in amendments and the application for reissuance. Examiners are obligated to reject new matters introduced into the abstract, specification, or drawings of a patent application once it is filed. The only changes permitted are rephrasing of a passage without changes in meaning and fixing obvious errors whose correction can be foreseen by a person skilled in the given art.[7] If so, the transmission of knowledge cannot take place in citation reversals.

---

[7] The literature on submarine patents discusses how inventors can keep their inventions secret for an extended period of time and change claims using continuation applications (Graham and Mowery, 2004; Reitzig et al., 2007). The change described by the literature pertains to claims rather than inventions and is thus consistent with our assumption that the underlying invention does not change.

4 Data

4.1 Sample

Our main sample is from the 2014 version of EPO Worldwide Patent Statistical Database (PatStat) with inventor distances extracted from Google Maps. We follow previous literature and limit our dataset to USPTO patents with inventors residing in the contiguous United States. Thus, we exclude Hawaii, Alaska, and offshore U.S. territories, such as Puerto Rico and Guam. The publication years of the citing patents range from 2001 to 2014. (Sample years are based on the availability of examiner citations, which as we later explain are an important part of our analysis.) Whenever multiple inventors are listed for a single patent, we take the city-state combination that occurs most frequently in our distance calculation. If there is an equal number of different city-state combinations, we randomly choose a location from them. Finally, we retain only the citations with a priority lag (difference in years between the priority dates of citing and cited patents) less than five years. This restriction is to account for the fact that citation reversals have short priority lags and in turn to ensure that reversals and non-reversals remain comparable. This procedure yields 1,356,738 actual citations.

For each pair of patents connected by a citation, we match the citing patent with another patent, to create a control pair. We construct the control group by matching each citing patent in the actual citations with a randomly selected, non-citing patent on four-digit IPC and publication year (Jaffe et al., 1993; Belenzon and Schankerman, 2013). These control patents are paired with the cited patents from the actual citations to form control citations. The sample includes 2,713,476 observations consisting of both actual and control citations.

4.2 Variable definitions

We proceed to describe the main variables used in the analysis. Online Appendix Table A1 summarizes their definitions and sources.

Priority date. A central piece in our analysis is identifying priority dates, which are used to determine whether a citation is a reversal. Priority date is the earliest application date of the patents that relate to the same underlying invention (also know as a patent family). It is the date when the invention was first recognized by a patent issuing authority to be in existence. A patent can claim as its priority date

the application filing date of an earlier patent if the two patents describe the same invention and share at least one inventor.

Priority dates can be claimed via different types of patent applications. Among the most common are applications taking advantage of a provisional application filed for the same invention; continuing applications that extend a prior application[8]; and applications filed in the United States within twelve months of filing a foreign application for the same invention. In all of these cases, to claim an application filing date of an earlier patent as the priority date, the specifications of the underlying invention described in the earlier and subsequent patents must be the same.

For example, US 7333597 is a patent on technology that enables telephone synchronization with software applications and documents. The inventors on the patent are Edward M. Silver, Linda A. Roberts, and Hong T. Nguyen. The application for this patent was filed on November 2, 2004. However, the priority date of the patent is March 29, 2002, based on an earlier patent, US 6873692, which relates to the same invention and has the same inventors. Because US 7333597 is a continuation of US 6873692 and because the underlying invention and at least one inventor are the same between the two patents, the more recent patent (US 7333597) can claim as priority date the application filing date of the earlier patent (US 6873692). Identifying an earlier patent application from which to take the priority date is not always trivial because an invention may be associated with multiple patent applications with different application filing dates. For this particular example, we went to the USPTO's patent application information retrieval website to look up all of the patents in the patent family of US 7333597 and looked for the patent application with the earliest filing date. (Onlnie Data Appendix describes how we find the priority dates of the patents in our sample using PatStat.)

In our sample, 61 percent of citing patents and 35 percent of cited patents claim priority dates from an earlier patent application. Of the 61 percent of citing patents, 59 percent claim priority based on a

---

[8]Continuing applications can be further broken down into continuation, divisional, and continuation-in-part. Continuation applications make additional claims based on an existing invention specified in an earlier patent application while divisional applications are filed to separate out distinct inventions from an earlier application usually because the earlier one fails to meet the "unity of invention"requirement. Continuation-in-part (CIP) applications can add extensions to an earlier invention, with claims on new subject matter taking as their priority date the application filing date of the CIP application. The prospect of adding extensions to an underlying invention is concerning since our test relies on the assumption that the underlying invention does not change over time. In Section 6, we run robustness tests after excluding CIP applications and applications claiming priority dates from a provisional application - two sources of potential changes in underlying inventions. Our main finding remains robust to these exclusions.

provisional application, 40 percent on a continuing application, and 1 percent on a foreign application. Of the 35 percent of cited patents, 59 percent claim priority based on a provisional application, 39 percent on a continuing application, and 2 percent on a foreign application. 39 percent of citing patents and 65 percent of cited patents that do not claim priority dates from an earlier application (Online Appendix Figure A1 presents a breakdown of applications based on the types of priorities they claim).

Citation reversal. We construct citation reversals for which transmission of knowledge from the cited to the citing patent is unlikely. The primary type of citation reversals occurs when the priority year of the citing patent is earlier than the priority year of the cited patent. We refer to this type of citations as "priority reversals". Priority reversals are not likely to reflect knowledge transmission because the cited invention does not exist at least until the citing invention is created. (We explore a secondary type called "disclosure citations" in Table 5 of the robustness section.) Priority reversals occur in about four percent (58,983) of all actual citations in our sample.

Priority reversals occur because priority dates are not always transparent to applicants without an extensive search. Although examiners are supposed to review applicant citations for accuracy and appropriateness, the examination process is not fully automated and thus a related patent may be incorrectly cited as prior art when it is not. According to our conversations with patent attorneys and patent examiners, priority reversals can arise when patent applicants include citations on the information disclosure statement (IDS) or when patent examiners add additional citations during the patent examination process that begins after the patent application is filed. Some patents have a priority date earlier than their filing date, for instance, because they are related to a foreign patent or they are continuations of an earlier patent. These types of patents require more effort in identifying priority dates, and thus applications involving them are more prone to having priority reversals.

An example of a priority reversal is patent US 8484303 citing patent US 8073918 (a graphical illustration and transaction highlights are also shown in Online Appendix Figure A4). Patent US 8484303 was filed on September 20, 2011, and US 8073918 was filed on August 6, 2010. However, the invention described in US 8484303 was created on February 17, 2000 while the invention described in patent US 8073918 was created on April 21, 2004. Thus, the priority year of the citing patent is prior to the priority year of the cited patent.

We illustrate priority reversals using the timeline in Figure 1. The horizontal line going from left to right represents the progression of time, and the vertical lines extending below the timeline represent the priority date and the earliest publication date of a cited patent. The two vertical lines extending above the timeline represent possible instances of the citing patent's priority date. A priority reversal occurs when an invention references another invention that has not been created, as demonstrated by the first instance of the citing patent's priority date coming before the priority date of the cited patent. A non-reversal occurs when an invention references another invention that has been created and disclosed to the public, as demonstrated by the second instance of the citing patent's priority date falling after the earliest publication date of the cited patent. In this case, knowledge is more likely to have been transmitted from the cited to the citing invention.

Insert Figure 1 here

Geographical distance. We calculate distances between citation pairs using PatStat's inventor addresses dataset and Google Maps API. We first extract inventor city and state information from PatStat's inventor addresses dataset and then use a custom software application that communicates with Google Maps Geocoding API to obtain geographical coordinates and straight line distances between inventors of citing and cited patents.[9] In addition to continuous distance, we construct dummy variables for distance ranges to examine the nonlinear effect of distance on citation probability. The reference range is 0–25 miles, and the rest of the distance ranges are as follows: 25–50, 50–100, 100–150, 150–250, 250–500, 500–1000, 1000–1500, 1500–2500, and greater than 2500.

5   Non-parametric Evidence

Table 2 presents summary statistics for the main variables used in the analysis. The average distance between patents linked by a citation is 950 miles with a standard deviation of 892 miles. 4 percent of the citations are priority reversals, and 33 percent of the citations are added by examiners. On average, a citing patent receives 6 and cited patents receive 48 forward citations. Of the citations, about 14 percent

---

[9]In cases with multiple inventor locations for a single patent, we use the city-state combination that occurs most frequently. If there is an equal number of different city-state combinations, we randomly choose a location among them.

are coast-to-coast citations, and thus we add dummies to control for citations between research clusters that are located in the opposite coasts.

Table 3 presents the mean comparisons of the main variables used in the analysis for citation reversals and non-reversals. The comparison of geographical distances shows that reversals are at least as localized as non-reversals. The average distance between citation pairs is 860 miles for reversals and 954 miles for non-reversals. The share of citations with citing patents that are within 50 miles of the cited patent is 33 percent for reversals and 25 percent for non-reversals. These findings are inconsistent with the view that localization of citations is driven by localized knowledge transmission.

Figure 2 presents comparisons of average distances between citing and cited inventors across various types of citations as well as the share of citations with citing inventors within 50 miles of cited inventors. The general pattern shows that the average distance is greater for non-reversals than for reversals. Furthermore, the fraction of citing inventors within 50 miles of the cited inventors is greater for reversals than for non-reversals. These results provide further evidence inconsistent with the notion that localization of citations reflect localized knowledge transmission.

Insert Table 2, Table 3, and Figure 2 here

6   Econometric analysis

Our empirical analysis tests the implication of the model from Section 3.2. We examine whether the effect of distance on citation probability is smaller in magnitude for reversals than for non-reversals. As shown in the model, this relationship will hold if localization of citations is driven by localized knowledge transmission.

We follow Jaffe et al. (1993) and match each citing patent with a control, non-citing, patent with the same four-digit IPC code and publication year. Our results are robust to matching at the six-digit IPC code and publication year, though a finer matching naturally reduces the sample size (cf. Online Appendix Table A5). We use a linear probability model to estimate the effect of distance on citation probability for non-reversals and reversals. Our main empirical specification is as follows:

$$Pr(C_{ij} = 1) = \beta_1 \ln D_{ij} + \beta_2 \ln D_{ij} \times reversal_{ij} + \beta_3 reversal_{ij} + \mathbf{Z}'\gamma + \eta_j + \epsilon_{ij}$$

Where i and j denote citing and cited patents, respectively, C_{ij} is a dummy variable that receives the value of one for an actual citation and zero for a control (non-) citation, D_{ij} is the distance in miles between the location of citing and cited inventors, and $Reversal_{ij}$ is a dummy variable that receives the value of one for a citation reversal (and for the matched control non-citing patent). $\mathbf{Z}$ is a vector of dyadic dummies indicating citations between leading research clusters (i.e., Austin, TX; Route 128, MA; Raleigh-Durham, NC; San Diego, CA; and Silicon Valley, CA). These dyadic research cluster dummies are important because patents produced in clusters specializing in similar inventions are likely to cite one another and the clusters are often located on opposite coasts. The stochastic components are represented by η_j, a cited patent fixed effect, and an *iid* error term ϵ_{ij}. Standard errors are always clustered at the cited patent level.

If localization of citations is driven by localized transmission of knowledge, we expect the effect of distance on citation probability to be larger in magnitude for non-reversals than for reversals. Thus, we expect $\widehat{\beta}_1 < 0$ to confirm previous evidence on localized citations, and $\widehat{\beta}_2 > 0$ to support the view that citations with potential knowledge transmission (non-reversal) are more localized than citations where knowledge transmission is unlikely (reversal).

6.1 Results

6.1.1 Reversal *vs.* non-reversal localization effect

Table 4 presents the results from our main test for localized transmission of knowledge in patent citations. Column 1 presents the estimation results for the effect of distance on the probability of citation. Consistent with previous findings in the literature, the results show that patent citations are localized. The coefficient estimate on the distance between citing and cited patents is negative and statistically significant, indicating that two inventors who are geographically close to each other are more likely to cite than inventors who are far away from each other. Column 2 explores the effect of distance on citation probability using distance dummies. The reference distance is 0–25 miles. Based on our estimates, moving from 0 to 50 miles between inventors lowers the probability of citation by about 19 percentage points, or close to 40 percent of the sample average.

