Spreading the Word: Geography, Policy and Knowledge Spillovers*

Sharon Belenzon
Fuqua School of Business, Duke University

Mark Schankerman
London School of Economics and CEPR

January 9, 2012

Abstract

Using new data on citations to university patents and scientific publications, we study how geography affects university knowledge spillovers. Citations to patents decline sharply with distance up to about 150 miles and are strongly constrained by state borders. Distance also constrains citations to scientific publications, but the impact is less sharp and persists over greater distances. The state border effect for publications is significant only for lower quality public universities. We show that the state border effect is heterogeneous, and is strongly influenced by university and state characteristics and policies. It is larger for public universities and those with strong local development policies. The border effect is larger in states with strong non-compete laws that facilitate intrastate labor mobility, states with greater reliance on in-state educated scientists and engineers, and states with lower rates of interstate scientific labor mobility. We also confirm the impact of non-compete statutes by studying a policy reform in Michigan.

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

JEL Classification: K41, L24, O31, O34

*We would like to thank two referees and the Editor for very constructive suggestions on an earlier version of the paper. We also thank Ajay Agrawal, Saul Lach, Jacques Mairesse, John Van Reenen and participants in the 2009 NBER Summer Institute for helpful comments. We gratefully acknowledge the financial support of the Centre for Economic Performance at the London School of Economics, and the Muzzy Chair for Entrepreneurship at the University of Arizona, which made it possible for us to construct the data sets used in this project. Finally, we thank Hadar Gafni and Kristy KyuSeon Lee for excellent research assistance during this project.

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 predicated on the assumption that knowledge spillovers are geographically localized and best exploited by agglomerating high technology activity. Thus it is important to understand how geography, and the characteristics and policies of universities and states, constrain knowledge spillovers.[2]

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 spillovers.

We focus on two channels through which state borders can affect university knowledge diffu-

[1] There is substantial evidence of R&D spillovers (e.g., Jaffe, 1989; Adams, 1990; Jaffe and Trajtenberg, 2002). 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). 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 (Adams, 2002; Agrawal and Goldfarb, 2008). The results in our paper point to an important role for labor mobility and policies that influence it.

sion: local information, and policies for commercializing university innovation. The first channel is important when dealing with tacit knowledge, which 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, and this encourages development of local information networks. 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 are more likely to have been educated at a local (in-state) university. State policies can influence the prevalence of such local information by affecting scientific labor mobility. One important example is 'non-compete' labor laws, which make it more likely that inventors who shift employers will leave the state. In addition, a variety of university and state policies can affect the retention of locally educated scientists and engineers.[3]

The second channel involves policies that promote local commercial development of university innovations. 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 that are often constrained, and informally influenced, by state government in their technology licensing decisions. One important manifestation is that public universities typically attach greater importance 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. Many 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

[3] Sumell, Stephan and Adams (2008) document that U.S. states differ widely in the extent to which locally educated scientists and engineers remain in-state for their first job, and they show that this varies with university and state characteristics.

[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.

borders, especially if the border effect is significantly influenced by technology commercialization policies.

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

[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). Another form of local internalization is highlighted by Mowery and Ziedonis (2001), who show that market-mediated spillovers in the form of licensing of university inventions tend to be strongly localized in nature.

[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 Fleming (2010) document a persistent state border effect while controlling for refined distance measures.

level of 'threshold distance' – corresponding as it does to an extended commuting distance – strongly 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 contrast, we find that state borders have essentially no impact on citations by patents to university scientific publications, except for lower quality public universities.

However, the impact of state borders on patent citations varies widely across states, and we show that this variation is consistent with the predictions of 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 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 studied by Marx et. al. (2007, 2010), who show that non-compete laws 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.[8]

Second, we show that the border effect is stronger in states which have a higher fraction of inventors educated at in-state universities, a greater density of scientists and engineers, and lower rates of interstate labor mobility for scientists and engineers.

Third, the state border effect is much stronger for citations to patents from public (as compared to private) universities, even after controlling for the academic quality of the university. 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

[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).

