# Spreading the Word: Geography, Policy and University Knowledge Diffusion*

## Sharon Belenzon
Fuqua School of Business, Duke University

## Mark Schankerman
London School of Economics and CEPR

# Abstract

Using new data on citations to university patents and scientific publications, and measures of distance based on Google maps, we study how geography affects university knowledge diffusion. We show that knowledge flows from patents are localized in two respects: they decline sharply with distance up to about 100 miles, and they are strongly constrained by state borders, controlling for distance. While distance also constrains knowledge spillovers from publications, the state border does not. We investigate how the strength of the state border effect varies with university and state characteristics. It is larger for patents from public, as compared to private, universities and this is partly explained by the local development policies of universities. The border effect is larger in states with stronger non-compete laws that affect intra-state labor mobility, and those with greater reliance on in-state educated scientists and engineers. We confirm the impact of non-compete statutes by studying a policy reform in Michigan that introduced such restrictions.

**Keywords:** knowledge spillovers, diffusion, geography, university technology transfer, patents, scientific publications

**JEL Classification:** K41, L24, O31, O34

## 1. Introduction

Innovation and knowledge spillovers are the key to economic growth, and universities play a central role. In the U.S., academic institutions spent $48 billion on R&D, accounting for 56 percent of basic research and 33 percent of total research in the U.S. (National Science Board, 2008). Academic research increases productivity growth in the economy and stimulates greater private sector R&D through spillovers, and through licensing university innovations to private firms for commercialization.[1] Academic research output takes two main forms: scientific publications and, increasingly since the 1980 Bayh-Dole Act, patents. Promoting university innovation and its diffusion, especially through science-based research clusters, is a major policy objective in industrialized countries. This policy focus is based in part on the assumption that knowledge spillovers are geographically localized and best exploited by agglomerating high technology activity.[2] Thus it is important to understand how geography, and the characteristics and policies of universities and states, constrain knowledge diffusion.[3]

This paper focuses on how state borders, and distance, influence the diffusion of knowledge from private and public American universities, and explores why the state may be a relevant geographical unit when analyzing knowledge flows. Whereas country borders typically demarcate zones with different cultures, languages, and political institutions, American states are not likely to vary much on these dimensions. Thus it is not immediately clear why state borders would matter in this context. Moreover, the difficulty of disentangling state border effects from pure distance effects makes it difficult to isolate and interpret whatever effects appear to be associated with state borders. Nonetheless, because state borders are not strongly associated with different linguistic, culture, or political institutions, they provide a clean framework for investigating how local policy, both at the state and university levels, influences knowledge diffusion.

---

[1] There is substantial evidence of R&D spillovers (e.g., Jaffe, 1989; Jaffe and Trajtenberg, 2002; Adams, 1990). Research spillovers tend to be geographically localized, as might be expected if direct knowledge transfers are important (Jaffe, Trajtenberg and Henderson, 1993; Audretsch and Stephan, 1996). There is also a growing empirical literature on university patenting and technology transfer policies (e.g.,Henderson, Jaffe and Trajtenberg, 1998; Lach and Schankerman, 2008; Belenzon and Schankerman, 2009), and university research productivity (Adams and Griliches, 1998).

[2] For a review of economic studies of links between universities, entrepreneurship, and regional development, see Astebro and Bazzazian (2010).

[3] Knowledge diffusion can be 'disembodied' (e.g. reading patents or publications) or transmitted through more direct interaction, such as collaborative research and consulting activity. Both forms of transmission may be constrained by geographic distance, and facilitated by improvements in information and communication technologies and other channels (Agrawal and Goldfarb, 2008; Adams, 2002). Some of our results point to an important role for labor mobility and policies that influence it.

We focus on two channels through which state borders can affect knowledge diffusion: local information and commercialization of university inventions. The first channel is important when dealing with tacit knowledge that is difficult to codify and transfer by simply reading patent documents or academic publications. This means that inventors located closer to the cited university have a greater potential for learning than those located further away. In such cases, the border effect should be stronger in states where inventors are more likely to remain in the state when they move jobs, and when inventors working in a state are more likely to have been educated in that state. State policies can influence the prevalence of such local information – e.g., by discouraging inter-state mobility of inventors by strongly enforcing 'non-compete' labor laws, or by more effectively retaining locally educated scientists and engineers. The second channel involves university and state characteristics and policies that promote the local commercial development of university research output. This is more likely to occur in states with a dense and vibrant community of scientists and engineers, who can potentially build on and cite university patents and publications. In addition, the state border is likely to be more important for public universities which are more exposed than private ones to a variety of constraints and influences by state government. One manifestation of this relationship is the greater importance that public universities typically attach to promoting local and regional development through their technology licensing policies (Belenzon and Schankerman, 2009).

To study these questions, we use two complementary measures of knowledge spillovers. The first is citations to university-owned patents. Citations have been widely used in the literature to trace spillovers from corporate R&D (Jaffe and Trajtenberg, 2002). However, citations to university patents are an imperfect measure of the reliance of corporate research on university knowledge. The reason is that many of the scientific contributions made by university faculty never find their way into patents.[4] The most important complementary measure of knowledge spillovers is the extent to which corporate patents cite university scientific publications. One might expect the geographic pattern of diffusion for 'open science' knowledge in publications to differ from the 'proprietary' knowledge embedded in university patents. In addition, if the information in scientific publications is more 'general', and thus multi-use in character, we would expect it to exhibit less sensitivity to distance and state borders, especially if the border effect is driven in part by technology licensing

---

[4]Only about one third of inventions disclosed by faculty to university technology transfer offices end up as patent applications (Lach and Schankerman, 2008). In addition, there are purely scientific discoveries by faculty that are not embodied in inventions with commercial applications, but which may contribute to subsequent corporate innovation.

and other university or state policy.

There is a substantial literature on the localization of knowledge spillovers using patent citations.[5] The basic idea is that a citation indicates that the later invention in some way builds on the earlier one, and that some knowledge transfer has occurred. The seminal paper in this area is Jaffe, Trajtenberg, and Henderson (1993). They compare the average distance of patents that cite another patent and a random control group of patents that do not cite (the control patent is drawn from the same technology field and patent cohort as the cited patent). They show that firms located in the same city as the inventor are much more likely than others to benefit from knowledge spillovers from that innovation. This approach has been used and refined by later studies.[6] Geography is typically summarized as a set of broad areas – identifying only whether inventors are in same city, state, or country. These studies do not use a measure of geographic distance, so they are not able to explore in more detail how distance affects citation rates – e.g., whether the effects of distance on spillovers dissipate after some point. We address this gap by using the actual distance between the locations of patent assignees (measured by Google Maps).

We adopt a similar econometric approach to study how geography shapes university knowledge spillovers, and how this impact varies with state and university characteristics and policies. We distinguish between two dimensions of localization: the relationship between spillovers and geographic distance, and the impact of state borders, controlling for distance. Using new data on citations to university patents and scientific publications, and measures of distance based on Google Maps, we show that spillovers are highly localized. Citations to both university patents and publications decline sharply with distance up to about 150 miles, but are essentially constant beyond that. This level of 'threshold distance' – corresponding as it does to an extended commuting distance – strongly

---

[5]Of course, not all university knowledge diffusion represents spillovers in the economic sense. The benefits are partially internalized when university inventors collaborate with private firms in the commercialization of their inventions (e.g. through consulting or participation in start-up companies). This is the argument that Zucker, Darby, and Brewer (1998), and Zucker, Darby, and Armstrong (1998) make with respect to the development of the U.S. biotechnology sector. However, it is unlikely that the social returns to knowledge diffused through university patents and scientific publications are fully internalized by the inventors. See also Audretsch and Stephan (1996).

[6]Leading examples of papers that document the state-(or other sub-national or national) border effect include Thompson (2006), Alcacer and Gittleman (2006), and Peri (2005). The first two papers uses the control group approach but exploit the distinction between citations by the patentee and those added by the patent examiner to help identify localized spillover effects. Peri (2005) uses the citation function approach developed by Jaffe and Trajtenberg (1998), which requires explicit functional form assumption on the probability to cite. The border effects' found by these studies are difficult to interpret, however. Thompson does not include a distance measure, which confounds the effects of distance and borders. Peri includes only a linear distance measure, and thus potentially confounds the border effect with nonlinear distance effects. In a more recent (unpublished) paper, Singh, Marx, and Flemming (2010) document a persistent state border effect while controlling for refined distance measures.

suggests that direct personal interaction plays an important role in knowledge flows. Controlling for distance, we find strong evidence of a state border effect for citations to university patents. Inventors located in the same state as the cited university are substantially more likely to cite one of the university's patents than an inventor located outside the state. In sharp contrast, we find that state borders have essentially no impact on citations by patents to university scientific publications (except for very low quality publications).

The state border effect is much stronger for citations to patents from public, as compared with private, universities. A substantial part, but not all, of this ownership effect is associated with the local development focus of the university technology transfer activity. This finding has a potentially important policy implication. Belenzon and Schankerman (2009) show that there is a cost to pursuing local development in this way – universities with strong local focus earn substantially less licensing income from their inventions. But there may be offsetting benefits, most importantly in the form of greater localization of knowledge spillovers. This issue is key to understanding whether it makes economic sense for universities (or state governments) to promote local development through local licensing. Our finding that strong local development objectives are associated with greater localization of knowledge flows shows that there is a genuine tradeoff which policymakers need to bear in mind.

However, the impact of the state border on patent citations is very heterogeneous across states. We show that the variations in the border effect are generally consistent with the local information and commercialization hypotheses. First, the border effect is larger in states that do not have, or do not strongly enforce, 'non-compete' labor laws. These statutes restrict employees from moving jobs to a competing (typically, same industry) firm within the same state for some period of time. By so doing, they should reduce within-state knowledge spillovers and thus weaken the state border effect on citation behavior.[7] We confirm the impact of non-compete statutes by studying a policy reform in Michigan that introduced such restrictions. This reform was first studied by Marx et. al. (2007, 2010), who show that non-complete laws do, in fact, increase out-migration for job movers. Our finding reinforces those studies by showing that non-compete statutes affect not only labor mobility directly, but also the knowledge diffusion that labor mobility generates. This finding is consistent with earlier work by Almeida and Kogut (1999), who document the link between patent citations

---

[7]The impact of non-compete statutes on growth is theoretically ambiguous. They intensify local knowledge spillovers by allowing intra-state job hopping, but reduce the incentives of employers and employees to invest in job-specific human capital. For discussion see Fallick, Fleishman, and Rebitzer (2006).

and labor mobility. The second prediction that is supported by the evidence is that intrastate citation is stronger in states with greater density of scientists and engineers, and a higher fraction of inventors who were educated at in-state universities.

Finally, we investigate how localization of knowledge diffusion varies across technology areas. The importance of tacit knowledge and the associated channels of information transmission may differ. In fields where information is less codified and thus harder to transmit, direct social relationships – e.g.. collaboration, seminars and so on– are likely to play a larger role, making knowledge spillovers more sensitive to geographic distance. We find that localization occurs mostly in biotechnology, pharmaceuticals, and chemicals, and much less so in electronics, information technology, and telecommunications. These differences imply that some of the variation we observe in the strength of the border effect across states may be attributable to differences in their technology specialization.

The rest of the paper is organized as follows. Section 2 presents the data. In Section 3 we describe the econometric specification. The results are reported and discussed in Section 4 reports the results. Section 5 summarizes the key findings and some directions for further research.

## 2. Data

For this paper we constructed several new, large-scale data sets that allow us to look at localization of knowledge flows in novel and more detailed ways. These are described briefly below. Details are provided in the Data Appendix.