Columns 3–5 present our key findings from comparing the estimated effect of distance on citation

probability between non-reversals and reversals. If localization of citations reflects localized transmission of knowledge, then we would observe a significantly larger effect of distance on citation probability for non-reversals than for reversals. Column 3 reveals that the effect of distance on citation probability is -0.08 for non-reversals and -0.10 for reversals, indicating that citation reversals are at least as localized as citation non-reversals. Columns 4–5 also show no difference in the effect of distance on citation probability for subsamples consisting separately of non-reversals and reversals.

Columns 6–7 explore the robustness of the results by allowing for non-linear distance effects. The same pattern of results emerges. For example, moving from 0 to 50 miles reduces the citation probability by 18.2 percentage points for non-reversals and 22.5 percentage points for reversals. These findings are inconsistent with the interpretation that localization of citations reflects localized knowledge transmission since citations do not become more localized as transmission of knowledge becomes more likely. They are consistent with the view that local inventors tend to work on similar technical problems, but not necessarily disproportionately learn from one another.

<div align="center">**Insert Table 4 here**</div>

6.1.2 Do inventor interactions drive localized reversals?

An important concern about our analysis is that citation reversals might be driven by interactions among local inventors and/or patent intermediaries. If citation reversals are driven by inventors sharing knowledge about new inventions before the inventions are publicly disclosed, then reversals would reflect highly localized knowledge transmission. For instance, it is possible that two local inventors work on related ideas and that knowledge is transmitted from one invention to the other. If, however, the invention building on the other invention is filed for a patent before the invention being built on and a citation is made from the former to the latter, then this priority reversal would capture knowledge transmission.[10] In such cases, reversals could not be used as a benchmark for the localization of non-learning citations against which citation non-reversals are compared. This section presents several tests to mitigate this concern.

---

[10]Under "Patent citation lag" in Section 6.1.3, we perform tests targeted specifically to address this timing issue. Given that the reversals in patent application filing time would be more likely to occur when inventions are created close in time, we test whether our main finding still holds if we re-categorize reversals with up to one-year priority lag as non-reversals and at the same time use different citation lags.

Self citations If local inventors interact with one another to share knowledge about unpublished inventions, such interactions are more likely to occur within firms than across firms. Thus, if citation reversals were driven by highly localized inventor interactions, self citations would be more prevalent within reversals than within non-reversals and more localized than external citations. This bias will prevent us from rejecting the null hypothesis that the localization effect of non-reversals is the same as that of reversals.

The share of self citations is 28 percent for reversals and 19 percent for non-reversals. Self citations are also more localized than external citations. The share of citations whose citing patent is within 50 miles of the cited patent is 75 percent for self citations and 14 percent for external citations. (These differences are statistically significant at the 1% level.) Within citation reversals, self citations are also more localized than external citations (80% of citing patents being with 50 miles of cited patents for self citations relative to 15% for external citation reversals). These findings are consistent with the concern that reversals might be driven by highly localized knowledge transmission.

To mitigate the potential bias caused by self citations, we exclude them from our sample. Columns 1 and 2 of Table 5 present the estimation results. The results show that the effect of distance on citation probability is still quite similar between citation non-reversals (-0.045) and reversals (-0.035) and thus confirm our main finding that citation reversals are as localized as non-reversals.

Patent attorneys Patent attorneys are another potential source of highly localized knowledge transmission that might generate citation reversals. Wagner et al. (2014) argue that patent attorneys tend to cite known patents from their knowledge repositories which they develop based on their interactions with their clients. Thus, if local inventors engage with the same attorneys who share knowledge about unpublished local inventions, then such interactions might result in citation reversals that reflect highly localized knowledge transmission. In this case, same-attorney citations (i.e. citing and cited patents are prepared by the same attorney) would be more prevalent within citation reversals than within non-reversals and more localized than different-attorney citations (i.e. citing and cited patents are prepared by different attorneys). This bias will prevent us from rejecting the null hypothesis that the localization effect of non-reversals is the same as that of reversals.

To perform this analysis, we extracted attorney information from the weekly compilations of patent

publications released by the USPTO for years 2001 through 2014. We standardized attorney names and removed any corporate legal offices. In our sample, the share of same-attorney citations is 5 percent for non-reversals and 11 percent for reversals. Same-attorney citations are also more localized than different-attorney citations. The share of citations that take place within 50 miles of the cited patent is 79 percent for same-attorney citations and 23 percent for different-attorney citations (statistically significant at the 1% level). These findings are consistent the concern that reversals might be driven by highly localized knowledge transmission.

To mitigate this concern, we exclude from our sample the same-attorney citations. Columns 1 and 3 in Table 5 show that even after excluding same-attorney citations, the effect of distance on citation probability is similar between citation non-reversals (-0.045) and reversals (-0.031). This finding is consistent with our main finding and mitigates the concern that citation reversals capture highly localized knowledge transmission driven by interactions among inventors and patent attorneys.

Examiner reversals To further mitigate the concern that citation reversals might capture highly localized knowledge transmission, we compare the localization of citation non-reversals to that of citation reversals added by patent examiners, which arguably are even less prone to biases that could arise from local inventor interactions. If reversals arise due to local interactions among inventors, we expect inventors to generate proportionally more reversals than examiners and that inventor reversals would be more localized than examiner reversals.

Despite the concern, the share of reversals for inventor and examiner citations in our sample are quite similar (4.4 percent for inventor citations and 4.2 percent for examiner citations). The share of citing inventors who are within 50 miles of the cited inventors is 33.7 percent for inventor reversals and 31.1 percent for examiner reversals, indicating that inventor-added reversals are somewhat more localized than examiner-added reversals. This observation is consistent with the possibility that reversals are driven by inventor interactions. To further explore this concern, we test whether citation non-reversals are more localized than examiner reversals in columns 1 and 4 of Table 5. The estimation results show that examiner reversals (-0.043) are as localized as non-reversals (-0.045), evidence that localization of patent citations is not likely to be driven by localized knowledge transmission.

Disclosure citations If citation reversals capture highly localized inventor interactions, we expect the highest degree of localization to exist for citations made to inventions that have been created but have not been published. To test this hypothesis, we introduce into our sample another citation type, "disclosure citation", which occurs when the priority year of the citing patent comes after the priority year but before the earliest publication year of the cited patent. This time sequence is due to overlapping patent examination periods where a prior-art patent is published while the citing patent application is being examined. (Online Appendix Figure A3 demonstrates disclosure citation on a timeline, and Online Appendix Figure A5 provides an example of a disclosure citation.) Because the cited patent was not known to the public at least until the citing inventor applied for a patent on his invention, the only way that the citing inventor would have known about the cited invention is through a local interaction with the cited inventor. Thus, if local inventors share knowledge about their inventions before the inventions are publicly disclosed, then we expect disclosure citations to be substantially more localized than both citation non-reversals and priority reversals.

Columns 1, 2 and 5 in Table 5 compare the localization of citations across citation non-reversals, priority reversals, and disclosure citations. The results show that disclosure citations (-0.044) are essentially as localized as priority reversals (-0.035) and non-reversals (-0.045). In Online Appendix Table A8, we further present non-linear effect of distance on the citation probability for citation non-reversals, priority reversals, and disclosure citations. The results continue to show that the effect of distance on citation probability is quite similar across non-reversals, priority reversals, and disclosure citations. These results are inconsistent with the notion that highly localized knowledge transmission via inventor interaction is a major concern.

Overall, the results in Table 5 show that, although citation reversals are more frequent for citation pairs filed by the same organization and for citation pairs with a common patent attorney, excluding them from the sample does not change the broad conclusion that localization of patent citations is unlikely to reflect transmission of knowledge from the cited to the citing invention. Additionally, focusing only on citation reversals added by examiners yields a similar effect of distance on citation propensity between reversals and non-reversals. Lastly, the results show that disclosure citations are not particularly more localized than priority reversals and non-reversals, evidence inconsistent with the concern that highly

localized knowledge transmission from local inventor interactions is responsible for citation reversals.

<center>**Insert Table 5 here**</center>

6.1.3 Other Robustness Checks

Changes to the underlying invention over time Our main test of comparing the localization of citation non-reversals with reversals relies on the assumption that the underlying invention does not change over time during the examination process. There are some instances where this assumption might be violated. They include filing continuation-in-part (CIP) applications and applications extending provisional applications. Thus, we test whether our main finding is robust to the exclusion of patents issued from these two types of applications.

CIP applications are filed when new subject matter is added to the original application for an invention, and thus the original material and the newly-added material may have different priority dates. The multiple priority dates might cause mis-categorization of reversals because it is difficult to determine whether a citation is made to (or from) the old or the new material. To mitigate this concern, we examine whether our main findings hold with the CIP applications excluded from our sample. Columns 1 and 3 in Online Appendix Table A2 show that, even with CIP applications excluded, the localization effects are similar between non-reversals (-0.044) and reversals (-0.035).

Another way by which the subject matter of an invention could change over the course of the patent application process involves provisional applications. Unlike regular patent applications, a provisional application can be filed with an incomplete invention, which can later be supplemented using a non-provisional application once the invention is completed. To mitigate the concern that the underlying invention could change when a non-provisional application is filed to supplement a provisional application, we test whether our main findings are robust to the exclusion of applications related to provisional applications. Columns 2 and 4 in Online Appendix Table A2 show that the extent of localization is similar between citation non-reversals (-0.042) and reversals (-0.029) even after excluding applications relating to provisional applications.

<center>21</center>

Alternative specifications We address the concern that the use of control patents might miss an important variation that is correlated with distance. Online Appendix Table A3 presents results from an alternative specification in which localization of citation reversals are compared directly with that of non-reversals without the control citations. For each citation reversal, we find a non-reversal with the same cited patent and with the citing patent in the same IPC and publication cohort. To control for potential differences across citing patents, we also add citing patent technology area and publication year fixed effects.

Columns 1–3 show that a citing patent in citation reversals is at least as likely to be within fifty mile radius of the cited patent as it is in citation non-reversals. The results are robust to using twenty-five mile radius (columns 4–6). Thus, our main finding, that localization of citations is not likely to be driven by localized knowledge transmission, continues to hold.

Extended sample period For our main results, we used a sample that contains citations with citing patents published over years 2001 through 2014 because examiner citations became identifiable only in 2001. Online Appendix Table A4 presents results from a test that uses a sample whose citing patents cover years 1977 through 2014 to make sure that our results are not biased by factors inherent to more recent citations. The results from this larger sample are consistent with our main finding and provide additional support that localization of citations is not likely to reflect localized knowledge transmission. For instance, columns 2 and 3 show that going from 0 to 50 miles reduces the citation probability by 16.8 percentage points for non-reversals and 21.8 percentage points for reversals.

Six-digit IPC We further examine robustness of our findings by replicating our test with a sample whose controls are matched on six-digit technology classification codes. This test addresses the concern raised by Thompson and Fox-Kean (2005) that matching on broad technology classification code might not be adequate to control for existing regional specialization.

Online Appendix Table A5 presents the results from the sample matched on six-digit technology classification code. As shown in column 1, patent citations are localized, consistent with the findings in Murata et al. (2014). Also, consistent with our main finding, the results show that citation reversals are at least as localized as non-reversals. For instance, columns 2–3 show that going from 0 to 50 miles reduces

citation probability by 18 percentage points for non-reversals and 21 percentage points for reversals.[11] These results are consistent with our main finding that, while citations are localized as indicated by prior studies, localization of citations is not likely to be driven by localized knowledge transmission.

Patent citation lag   When two inventions are generated during the same time period, it is possible that the patent application of the invention building on the other invention gets filed before the patent application of the invention being built on. Under this scenario, citation reversals with short citation lags would reflect knowledge transmission and our test would produce spurious results. However, the possibility of reversals being generated by this process should be reduced as the reversal lag (i.e. difference between the priority years of citing and cited patents in citation reversals) widens. We address this concern by testing our results after re-categorizing reversals with up to one-year priority lag as non-reversals and performing separate tests for reversals with different reversal lags.

Online Appendix Table A6 presents results from the tests performed after re-categorizing reversals with up to one-year priority lag as non-reversals. The pattern of results is consistent with our main finding and thus mitigates the concern that our main results might be driven by citation reversals generated by reversed timing of patent application filings.