[8] This finding is consistent with earlier work by Almeida and Kogut (1999), who document the link between patent citations and labor mobility.

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.

Finally, we examine how localization of knowledge spillovers varies across technology areas, reflecting differences in the importance of tacit knowledge and the associated channels of information transmission. 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 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. 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.

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 which we will use as an additional control variable.[9] The final data set includes 26,914 university patents granted during the period 1975-2006. These patents receive 191,043 citations from patents that have at least one American-based inventor. With a matched (non-citing, control) patent for each of these, the final data set has 382,086 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 (or individuals) 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 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 Office.

[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.

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 publication. In total, 365,205 patents in the complete sample make at least one citation to academic publications. Of these citations, 34,714 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 69,428 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 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, 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 impact of this policy variable, it will be important to control

[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).

for university ownership status in the regressions.

Finally, we construct measures of university quality based on NSF rankings of 23 different academic departments in the hard sciences (National Research Council (1993); for details, see Lach and Schankerman, 2008). 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.

We use linear probability models that relate a dummy variable for whether the patent or publications is cited to a set of control variables.[11] 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

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

where $C_{i(u,s),j(s')}$ is a dummy variable equal to one if patent j located in state $s\prime$ cites a patent (or publication) i from university u located in state s,and zero otherwise. 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}, equal to one if the citing patent is in the same state as the cited patent (or publication), 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, η_u. We compute standard errors clustered at the level of the cited patent, which allows the disturbance ε_{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

[11] The marginal effects implied by Probit models for all of the main specifications are nearly identical to those from the linear probability model (LPM). We use the LPM because it is computationally much easier to accommodate large number of fixed effects. In addition to the fixed effects for universities and high-tech cluster pairs used in all specifications, we later introduce a complete set of technology field-state interaction dummies.

– geographic distance of the citing patent, university and state level characteristics, and university local development focus – are exogenous factors, unrelated to the disturbance ε_{ij} in the citation equation. The main concern is unobserved quality of a patent, which might affect both the probability that it is cited and 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* the true localization effects, i.e., 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, and interpret this result as saying 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 cannot 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, we will be able to reject the null hypothesis that the state border effect is solely driven by endogeneity because we show that it varies systematically with both university and state policies and characteristics. If the state border effect were driven only by '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 (holding constant both patent and university quality) and for universities that are located in states that more strongly enforce 'non-compete' labor laws.

Turning to the key control variables, we measure the distance between the inventor(s) of the citing patent and the university whose patent with a flexible specification that allows for nonlinear effects of distance. Specifically, we use a set of nine dummy variables for intervals of distance (in miles): 25-50, 50-100, 100-150, 150-250, 250-500, 500-1000, 1000-1500, 1500-2500 and greater than 2500; the reference category is 0-25 miles (which might be interpreted as a metropolitan effect). This specification is more flexible than existing studies (e.g., Peri, 2005) and we choose it for two reasons. First, it is of interest to know how the impact of distance on knowledge spillovers dissipates because it may give insight into how information diffuses. 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. The second reason to use a flexible specification of distance is to avoid any risk of confounding the effects of distance and the state border.

To examine the effect of state borders on citation, 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 control for distance in flexible way, the within-state dummy will identify whether there is any pure 'border effect' on knowledge spillovers.

Of course, the probability a university patent (or scientific publication) is cited may depend on a variety of university characteristics, including the quality and visibility of its faculty, entrepreneurial orientation, and high-tech density and specialization of the university location, as well as the university policies for promoting technology transfer and academic interaction (conference attendance, consulting activities and so on). To capture these factors in a flexible way, we introduce a complete set of university fixed effects for the cited patent.[13]

Finally, we include a complete set of dummy variables for 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 twenty 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.[14]

[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] 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, quality and other characteristics of the university and state to interact with geographic distance and/or the state border effect.