### 2.1. Patent Citations to University Patents

The sample covers 184 research-oriented (Carnegie I) universities in the United States, which account for the vast bulk of academic R&D in the United States. We follow the conventional approach of using patent citations to trace knowledge spillovers. In order to identify the population of university patents, we matched the names of the assignees of U.S. patents to universities, using a wide range of possible appellations for the university (e.g. the names of the technology licensing office, the university, and relevant abbreviations). This allows us to identify all patents applied for by each university in the sample, and then to identify the set of all U.S. patents that subsequently cite these university patents. The standard data source for U.S. patents is the 2002 version of the NBER patents and citations data archive. We updated the patent data to 2007 by extracting all information, including inventor address and citations, for all patents granted between 2002 and

2007 directly from the USPTO website.[8] Updating the patent data improves our ability to study patterns of knowledge spillovers for relatively new technology areas, such as Information Technology and Biotechnology.

We construct a control group to compare to this set of citing patents. Self-citations and citations by foreign patents are excluded from this analysis. For each citation to a university patent, we randomly draw another (non-citing) patent in the same three-digit U.S. patent class and patent grant year. Thompson and Fox-Keene (2005) argue that findings of localized knowledge spillovers using patent citations may be sensitive to the technology classification – specifically, that more detailed disaggregation is essential – so as a further step we also collected the more detailed, six-digit assignment using the International Patent Classification for each patent.[9] The final data set includes 26,914 university patents granted during the period 1976-2006. These patents receive a total of 383,096 citations during the sample period 1976-2007. With a matched (non-citing, control) patent for each of these, the final data set has 258,966 observations.

## 2.2. Geographical distance of spillovers

To examine the relationship between distance and knowledge spillovers, we constructed a novel data set on the distance between the cited university and all of the firms that cite its patents over the period 1976-2007. The distance is measured on the basis of the address of the inventor on the citing patent and the address of the university whose patent is cited (i.e. where the patent assignee is the university). To do this, we developed new data extraction software that uses Google Map as the source of information for the geographical (driving) distance in miles between each university and the citing inventor's location. In cases where there are multiple (domestic) inventors on the citing patent, we take the average geographic distance between the addresses of the various inventors and the university whose patent is cited. The econometric results are robust to using the alternative approach of taking the minimum distance when there were multiple inventors.

## 2.3. Patent Citations to University Scientific Publications

We constructed a new data base on citations by corporate-assigned patents to scientific publications by university faculty. For each patent granted in the period 1975-2007, we extract the citations it makes to non-patent literature directly from the patent document as it appears in the U.S. Patent

---

[8] http://patft.uspto.gov/netahtml/PTO/srchnum.htm

[9] For this purpose we adopt the IPC because concerns have been raised about the accuracy of the more detailed U.S. patent sub-classes.

Office. We then identify the author(s) and her affiliation from the citation text and determine the name of the cited university. In cases where the citation has incomplete information about the authors or affiliations, we use the Web of Science data base to track the name of the publication and determine the university to which it belongs. The output of this procedure is a comprehensive data set that maps the link between corporate innovation and university scientific discoveries. We then use Google Map to calculate the distance between the location of the citing inventors and the cited university, similar to the patent citations data. Finally, we construct a control group of patents – for each patent citing an academic publication from one of the universities in our sample, we randomly draw another patent with the same technology (patent) sub-class and cohort that does not cite the university's publications. In total, 365,205 patents in the complete sample make at least one citation to academic publications. Of these citations, 35,043 involve (matched) publications from our sample of universities. With a matched (non-citing, control) patent for each of these, the final data set for publication citations has 70,086 observations.

## 2.4. University characteristics and local development objectives

For each university in the sample, we have information about whether the university is public or private, and about the extent to which its technology licensing activity is aimed at promoting local development. The latter information is based on a survey of university technology licensing offices (TLO's) developed by Lach and Schankerman (2008).[10] Among other things, this survey (conducted in 2001) asks about the importance the TLO attaches to promoting 'local and regional development' (i.e., a preference for licensing to local firms), using a four point Likert scale – very important, moderately important, relatively unimportant, or unimportant. We define a dummy variable that is set equal to one if the university TLO a answers 'relatively important' or 'very important'; the reference category corresponds to the other two categories. This survey covers only 75 universities (the patents and publications data cover 184), but these universities account for about 68 percent of the total number of patent citations in the overall sample. Of these 75 universities, 57 rank local development objectives as either relatively or very important. Not surprisingly, public universities typically rank local development highly, though there are both public institutions that do not and private ones that do (Belenzon and Schankerman, 2009). Therefore, in examining the

---

[10] The survey of TLO directors was developed in late 2001. It was sent to about 200 U.S. and Canadian research universities that belong to the Association of University Technology Managers, with 102 responses. After matching to other data for the empirical analysis, the final sample consists of 84 universities. In this analysis we exclude the nine Canadian universities because we only use patent citations by U.S.-based inventors. For more details, see Lach and Schankerman (2008) and Belenzon and Schankerman (2009).

impact of this policy variable, it will be important to control for university ownership status in the regressions.

In addition to these data sets, we use a set of state-level control variables in some of the regressions. The variables will be introduced later when we use them.

## 3. Econometric specification

We follow the empirical methodology of Jaffe, Trajtenberg and Henderson (1993), comparing the characteristics of corporate patents that cite university patents and a control group that does not. The control group is constructed as follows: for each citation received by a university patent (excluding self-citations), we randomly select another patent that does not cite but which is in the same cohort (patent grant year) and four-digit patent class. Essentially the methodology involves comparing the geographic distance, and other patent characteristics, between the citing patents and the control group. Specifically, we estimate Probit regressions of the probability of citation against a set of control variables. Since the control group is matched on the patent application date and technology field, the methodology automatically controls for these factors in the regressions.

The general empirical specification is

$$\mathrm{Pr}\,ob(C_{i(u,s),j(s')} = 1) = F(\alpha'D_{ij} + \beta'X_{ij} + \gamma D_{ws} + \lambda Z_u D_{ws} + \delta W_s D_{ws} + \eta_u + \varepsilon_{ij})$$

where $C_{i(u,s),j(s')}$ is a citation by patent $j$ (located in state $s\prime$) to a patent (or academic publication) $i$ from university $u$ (located in state $s$), and $F$ denotes the cumulative normal distribution. The control variables (discussed more fully below) include measures of geographic distance between the citing and cited patent, $D_{ij}$, a set of other controls $X_{ij}$, a within-state dummy (border effect), $D_{ws}$, interactions between university and state level variables with the within-state dummy, $Z_u D_{ws}$ and $W_s D_{ws}$, and a set of university fixed effects, $\eta_u$. We compute standard errors clustered at the level of the cited patent, which allows the disturbance $\varepsilon_{ij}$ to be correlated across citing patents for the same cited patent.

The identification assumption in this analysis is that the key observed characteristics of interest – geographic distance of the citing patent, university and state level characteristics, and university local development focus – are exogenous factors, unrelated to the disturbance $\varepsilon_{ij}$ in the citation equation. The main concern that might arise here is unobserved quality of a patent, which might affect both the probability that it is cited and (possibly) the distance of the citing patent. But here one would expect that higher (unobserved) quality would be *positively* correlated with the distance

of citing patents (i.e., weak patents tend to be cited more locally). Such correlation would induce a positive bias in our coefficient on distance, and thus cause us to *understate* localization effects, that is, to understate the negative impact of distance on citation behavior.

One important issue to bear in mind is the endogeneity of location. We treat distance between the citing firm (inventor) and the cited university as exogenous. We find that citation dissipates with distance. One interpretation of this result is that inventors learn less the further they are from the cited patent. But it could also be a reflection of an endogenous spatial distribution of inventors, driven by an attempt to exploit knowledge spillovers. The extreme version of this is what we might call 'pure assortative matching' – inventors learn only from their own types (e.g. those in their specific technology area), and distance does not affect this learning per se. One way to distinguish between these interpretations is to use more disaggregated controls for technology fields (as we do in this paper), but one can not entirely rule out endogenous location as part of the explanation. In an important sense, however, this is not so much an identification issue as an interpretational one. Nonetheless, our paper can rule out the null hypothesis that the state border effect is solely driven by endogeneity because we show that it varies systematically across university and state policy and characteristics. If the state border effect were driven only be 'assortative matching' by technology specialization, or the desire of inventors to locate closer to higher quality universities, it would be hard to explain why this effect is weaker for private than for public universities (conditional on patent quality) and for universities that are located in states that more strongly enforce 'non-compete' labor laws.

Turning to the control variables, we measure the distance between the inventor(s) of the citing patent and the university whose patent is cited in two ways. The first is a simply to use the driving mileage in logs. The logarithmic specification seems preferable on a priori grounds to the simple linear version in Peri (2005) (the marginal impact of distance is not likely to be constant), even this is restrictive because there could be highly nonlinear impacts of distance, e.g. after some level, distance may not matter at all. The way distance affects knowledge diffusion depends on how information spreads. If knowledge is primarily transferred through personal contact in research collaborations, participation of university inventors in the development of licensed technologies (including start-ups) and so on, then we might expect diffusion to be highly localized and distance not to matter after some point. But if information is spread more though information technology, or inventor participation in scientific conferences, the effects of distance should be less local. Therefore, we adopt a second, more flexible specification that allows for nonlinear effects of distance. To do that,

we use a set of seven dummy variables for intervals of distance (in miles): 25-50, 50-100, 100-150, 150-250, 250-500, 500-1000 and greater than 1000; the reference category is 0-25 miles (which might be interpreted as a metropolitan effect).

Of course, the probability a university patent is cited might be expected to depend on a variety of characteristics of the university (e.g. the quality and visibility of its faculty, its general commercial orientation, and the high-technology density and specialization of the university location), and its policies in promoting dissemination through technology transfer and academic interaction (conference attendance, consulting activities and so on). To capture these factors, we introduce a complete set of university fixed effects for the cited patent.[11]

To allow for the state border to affect the citation probability, we define a 'within-state' dummy variable that is set equal to one if the inventor of the citing patent is located in the same state as the university whose patent is cited (zero otherwise)[12] Since we are controlling separately for distance with a very flexible non-linear specification, this within-state dummy will identify whether there is a pure 'border effect'. We investigate whether there is a significant border effect on citation behavior, and whether the strength of this impact is related to university and state government policies.

Finally, we include a complete set of dummy variables for bilateral effects between pairs of the five leading high-tech clusters in the U.S.: Austin, Boston, Raleigh-Durham, San Diego, and Silicon Valley. We allow for the ordering of the location of the cited and citing inventor to matter (e.g. the San Diego-Boston link may differ from Boston-San Diego). This gives a total of 20 dummy variables for the high-tech city pairs. These controls are introduced to account for the possibility of higher citation rates between high-technology clusters, independent of distance.[13]

## 4. Non-parametric Evidence

Table 1 presents descriptive statistics on the locational characteristics of citations to university patents (Panel A) and scientific publications (Panel B). The mean share of citations that are from the same state as the inventor is 12 percent but it varies widely across patents (from 0 to 100

---

[11] This additive specification will not pick up characteristics of universities that affect the geographic profile of citations (i.e., the way they depend on distance). In the empirical analysis we will allow for the ownership type and other characteristics of the university and state to interact with geographic distance and/or the state border effect.

[12] If there are multiple inventors, the state dummy is set equal to one if any of the inventors on the citing patent is located in the same state as the cited university patent.

[13] Almeida and Kogut (1999) show that localization effects are stronger in certain high-technology regions in the U.S. than other regions. This is not surprising, given the agglomeration of technologically related activity in those areas. In our analysis, we control for university fixed effects, which should pick up much of this effect. Our dyad dummies for high-tech cluster pairs should pick up links between clusters with similar technological focus.

percent). The average distance between citing and cited patent is 1,218 miles (not reported), but citations are geographically concentrated – overall, 13 percent of all citations originate within 150 miles, and 28 percent within 500 miles, from the cited university patent.[14] At the same time, 56 percent of citations originate at a distance exceeding 1,000 miles from the cited university. The locational pattern for citations to publications is very similar. However, nothing can be concluded about the localization of knowledge diffusion from these facts alone. For that, we need to compare the geography of citing and a control group of non-citing patents. We do this non-parametrically in the next table, and econometrically in Section 4.