Technology areas   The importance of inventor interactions for learning from invention is likely to vary across technology areas. For example, inventions in complex technology areas such as telecommunications are likely to require more tacit knowledge than those in discrete technology areas such as chemicals. Such tacit knowledge might be more localized. To test whether reversals are driven by localized inventor interactions, we first compare the share of reversals across six technology areas: Chemistry, Pharmaceutical, Biotechnology, Medical Technology, Computer Technology, and Telecommunications. If technology areas characterized by tacit knowledge are more prone to inventor interactions, then we expect to see higher shares of citation reversals for complex technology areas than for discrete technology areas. Inconsistent with the view that reversals are driven by transmission of tacit knowledge, we find that the share of

---

[11]For the sample matched on six-digit technology classification code, the average distance between citing and cited patents is 977 miles for citation non-reversals and 904 miles for reversals, with the difference statistically significant at the 1% level. Furthermore, the share of citing patents within 50 miles of cited patents is higher for citation reversals at 30 percent than for non-reversals at 24 percent.

reversals are fairly consistent across the six technology areas, ranging from 4 (Pharmaceuticals) percent to 10 percent (Medical Technology).

We also examine whether the difference between the localization effect for non-reversals and reversals varies by technology area. If reversals in complex technology areas are more localized because inventor interactions are more important for learning, we expect our test to bias against finding a difference in the localization effect between non-reversals and reversals in complex technology areas, but not in discrete technology areas.

Online Appendix Tables A7a and A7b present the results from comparing localization of citation non-reversals and that of reversals. Inconsistent with the expectation, the overall pattern does not show any systematic variation according to our expectation. For instance, going from 0 to 50 miles reduces the citation probability by 22 percentage points for non-reversals (column 5 of Online Appendix Table A7a) and 51 percentage points for reversals (column 6 of Online Appendix Table A7a) in Biotechnology and by 21 percentage points for non-reversals (column 1 of Online Appendix Table A7b) and 23 percentage points for reversals (column 2 of Online Appendix Table A7b) in Medical Technology. These results provide evidence that localization of citations is not likely to be driven by localized knowledge transmission in complex or discrete technology areas.

7 Concluding Remarks

This study examines whether localization of patent citations is driven by localized knowledge transmission. Our results show that the effect of distance on citation probability is similar for citation non-reversals and reversals, implying that localized knowledge transmission from the cited to the citing invention is not likely to be a major driver of localized citations. The concern that citation reversals might reflect highly localized knowledge transmission between inventors, either through direct interactions among themselves or through intermediaries, is addressed by comparing localization of non-reversals with various subsets of reversals that are even less likely to be driven by localized knowledge transmission.

Our findings imply that either patent citations do not measure knowledge transmission, or that knowledge transmission is not localized, or both. Our findings are consistent with the growing evidence pointing to the inadequacy of using patent citations as a measure of knowledge flows. Patents may cite other patents

because they are solving common problems or drawing upon similar techniques. It may well be that the bulk of citations arise from such circumstances rather than from knowledge flows between inventions linked by citations. If so, we need better ways to track such knowledge flows. Fortunately, the growth in computing power and machine learning methods offer new possibilities. For instance, measuring textual similarity between patents is be a promising way to infer overlap between patents, and perhaps a way to infer knowledge transmission. Other methods include in-text citations to patents instead of relying upon front page citations (Bryan and Ozcan, 2017).

We are agnostic about the importance of knowledge spillovers across space. Our findings do point to the potentially important role that organizational links play in knowledge transmission, and the important interactions between geographical and organizational proximity. Further, the role of specialized intermediaries, such as patent agents, and of the providers of specialized technical inputs, such R&D and engineering services, may be promising avenues for future research.

Methodologically, this study provides a way to isolate citations that are unlikely to be associated with knowledge transmission. We identify these special citations by checking whether the citing patent's priority date comes before the priority date of the cited patent. We use these citations to benchmark localization of knowledge transmission as reflected in patent citations.

Patent citations could reflect direct transmission of knowledge from the citing to the cited invention. However, it could also reflect commonalities in domain or solution to a specific technical problem. That is, the patents linked by citation may be linked not by knowledge transmission but by a body of knowledge that is common ground for both inventions.

The distinction between an invention building on the cited invention and inventions drawing upon common, or background knowledge is important for both policy and firms. An inventor whose invention is built upon by another inventor might extract licensing revenues from the latter. However, if the subsequent inventor draws upon the background knowledge, then it would be difficult even to identify inventions that build on such knowledge.

The distinction also provides insights into entrepreneurial spin-offs and regional clusters. If founders of spin-offs draw knowledge from specific discoveries during their employment at the parent company, then employers can develop contracts to prevent loss of rents. However, if knowledge drawn by the

former employees is more general and cannot specifically be identified, then it would be more difficult for employers to protect themselves from former employees utilizing the knowledge. Furthermore, it would be difficult for firms to prevent background knowledge from spilling over to competitors even if they try to disperse inventors or R&D operations across different geographies. Indeed, to the extent that the background knowledge is useful for invention, such dispersion may be counter-productive.

In summary, our findings force an uncomfortable choice between two very appealing and widely accepted beliefs, namely that knowledge transmission is less likely over longer distances and that patent citation is a good measure of knowledge transmission. Our empirical setup does not allow us to weigh in on this tradeoff. We look forward to future research for help.

References

[1] Agrawal, A., Cockburn, I., Galasso, A., Oettl, A., 2014. Why are some regions more innovative than others? The role of small firms in the presence of large labs. *Journal of Urban Economics* 81: 149–165.

[2] Alcácer, J., Chung, W., 2007. Location strategies and knowledge spillovers. *Management Science* 53(5): 760–776.

[3] Alcácer, J., Gittelman, M., 2006. Patent citations as a measure of knowledge flows: The influence of examiner citations. *The Review of Economics and Statistics* 88(4): 774–779.

[4] Alcácer, J., Zhao, M., 2012. Local R&D strategies and multilocation firms: The role of internal linkages. *Management Science* 58(4): 734–753.

[5] Almeida, P., Kogut, B., 1999. Localization of knowledge and the mobility of engineers in regional networks. *Management Science* 45(7): 905–917.

[6] Anton, J.J., Yao, D.A., 1995. Start-ups, spin-offs, and internal projects. *Journal of Law, Economics, & Organization* 11(2): 362–378.

[7] Audretsch, D.B., 1998. Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy* 14(2): 18–29.

[8] Audretsch, D.B., Feldman, M., 1996. R&D spillovers and the geography of innovation and production. *The American Economic Review* 86(3): 630–640.

[9] Belenzon, S., 2012. Cumulative Innovation and Market Value: Evidence from Patent Citations. *Economic Journal* 559(122): 265-285.

[10] Belenzon, S., Schankerman, M., 2013. Spreading the word: Geography, policy, and knowledge spillovers. *The Review of Economics and Statistics* 95(3): 884–903.

[11] Bessen, J., Maskin E., 2009. Sequential innovation, patents, and imitation. *The RAND Journal of Economics* 40(4): 611–635.

[12] Breschi, S., Lissoni, F., 2001. Knowledge spillovers and local innovation systems: A critical survey. *Industrial and Corporate Change* 10(4): 975–1005.

[13] Breschi, S., Lissoni, F., 2009. Mobility of skilled workers and co-invention networks: An anatomy of localized knowledge flows. *Journal of Economic Geography* 9(4): 439–468.

[14] Chatterji, A.K., 2009. Spawned with a silver spoon? Entrepreneurial performance and innovation in the medical device industry. *Strategic Management Journal* 30(2): 185–206.

[15] Delgado, M., Porter, M.E., Stern, S., 2008. Convergence, clusters, and economic performance. *Institute for Strategy and Competitiveness.*

[16] Delgado, M., Porter, M.E., Stern, S., 2014. Clusters, convergence, and economic performance. *Research Policy* 43(10): 1785–1799.

[17] Duguet, E., MacGarvie, M., 2005. How well do patent citations measure flow of technology? Evidence from French innovation surveys. *Economics of Innovation and New Technology* 14(5): 375–393.

[18] Galasso, A., Schankerman, M., 2014. Patents and cumulative innovation: Causal evidence from the courts. *The Quarterly Journal of Economics* 130(1): 317—369.

[19] Graham, S.J.H., Mowery, D.C., 2004. Submarines in software? Continuations in U.S. software patenting in the 1980s and 1990s. *Economics of Innovation and New Technology* 13(5): 443–456.

[20] Green, J.R., Scotchmer, S., 1995. On the division of profit in sequential innovation. *The RAND Journal of Economics* 26(1): 20–33.

[21] Giuliani, E., 2007. The selective nature of knowledge networks in clusters: Evidence from the wine industry. *Journal of Economic Geography* 7(2): 139–168.

[22] Henderson, R., Jaffe, A.B., Trajtenberg, M., 2005. Patent citations and the geography of knowledge spillovers: A reassessment: Comment. *The American Economic Review* 95(1): 461–464.

[23] Jaffe, A.B., Trajtenberg, M., Fogarty, M.S., 2000. The meaning of patent citations: Report on the NBER/Case-Western Reserve Survey of Patentees (No. w7631). National Bureau of Economic Research.

[24] Jaffe, A.B., Trajtenberg, M., Henderson, R., 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics* 108: 577–598.

[25] Krugman, P., 1991. *Geography and Trade.* MIT Press, Cambridge, M.A.

[26] Laitner, J., Stolyarov, D., 2013. Derivative ideas and the value of intangible assets. *International Economic Review* 54(1): 59–95.

[27] Martin, R., Sunley, P., 2003. Deconstructing clusters: chaotic concept or policy panacea? *Journal of Economic Geography* 3(1): 5–35.

[28] Menon, C., 2015. Spreading big ideas? The effect of top inventing companies on local inventors. *Journal of Economic Geography* 15(4): 743–768.

[29] Moser, P., Ohmstedt, J., Rhode, P., 2017. Patent citations — An analysis of quality differences and citing practices in hybrid corn. *Management Science*.

[30] Murata, Y., Nakajima, R., Okamoto, R., Tamura, R., 2014. Localized knowledge spillovers and patent citations: A distance-based approach. *The Review of Economics and Statistics* 96(5): 967–985.

[31] Nomaler, O., Verspagen, B., 2016. River deep, mountain high: of long run knowledge trajectories within and between innovation clusters. *Journal of Economic Geography* 16(6): 1259–1278.

[32] O'Donoghue, T., Scotchmer, S., Thisse, J., 1998. Patent breadth, patent life, and the pace of technological progress. *Journal of Economics & Management Strategy* 40(4): 611–635.

[33] Porter, M.E., 1990. *The Competitive Advantage of Nations.* Free Press, New York, N.Y.

[34] Porter, M.E., 2003. The economic performance of regions. *Regional Studies* 37: 549–578.

[35] Porter, M.E., 2007. Clusters and economic policy: Aligning public policy with the new economics of competition. White Paper. Institute for Strategy and Competitiveness, Harvard Business School.

[36] Reitzig, M., Henkel, J., Heath, C., 2007. On sharks, trolls, and their patent prey: Unrealistic damage awards and firms' strategies of "being infringed." *Research Policy* 36(1): 134–154.

[37] Roach, M., Cohen, W.M., 2013. Lens or prism? Patent citations as a measure of knowledge flows from public research. *Management Science* 59(2): 504–525.

[38] Rosenthal, S.S., Strange, W.C., 2001. The determinants of agglomeration. *Journal of Urban Economics* 50(2): 191–229.

[39] Saxenian, A., 1994. *Regional Advantage.* Harvard University Press, Cambridge, M.A.

[40] Singh, J., Marx, M., 2013. Geographic constraints on knowledge spillovers: Political borders vs. spatial proximity. *Management Science* 59(9): 2056–2078.

[41] Shaver, J. M., Flyer, F., 2000. Agglomeration economies, firm heterogeneity, and foreign direct investment in the United States. *Strategic Management Journal* 21(12): 1175–1193.

[42] Suire, R., Vicente, J., 2009. Why do some places succeed when others decline? A social interaction model of cluster viability. *Journal of Economic Geography* 9(3): 381–404.