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

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). On average 12 percent of citations originate from the same state as the inventor, but the share varies widely across patents (from 0 to 100 percent). The average distance between citing and cited patent is 1,218 miles (not reported), but citations are geographically concentrated – overall, 15 percent of all citations originate within 150 miles, and 29 percent within 500 miles, from the cited university patent.[15] At the same time, 53 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

Table 2 presents non-parametric comparisons of citing and control patents (Panel A) and scientific publications (Panel B). Column (2) in Panel A compares 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 important conclusions emerge. 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. This confirms that distance constrains university knowledge spillovers. Second, the degree of localization is more than twice as large for public institutions than for private ones – the differences are –9.2 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 the degree of localization is much smaller, and for private universities there is no statistically significant localization.

Insert Table 2 here

[15] The distribution of citations across different distance intervals is similar for public and private universities (not shown).

Column (3) compares citing and control patents in terms of the fraction of citations originating from within-state inventors, another dimension of localization. The pattern is broadly similar to those in column (2). 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, this within-state citation bias is stronger for public universities than for private ones, and it is more pronounced for 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 striking, 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 5, where we discuss the more detailed econometric results.

Figure 1a provides evidence on how the effects of distance on knowledge spillovers dissipate 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 that university knowledge spillovers are strongly localized. For example, the first 'distance bar' shows that the probability that an inventor within 25 miles cites a university patent is 34 percentage points greater than for inventors located beyond 25 miles from the university. Since the mean citation probability is 50 percent by construction, this effect is huge – equivalent to a 65 percent decline in the mean citation probability. There is a further steep decline 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. After that, it appears that distance exerts no further effect.[16]

Figure 1b shows the impact of state borders on citation, for different intervals of distance between the citing and cited patents. The dark colored bars depict the difference between the citation

[16] The last bar suggests 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 are at these greater distances from each other (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.

probability for inventors located within the same state as the university and those outside the state, for each distance interval. For example, the first bar shows that inventors located within the state and within 25 miles of the university are 22 percentage points more likely to cite than inventors located at that distance but outside the state border. In contrast to the distance gradient, Figure 1b shows that the impact of the state border persists over much longer distances (the maximum within-state distance is 707 miles, in California). This finding is consistent with the hypothesis that the state border effect is determined (at least in part) by university and/or state policies, whose effects we would not expect to disappear with distance.[17]

Insert Figures 1a and 1b here

5. Estimation results

5.1. State-border effect

Table 3 presents the baseline linear probability 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. Standard errors are clustered at the level of the cited patent.

Column (1) presents the specification with the dummy variables for different distance intervals, but no within-state dummy. The results show that geography sharply constrains knowledge spillovers. Moving from 0-25 to 25-50 miles reduces the citation probability by 16.5 percentage points (which is about a third of the mean citation rate), and moving out to 50-100 miles further reduces it by another 8.7 (= 25.2 - 16.5) percentage points. There is a further drop of 5.0 percentage points up to 150 miles, but thereafter distance has no appreciable effect on citation. These econometric results confirm what we observe in Figure 1a.

In column (2) we report a specification with the within-state dummy but without distance effects. The results show that citation is much more likely from inventors located within the same state – the marginal effect of being within-state (0.196) is very large, nearly 40 percent of the mean citation rate. However, the estimated state border effect is likely to be overstated because we do not control for pure distance effects in this regression.

[17] While there is some variation in the border effect at different distance intervals, this reflects the fact that these intervals correspond to different subsets of states (it depends on the distance of the university from the border) and, as we show later, there is substantial heterogeneity in the state border effects across states.

Insert Table 3 here

To address this, in column (3) we introduce both nonlinear distance effects and the within-state dummy. We refer to this as the baseline specification. Two key findings emerge. First, both distance and the state border effect are statistically significant, and it is important to include both variables in the specification. Including flexible distance effects reduces the estimated effect of the state border by more than 50 percent, from 0.196 to 0.084, but this still represents nearly 20 percent of the mean citation probability. This result confirms that the state border effect is not simply a proxy for geographic distance. At the same time, the estimated distance effects are robust to allowing for a state border effect, and they confirm that the impact of distance dies out after about between 100-150 miles.[18]

Second, the coefficient on the technology matching dummy is large and statistically significant, confirming that citation is much more likely between patents in the same technology area. This is not surprising, but the interesting fact is that we still find strong geographic localization even after controlling for this matching dummy at the disaggregated 6-digit IPC level. This finding suggests that localization is not just a reflection of the spatial distribution of technological activity.[19] This conclusion is robust across all specifications we estimate.