**Insert Table 1 here**

In Table 2 we present a non-parametric comparison of citing and control group patents (Panel A) and scientific publications (Panel B). In column (2) of Panel A, we compare the average difference between the distance of patents that cite and those that do not (control group), broken down by university ownership type and patent quality. Several conclusions are worth noting. First, in the overall sample citing patents are systematically closer to the cited university than the control group – the difference is -6.9 percent – and we easily reject they hypothesis that there is no difference. Distance does constrain university knowledge diffusion. Second, the degree of localization is more than twice as large for public institutions than for private ones – the differences are –9.3 and -4.3 percent, respectively. Third, the degree of localization is much more pronounced for the lowest quartile of patent quality, both for public and private institutions. For the upper quartile, there is much less localization and, for private universities, there is actually no statistically significant localization.

Column (3) presents the comparisons between citing and control group patents on the fraction of citations originating from within-state inventors, another measure of localization. The pattern is broadly similar to those in column (1). First, inventors that cite university patents are significantly more likely to be located in the same state. We decisively reject the null hypothesis that there is no difference between citing and non-citing patents. Second, the 'within-state citation bias' is stronger for public universities than for private ones. Finally, the within-state bias is more pronounced for

---

[14]The number of universities with citations represented in a given bracket, and the share from private universities, are shown in the second and last columns, respectively. We have a fairly even representation of private universities across the distance brackets.

the lowest quartile of patents – the difference with the upper quartile is especially large for public universities.

Overall, the pattern for publications is very similar to patents, so we will not go through it in detail. The similarity is noteworthy, and perhaps a little surprising, because publications correspond to an open science regime, where dissemination is encouraged by the norms of the profession and the academic reward structure. In contrast, patents are proprietary knowledge apart from the information disclosure mandated in the patent document. The fact that the two knowledge regimes exhibit similar characteristics suggests that there are common, geographically mediated determinants of information dissemination. We return to this point in Section 4, where we discuss the more detailed econometric results.

**Insert Table 2 here**

Figure 1 provides additional evidence on the relationship between citations, distance and state borders. In this graph we measure the effect of state borders on citation, holding constant the distance between the citing and cited patents. The graph depicts the extent to which the effects of distance and state border die out as we extend the distance. The light colored bars show the difference between the average citation probability for an inventor in the specified distance interval and those at greater distances (the 95 percent confidence interval is given at the top of each bar). These bars show clearly the very significant localization of university knowledge spillovers. For example, the first 'distance bar' shows that the probability that an inventor within 25 miles cites the university patent is 34 percentage points greater than for inventors located beyond 25 miles. Since the probability of citation is 50 percent by the construction of the control group, this effect is huge – equivalent to a 65 percent decline in the mean citation probability. We observe a further steep decline in citation probability as we move from 25-50 to 50-100 miles – there is still a small, but statistically significant, distance effect at 50-100 miles, equivalent to a 10 percent higher citation probability (relative to the mean) than at greater distances. But after that, it appears that distance exerts no further effect.[15]

The dark colored bars depict the difference between the citation probability for inventors located within the same state as the university and those outside the state, for each distance interval. These

---

[15] The last two bars suggest that the citation probability appears to rise slightly with distance at distances beyond 500 miles. This is an artifact of the higher citation probabilities between high-technology clusters which at greater distances in the U.S. (e.g. Boston, Silicon Valley, San Diego, Raleigh-Durham and Austin). When we control for cluster pairs and other factors in the econometrics, this anomaly disappears.

bars highlight the distinction between the impacts of distance and state borders on citations. While we found that effects of distance die out after 100 miles, we see that the border effect persists over much longer distances (the maximum within-state distance is 707 miles, in California). This finding is interesting because it is consistent with the hypothesis that the border effect is determined (at least in part) by university and/or state policies, whose effects we would not expect it to disappear with distance.

<div align="center">**Insert Figure 1 here**</div>

## 5. Estimation results

### 5.1. State-border effect

Table 3 presents the baseline regressions relating patent citation to distance and state borders. In all regressions, we include university fixed effects, dummy variables for pairs of five high-technology clusters, and a dummy variable for whether the citing and cited patents are in the same 6-digit IPC patent class.[16] The reported coefficients are the estimated marginal effects from Probit regressions, and standard errors are clustered at the level of the cited patent.

In column (1) we begin with the simplest specification relating the citation probability to distance measured in the log of miles between the citing and cited inventors. Distance has a statistically significant but small impact in dampening citations. A ten percent increase in distance – which corresponds to 120 miles, evaluated at the sample mean – is associated with a 0.43 percentage point increase in the probability of citation. This is equivalent to only a 1.9 percent increase relative to the mean citation probability.[17] It is also worth noting that the coefficient on the technology matching dummy is large and statistically significant, confirming that citation is more likely between patents in the same technology area. Yet the fact that we find localization, even when we control for this dummy at the disaggregated, 6-digit IPC level, suggests that localization is not just a reflection of

---

[16] Including university fixed effects in the Probit regressions does not cause an incidental parameters problem because the limiting dimension for consistency here is the number of patent citations, not the number of universities.

[17] Our estimate is larger than the one obtained by Peri (2005). He estimates that an increase in distance of 1000 Km (600 miles) is associated with a reduction in citations of about 3 percent, whereas our estimate (evaluated at the sample means) implies a 9.5 percent decline. Part of this difference disappears when we include a state border effect (column (3) in Table 2), but our finding of greater localization may also be due to our focus on *university* patents. Evidence in Adams (2002) suggests that university spillovers are more localized than corporate-generated spillovers, but a full examination of this interesting question is left for future research.

the spatial distribution of technological activity.[18] This conclusion is robust across all specifications we estimate.

In column (2) we replace the distance measure by a within-state dummy. The estimated parameter shows that citation is much more likely from inventors located within the same state – the marginal effect of being within-state is very large, 0.225, which is nearly half of the mean citation probability. Column (3) reports results for the specification that includes both the distance measure and the within-state dummy. The results confirm that both distance and the state border effect are statistically significant, and that it is important to include both variables. Including distance reduces the estimated effect of the state border from by more than 50 percent, from 0.225 to 0.122. At the same time, including the within-state dummy also reduces the estimated impact of distance by half, from -0.043 to -0.024.

There is the further concern that part of the reason there appears to be a state border effect is that we have not allowed for non-linear distance effects. To address this, in column (4) we introduce a set of dummy variables for different distance intervals. We will refer to this as the baseline specification. Two key findings emerge. First, the estimated state border effect is robust to allowing for flexible distance specification. The estimated marginal effect of crossing the state border is 0.089 – this represents about 20 percent of the mean citation probability, which is close to (and not statistically different from) the estimate of 0.122 obtained with the more restrictive distance specification. This result confirms that the border effect is not simply a proxy for geographic distance.

The second important result in column (4) is that geographic distance sharply constrains knowledge spillovers – moving from 0-25 to 25-50 miles reduces the citation probability by 20.2 percentage points, and moving out to 50-100 miles further reduces it by another 5.6 (=25.8-20.2) percentage point. But after that, distance has no appreciable effect on citation. We can test the hypothesis that there is no incremental distance effect beyond 100 miles by constraining the coefficients on those dummy variables to be the same as the coefficient for the 50-100 mile dummy. We do not reject this hypothesis if we exclude the last dummy, which captures mostly bi-coastal effects – the p-value of the test is 0.73. These econometric results confirm what we saw in Figure 1.

One concern is that the pattern of knowledge diffusion for patents that represent important

---

[18] If localization is driven only be spatial agglomeration of technologically similar innovation, we would expect to find no (or much weaker) localization when we control in a more refined way for matching on technology class. This concern was originally raised by Thompson and Fox-Keene (2005) in the context of the classic paper by Jaffe, Trajtenberg and Henderson (1993).

technological or economic advances may be very different than for marginal improvements. In particular, while unimportant patents are likely to be cited only by locals, we would expect important ideas to diffuse more widely. To investigate this hypothesis, we use the total number of citations that a patent receives over its lifetime as a proxy for its importance.[19] We re-estimated the regression for the lowest quality quartile of patents (column 5), and for the upper quartile (column 6). The results show clearly that knowledge diffusion is more localized for low value patents. The estimated coefficients on the distance dummies show a sharper distance gradient for the lower quartile. Moving from 0-25 to 25-50 miles reduces the citation probability by nearly twice as much for lowest quartile than for upper quartile (-0.247 versus -0.130). But it also interesting that for both categories of patents, the effect of distance dies out relatively quickly – it is exhausted after 100 miles for the lower quartile of patents, and 150 miles for the upper quartile.

However, while knowledge diffusion of low value patents drops off more sharply with distance, the state border effect is *weaker* for low value patents (0.078 versus 0.118). If the state border effect is due, at least in part, to the local development policies of the university, as we show in the next section, this evidence suggests that these policies target high valued innovations.

There is a concern that the results might be driven by a small number of leading universities which dominate patenting activity. In order to address this issue, we drop the top five universities in terms of their total number of patents, and re-estimate the baseline specification in column (4). These top universities, in descending order, are MIT, University of California at Berkeley, Stanford, California Institute of Technology, and the University of Wisconsin, and together they account for nearly a quarter of the citations in our sample. Nonetheless, when we drop these universities, the parameter estimates (reported in column 7) are very similar to those using the entire sample. This confirms that our key findings about the pattern of localization are robust, and are not driven by these top performers.

Finally, we also checked whether the geographic profile of knowledge spillovers changed over time. To do this, we re-estimate the baseline specification in column (4) for two sub-periods: 1976-

---

[19]There is a large empirical literature showing that such citation measures are correlated with measures of economic value (for extensive discussion, see Jaffe and Trajtenberg, 2002). We observe patents granted up to 2006 and citations through the year 2007, so there is an issue of truncation for the more recent patents. However, since we study the relationship between citation and distance, and not the number of citations per se, truncation would only cause a problem to the extent that the timing of citations is correlated with distance (e.g. earlier citations to a patent are from less distant inventors). Since that is possible, we checked robustness of results by re-estimating the specification in column (4) in Table 3, using only patents granted before 2000. The results are very similar to those in the table. For example, the coefficient on the within-state dummy is 0.096, which is nearly identical to the one from the full sample in column (4).

1993 and 1994-2006 (1993 is the median year for patent citations). The breakdown by period is based on the date of the cited patent, i.e. the 'vintage' of the technology, not the date at which the citation occurs. The results (not reported, for brevity) are broadly similar for the two sub-periods – specifically, we find no evidence that the degree of localization declined. In fact, the coefficients on the distance dummies shows somewhat stronger localization for the later period, but in both periods the distance gradient is essentially flat after 100 miles and the estimated state border effect is similar.

<center>**Insert Table 3 here**</center>

### 5.2. Public versus private universities and the border effect

In this section we examine the differences in knowledge diffusion from public and private universities. We begin by estimating the baseline specification separately for each ownership type, allowing for all coefficients to differ. Table 4 presents the results. A comparison of columns (1) and (2) shows that there is significantly stronger localization of knowledge spillovers for public universities. This takes two forms. First, patent citations drop off more sharply with distance for public universities. For example, moving from 0-25 to 25-50 miles reduces the citation probability by 26.5 percentage points for public institutions, and moving out to 50-100 miles further reduces it by another 3.2 percentage points (= 29.7-26.5). For private universities, the corresponding (incremental) declines are 17.0 and 8.4 percentage points, respectively. Yet for both types of universities, we observe that distance has no appreciable effect on citation beyond 100 miles.