[43] Swann, P., 1998. Clusters in the US computing industry. In: Swann, P., Prevezer, M., Stout, D. (Eds.), *The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology.* Oxford University Press, Oxford, pp. 76–105.

[44] Thompson, P., 2006. Patent citations and the geography of knowledge spillovers: Evidence from inventor- and examiner-added citations. *The Review of Economics and Statistics* 88(2): 383–388.

[45] Thompson, P., Fox-Kean, M., 2005. Patent citations and the geography of knowledge spillovers: A reassessment. *American Economic Review*: 450–460.

[46] Verspagen B., Schoenmakers W., 2004. The spatial dimension of patenting by multinational firms in Europe. *Journal of Economic Geography* 4(1): 23–42.

[47] Wagner S., Hoisl, K., Thoma G., 2014. Overcoming localization of knowledge: The role of professional service firms. *Strategic Management Journal* 35(11): 1671–1688.

[48] Williams, H.L., 2013. Intellectual property rights and innovation: Evidence from the human genome. *Journal of Political Economy* 121(1): 1–27.

[49] Zhao, M., 2006. Conducting R&D in countries with weak intellectual property rights protection. *Management Science* 52(8): 1185–1199.

8 Online Data Appendix

8.1 Priority year

Priority reversals are determined by examining the priority years of the citing and cited patents, and disclosure citations are determined by examining the priority year of the citing patent and the earliest publication year of the cited patent. We use the priority and continuation dataset in PatStat to determine priority years. We take each of the citing and cited patents and find other patents linked to it via priority established by the Paris Convention of 1883 (TLS 204), provisional applications (TLS 201), and continuing applications (TLS 216). From this patent set, we keep the patent with the earliest application filing year as the priority year.

8.2 Earliest publication year

The earliest publication year is the year in which the first patent from a patent family is publicly disclosed, indicating the first time an invention becomes known to the public. Since the enactment of American Inventors Protection Act of 1999, a patent application or a patent is published usually eighteen months from its earliest filing date, i.e., priority date, unless the applicant requests to have it disclosed earlier or to delay its disclosure until patent issuance. (Online Appendix Figure A2 presents the distribution of patents over number of months it takes for them to be published.)

The earliest publication year of an invention is determined by identifying the earliest publication year from all of the patents within a patent family. The European Patent Office recognizes two types of patent families, simple and extended. We use the simple patent family notion, which groups patents claiming exactly the same set of priorities into a family. We construct this variable by linking the patent family dataset (TLS 218) with application (TLS 201), publication (TLS 211), priority (TLS 204), and continuation (TLS 216) datasets in PatStat.

8.3 Application year

Application year is the year in which an application for a patent is filed. Application year is different from the priority year as it may not indicate the year in which the invention is created, but merely the year in which an application for a patent is filed. We use the applications dataset (TLS 201) in PatStat to find application years of all citing, cited, and control patents.

8.4 Publication year

Publication year is the year in which an application or a patent is disclosed to the public. Publication year is different from the earliest publication year as it may not indicate the year in which the invention is first disclosed to the public, but merely the year in which a patent application or patent is published. Since November 2000, the USPTO has started publishing patent applications 18 months after the priority date. This means that a patent application claiming an earlier priority date may be published less than 18 months from its application date. On the other hand, an applicant may request publication of its application to be delayed until patent issuance if the applicant does not intend to file for a patent in a

foreign country. We use the publication dataset (TLS 211) in PatStat to find publication years of citing, cited, and control patents.

8.5 Inventor distance

We extract inventor cities and states for the citing and cited patents in our sample by linking several datasets in PatStat (TLS 206, TLS 226, TLS 227). We first identify inventors of each patent in our sample from TLS 227, which is merged with TLS 226 to determine their cities and states. We use a software application developed in-house that communicates with Google Maps Geocode API to find geographic coordinates of each city and state pair and to find a straight line distances between inventors of citation and control patent pairs. An advantage of using Google Maps Geocoding API to obtain geographic coordinates is that the Google API can intelligently handle several misspellings of city or state names. This is beneficial because spelling errors are frequent in the inventor city and state fields.

Table 1. Types and Sources of Knowledge

Source / Type	1 Patent	2 Inventor	3 Intermediary (e.g. Patent Attorney)
Invention	Not Localized	Localized	Localized
Background Knowledge	Not Localized	Localized	Localized

Notes: This figure shows the types and sources of knowledge distinguished in the paper. Knowledge about invention is knowledge directly related to the invention while background knowledge is knowledge that might underlie the invention and can be drawn upon from a common knowledge pool. Knowledge can also have multiple sources that can influence temporal and geographical characteristics of knowledge transmission. The six inner cells indicate whether knowledge tends to be localized given its type and source.

Table 2. Summary Statistics for Main Variables

VARIBALES	No. Obs.	Mean	Std. Dev.	Distribution 10th	50th	90th
Citing priority year	1,356,738	2004	3.3	1999	2004	2008
Cited priority year	1,356,738	2000	3.2	1996	2000	2005
Citations per citing patent	1,356,738	6.4	14.1	0	2	16
Citations per cited patent	1,356,738	48.0	80.6	4	22	115
Priority lag in years	1,356,738	3.2	1.6	1	3	5
Dummy for priority reversal	1,356,738	0.04	0.20	0	0	0
Dummy for examiner citation	1,356,738	0.33	0.47	0	0	1
Dummy for self-citation	1,356,738	0.20	0.40	0	0	1
Distance (mi)	1,356,738	950.4	891.6	6	707	2,414
Dummy for 0 ≤ Distance < 25 miles	293,769	7.9	7.6	0	7	19
Dummy for 25 ≤ Distance < 50 miles	53,786	34.1	6.6	26	33	44
Dummy for 50 ≤ Distance < 100 miles	23,256	73.3	14.4	54	73	94
Dummy for 100 ≤ Distance < 150 miles	21,960	125.1	15.2	104	125	146
Dummy for 150 ≤ Distance < 250 miles	48,873	201.7	28.9	161	202	241
Dummy for 250 ≤ Distance < 500 miles	121,888	370.5	66.3	280	362	466
Dummy for 500 ≤ Distance < 1000 miles	250,952	737.4	144.6	545	715	944
Dummy for 1000 ≤ Distance < 1500 miles	147,507	1270.3	165.0	1,043	1,271	1,473
Dummy for 1500 ≤ Distance < 2500 miles	300,460	2026.6	326.2	1,572	2,057	2,430
Dummy for Distance ≥ 2500 miles	94,287	2593.0	60.3	2,527	2,568	2,682

Notes: This table provides summary statistics for the main variables used in the econometric analysis of the effect of distance on citation probability for the main sample. The sample consists only of actual patent citations. The publication years of citing patents in the sample covers years 2001 through 2014. *Priority lag* is the difference between the priority years of the citing and cited patents. *Dummy for reversal* is a variable that takes 1 if the priority date of the citing patent is earlier than the priority date or the earliest publication date of the cited patent and indicates that knowledge transmission is unlikely. *Dummy for examiner citation* is a variable that takes 1 if a citation was added by a patent examiner. *Dummy for self-citation* is a variable that takes 1 if the citing and cited patents are assigned to the same assignee.

Table 3. Comparisons of Main Citation Characteristics: Non-Reversals *vs*. Reversals

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Difference in means		Non-Reversals			Reversals	
VARIBALES	(3) minus (6)	No. Obs.	Mean	Std. Dev.	Obs.	Mean	Std. Dev.
Distance (mi)	94.21**	1,297,755	954.49	890.89	58,983	860.28	901.55
% of citations within 50 miles	-0.08**	1,297,755	0.25	0.43	58,983	0.33	0.47
Citing priority year	4.05**	1,297,755	2004	3.17	58,983	2000	3.41
Cited priority year	-1.20**	1,297,755	2000	3.18	58,983	2001	3.26
Citations per citing patent	0.75**	1,297,755	6.41	14.06	58,983	5.66	15.02
Citations per cited patent	10.54**	1,297,755	48.43	81.18	58,983	37.89	64.69
Citation lag in years	5.25**	1,297,755	3.39	1.23	58,983	-1.86	1.12
Dummy for examiner citation	0.01**	1,297,755	0.33	0.47	58,983	0.32	0.47
Fraction of Self Citations	-0.09**	1,297,755	0.19	0.40	58,983	0.28	0.45
Fraction of Citations in the Same IPC	-0.05**	1,297,755	0.20	0.40	58,983	0.25	0.43

Notes: This table presents mean comparisons of main variables between non-reversals and reversals. The sample consists of actual patent citations. The publication years of citing patents covers years 2001 through 2014. ** $p<0.01$, * $p<0.05$

Table 4. The Effect of Distance on the Probability of Citation for Citation Non-Reversals *vs*. Reversals

Dependent Variable:	*Dummy for an actual citation*						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
				Distance effect		Nonlinear distance effect	
VARIBALES	All	All	All	Non-reversal	Reversal	Non-reversal	Reversal
Log(*Distance*)	-0.083**		-0.080**	-0.081**	-0.079**		
	(0.000)		(0.000)	(0.000)	(0.002)		
Log(*Distance*) x Dummy for Reversal			-0.020**				
			(0.001)				
Dummy for Reversal			-0.138**				
			(0.006)				
Dummy for 25 ≤ Distance < 50 miles		-0.189**				-0.182**	-0.225**
		(0.003)				(0.003)	(0.023)
Dummy for 50 ≤ Distance < 100 miles		-0.396**				-0.386**	-0.414**
		(0.004)				(0.005)	(0.030)
Dummy for 100 ≤ Distance < 150 miles		-0.439**				-0.424**	-0.495**
		(0.004)				(0.005)	(0.029)
Dummy for 150 ≤ Distance < 250 miles		-0.455**				-0.443**	-0.478**
		(0.003)				(0.003)	(0.022)
Dummy for 250 ≤ Distance < 500 miles		-0.482**				-0.467**	-0.489**
		(0.003)				(0.003)	(0.016)
Dummy for 500 ≤ Distance < 1000 miles		-0.476**				-0.464**	-0.471**
		(0.002)				(0.002)	(0.014)
Dummy for 1000 ≤ Distance < 1500 miles		-0.521**				-0.507**	-0.508**
		(0.002)				(0.003)	(0.015)
Dummy for 1500 ≤ Distance < 2500 miles		-0.516**				-0.501**	-0.510**
		(0.002)				(0.002)	(0.012)
Dummy for Distance ≥ 2500 miles		-0.470**				-0.455**	-0.476**
		(0.003)				(0.003)	(0.016)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Reversals	196,256	196,256	196,256	0	196,256	0	196,256
Observations	2,713,476	2,713,476	2,713,476	2,517,220	196,256	2,517,220	196,256
R-squared	0.079	0.083	0.093	0.096	0.728	0.100	0.730

Notes: This table presents the effect of distance on citation probability for non-reversals and priority reversals. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. Publication years of the citing patents range from 2001 and 2014. Distance dummies are included to show non-linear effect of distance for different distance ranges. (The reference category is 0-25 miles.) The time lag between priority years of the citing and cited year is limited to +-5 years. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent application level to allow for correlation among patents citing the same patent. ** p<0.01, * p<0.05

Table 5. Priority Reversals and Local Inventor Interactions

Dependent Variable:	*Dummy for an actual citation*				
	(1)	(2)	(3)	(4)	(5)
					Disclosure Citations
	Non-reversals	Priority Reversals			
	Excl. self citations	Excl. self citations	Excl. PA cites	Examiner reversals	Excl. self citations
VARIBALES					
Log(*Distance*)	-0.045**	-0.035**	-0.031**	-0.043**	-0.044**
	(0.001)	(0.003)	(0.003)	(0.010)	(0.001)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2,237,179	177,682	176,758	62,150	735,519
R-squared	0.098	0.720	0.721	0.795	0.391

Notes: This table presents results from comparing localization of citation non-reversals with different subsets of priority reversals and disclosure citations. Disclosure citations are citations that occur when the priority date of the citing patent comes before the priority date but after the earliest publication date of the cited patent. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. All columns exclude self citations. Publication years of the citing patents range from 2001 and 2014. The time lag between priority years of the citing and cited year is limited to +-5 years. Standard errors are clustered at the cited patent level to allow for correlation among patents citing the same patent. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors, in paranthesis, are robust to heteroskedasticity. ** p<0.01, * p<0.05