To investigate the impact of technological clustering on our findings more fully, we introduce into the baseline specification a complete set of fixed effects for technology field- state pairs. We use the 3-digit level of aggregation for this purpose, which gives us a total of 3,995 dummy variables. This is the most flexible way of allowing for state-specific technological specialization in order to check whether the distance or state border effects are simply reflections of such specialization. Column (4)

[18] We conducted two additional robustness checks on these baseline findings. First, we re-estimated the baseline specification with a more refined distance breakdown of the 0-100 mile bracket using ten mile intervals. The key findings of localized distance effects and the importance of the state border are robust. The main additional insight from this exercise is that the distance gradient appears to flatten out at somewhere between 70 and 100 miles, but given the standard errors we would not make too much of this difference.

Second, we checked whether the results are robust to a different rule for selecting the control group of patents. Instead of choosing only one random control (non-citing) patent for each citation to a university patent in our sample, we randomly selected five control patents for each citation and re-estimated the baseline specification. The estimated distance effects are broadly similar and have the same pattern, dissipating after 150 miles. The estimate (standard error) of the state border effect is 0.052 (0.004), which is about 30 percent of the mean citation rate for this sampling frame (with one citing patent and five control patents).

[19] If localization were driven by agglomeration based on technological specialization, we would expect to find 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). As an additional robustness check, we re-estimated the baseline specification with a full set of 6-digit IPC fixed effects – not just a matching dummy at this level of aggregation. The estimates of the distance and state border effects are very similar to those reported in Table 3.

presents the results. The estimated coefficients for the state border effect and distance dummies are almost unchanged (compare columns 3 and 4). This important result shows that the localization of knowledge spillovers is not an artifact of state-specific technological agglomeration. We maintain these technology-state dummies in all the subsequent regressions.[20]

Finally, 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 5) 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-estimated the baseline specification in column (4) for two sub-periods: 1976-1993 and 1994-2006 (1993 is the median year for patent citations). We do this in two ways: first, using the date of the cited patent, i.e. the 'vintage' of the technology; and second, using the date at which the citation occurs (the second approach is designed to check for changes in diffusion associated with improvements in information technology and the internet). The point estimates of the state border effect are similar, and not statistically different, between the two periods (0.080 versus 0.087 using the cited patent to date; 0.061 versus 0.086 using the citing patent to date). The coefficients on the distance dummies show somewhat stronger localization for the later period, using both dating methods, but in both periods the distance gradient is essentially flat after 150 miles.[21] Overall, we do not find evidence that the degree of localization changed substantially over time.

5.2. Public and private universities and the state 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

[20] We go one step further by allowing these fixed effects for technology field-state pairs to change over time, by including dummies at the technology-state-year level (using the grant year of the citing patent). The results on the state border effect, and the distance gradients, remain robust (results not reported for brevity).

[21] For example, using the date of the cited patent to split the sample, for the pre-1993 period the coefficients (s.e.) on the first three distance intervals are -0.152 (.012), -0.187 (.016) and -0.224 (.016). The corresponding numbers for the later period are -0.158 (.011), -0.245 (.014) and -0.272 (.014). There is even less difference between the two periods when we use the citing year of the patent.

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 19.1 percentage points for public institutions, and moving out to 50-100 miles further reduces it by another 3.9 percentage points (= 23.0-19.1). For private universities, the corresponding incremental declines are 13.9 and 9.4 percentage points, respectively. Yet for both types of universities, we observe that distance has no appreciable effect on citation beyond 150 miles. The second important difference is that the state border more strongly constrains knowledge diffusion for public universities – the estimates are 0.065 for public and 0.102 for private institutions.