The second important difference is that the state border much more strongly constrains knowledge diffusion for public universities – the estimates are 0.065 for public and 0.100 for private institutions. In column (3) we pool the two types of universities but continue to allow the distance gradient and state border effect to differ. This specification yields similar results – i.e constraining the other coefficients to be the same for public and private universities does not change our main conclusion that spillovers are more distance sensitive and more constrained by state borders for public universities. In this constrained version, the gap between the state border effect for public and private universities is even larger – 0.132 and 0.032, respectively.

In our sample, private universities tend to have somewhat higher quality patents as measured by the total number of subsequent citations received. The median number of patent citations is 37 (a median of 22) for public, and 48 (a median of 28) for private institutions. But the difference in

<center>17</center>

the border effect is not due to these differences in patent quality. To check this, we re-estimated the specification in column (3) separately for patents in the bottom and upper quartiles of the distribution of total citations received, our measure of patent quality. The results are presented in columns (4) and (5). For the lower quartile, the estimate of the border effect for public universities is 0.133 and essentially zero for private ones. For the upper quartile, the estimates are 0.157 and 0.096, respectively.

The evidence shows clearly that state borders are more important for public universities.[20] Does this reflect something intrinsic to ownership, or is it associated with university policy that is correlated with ownership? To examine this key question, in column (5) we add a control for the importance the university technology licensing office attaches to promoting local and regional development (interacted with the within-state dummy). This variable is only available for a subset of the universities (but they account for the majority of the sample patents), so the sample size drops by about a third. The results confirm that university policy matters for knowledge diffusion. The state border effect is more important when universities have strong local development objectives, and the size of the effect is large and statistically significant. For example, the point estimates imply that for a public university with strong local development objectives, the state border effect is 0.159. For a public university that places little weight on this objective, the border effect is reduced by about a quarter, to 0.128. For private universities, the corresponding state border effects are 0.097 and 0.066.

Nevertheless, while university policy makes a real difference to the degree of within-state knowledge spillovers, it only accounts for only about a third of the marginal effect of private ownership. To see this, note that the estimated coefficient on the private ownership (interaction) dummy in column (3) is -0.103, as compared to the point estimate of 0.031 on the dummy for strong local development objectives in column (7). It remains an important task for future research to understand more fully how private ownership affects university knowledge diffusion through other channels.

**Insert Table 4 here**

---

[20]In addition to the public-private distinction, we also examined whether the state border effect was different for land-grant universities. These are (mostly public) universities established by the federal government in the 19th century to promote research and technology diffusion. The coefficient on the interaction between land grant status and the within-state dummy was not statistically significant.

**5.3. State characteristics, policy and the border effect**

In the previous section we found that the size of the state border effect on patent citation differs between public and private universities, and is influenced by university technology transfer policy. In this section we examine the extent to which the border effect varies across states, and how state characteristics and policy affect it. Questions of interest include: Does the border matter less for knowledge diffusion in geographically larger states, like California and Texas? Does the intensity of high technology activity in the state affect the importance of the border? Do policies that affect intrastate mobility of scientists and engineers play a role?

We begin with the most flexible specification, allowing each state to have a different border effect.[21] Column (1) in Table 5 reports the results. The first thing to notice is that allowing for this general specification of the border effect does not materially affect our earlier results on the impact of distance. The citation probability declines sharply, and the effect of distance is exhausted after 100 miles. We find substantial variation in the estimated state border effects (not reported, for brevity), and decisively reject the null hypothesis that they are the same across states. The mean border effect is 0.218, but it ranges from a low of 0.024 to 0.472 (the standard deviation across states is 0.120).

Given the size of this variation, we want to understand the factors that determine when state borders are important for knowledge diffusion. We examine two main sets of factors: 1) factors that affect the in-state commercialization of university inventions, and 2) factors that influence the flow of local information. We explain these factors and their testable implications below.

*Commercialization hypothesis:* We expect that the state border effect will be stronger when the potential for in-state commercialization of university inventions is larger. This is more likely in states with a higher density of scientists and engineers who can potentially cite the university patent. However, controlling for the average density in the state, we expect the border effect to be smaller in states where the high-tech activity is concentrated at the location of the cited university, since this implies there are fewer potential citing inventors near the state border. To test these hypotheses, we use two variables interacted with the within-state dummy. The first is the number of scientists and engineers (S&E) per capita in the state (in 1995). The second is the Milken Institute 'TechPole' index of high-tech density in the city where the university is located (Devol and Wong, 1999). The index is a composite of the share of national high-tech real output and the concentration

---

[21]University fixed effects are also used in this regression, except in cases where there is only one university in a state.

of high-tech industries for each U.S. metropolitan area. In the patent citation equation, we expect a positive coefficient on the interaction with S&E density and a negative coefficient on the interaction with TechPole.

*Local information hypothesis:* The state border effect is simply the within-state citation bias controlling for distance. We expect this to be larger the more information that inventors have about the patents generated by the universities in the state. If information flows are in fact localized, the border effect should be stronger in states where 1) inventors are more likely to remain in the state when they move jobs (*'in-state mobility'*), and 2) inventors working in a state are more likely to have been educated (at the graduate level) in that state *'local education'*).[22]

To examine the in-state mobility hypothesis, one would like information on the probability that job movers among S&E remain within the same state. To our knowledge, no such information is available. Instead, we build on the recent literature on the economic impact of non-compete labor laws. These statutes restrict employees from taking jobs, for some period, with competing (same industry) companies within some geographic boundaries, typically the state. Exploiting the fact that the scope, and enforcement, of non-compete statutes vary across states, recent studies have shown that non-complete laws are associated with less intrastate job mobility, among other economic impacts (Marx et. al. 2007; 2010). We use the 'non-competition enforceability index' for each state constructed by Garmais (2009).[23]

To test the local education hypothesis, we need a measure of the fraction of S&E working in a state who were educated in that state. Unfortunately there is no information we are aware of that link the location of high-tech employees and their graduate education. The only available source is a single cross-sectional survey on new Ph.D graduates in the hard sciences conducted by the National Science Foundation (for details, see Sumell, Stephan and Adams, 2008). We use the percentage of new Ph.D. hires in a state who received their degree from universities in their state of employment (which we call S&E Inflow).[24] The samples in this survey are relatively small, and the variable is certainly measured with substantial error. The resulting attenuation bias will cause

---

[22] Of course, scientists who migrate out of state may maintain enduring professional links with local colleagues, and thus ongoing familiarity with and citation of, their research. In an interesting paper, Agrawal, Cockburn and McHale (2006) present evidence using patent citations that support this argument.

[23] This index is based on a count of twelve different dimensions of the scope and enforcement of these statutes (thus can range from zero to twelve). In the sample it from a low of zero (no enforcement) in California to a high of nine in Florida. We also tried the simple binary classification used by Marx et. al. (2007, 2010), but the empirical results were much weaker with this measure.

[24] The fraction of new hires educated in-state varies widely, from a low of 42.6 percent in Utah and Iowa to 91.7 in the District of Columbia and 81.5 percent in New Jersey.

us to underestimate the true impact of local education on the border effect. In addition to these variables, we include controls for the geographic size of the state and the level of economic activity (gross state product per capita).

The results are presented in column (1) in Table 5, and they are generally consistent with the hypotheses about the role of local information and commercialization (high-tech density).[25] Turning first to the commercialization hypotheses, we find that the density of educated inventors (scientists and engineers) in the state has a large impact on the state border effect. Evaluated at sample means, the point estimate of 0.110 implies that a one standard deviation increase in density raises the state border effect by 0.155, which is about 72 percent of the mean value of the individual state border effects (not reported in the table). Moreover, controlling for this density, we find that the state border effect is smaller when there is greater concentration of high-tech activity (TechPole) at the university location. A one standard deviation increase in TechPole (corresponding to a move from Phoenix to Boston) reduces the estimated border effect by 0.026, which is about 12 percent of the average border effect.

Turning to the role of local information, the results show that stronger enforcement of non-compete statutes is associated with less within-state knowledge spillovers, and the effect is large. The estimated coefficient of -0.014 implies that a one standard deviation increase in the enforcement index reduces the state border effect by 0.032. To put this another way, moving from a regime of complete non-enforcement (California, index=0) to the maximum enforcement state in our sample (Florida, index=9) reduces the border effect by 0.126, which is 57 percent of the average border effect. In addition, the evidence provides mixed support for our hypothesis that the border effect is larger in states with a larger fraction of locally educated scientists and engineers. The point estimate is negative, as predicted, but not statistically significant. However, when we include the control for the local focus of university technology transfer policy (column 2), the point estimate is both larger and statistically significant. While this suggests that retention of local university graduates increases the localization of knowledge spillovers, more research with better measures is needed to give a more conclusive answer.

In column (2) we add the control for local development objectives in the university technology

---

[25] As before, the standard errors are clustered at the patent level in these regressions. The state characteristics here are interacted with the dummy variable for whether the citation is within-the same state as the cited patent, so this regressor varies at the micro (citation) level. This makes it different from the case studied by Moulton (1990), where a micro regression includes an aggregate regressor that has no variation over a subset of micro observations, and thus requiring an adjustment to the standard errors.

licensing policy. As we found in Table 4 (where we did not include state characteristics and policies), university policy strongly affects in-state knowledge diffusion. The estimated coefficient of 0.039 implies that the state border effect is about 18 percent larger when universities have strong local objectives. The other results are robust (the coefficient on S&E density rises, but the difference is not statistically significant). Finally, controlling for these other factors, we find that larger states have smaller border effects. A one standard deviation increase in size reduces the border effect by about a third, evaluated at sample means. In addition, the border is less important in higher income states, but the effect is not large (a standard deviation change moves the border effect by 8 percent).

### 5.3.1. The Michigan 'Natural Experiment'

In the previous section we exploited the cross-state variation in characteristics and policy to identify the effects of interest. There is, of course, always the concern that unobserved state characteristics may be correlated with these variables, especially the enforcement of non-compete statutes. Fortunately, we are able to examine the impact of non-compete statutes on the state border effect in another way, by exploiting a policy reform in Michigan. Prior to 1985 Michigan outlawed non-compete agreements, but in 1985 it passed legislation that enforced them. In a series of recent papers, Marx et. al. (2007, 2010) exploit this reform as a 'natural experiment' and show that the introduction of non-compete legislation induced out-migration from Michigan, and that this was particularly strong for top-performing inventors. Building on their work, we use the Michigan reform to examine the effect of this statute on intrastate knowledge diffusion – i.e. on the importance of the state border effect on patent citation.

Specifically, we re-estimate the baseline specification with a full set of within-state dummies, allowing for a discontinuity in the border effect in Michigan after the reform. We would not expect an immediate impact of the reform, since labor mobility and new citing patents occur with some lag. To capture this, we estimate four distinct Michigan border effects: the pre-reform period (up to and including 1985), 1986-89, 1990-95 and post-1995. The prediction is that the state border effect should decline after the reform. The results are presented in column (3) in Table 5, and they confirm this prediction. We observe a sharp, and statistically significant, drop in the coefficient after 1989, and essentially no change thereafter. [26] Moreover, the estimated size of this shift in the state

---

[26] This conclusion holds up for different variants where we modify the timing of the dummies. We also estimated a specification that allows for different effects in each year during the period 1985-1990 and found similar (but noisier) changes.

border effect is consistent with the change implied by the estimates obtained from columns (1) and (2), where we identify the effect from the cross-state variation. Using the estimated marginal effect of the enforcement index in column (1), and assuming that Michigan moved from zero enforcement to the maximum level in the sample, we get an implied decline in the state border effect of 0.126. This is surprisingly close to the estimate using the natural experiment, which yields a decline of 0.121 (= 0.205 - 0.084).