Online Appendix Table A1. Description of Variables and Concepts

VARIBALE	DESCRIPTION	DATA SOURCE
Actual citation	A reference recorded on the front page of a patent that indicates its use of prior knowledge contained in another patent	EPO Worldwide Patent Statistical Database
Control citation	A non-citing patent pair constructed by matching citing patents on four-digit IPC and publication year	EPO Worldwide Patent Statistical Database
Examiner citation	A citation added by a patent examiner during the patent examination process	EPO Worldwide Patent Statistical Database
Self-citation	A citation where the citing and cited patents have the same original assignee	Thomson Innovation Patent Database
In-text citation	A reference to an earlier patent extracted from the specifications of the invention described on a patent	EPO Worldwide Patent Statistical Database
Patent family	A set of patents that relate to the same underlying invention and thus have the same priority date	EPO Worldwide Patent Statistical Database
Priority year	Earliest year in which an invention's novelty, usefulness, and non-obviousness is recognized by a patent-issuing authority (i.e. invention year)	EPO Worldwide Patent Statistical Database
Publication year	Year in which an application or a patent is disclosed to the public	EPO Worldwide Patent Statistical Database
Earliest publication year	Year in which an invention underlying a patent family is disclosed to the public for the first time	EPO Worldwide Patent Statistical Database
Priority reversals	A citation where the priority year of the citing patent is earlier than the priority year of the cited patent	EPO Worldwide Patent Statistical Database
Disclosure reversals	A citation where the priority year of the citing patent is later than the priority year but earlier than the earliest publication year of the cited patent	EPO Worldwide Patent Statistical Database

Online Appendix Table A2. Potential Changes to the Underlying Invention

Dependent Variable:	Dummy for an actual citation			
	(1)	(2)	(3)	(4)
	Non-reversals		Priority Reversals	
VARIBALES	Excl. CIP applications	Excl. prov. applications	Excl. CIP applications	Excl. prov. applications
Log(Distance)	-0.044**	-0.042**	-0.035**	-0.029**
	(0.001)	(0.001)	(0.005)	(0.001)
Tech Cluster Controls	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes
Observations	1,786,120	1,145,930	92,945	81,650
R-squared	0.135	0.2287	0.794	0.810

Notes: This table presents results from comparing localization of citation non-reversals with priority reversals after excluding continuation-in-part applications and applications taking priorities from provisional applications, two sources of potential changes to the underlying invention. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. All columns exclude self citations. Publication years of the citing patents range from 2001 and 2014. The time lag between priority years of the citing and cited year is limited to +-5 years. Standard errors are clustered at the cited patent level to allow for correlation among patents citing the same patent. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors, in paranthesis, are robust to heteroskedasticity. ** p<0.01, * p<0.05

Online Appendix Table A3. Alternative Specification: Within-patent Comparision of Citation Reversals and Non-reversals

Dependent Variable:	Dummy for a citation w/ distance <= 50 miles			Dummy for a citation w/ distance <= 25 miles		
	(1)	(2)	(3)	(4)	(5)	(6)
VARIBALES	All Reversals	Examiner Reversals	Inventor Reversals	All Reversals	Examiner Reversals	Inventor Reversals
Dummy for non-reversal	-0.012	-0.014	-0.006	-0.006	-0.005	0.001
	(0.007)	(0.018)	(0.008)	(0.006)	(0.016)	(0.008)
Citing Patent IPC Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Citing Publication Year	Yes	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	13,926	2,028	8,276	13,926	2,028	8,276
R-squared	0.646	0.624	0.705	0.624	0.619	0.664

Notes: This table presents a comparison of localization of non-reversals and priority reversals consisting of the same cited patent and whose citing patents are in the same publication year and IPC. The publication years of citing patents in the sample cover from 2001 to 2014. Cited patent fixed effects are included to control for invariant patent-level characteristics of both citing and cited patents that may influence the outcome. Citing patent technology area (IPC) fixed effects and citing patent publication year fixed effects are included to control for different factors inherent to technology fields and publication years that may influence the outcome. Standard errors are robust to heteroskedasticity and clustered at the cited patent level to allow for correlation among patents citing the same patent. ** $p<0.01$, * $p<0.05$

Online Appendix Table A4. Replicating the Results for Period 1977-2014

Dependent Variable:	*Dummy for an actual citation*		
	(1)	(2)	(3)
		Nonlinear distance effect	
VARIBALES	All	Non-reversal	Reversal
Log(*Distance*)			
Dummy for 25 ≤ Distance < 50 miles	-0.173**	-0.168**	-0.218**
	(0.002)	(0.002)	(0.017)
Dummy for 50 ≤ Distance < 100 miles	-0.352**	-0.344**	-0.371**
	(0.003)	(0.003)	(0.022)
Dummy for 100 ≤ Distance < 150 miles	-0.386**	-0.374**	-0.448**
	(0.003)	(0.003)	(0.023)
Dummy for 150 ≤ Distance < 250 miles	-0.409**	-0.399**	-0.420**
	(0.002)	(0.003)	(0.017)
Dummy for 250 ≤ Distance < 500 miles	-0.424**	-0.412**	-0.434**
	(0.002)	(0.002)	(0.012)
Dummy for 500 ≤ Distance < 1000 miles	-0.419**	-0.409**	-0.423**
	(0.002)	(0.002)	(0.011)
Dummy for 1000 ≤ Distance < 1500 miles	-0.455**	-0.443**	-0.456**
	(0.002)	(0.002)	(0.012)
Dummy for 1500 ≤ Distance < 2500 miles	-0.451**	-0.439**	-0.458**
	(0.002)	(0.002)	(0.010)
Dummy for Distance ≥ 2500 miles	-0.408**	-0.395**	-0.426**
	(0.002)	(0.002)	(0.012)
Tech Cluster Controls	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes
Number of Reversals	307,529	0	307,529
Observations	5,253,936	4,946,407	307,529
R-squared	0.180	0.189	0.754

Notes: This table presents the effect of distance on citation probability for non-reversals and priority reversals. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. Publication years of the citing patents range from 1977 and 2014. Distance dummies are included to show non-linear effect of distance for different distance ranges. (The reference category is 0-25 miles.) Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent application level to allow for correlation among patents citing the same patent. ** p<0.01, * p<0.05

Online Appendix Table A5. The Effect of Distance on the Probability of Citation using a Sample Matched on Six-digit Classification Code

Dependent Variable:	Dummy for an actual citation		
	(1)	(2)	(3)
VARIBALES	All	Non-revresal	Reversal
Log(*Distance*)	-0.071**		
	(0.000)		
Log(*Distance*) x Dummy for Reversal	-0.013**		
	(0.001)		
Dummy for Reversal	-0.221**		
	(0.008)		
Dummy for 25 ≤ Distance < 50 miles		-0.179**	-0.207**
		(0.004)	(0.029)
Dummy for 50 ≤ Distance < 100 miles		-0.331**	-0.328**
		(0.006)	(0.040)
Dummy for 100 ≤ Distance < 150 miles		-0.377**	-0.362**
		(0.006)	(0.046)
Dummy for 150 ≤ Distance < 250 miles		-0.388**	-0.363**
		(0.004)	(0.033)
Dummy for 250 ≤ Distance < 500 miles		-0.393**	-0.356**
		(0.003)	(0.023)
Dummy for 500 ≤ Distance < 1000 miles		-0.403**	-0.389**
		(0.003)	(0.020)
Dummy for 1000 ≤ Distance < 1500 miles		-0.442**	-0.433**
		(0.003)	(0.022)
Dummy for 1500 ≤ Distance < 2500 miles		-0.432**	-0.405**
		(0.003)	(0.018)
Dummy for Distance ≥ 2500 miles		-0.406**	-0.411**
		(0.003)	(0.023)
Tech Cluster Controls	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes
Number of Reversals	128,632	0	128,632
Observations	1,901,510	1,772,878	128,632
R-squared	0.072	0.080	0.728

Notes: This table presents findings on localization of patent citations using a dataset matched on six-digit IPC code constructed from PatStat. The publication years of citing patent in the sample matched on six-digit IPC code range from 2001 to 2014. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent level to allow for correlation among patents citing the same patent. ** p<0.01, * p<0.05

Online Appendix Table A6. Timing of Patent Application Filing

Dependent Variable:	*Dummy for an actual citation*				
	(1)	(2)	(3)	(4)	(5)
VARIBALES	Reversal lag 1-2 year	Reversal lag 2-3 years	Reversal lag 3-4 years	Reversal lag 4-5 years	Reversal lag 5-6 years
Log(*Distance*)	-0.081**	-0.081**	-0.081**	-0.081**	-0.081**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log(*Distance*) *x Dummy for Reversal*	-0.021**	-0.018**	-0.015**	-0.009**	-0.011**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.004)
Dummy for Reversal	-0.108**	-0.153**	-0.217**	-0.282**	-0.302**
	(0.013)	(0.012)	(0.016)	(0.022)	(0.025)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes
Number of Reversals	33,448	48,482	30,286	18,809	12,253
Observations	2,603,646	2,618,680	2,600,484	2,589,007	2,582,451
R-squared	0.091	0.091	0.092	0.093	0.093

Notes: This table presents results from testing the possibilty that citations get reversed because the patent application of the invention building on another invention gets filed before the patent application of the invention that it builds on. Priority reversals are re-categorized to be non-reversals unless the priority date of the citing patent is at least 1 year before the priority date of the cited patent. A *Reversal lag* is a negative priority lag . The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. Publication years of the citing patents range from 2001 and 2014. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent level to allow for correlation among patents citing the same patent. ** p<0.01, * p<0.05

Dependent Variable:	*Dummy for an actual citation*					
	(1)	(2)	(3)	(4)	(5)	(6)
	Chemical		Pharmaceutical		Biotechnology	
VARIBALES	Non-reversal	Reversal	Non-reversal	Reversal	Non-reversal	Reversal
Dummy for 25 ≤ Distance < 50 miles	-0.295**	-0.267	-0.337**	-0.377**	-0.215**	-0.508**
	(0.036)	(0.485)	(0.020)	(0.135)	(0.032)	(0.186)
Dummy for 50 ≤ Distance < 100 miles	-0.570**	-0.674	-0.565**	-0.511**	-0.462**	-0.393*
	(0.048)	(0.751)	(0.025)	(0.151)	(0.042)	(0.197)
Dummy for 100 ≤ Distance < 150 miles	-0.643**	-0.407	-0.600**	-0.707**	-0.517**	-0.655**
	(0.055)	(0.699)	(0.028)	(0.143)	(0.041)	(0.142)
Dummy for 150 ≤ Distance < 250 miles	-0.728**	-0.501	-0.663**	-0.536**	-0.547**	-0.647**
	(0.033)	(0.515)	(0.019)	(0.116)	(0.031)	(0.152)
Dummy for 250 ≤ Distance < 500 miles	-0.699**	-0.527**	-0.677**	-0.619**	-0.470**	-0.455**
	(0.022)	(0.200)	(0.014)	(0.086)	(0.022)	(0.120)
Dummy for 500 ≤ Distance < 1000 miles	-0.772**	-0.357*	-0.705**	-0.598**	-0.549**	-0.487**
	(0.018)	(0.146)	(0.012)	(0.075)	(0.019)	(0.120)
Dummy for 1000 ≤ Distance < 1500 miles	-0.804**	-0.551**	-0.733**	-0.694**	-0.533**	-0.482**
	(0.020)	(0.201)	(0.016)	(0.078)	(0.023)	(0.134)
Dummy for 1500 ≤ Distance < 2500 miles	-0.774**	-0.514**	-0.725**	-0.741**	-0.570**	-0.533**
	(0.020)	(0.144)	(0.011)	(0.052)	(0.016)	(0.086)
Dummy for Distance ≥ 2500 miles	-0.693**	-0.509	-0.664**	-0.549**	-0.531**	-0.729**
	(0.038)	(0.274)	(0.017)	(0.082)	(0.022)	(0.112)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	25,898	1,850	65,657	7,629	38,196	4,976
R-squared	0.250	0.876	0.222	0.811	0.152	0.769