As a further check, we conduct a set of 'placebo' tests by examining whether there is a similar effect in other states that did not introduce any reform. Finding an effect in those states would suggest that the change is being driven by some unobserved common factor other than the reform. We use three variants, based on different definitions of the placebo group of states. In column (4) we choose two neighboring states, Illinois and Indiana, in order to control for similar industrial structure (in particular, reliance on the automobile sector) and demand shocks. In column (5) we use the ten states whose individual, estimated state border effects were closest to the one for Michigan. Finally, column (6) treats the placebo group as all states other than Michigan. In each case, the states in the placebo group have their individual state border effects plus a common incremental effect for the different subperiods. In all three experiments, we find the large decline in the estimated border effect for Michigan, but no statistically significant drop for the placebo group of states. This gives us confidence that the Michigan reform did in fact have the impact we attribute to it.

**Insert Table 5 here**

### 5.4. Technology fields

The analysis to this point was based on pooling data for different technology areas. It is important to pin down whether our findings of localized knowledge spillovers is common to all fields, or are driven by only a few technology areas. Table 6 presents parameter estimates of the baseline specification for nine broad technology areas we constructed, based on the IPC patent class of the cited patent. These areas are: Biotechnology, Pharmaceuticals, Chemicals, Medical Instruments, Engineering, Electronics, Information Technology, and Telecommunications.[27]

We find substantial variation in the localization of knowledge diffusion across fields, both in terms of the distance gradient and the state border effect. While distance strongly mediates spillovers in all technology areas, localization is considerably less sharp in Biotechnology, Information Technology

---

[27] The international patent classes that are included in each technology field are given in the appendix.

and Telecommunications. The estimated coefficients on the distance dummies, up to 150 miles, are only about half as large for patents in these relatively younger fields, as compared to the more traditional areas. For example, citation declines by 15 percentage points (30 percent of the mean citation probability) after 100 miles for the newer fields, but by more than 25 percentage points for the others. At the same, however, the distance effects largely die out beyond 150 miles in *all* of the technology areas.

The second important finding is that the state border effect is not present in all fields. The statistically, and economically, significant state border effects are in the biomedical-related fields – Biotechnology, Pharmaceuticals, Chemicals and Medical Instruments. It remains an open question whether this finding is due to technology specialization in universities or state policies to promote local development of innovations from in-state universities. Either way, the technology field variation we observe implies that some of the variation we observe *across states* in the strength of the border effect may be attributable to differences in technology specialization.

<center>**Insert Table 6 here**</center>

## 5.5. Citations to university publications

Thus far we have traced knowledge spillovers by using citations to university patents. In this section we present a similar analysis using citations to university scientific publications. We are particularly interesting in knowing whether the geography of diffusion differs in 'open science' (publication) regime and proprietary (patent) knowledge regimes, as emphasized by Dasgupta and David (1994). It is worth bearing in mind, however, that our analysis can only partially inform on this distinction because we focus exclusively on citations to scientific publications by patents, not by other academic publications.

Table 7 presents the regressions results for the baseline specifications used for patent citations.[28] As with patents, we find that there is both negative distance effect and a positive state border effect on citations to publications. This holds both with the simple (log) linear specification of distance (column 1), and when use a more flexible specification for the distance effects (column 2). The estimated coefficients on the distance dummies are very similar to those obtained for citations to patents. The localization effects are pronounced – moving from 0-25 to 25-50 miles reduces the citation probability by -0.217, which is 40 percent of the mean citation rate. As with patent

---

[28]The technology match variable is omitted in these regressions as it does not apply to scientific publications.

citations, the distance effects die out after 150 miles (roughly the distance for channels of direct contact), by which point the citation probability has dropped by -0.286, or nearly 60 percent of the mean citation rate. This finding is striking, as it shows that information flows in 'open science' appear to be as constrained by distance as those from more proprietary knowledge in patents. It seems that the proprietary regime is not the dominant factor constraining information flows (though it may, of course, limit the way in which such information can be used). More fundamental factors, related to the channels of interaction and their relationship to distance, appear to be at work.

However, we find that knowledge spillovers from publication differ significantly from patents in the other dimension of localization – the role of the state border. In the pooled regression including both public and private universities (column 2), the estimated border effect for citations to publications is 0.034, which is much smaller than the coefficient of 0.089 for patents. This difference is even pronounced when we estimate the regression separately for public and private universities (columns 3 and 4). We find no statistically significant state border effect for publications from private universities, in sharp contrast to patents. Moreover, while the estimated impact is statistically significant for public universities, it is much smaller than its counterpart for patents (0.040 for publications, compared to 0.100 for patents, based on column (2) in Table 4).

We also examine whether the pattern of diffusion varies with the quality of scientific publications. We measure quality by the total number of (journal) citations received by the scientific publication, analogous to our measure for patents. Columns (5-7) present the parameter estimates for different quartiles of publication quality. We find that the distance gradients are quite similar for low and high quality publications, and they both have the characteristic that distance effects are exhausted after only 50 miles. The role of the state border, however, is very different for the proprietary knowledge regime of patents and the open science regime of publications. There is a statistically significant border effect only for low quality publications, and even here it is not large (the point estimate of 0.074 represents about 15 percent of the mean citation rate). For the other three quartiles the border effect is essentially zero (in the table we aggregate the middle two quartiles, but it holds for each separately too).[29]

**Insert Table 7 here**

---

[29] We also examined whether the geographic profile of knowledge spillovers changed over time. To do this, we re-estimated the baseline specification in column (2) for two sub-periods, 1976-1993 and 1994-2006, based on the date of the cited publication. The results for the two sub-periods are broadly similar – there is no evidence of any decline in localization for later scientific publications. This is consistent with our earlier finding for citation to patents.

## 6. Concluding Remarks

This study examines how geography, and university and state policies, affect knowledge spillovers from university innovation. We use patent citations both to university patents and scientific publications to trace these knowledge flows. Our main empirical findings are as follows. First, university knowledge spillovers are strongly localized. They are very sensitive to distance up to about 150 miles, and constant thereafter. Controlling for distance, we find strong evidence of a state border effect. Inventors located in the same state as the cited university are substantially more likely to cite one of the university patents than an inventor located outside the state. In sharp contrast, we find essentially no state border effect for patent citations to university scientific publications (except for the lowest quartile of quality). The distinction between the open science regime of scientific publications and the proprietary regime of patents seems to be important in regard to the geography of knowledge spillovers.

The state border effect is influenced by the characteristics and policies of the university and state. It is significantly larger public universities, and in particular those (both public and private) universities that pursue local and regional development in their technology licensing policies. The magnitude of the state border effect varies widely across states, and these variations are related to the high-tech density and state policy toward non-compete laws that constrain intrastate labor mobility. Finally, we show that there are differences across technology areas in how distance and state borders affect knowledge diffusion.

The important challenge for future research is to understand better the channels through which distance and state borders mediate knowledge spillovers. Is the border effect stronger for universities with greater technological specialization? What other aspects of state policy play a role – e.g. policies related to oversight and funding of public universities, provision of incentives to local companies using public university inventions, and so on? And, perhaps most importantly, what is the ultimate impact of intrastate knowledge spillovers on growth in the state?

# References

[1] Adams, James (1990), "Fundamental Stocks of Knowledge and Productivity Growth," *Journal of Political Economy*, 98(4), 673-702

[2] Adams, James (2002), "Comparative Localization of Academic and Industrial Spillovers," *Journal of Economic Geography*, 2: 253-278

[3] Adams, James and Zvi Griliches (1998), "Research Productivity in a System of Universities," *Annales D'Economie et de Statistique*, No. 49/50.

[4] Agrawal, Ajay, Iain Cockburn and John McHale (2006), "Gone But Not Forgotten: Knowledge Flows, Labor Mobility and Enduring Social Relationships," *Journal of Economic Geography*, 6(5): 571-591

[5] Agrawal, Ajay and Avi Goldfarb (2008), "Restructuring Research: Communication Costs and the Democratization of University Innovation," *American Economic Review* 98(4): 1578-90

[6] Almeida, Paul and Bruce Kogut (1999), "Localization of Knowledge and the Mobility of Engineers in Regional Networks," *Management Science*, 45(7): 905-917

[7] Alcacer, Juan and Michelle Gittelman (2006), "Patent Citations as a Measure of Knowledge Flows: The Influence of Examiner Citations," *Review of Economics and Statistics*, 88(4): 774-779

[8] Astebro, Thomas and Navid Bazzazian (2010), "Universities, Entrepreneurship and Local Economic Development," in M. Fritsch, ed., *Handbook of Research on Entrepreneurship and Regional Development* (Edward Elgar Publishing, forthcoming)

[9] Audretsch, David and Paula Stephan (1996), "Company-Scientist Locational Links: The Case of Biotechnology," *American Economic Review*, 86(3), 641-652.

[10] Belenzon, Sharon and Mark Schankerman (2009), "University Knowledge Transfer: Private Ownership, Incentives and Local Development Objectives," *Journal of Law and Economics*, 52(1): 111-144

[11] Dasgupta, Partha and Paul David (1994), "Toward a New Economics of Science," *Research Policy,* 23: 487-521

[12] Devol, Ross and Perry Wong (1999), America's High-Tech Economy: Growth, Development and Risks for Metropolitan Areas (Santa Monica: Milken Institute)

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

[14] Garmaise, M. (2009), "The Ties That Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment," *Journal of Law, Economics, and Organization*

[15] Henderson, Rebecca, Adam Jaffe and Manuel Trajtenberg (1998), "Universities as a Source of Commercial Technology: A Detailed Analysis of University Patenting, 1965-1988," *Review of Economics and Statistics*, 119-127

[16] Jaffe, Adam (1989), "Real Effects of Academic Research," *American Economic Review*, 79(5), 957-970

[17] Jaffe, Adam and Manuel Trajtenberg (1998), "International Knowledge Flows: Evidence from Patent Citations," *Economics of Innovation and New Technology*, 8: 105-136

[18] Jaffe, Adam and Manuel Trajtenberg (2002), *Patents, Citations and Innovations: A Window on the Knowledge Economy* (Cambridge, MA: MIT Press)

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

[20] Marx, Matt, Deborah Strumsky and Lee Fleming (2007), "Noncompetes and Inventor Mobility: Specialists, Stars and the Michigan Experiment," Harvard Business School paper

[21] Marx, Matt, Jasjit Singh and Lee Flemming (2010), "Regional Disadvantage? Non-Compete Agreements and Brain Drain," Harvard Business School paper

[22] Moulton, Brent (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units," *Review of Economics and Statistics,* 72(2): 334-338

[23] National Science Board (2008), *Science and Engineering Indicators* (Washington D.C.: National Science Foundation)

[24] Peri, Giovanni (2005), "Determinants of Knowledge Flows and their Effect on Innovation," *Review of Economics and Statistics*, 87(2): 308-322

[25] Singh, Jasjit, Matt Marx and Lee Flemming (2010), "Patent Citations and the Geography of Knowledge Spillovers: Disentanglig the Role of State Borders, Metropolitan Boundaries and Distance," Harvard Business School paper

[26] Sumell, Albert, Paula Stephan and James Adams (2008), "Capturing Knowledge: The Location Decision of New Ph.D.s Working in Industry," in Richard Freeman and Daniel Goroff, eds., *Science and Engineering Careers in the United States: An Analysis of Markets and Employment* (Chicago: University of Chicago Press, for the NBER): 257-87

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

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

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

[30] Zucker, Lynne, Michael Darby and Jeff Armstrong (1998), "Geographically Localized Knowledge: Spillovers or Markets?" *Economic Inquiry*, 36: 65-86

## A. Data Appendix

### A.1. Matching patents to universities

Our patent sample includes 3,309,736 patents that were granted between 1975 and 2007. Patents data are taken from the NBER patent file for the period 1975-2002 (2,630,106 patents). We directly extract from the USPTO website all granted patents for the period 2003-2007 (679,630 patents). We exclude patents that do not include at least one domestic assignee, losing 1,508,612 patents.[30] University patents can be assigned directly to the University, or to affiliate institutions. We manually explore the websites of all universities in our sample to identify the different legal entities to which the university patents can be assigned. For example, M. D. Anderson Cancer Center is an affiliate of the University of Texas. The matching procedure consists of the following steps:

1. Standardizing names of patent university assignees. This involves erasing phrases which comes before the name of the university, e.g. "The Boards of Regents of", "Trustees of", "A Governing Body of the", or after the name, e.g. "Research Foundation". As an example, the name "The Board of Trustees of The Stanford University - Office Of Technology" becomes "Stanford University".

2. Name matching: match the standard names of the patent applicants with our university sample, including the affiliated assignees we have identified for each university.

In total, we match 46,536 patents to 234 universities. The average number of patents per university is 211 but this varies widely, from a low of one for Oklahoma State University (Tulsa) to a high of 2,704 for MIT. The patents sample receives 408,155 citations. Of these citations, 19 percent do not include at least one American inventor and are thus excluded from the analysis.

#### A.1.1. Multiple assignments

**Co-assignees** In some cases, a patent has more than one assignee (72,714 patents in the complete sample patents). In case of co-assignment, we make the following assumptions. If the patent is assigned to two universities, then the patent is counted twice in our sample, once for each university. If the patent is assigned to a university and a company, then it is included in our sample as a university patent. Importantly, when selecting the random control sample, we ensure that the citing patent does not below to the same university or companies that are listed as co-assignees on the patent. Multiple assignments have important implications for the way we measure distance between the citing inventor and cited university. In the case of multiple assignments, we assume a citation from each assignee to the same university patent. We check the sensitivity of our results to different ways of dealing with co-assignments. We compute distance as the average, median, and maximum distance between the location of the citing inventors and cited universities. In all cases, the results are not sensitive to the way we deal with co-assignments.

#### A.1.2. Multiple campuses and central assignments

Patents may be assigned to a university system, rather than to a specific campus (e.g. University of California). In order to compute the correct distance between the inventor and the university, we

---

[30]The addresses of USPTO assignees may be ambiguous is certain cases; the address format limits either the US state name or the non-US country at 2 letters, e.g. "Los Angeles, CA" and "Toronto, CA". The ambiguity appears also for DE (Delaware/Germany), IL (Illinois/Israel), AR (Arkansas/Argentina) and IN (Indianapolis/India). We ensure we keep only US assignees by identifying the cities, and company prefix (e.g. GMBH firms are German and not from Delaware).

have to match the patent to the relevant campus. The matching procedure consists of the following steps: 1. We generate a list of the different campuses of the samples universities (e.g., University of Califronia-Berkeley, University of California-San Francisco etc.) where that information is available from the university websites 2. In cases where the relevant city is stated in the assignee address field rather than the city of the system's main campus, the patent is reassigned to the campus in that city. 3. The remaining 'system' patents are matched by the addresses of their inventors: the distance between each of the inventors which live in the local state to each of the university's campus is computed, and the closest university is affiliated to each inventor. In total, 12,116 patents were adjusted using this procedure (details available on request).

## A.2. Matching scientific publications to universities

Patent documents usually include citations to non-patent literature, such as scientific papers. In total, 365,205 patents cite non-patent literature (the average number of non-patent references is 4.7). We develop specialized extraction software that scans patent documents and systematically extracts the citations to non-patent literature section. We then match the articles to our university sample. The matching procedure is quite complex because the name of the university where the publication's authors are employed is almost never listed. To assign universities to publications, we use the Web of Science database by Thomson, which is the largest source of information on scientific publications in "hard-science" journals (covers more than 20 million records). These data include the publication title, authors, and university name (affiliation). We develop additional specialized software that extracts this information from the Web of Science articles where at least one of university in our sample appears in the affiliation field.

Having constructed this list of publications, we match the non-patent citations from the patents documents to the list of university publications. Identifying the title, author, journal name, and publication year out of the citation line is extremely difficult, as there are many different formats. We follow a similar procedure as we did for patent matching. However, here we apply more manual checks and rely less on generalized, automated rules. The following examples illustrate the varying formats of these citations:

1. Greenwalt et al., "Evaluation of fructose diphosphate in RBC preservation", Transfusion 42: 384-5 (2002).

2. Quality of Service Protocols Use a Variety of Complementary Mechanisms to Enable Deterministic End-to-End Data Delivery, QoS Protocols & Architectures, QoS Forum White Paper, Stardust.com, Inc., pp. 1-25, Jul. 8, 1999.

3. Swan, "Properties of Direct AVO Hydrocarbon Indicators", Offset-Dependent Reflectivity–Theory and Practice of AVO Analysis (Castagna, J.P. & Backus, M.M., eds., Soc. Expl. Geophys., 1993), pp. 78-92.

4. T.J. Kostas, M.S. Borella, I. Sidhu, G.M. Schuster, J. Mahler, J. Grabiec, "Real-time voice overpacket-switched networks," IEEE Network, vol. 12, No. 1, pp.1987, Jan./Feb. 1998.

5. A fast blind source separation for digital wireless applications Toriak, M.; Hansen, L.K.; Xu, G.; Acoustics, Speech, and Signal Processing, 1998.

Our matching algorithm tries to capture all the different variants in which citations may appear, by effectively running the matching procedure for a wide variety of possible formats. For example, we first assume citations appear, as in the first example above. We run the whole matching procedure according to this format, where the authors' names appear first, then the name of the article, followed by the journal where it was published (and year of publication in brackets). We then keep all unmatched citations, and repeat the matching by assuming all formats are as in the second example. For the unmatched citations, we proceed to the format in the third example, and so forth. The intensive manual checking is used to identify all possible formats in which citations can appear. We manually go over close to 75 percent of all citations to ensure we cover all possible citation structures.

The way authors' names are listed within different formats varies widely. The first example shows that names can be listed by indicating the last name of the first author followed by "et al." The fourth example, however, shows a case where all authors are listed by indicating their last names and their first initial. While the Web of Science database has less variation in the citation formats (which makes matching easier), citations in the patent document do not follow specific rules. Thus, when matching by authors' names we allow for a wide range of formats according to what we find in our vast manual inspection. For quality assurance, we manually checked the matched sample by comparing the full reference in the Web of Knowledge to the citation in the patent document. For a small percentage of the matched sample, we also checked that the publication record appears in the curriculum vitas of the authors, which were downloaded directly from their personal websites.

In total, we match 26,533 publications to our university sample.[31] To compute the distance between the citing inventor and cited university, we follow the same procedure as for patent citations. However, there is an important difference between matching citations to university patents and scientific publications. While the assignment of university patents tends to be complex, especially for public university that in some cases centrally assign patents, scientific publications do not have the same problem, as authors' affiliation is indicated at the university and campus level.

### A.3. Measuring geographic distance for citations

We develop specialized software that extracts driving distance information between city pairs directly from Google Maps (http://maps.google.com). We generate a list of all American cities and states (excluding Hawaii) that appear on all USPTO patent documents before selecting the sample of control patents. This list includes 33,127 citing inventor's cities. We add to this list the location of our sample of cited universities – 205 cities. Our distance software computes the distance for all city pairs.

### A.4. Definition of Patent Technology Fields (IPC codes)

Biotechnology: A01H C02F3/34 C07G11 C07G13 C07G15 C07K4 C07K14 C07K16 C07K17 C07K 19/00 C12M C12N C12P C12Q C12S G01N 27/327 G01N 33/53 G01N 33/54 G01N 33/55 G01N 33/57 G01N 33/68 G01N 33/74 G01N 33/76 G01N 33/78 G01N 33/88 G01N 33/92
    Chemicals: C0 C1 B01 D01F A62D (excluding Biotechnology)
    Pharmaceuticals: A61K, A61P

---

[31] Matches are dropped if one of inventors' names and one of the authors' names share the same family name, which might indicate that the inventor of the patents cites his own publication. This procedure is deliberately conservative in avoiding possible self-cites (which could give a false appearance of localized spillovers).

Medical Equipment: A61B, A61C, A61D, A61F, A61G, A61H, A61J, A61L, A61M, A61N, A01K, A01N

Engineering: A01B, A01C, B021 D21, B06B, B09, B21, B22, B23, B25, B29, B60, B62, B65, B81, B82, D01D, D02, D03, D04, D05, D06M, D21, E21, F04, F25, G05G, G07

Electronics: H01L H03 G11C G06C G06D G06E G06F11 G06F15 G06F17 G06G H01(excluding H01L) H02 H04 (excluding H04N H04L H04M) H05 B03C

Information Technology: G05B G05D G06F (excluding G06F17,G06F15,G06F11) G06J G06K G06N G06T G11B

Telecommunications: H04L H04M H04N

# Figure 1. Distance, State Borders and Patent Citation Probability
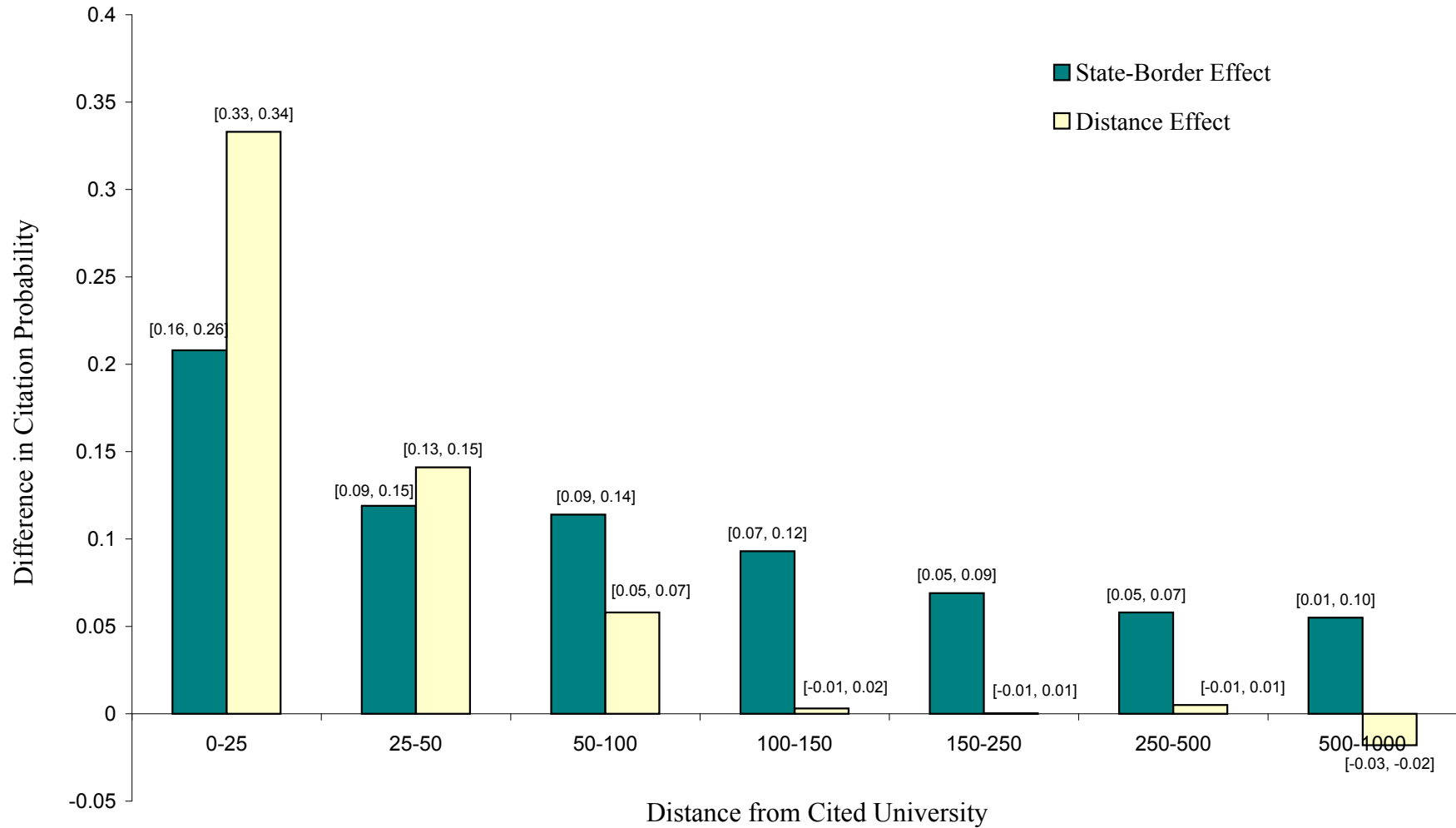


*Notes:* This figure presents the effect of distance and state-border on citations probability. For distance, each bracket compares the difference in the probability to cite between citing inventors that are located in that bracket, and all other inventors that are located at a greater distance from the University. For state-border, each bracket compares the difference in the probability to cite between inventors that are located in the University state, and inventors that are located in different states. The sample average probability of citations is 50 percent, by construction. 5 percent confidence intervals are reported for each distance bracket.