Notes: This table presents the effect of distance on the probability of citations for non-reversals and priority reversals across technology areas. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. Publication years of the citing patents range from 2001 and 2014. Distance dummies are included to show non-linear effect of distance for different distance ranges. (The reference category is 0-25 miles.) The time lag between priority years of the citing and cited year is limited to +-5 years. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent level to allow for correlation among patents citing the same patent. ** $p<0.01$, * $p<0.05$

| Dependent Variable:	*Dummy for an actual citation*					
	(1)	(2)	(3)	(4)	(5)	(6)
	Medical Technology		Computer Technology		Telecommunications	
VARIBALES	Non-reversal	Reversal	Non-reversal	Reversal	Non-reversal	Reversal
Dummy for 25 ≤ Distance < 50 miles	-0.210**	-0.225**	-0.150**	-0.173*	-0.199**	-0.496*
	(0.011)	(0.078)	(0.008)	(0.077)	(0.017)	(0.248)
Dummy for 50 ≤ Distance < 100 miles	-0.383**	-0.306**	-0.358**	-0.533**	-0.374**	-0.474
	(0.015)	(0.093)	(0.013)	(0.099)	(0.021)	(0.382)
Dummy for 100 ≤ Distance < 150 miles	-0.488**	-0.442**	-0.350**	-0.460**	-0.352**	-0.690**
	(0.016)	(0.103)	(0.011)	(0.086)	(0.025)	(0.243)
Dummy for 150 ≤ Distance < 250 miles	-0.500**	-0.340**	-0.413**	-0.477**	-0.376**	-0.362
	(0.012)	(0.091)	(0.009)	(0.067)	(0.017)	(0.196)
Dummy for 250 ≤ Distance < 500 miles	-0.458**	-0.385**	-0.491**	-0.457**	-0.370**	-0.420*
	(0.008)	(0.056)	(0.006)	(0.052)	(0.013)	(0.170)
Dummy for 500 ≤ Distance < 1000 miles	-0.457**	-0.393**	-0.381**	-0.398**	-0.405**	-0.457**
	(0.007)	(0.051)	(0.005)	(0.043)	(0.011)	(0.128)
Dummy for 1000 ≤ Distance < 1500 miles	-0.458**	-0.476**	-0.479**	-0.481**	-0.404**	-0.465**
	(0.008)	(0.051)	(0.006)	(0.046)	(0.011)	(0.136)
Dummy for 1500 ≤ Distance < 2500 miles	-0.471**	-0.401**	-0.452**	-0.441**	-0.406**	-0.501**
	(0.006)	(0.047)	(0.005)	(0.039)	(0.010)	(0.117)
Dummy for Distance ≥ 2500 miles	-0.500**	-0.489**	-0.375**	-0.361**	-0.441**	-0.534**
	(0.009)	(0.056)	(0.006)	(0.046)	(0.013)	(0.142)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	238,348	17,792	555,193	36,209	138,100	8,212
R-squared	0.087	0.745	0.086	0.803	0.068	0.860

Notes: This table presents the effect of distance on the probability of citations for non-reversals and priority reversals across technology areas. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. Publication years of the citing patents range from 2001 and 2014. Distance dummies are included to show non-linear effect of distance for different distance ranges. (The reference category is 0-25 miles.) The time lag between priority years of the citing and cited year is limited to +-5 years. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors are robust to heteroskedasticity and clustered at the cited patent level to allow for correlation among patents citing the same patent. ** p<0.01, * p<0.05

Online Appendix Table A8. Citation Reversals and Local Inventor Interactions (incl. disclosure citations)

| Dependent Variable:	*Dummy for an actual citation*					
	(1)	(2)	(3)	(4)	(5)	(6)
	Non-reversals	Priority reversals and disclosure citations			Priority reversals	Disclosure citations
VARIBALES	Excl. self cites	Excl. self cites	PA cites	Examiner	Excl. self cites	Excl. self cites
Dummy for 25 ≤ Distance < 50 miles	-0.094**	-0.116**	-0.102**	-0.108**	-0.089**	-0.102**
	(0.004)	(0.007)	(0.007)	(0.012)	(0.021)	(0.008)
Dummy for 50 ≤ Distance < 100 miles	-0.211**	-0.232**	-0.210**	-0.206**	-0.181**	-0.222**
	(0.005)	(0.009)	(0.009)	(0.016)	(0.026)	(0.011)
Dummy for 100 ≤ Distance < 150 miles	-0.220**	-0.252**	-0.227**	-0.211**	-0.183**	-0.246**
	(0.005)	(0.009)	(0.009)	(0.015)	(0.026)	(0.011)
Dummy for 150 ≤ Distance < 250 miles	-0.242**	-0.276**	-0.251**	-0.259**	-0.200**	-0.264**
	(0.004)	(0.007)	(0.007)	(0.011)	(0.020)	(0.008)
Dummy for 250 ≤ Distance < 500 miles	-0.256**	-0.296**	-0.271**	-0.284**	-0.198**	-0.291**
	(0.003)	(0.005)	(0.005)	(0.009)	(0.015)	(0.006)
Dummy for 500 ≤ Distance < 1000 miles	-0.245**	-0.250**	-0.225**	-0.248**	-0.180**	-0.242**
	(0.003)	(0.005)	(0.005)	(0.007)	(0.014)	(0.005)
Dummy for 1000 ≤ Distance < 1500 miles	-0.281**	-0.302**	-0.276**	-0.294**	-0.207**	-0.297**
	(0.003)	(0.005)	(0.005)	(0.008)	(0.015)	(0.006)
Dummy for 1500 ≤ Distance < 2500 miles	-0.278**	-0.297**	-0.272**	-0.291**	-0.207**	-0.291**
	(0.003)	(0.004)	(0.005)	(0.007)	(0.013)	(0.005)
Dummy for Distance ≥ 2500 miles	-0.241**	-0.250**	-0.226**	-0.241**	-0.185**	-0.235**
	(0.003)	(0.005)	(0.005)	(0.008)	(0.015)	(0.006)
Tech Cluster Controls	Yes	Yes	Yes	Yes	Yes	Yes
Cited Patent Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,661,472	985,296	977,560	451,344	249,777	735,519
R-squared	0.144	0.362	0.362	0.378	0.712	0.392

Notes: This table presents results from comparing localization of citation non-reversals with different subsets of citation reversals, including disclosure citations. The sample consists of actual USPTO citations and non-citing, control citations that are randomly matched to citing patents on publication year and 4-digit IPC code. All columns exclude self citations. Publication years of the citing patents range from 2001 and 2014. Distance dummies are used to show non-linear effect of distance for different distance ranges. (The reference category is 0-25 miles.) The time lag between priority years of the citing and cited year is limited to +-5 years. Standard errors are clustered at the cited patent level to allow for correlation among patents citing the same patent. Cited patent fixed effects are included to control for invariant patent-level characteristics that may influence citation probability. Dyadic dummies indicating citations between leading tech / research clusters (i.e. Austin, MA Route 128, Raleigh-Durham, San Diego, and Silicon Valley) are included. Standard errors, in paranthesis, are robust to heteroskedasticity. ** $p<0.01$, * $p<0.05$