# Table 1. Descriptive Statistics on Geography of Citations to University Patents and Publications

| | *Panel A. Patents* | | | | | *Panel B. Scientific Publications* | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Share of citations | No. of universities | No. cited patents | No. citing patents | Share private universities | Share of citations | No. of universities | No. cited articles | No. citing patents | Share private universities |
| Dummy for within-state citation | 0.12 | 183 | 14,180 | 41,057 | 0.49 | 0.13 | 165 | 5,806 | 8,411 | 0.46 |
| Distance < 25 | 0.06 | 180 | 9,419 | 20,501 | 0.57 | 0.10 | 154 | 2,548 | 3,311 | 0.53 |
| 25 ≤ Distance < 50 | 0.03 | 131 | 4,075 | 8,390 | 0.48 | 0.03 | 82 | 720 | 960 | 0.45 |
| 50 ≤ Distance < 100 | 0.02 | 144 | 3,351 | 6,205 | 0.49 | 0.02 | 85 | 532 | 701 | 0.45 |
| 100 ≤ Distance < 150 | 0.02 | 147 | 3,281 | 5,604 | 0.56 | 0.02 | 88 | 448 | 528 | 0.47 |
| 150 ≤ Distance < 250 | 0.04 | 164 | 6,497 | 14,111 | 0.64 | 0.04 | 110 | 1,033 | 1,207 | 0.64 |
| 250 ≤ Distance < 500 | 0.11 | 182 | 13,374 | 38,438 | 0.54 | 0.11 | 156 | 2,744 | 3,265 | 0.53 |
| 500 ≤ Distance <1000 | 0.19 | 184 | 17,323 | 61,696 | 0.37 | 0.16 | 170 | 4,075 | 4,541 | 0.36 |
| 1000 ≤ Distance < 1500 | 0.15 | 184 | 14,952 | 47,143 | 0.36 | 0.13 | 167 | 3,330 | 3,653 | 0.32 |
| 1500 ≤ Distance < 2500 | 0.21 | 184 | 17,441 | 66,972 | 0.35 | 0.20 | 171 | 4,838 | 5,673 | 0.33 |
| Distance ≥ 2500 | 0.20 | 150 | 13,466 | 61,366 | 0.62 | 0.20 | 113 | 4,529 | 5,627 | 0.59 |

*Notes:* Distance refers to the mileage between the location of the citing inventor and the cited university. The values include only actual citations (not control group patents).The within-state dummy is one if the citing inventor resides in the same state as the cited university.

## Table 2. Distance and State Borders, by University Type and Patent/Publication Quality

| Universities: | Panel A: Patents | | | Panel B: Scientific Publications | | |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| | # Pairs | % Diff. in means Distance | % Diff. in means Within-State | # Pairs | % Diff. in means Distance | % Diff. in means Within-State |
| All | 383,096 | -6.9** | 53.4** | 35,043 | -8.3** | 44.5** |
| Private | 176,292 | -4.3** | 48.5** | 15,645 | -6.2** | 40.8** |
| Public | 206,804 | -9.3** | 56.7** | 19,398 | -10.1** | 47.1** |
| *Cites received ≤ 25th* | | | | | | |
| All | 98,495 | -13.1** | 62.5** | 8,911 | -3.3** | 46.5** |
| Private | 39,016 | -10.8** | 55.1** | 3,431 | -4.5** | 41.8** |
| Public | 59,479 | -14.8** | 68.1** | 5,480 | -8.5** | 49.1** |
| *Cites received ≥ 75th* | | | | | | |
| All | 95,435 | -1.0* | 45.1** | 8,680 | -6.8* | 39.9** |
| Private | 50,459 | -1.0 | 46.3** | 3,984 | -8.5 | 35.3** |
| Public | 44,976 | -3.4** | 42.8** | 4,696 | -5.4** | 45.3** |

*Notes:* Panel A reports mean comparison tests between cited and control patents for distance from, and fraction that are in the same state as, the cited university. Panel B reports the corresponding figures for scientific publications.  * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

## Table 3. Baseline Specifications for Citations to Patents

*Dependent variable: Citation Dummy (Marginal Effects)*

| University cited patents: | (1)<br>All | (2)<br>All | (3)<br>All | (4)<br>All | (5)<br>Cites received $\leq 25^{th}$ | (6)<br>Cites received $>75^{th}$ | (7)<br>Exc. Top Patenting Universities |
|---|---|---|---|---|---|---|---|
| Dummy Intra-State Citation | | 0.207**<br>(0.004) | 0.122**<br>(0.007) | 0.089**<br>(0.007) | 0.078**<br>(0.011) | 0.118**<br>(0.018) | 0.096**<br>(0.009) |
| log(Distance), Miles | -0.043**<br>(0.001) | | -0.024**<br>(0.001) | | | | |
| Match on 6-digit IPC | 0.307**<br>(0.003) | 0.306**<br>(0.003) | 0.306**<br>(0.003) | 0.303**<br>(0.003) | 0.323**<br>(0.004) | 0.282**<br>(0.009) | 0.301**<br>(0.004) |
| Dummy 25 ≤ Distance < 50 | | | | -0.202**<br>(0.009) | -0.247**<br>(0.013) | -0.130**<br>(0.022) | -0.247**<br>(0.011) |
| Dummy 50 ≤ Distance <100 | | | | -0.258**<br>(0.009) | -0.326**<br>(0.012) | -0.186**<br>(0.026) | -0.300**<br>(0.010) |
| Dummy 100 ≤ Distance <150 | | | | -0.284**<br>(0.009) | -0.320**<br>(0.013) | -0.259**<br>(0.024) | -0.318**<br>(0.010) |
| Dummy 150 ≤ Distance <250 | | | | -0.278**<br>(0.008) | -0.329**<br>(0.011) | -0.214**<br>(0.023) | -0.315**<br>(0.009) |
| Dummy 250 ≤ Distance <500 | | | | -0.282**<br>(0.007) | -0.341**<br>(0.010) | -0.215**<br>(0.020) | -0.317**<br>(0.009) |
| Dummy for 500 ≤ Distance <1000 | | | | -0.288**<br>(0.008) | -0.360**<br>(0.012) | -0.222**<br>(0.022) | -0.330**<br>(0.010) |
| Dummy 1000 ≤ Distance <1500 | | | | -0.282**<br>(0.008) | -0.351**<br>(0.012) | -0.197**<br>(0.023) | -0.322**<br>(0.010) |
| Dummy 1500 ≤ Distance < 2500 | | | | -0.277**<br>(0.009) | -0.357**<br>(0.012) | -0.190**<br>(0.023) | -0.315**<br>(0.010) |
| Dummy Distance ≥ 2500 | | | | -0.247**<br>(0.009) | -0.332**<br>(0.012) | -0.190**<br>(0.023) | -0.285**<br>(0.011) |
| University Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| High-tech Pair Dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.055 | 0.055 | 0.057 | 0.061 | 0.078 | 0.051 | 0.062 |
| Observations | 383,096 | 383,096 | 383,096 | 383,096 | 98,495 | 95,435 | 283,476 |

*Notes:* This table reports the results of Probit regressions relating the probability of citing a university patent and the distance of citing inventor from the cited university. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

## Table 4. Public and Private Ownership and the State Border Effect

*Dependent variable: Citation Dummy (Marginal Effects)*

| University cited patents: | (1) Private | (2) Public | (3) All | (4) Cites received ≤25th | (5) Cites received >75th | (6) Exc. Top Patenting Universities | (7) All |
|---|---|---|---|---|---|---|---|
| Dummy Intra-State Citation | 0.065** | 0.100** | 0.135** | 0.133** | 0.157** | 0.128*** | 0.128** |
| | (0.010) | (0.011) | (0.009) | (0.013) | (0.023) | (0.010) | (0.012) |
| Dummy Intra-State Citation × Dummy Private | | | -0.103** | -0.137** | -0.061** | -0.102*** | -0.062** |
| | | | (0.010) | (0.014) | (0.025) | (0.012) | (0.013) |
| Dummy Intra-State Citation × Dummy for High Local Objectives | | | | | | | 0.031** |
| | | | | | | | (0.011) |
| Match on 6-digit IPC | 0.290** | 0.313** | 0.303** | 0.323** | 0.282** | 0.301*** | 0.300** |
| | (0.005) | (0.004) | (0.003) | (0.004) | (0.009) | (0.004) | (0.004) |
| Dummy 25 ≤ Distance < 50 | -0.170** | -0.265** | -0.234** | -0.294** | -0.140** | -0.246** | -0.211** |
| | (0.011) | (0.012) | (0.012) | (0.016) | (0.033) | (0.016) | (0.014) |
| Dummy 50 ≤ Distance < 100 | -0.254** | -0.297** | -0.263** | -0.334** | -0.185** | -0.278** | -0.258** |
| | (0.013) | (0.012) | (0.013) | (0.015) | (0.038) | (0.014) | (0.015) |
| Dummy 100 ≤ Distance < 150 | -0.258** | -0.340** | -0.304** | -0.337** | -0.315** | -0.327** | -0.311** |
| | (0.013) | (0.011) | (0.012) | (0.016) | (0.030) | (0.012) | (0.015) |
| Dummy 150 ≤ Distance < 250 | -0.261** | -0.321** | -0.277** | -0.339** | -0.198** | -0.304** | -0.263** |
| | (0.011) | (0.011) | (0.011) | (0.013) | (0.037) | (0.011) | (0.014) |
| Dummy 250 ≤ Distance < 500 | -0.249** | -0.345** | -0.300** | -0.356** | -0.231** | -0.328** | -0.284** |
| | (0.010) | (0.010) | (0.009) | (0.012) | (0.025) | (0.010) | (0.010) |
| Dummy 500 ≤ Distance < 1000 | -0.267** | -0.348** | -0.289** | -0.373** | -0.191** | -0.323** | -0.268** |
| | (0.011) | (0.012) | (0.009) | (0.012) | (0.026) | (0.011) | (0.011) |
| Dummy 1000 ≤ Distance < 1500 | -0.252** | -0.345** | -0.287** | -0.367** | -0.156** | -0.322** | -0.266** |
| | (0.011) | (0.012) | (0.009) | (0.012) | (0.028) | (0.011) | (0.011) |
| Dummy 1500 ≤ Distance < 2500 | -0.250** | -0.343** | -0.281** | -0.368** | -0.163** | -0.313** | -0.261** |
| | (0.012) | (0.013) | (0.010) | (0.013) | (0.027) | (0.011) | (0.012) |
| Dummy Distance ≥ 2500 | -0.212** | -0.324** | -0.260** | -0.343** | -0.138** | -0.299** | -0.247** |
| | (0.012) | (0.013) | (0.009) | (0.012) | (0.025) | (0.011) | (0.011) |
| Interactions between Distance Dummies and Dummy for Private | No | No | Yes | Yes | Yes | Yes | Yes |
| University Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| High-tech Pair Dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.057 | 0.067 | 0.062 | 0.079 | 0.052 | 0.063 | 0.060 |
| Observations | 176,292 | 206,801 | 383,096 | 98,483 | 95,414 | 283,476 | 259,902 |