Figure 1. Priority Reversal
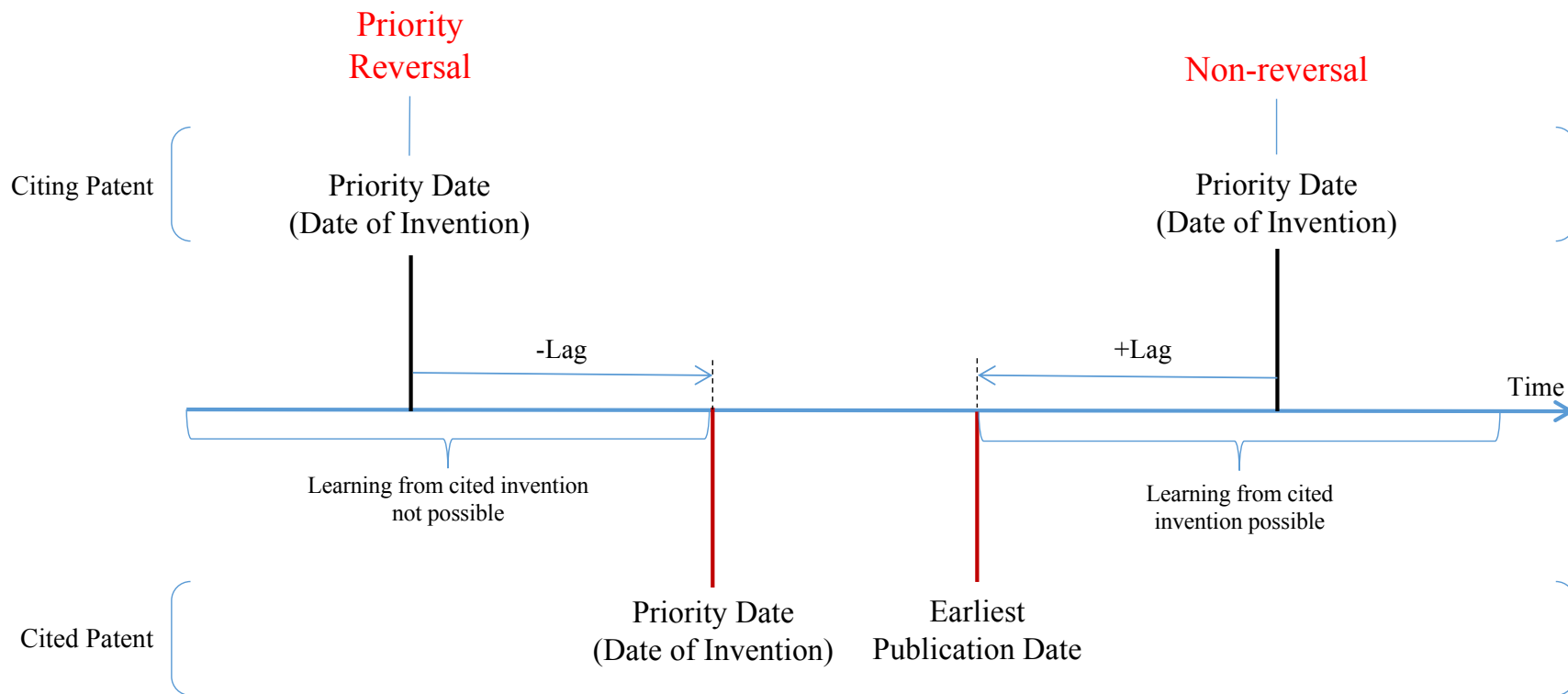
Notes: This figure presents how priority reversals occur. Priority reversals are citations where the priority date of the citing patent comes before the priority date of the cited patent. The priority date of the citing patent in non-reversals comes after both the priority date and the earliest publication date of the cited patent.
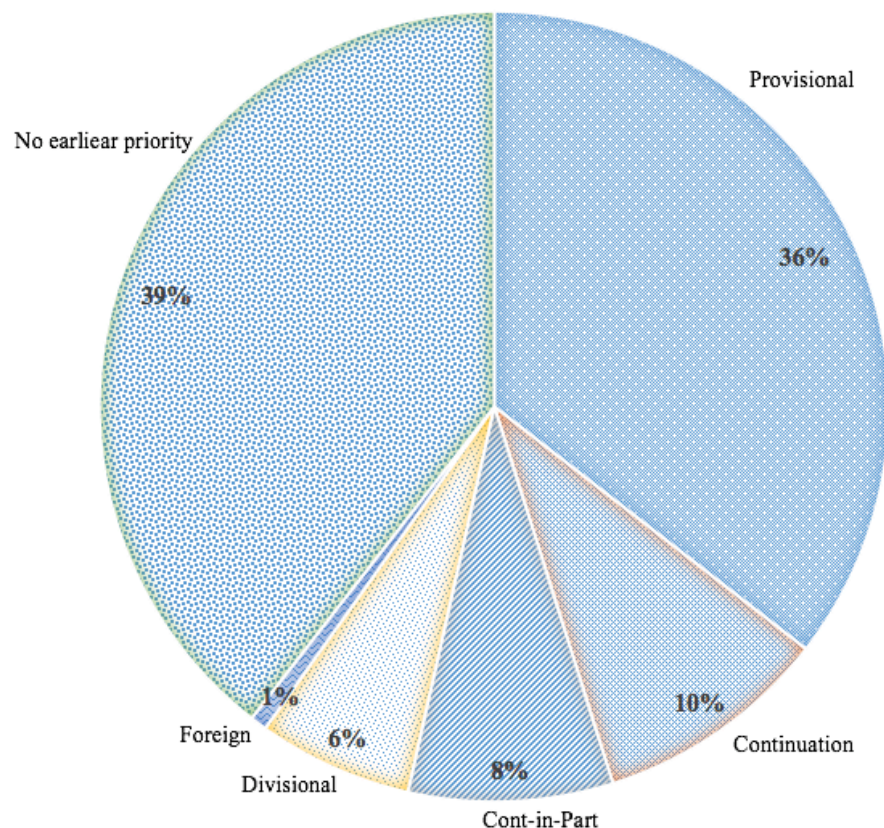
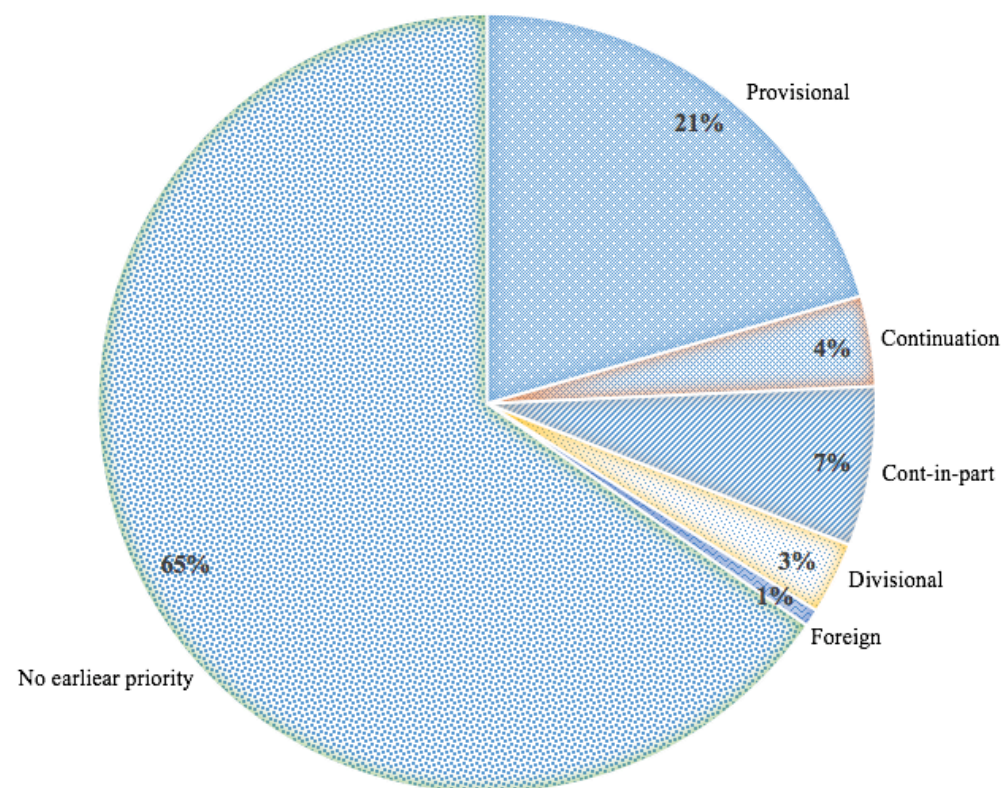Figure 2. Average Distance Between Inventors by Citation Type
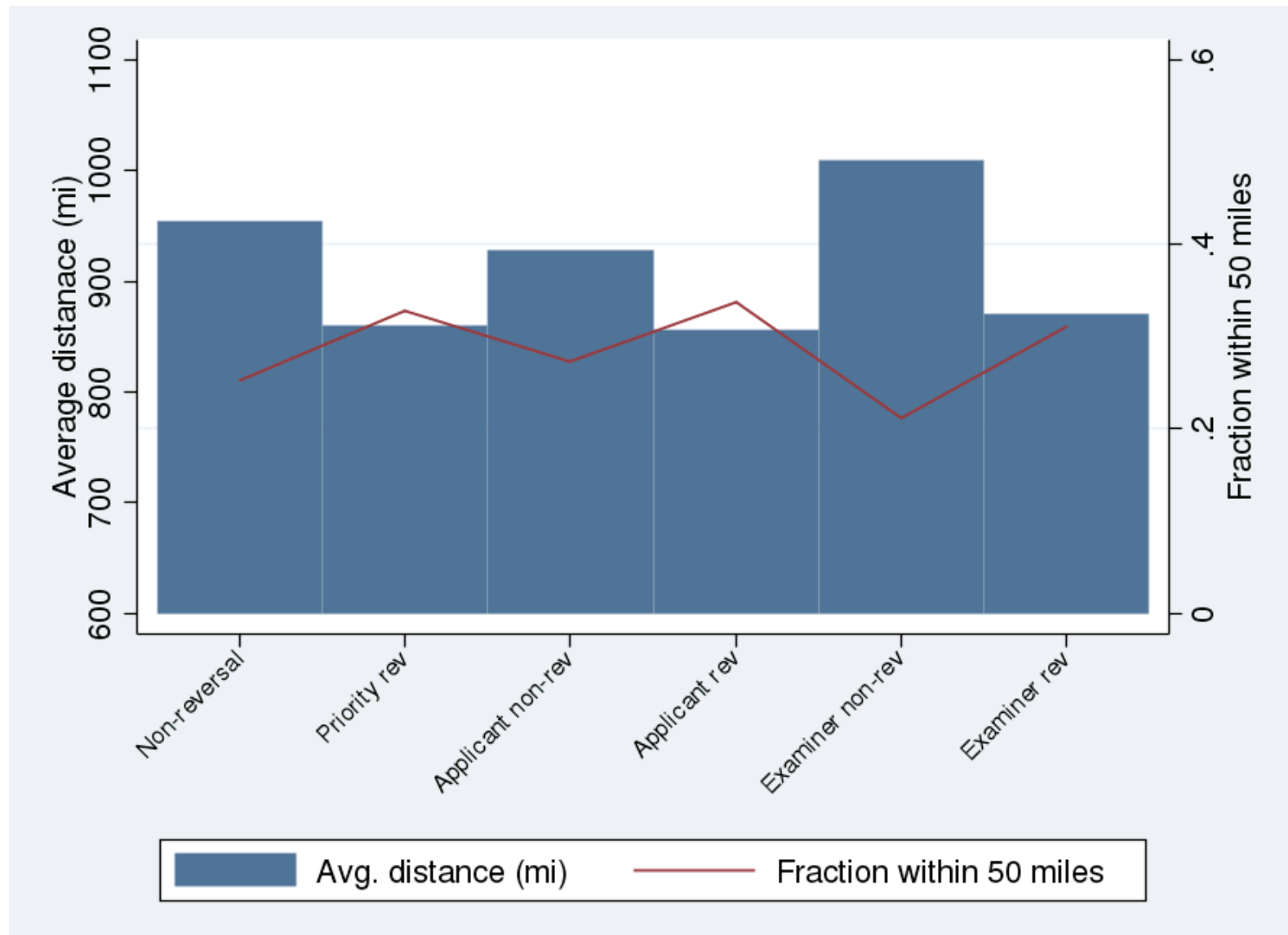
Notes: This figure compares localization of citations across different citation types. "Fraction within 50 miles" is the fraction of citations whose inventors reside within 50 miles of each other. The sample contains actual citations with citing patents covering years 2001 through 2014, and includes only priority reversals.

Online Appendix Figure A1. Breakdown of Applications by Priority Type

Citing Patents

Provisional 36%

No earliear priority 39%

Foreign 1%

Divisional 6%

Cont-in-Part 8%

Continuation 10%

Cited Patents

Provisional 21%

Continuation 4%

Cont-in-part 7%

Divisional 3%

Foreign 1%

No earliear priority 65%

Notes: This figure presents the breakdown of the types of applications that citing and cited patents claim priorities based on. Provisional applications are those that are filed prior to non-provisional (regular) applications to establish an earlier priority date. Continuation applications make additional claims based on an existing invention specified in an earlier patent application. Divisional applications are filed to separate out distinct inventions from an earlier application. Continuation-in-part applications add extensions to an earlier invention. (The claims based on new subject matter are assigned as their priority date the application filing date of the continuation-in-part application.)

Online Appendix Figure A2. Distribution of Citing and Cited Patents over Publication Lag
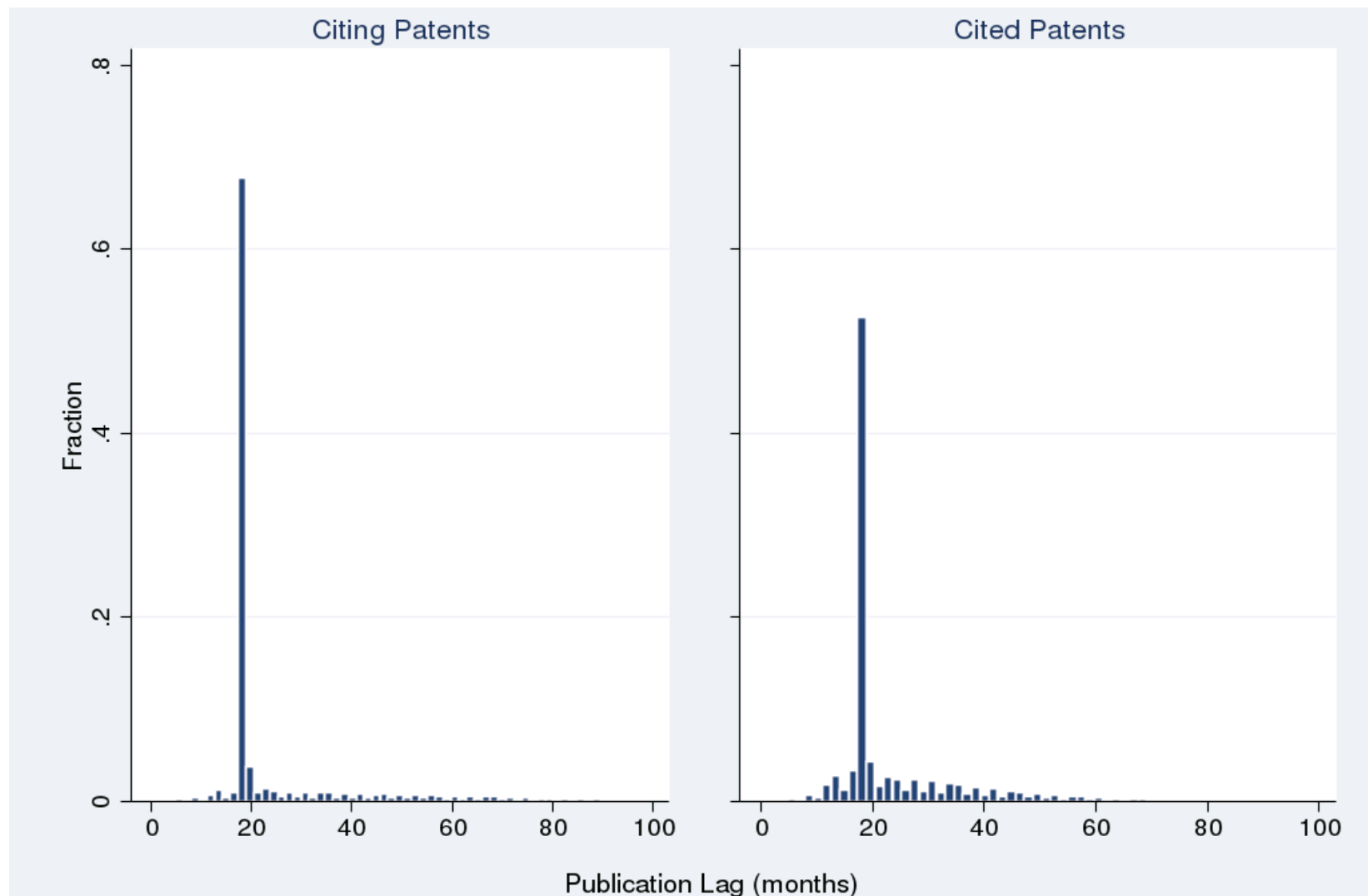
Notes: This figure presents the distribution of citing and cited patents over publication lag. "Publication lag" is time (in months) it takes for a patent or an application to be published and is measured as the earliest publication date minus the priority date.

Online Appendix Figure A3. Disclosure Citation
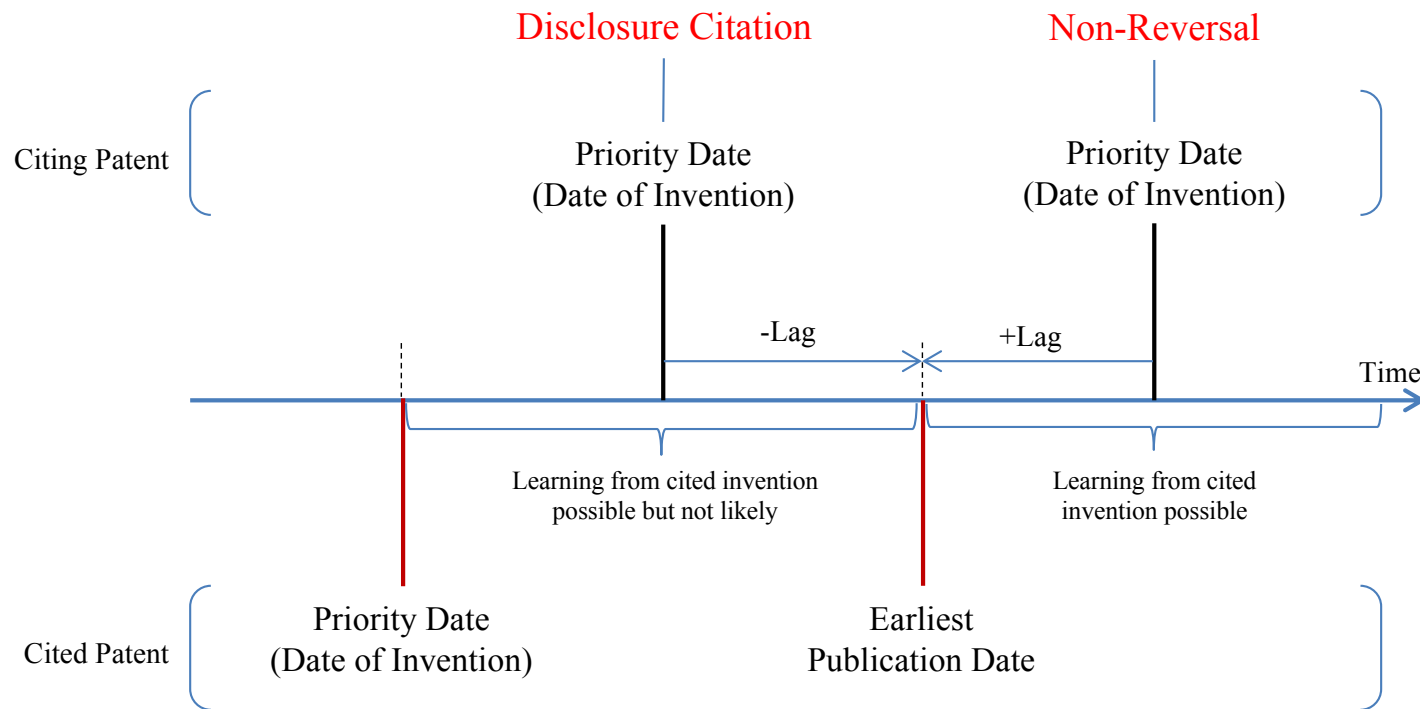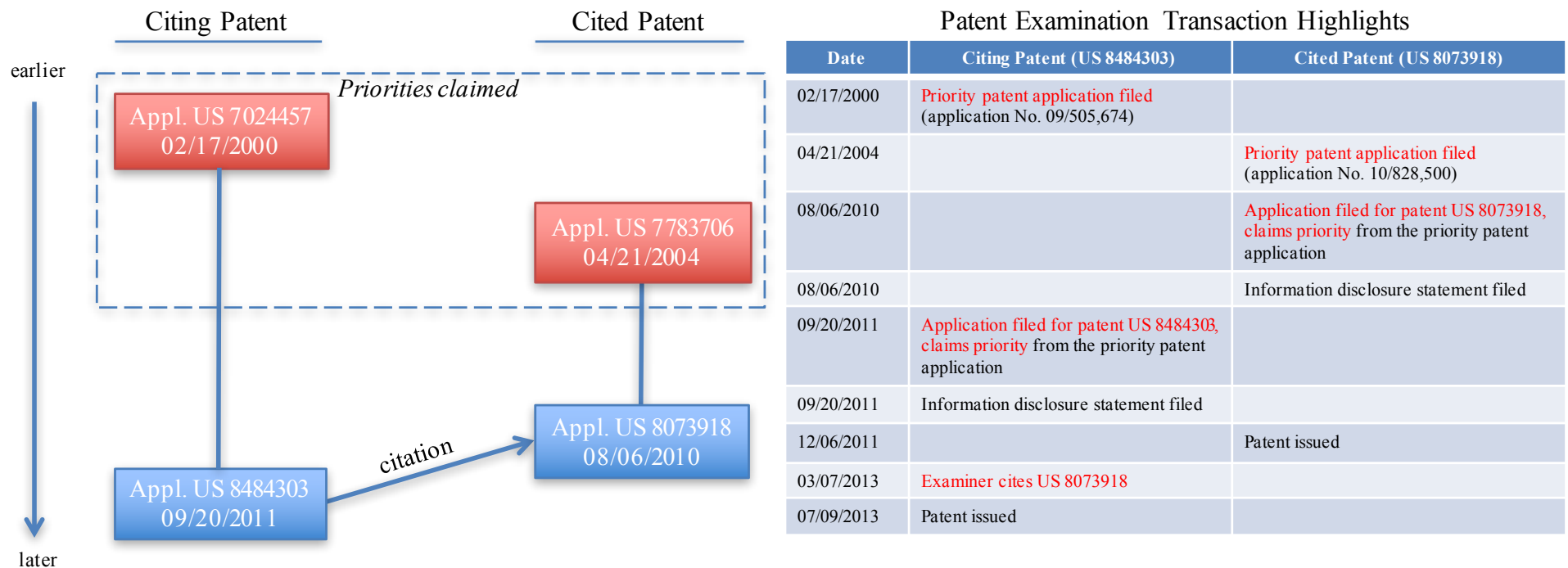
Notes: This figure presents how disclosure citations occur. Disclosure citations occur when the priority date of the citing patent comes after the priority date of the cited patent but before the earliest publication date of the cited patent. The priority date of the citing patent in non-reversals comes after both the priority date and the earliest publication date of the cited patent.

Online Appendix Figure A4. Priority Reversal Example

Citing Patent (J2 Global Communications): E-mail synchronization between heterogeneous mail servers
Cited Patent (Aristotle.Net): Filtering and managing electronic mail

Patent Examination Transaction Highlights

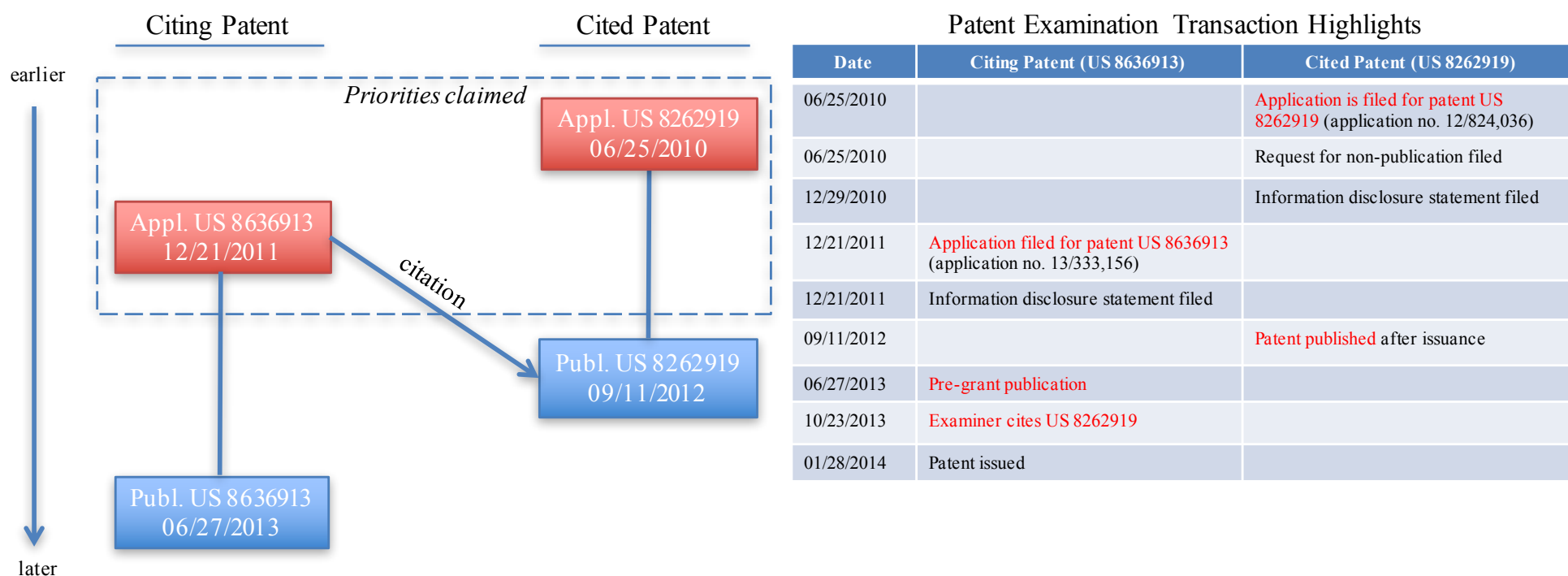| Date	Citing Patent (US 8484303)	Cited Patent (US 8073918)
02/17/2000	Priority patent application filed (application No. 09/505,674)	
04/21/2004		Priority patent application filed (application No. 10/828,500)
08/06/2010		Application filed for patent US 8073918, claims priority from the priority patent application
08/06/2010		Information disclosure statement filed
09/20/2011	Application filed for patent US 8484303, claims priority from the priority patent application	
09/20/2011	Information disclosure statement filed	
12/06/2011		Patent issued
03/07/2013	Examiner cites US 8073918	
07/09/2013	Patent issued	

Notes: The application for the citing patent (09/20/2011) was filed after the application for the cited patent (08/06/2010), but the priority date of the citing patent (02/17/2000) comes before the priority date of the cited patent (04/21/2004), indicating that the citing invention was created before the cited invention.

Online Appendix Figure A5. Disclosure Citation Example

Citing Patent (HGST Netherlands):   Removing residues in magnetic head fabrication

Cited Patent   (Western Digital):   Method and system for providing a perpendicular magnetic recording pole using multiple chemical mechanical planarization

Citing Patent

Cited Patent

Patent Examination Transaction Highlights

earlier

Priorities claimed

Appl. US 8262919
06/25/2010

Appl. US 8636913
12/21/2011

citation

Publ. US 8262919
09/11/2012

Publ. US 8636913
06/27/2013

later

Date	Citing Patent (US 8636913)	Cited Patent (US 8262919)
06/25/2010		Application is filed for patent US 8262919 (application no. 12/824,036)
06/25/2010		Request for non-publication filed
12/29/2010		Information disclosure statement filed
12/21/2011	Application filed for patent US 8636913 (application no. 13/333,156)	
12/21/2011	Information disclosure statement filed	
09/11/2012		Patent published after issuance
06/27/2013	Pre-grant publication	
10/23/2013	Examiner cites US 8262919	
01/28/2014	Patent issued	

Notes: The priority date of the citing patent (12/21/2011) comes after the priority date (6/25/2010) but before the earliest publication date (9/11/2012) of the cited patent, indicating that the citing invention was created before the cited invention was disclosed to the public.