Notes: Local Objectives measures the weight the university attaches to local/regional development objectives in its licensing activity. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

## Table 5. Determinants of the State Border Effect

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| _Dependent variable: Citation Dummy (Marginal Effects)_ | | | | | | |
| | _State effects_ | | _Michigan "experiment"_ | | | |
| | | | | Controls are bordering states | Controls have similar border effect | All non-Michigan states |
| Dummy Intra-State Citation | 0.528** | 0.518** | | | | |
| | (0.028) | (0.044) | | | | |
| Dummy Intra-State Citation × Dummy Private | -0.046** | -0.049** | | | | |
| | (0.012) | (0.016) | | | | |
| Dummy Intra-State Citation × Dummy for High Local Objectives | | 0.039** | | | | |
| | | (0.015) | | | | |
| Dummy Intra-State Citation × Techpole | -0.005** | -0.006** | | | | |
| | (0.001) | (0.001) | | | | |
| Dummy Intra-State Citation × S&E Density | 0.110* | 0.247** | | | | |
| | (0.058) | (0.076) | | | | |
| Dummy Intra-State Citation × Index of Non-Compete Laws | -0.014** | -0.016** | | | | |
| | (0.004) | (0.005) | | | | |
| Dummy Intra-State Citation × S&E Inflow | -0.001 | -0.003** | | | | |
| | (0.001) | (0.001) | | | | |
| GSP/Capita | -0.022** | -0.016** | | | | |
| | (0.003) | (0.005) | | | | |
| State Size | -0.024** | -0.013 | | | | |
| | (0.004) | (0.008) | | | | |
| _Dummy for Michigan × :_ | | | | | | |
| Dummy for Pre-1985 | | | 0.205 | 0.233 | 0.239 | 0.232 |
| | | | (0.147) | (0.141) | (0.141) | (0.141) |
| Dummy for 1986-1989 | | | 0.328** | 0.341 | 0.312** | 0.341** |
| | | | (0.098) | (0.093) | (0.105) | (0.093) |
| Dummy for 1990-1995 | | | 0.084* | 0.113** | 0.076 | 0.113** |
| | | | (0.043) | (0.042) | (0.045) | (0.042) |
| Dummy for Post-1995 | | | 0.067** | 0.097** | 0.030 | 0.097** |
| | | | (0.021) | (0.021) | (0.023) | (0.021) |
| _Dummy for Control States × :_ | | | | | | |
| Dummy for Pre-1985 | | | | 0.165* | -0.008 | 0.095** |
| | | | | (0.078) | (0.030) | (0.030) |
| Dummy for 1986-1989 | | | | 0.173 | 0.046* | 0.131** |
| | | | | (0.099) | (0.021) | (0.025) |
| Dummy for 1990-1995 | | | | 0.201** | 0.038* | 0.133** |
| | | | | (0.053) | (0.016) | (0.021) |
| Dummy for Post-1995 | | | | 0.240** | 0.068** | 0.148** |
| | | | | (0.039) | (0.013) | (0.020) |
| $R^2$ | 0.063 | 0.061 | 0.062 | 0.062 | 0.062 | 0.062 |
| Observations | 383,096 | 259,902 | 383,096 | 383,096 | 383,096 | 383,096 |

_Notes:_ This table reports the results of Probit regressions of the determinants of the state border effect for citations to university patents. TechPole is a measure of high-tech density constructed by the Milken Institute (Devol and Wong, 1999). All columns include a complete set of distance dummies, and a dummy for same IPC. The control states in column 4 are IN and IL. The control states in column 5 are NY, PA, MA, CA, NJ, MI, WA, MD, MS, and CT. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

# Table 6. The Effects of Distance and State Border on Patent Citations, by Technology Area

## Dependent variable: Citation Dummy (Marginal Effects)

| Technology area: | (1) Biotechnology | (2) Chemicals | (3) Pharma | (4) Medical Equipment | (5) Engineering | (6) Electronics | (7) Information Technology | (8) Telecommuni-cations |
|---|---|---|---|---|---|---|---|---|
| Dummy for Intra-State Citation | 0.168** | 0.151** | 0.125** | 0.101** | 0.059** | 0.015 | 0.029 | 0.037 |
| | (0.024) | (0.022) | (0.022) | (0.021) | (0.016) | (0.020) | (0.036) | (0.046) |
| Matched on six-digit IPC | 0.346** | 0.278** | 0.345** | 0.288** | 0.370** | 0.265** | 0.203** | 0.255** |
| | (0.011) | (0.009) | (0.009) | (0.008) | (0.007) | (0.008) | (0.012) | (0.017) |
| Dummy 25 ≤ Distance < 50 | -0.115** | -0.289** | -0.223** | -0.166** | -0.187** | -0.204** | -0.130** | -0.108* |
| | (0.031) | (0.021) | (0.024) | (0.026) | (0.020) | (0.024) | (0.049) | (0.049) |
| Dummy 50 ≤ Distance < 100 | -0.145** | -0.312** | -0.302** | -0.228** | -0.263** | -0.273** | -0.165** | -0.123 |
| | (0.046) | (0.023) | (0.026) | (0.025) | (0.020) | (0.026) | (0.045) | (0.087) |
| Dummy 100 ≤ Distance < 150 | -0.187** | -0.334** | -0.320** | -0.246** | -0.282** | -0.330** | -0.137** | -0.156** |
| | (0.035) | (0.022) | (0.026) | (0.027) | (0.021) | (0.022) | (0.061) | (0.068) |
| Dummy 150 ≤ Distance < 250 | -0.210** | -0.332** | -0.307** | -0.239** | -0.301** | -0.287** | -0.125** | 0.008 |
| | (0.029) | (0.020) | (0.024) | (0.024) | (0.017) | (0.023) | (0.052) | (0.060) |
| Dummy 250 ≤ Distance < 500 | -0.181** | -0.333** | -0.309** | -0.242** | -0.284** | -0.311** | -0.158** | -0.121** |
| | (0.029) | (0.020) | (0.023) | (0.023) | (0.016) | (0.019) | (0.039) | (0.046) |
| Dummy 500 ≤ Distance < 1000 | -0.181** | -0.362** | -0.323** | -0.229** | -0.298** | -0.322** | -0.170** | -0.109 |
| | (0.031) | (0.022) | (0.025) | (0.025) | (0.018) | (0.022) | (0.043) | (0.060) |
| Dummy 1000 ≤ Distance < 1500 | -0.164** | -0.339** | -0.268** | -0.222** | -0.294** | -0.334** | -0.192** | -0.083 |
| | (0.034) | (0.022) | (0.027) | (0.025) | (0.019) | (0.021) | (0.042) | (0.063) |
| Dummy 1500 ≤ Distance < 2500 | -0.174** | -0.324** | -0.268** | -0.202** | -0.293** | -0.341** | -0.166** | -0.122* |
| | (0.032) | (0.023) | (0.030) | (0.026) | (0.020) | (0.022) | (0.045) | (0.061) |
| Dummy Distance ≥ 2500 | -0.154** | -0.290** | -0.234** | -0.185** | -0.273** | -0.304** | -0.131** | -0.064 |
| | (0.033) | (0.025) | (0.030) | (0.027) | (0.020) | (0.024) | (0.044) | (0.063) |
| $R^2$ | 0.079 | 0.074 | 0.078 | 0.059 | 0.078 | 0.051 | 0.026 | 0.036 |
| Observations | 25,804 | 45,778 | 35,257 | 70,427 | 62,851 | 43,065 | 16,118 | 8,076 |

*Notes:* This table reports the results of Probit regressions relating the probability of citing a university patent and the distance of the citing inventor from the cited university. All regressions include a complete set of university dummies and high-tech pair dummies. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

## Table 7. Baseline Specifications for Scientific Publications

*Dependent variable: Citation Dummy (Marginal Effects)*

| University cited publications: | (1) All | (2) All | (3) Private | (4) Public | (5) Cites received ≤25th | (6) Cites received 25th-75th | (7) Cites received >75th |
|---|---|---|---|---|---|---|---|
| Dummy Intra-state citation | 0.048** | 0.034** | 0.021 | 0.040* | 0.074** | 0.033 | 0.004 |
| | (0.012) | (0.013) | (0.018) | (0.018) | (0.026) | (0.017) | (0.027) |
| log(Distance), Miles | -0.036** | | | | | | |
| | (0.002) | | | | | | |
| Dummy 25 ≤ Distance < 50 | | -0.217** | -0.166** | -0.302** | -0.202** | -0.240** | -0.193** |
| | | (0.015) | (0.021) | (0.018) | (0.030) | (0.020) | (0.032) |
| Dummy 50 ≤ Distance < 100 | | -0.247** | -0.195** | -0.331** | -0.253** | -0.275** | -0.194** |
| | | (0.017) | (0.025) | (0.021) | (0.031) | (0.021) | (0.048) |
| Dummy 100 ≤ Distance < 150 | | -0.286** | -0.258** | -0.346** | -0.261** | -0.316** | -0.252** |
| | | (0.016) | (0.024) | (0.019) | (0.036) | (0.020) | (0.034) |
| Dummy 150 ≤ Distance < 250 | | -0.277** | -0.225** | -0.356** | -0.250** | -0.293** | -0.274** |
| | | (0.014) | (0.020) | (0.017) | (0.031) | (0.019) | (0.028) |
| Dummy 250 ≤ Distance < 500 | | -0.299** | -0.242** | -0.385** | -0.273** | -0.320** | -0.287** |
| | | (0.012) | (0.018) | (0.014) | (0.028) | (0.016) | (0.026) |
| Dummy 500 ≤ Distance < 1000 | | -0.313** | -0.261** | -0.399** | -0.271** | -0.340** | -0.300** |
| | | (0.014) | (0.020) | (0.018) | (0.031) | (0.018) | (0.029) |
| Dummy 1000 ≤ Distance < 1500 | | -0.307** | -0.270** | -0.385** | -0.284** | -0.326** | -0.294** |
| | | (0.014) | (0.020) | (0.018) | (0.030) | (0.019) | (0.028) |
| Dummy 1500 ≤ Distance < 2500 | | -0.314** | -0.257** | -0.406** | -0.276** | -0.341** | -0.299** |
| | | (0.014) | (0.021) | (0.019) | (0.032) | (0.019) | (0.031) |
| Dummy Distance ≥ 2500 | | -0.279** | -0.215** | -0.381** | -0.231** | -0.308** | -0.269** |
| | | (0.015) | (0.021) | (0.018) | (0.032) | (0.020) | (0.031) |
| University Fixed Effects | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| High-tech Pair Dummies | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| $R^2$ | 0.014 | 0.018 | 0.015 | 0.023 | 0.018 | 0.021 | 0.014 |
| Observations | 70,086 | 70,086 | 31,290 | 38,784 | 17,822 | 34,904 | 17,360 |

*Notes:* This table reports the results of Probit regressions relating probability of citing a university scientific publication and the distance of the citing inventor from the cited university. Standard errors (in brackets) are clustered by cited publication. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.