Insert Table 4 here

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 (0.125 and 0.045, respectively).

One concern is that the localization of knowledge diffusion for patents representing important advances may be very different than for marginal improvements. In particular, we would expect important ideas to diffuse more widely. Moreover, in our sample private universities tend to have somewhat higher quality patents, as measured by the total number of subsequent citations received (mean numbers of patent citations for public and private universities are 37 and 48, respectively). Thus part of the difference in the state border effects we find might also reflect these differences in patent quality.

To check this hypothesis, we re-estimate 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.[22] The results are presented in columns (4) and (5). Three conclusions are worth noting. First, the estimated coefficients on the distance dummies show a sharper distance gradient for the

[22] There is a large empirical literature showing that such citation measures are correlated with measures of economic value (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

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 (-0.207 versus -0.114). This finding confirms that knowledge diffuses more widely for important patents. But it also interesting that for both categories of patents, the effect of distance dies out relatively quickly.

Second, the state border effect is larger for public universities, both for low and high value patents. For the lower quartile, the estimate of the border effect for public universities is 0.113 and 0.017 for private ones. For the upper quartile, the estimates are 0.158 and 0.102, respectively.

Third, while distance constrains knowledge diffusion more strongly for low value patents, the state border is more important for high value patents, both for public and private universities. If the state border effect is, at least partly, due to a local preference in university technology commercialization policies, as we show in the next section, this evidence suggests that universities target high valued innovations for local development.

As a further check, we investigate whether the quality of the university (as opposed to the specific patent) partly explains the localization of knowledge spillovers, and the difference between public and private institutions. If there is imperfect information about the quality of the individual patent, the quality of the institution with which the patent is affiliated might be an informative signal and thus affect diffusion. To do this, we construct quality measures of each university using data from the National Research Council (1993) on the academic quality of individual departments in the hard sciences and aggregating them to the university level.[23] Column (6) reports the pooled regression using dummy variables for quality quartiles interacted with the within-state citation dummy. Allowing for university quality does not affect the distance gradient (compare columns 3 and 6). The new finding here is that the state border effect is unaffected by institutional quality except for the highest quartile, where it is weaker, but the difference between public and private universities remains robust.

The evidence in this section shows clearly that state borders are more important for public universities.[24] Does this reflect something intrinsic to ownership, or is it associated with university

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 by re-estimating the baseline 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 0.080, which is similar to the one obtained with the full sample, which is 0.089.

[23] We use three alternative measures from the National Research Council data: ranking of faculty quality, publications per faculty, and citations per faculty. In each case, we assign each university to a quality quartile and interact quartile dummies with the within-state dummy. The results in the text are based on faculty rankings, but we get qualitatively similar results with the publications and citation measures.

[24] 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

policy that is correlated with ownership? We examine this question, and how other state policies and characteristics affect the state border effect, in the next section.

5.3. University and state policies and the state border effect

In this section we examine how the strength of the border effect varies across states, and how policy influences it. In particular, we are interested in the role played by university policies toward technology transfer, and state policies that affect the mobility of scientists and engineers and thus circulation of local information. We now turn to a discussion of our specific hypotheses and how we test them empirically.

Local Commercialization of Innovation: The main idea is that the state border effect is likely to be stronger when universities have a policy to promote local and regional development through their technology transfer activity. These policies take the form of a preference for licensing university inventions to local firms, or establishing local start-up companies. In addition, there are sometimes formal constraints (or informal pressure) on universities – particularly public institutions – that influence licensing behavior (Belenzon and Schankerman, 2009). Such local licensing creates an information base on which the licensee and other local inventors are likely to build on, and thus a greater probability of within-state citation (stronger state border effect).

In addition, the state border effect should be larger when the potential for exploiting the knowledge spillovers within the state is greater, and this is more likely in states with a higher density of scientists and engineers (S&E). However, controlling for the average S&E 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 interact the within-state dummy with three measures. The first is a dummy variable indicating whether the university reports having a strong local/regional development objective in its technology transfer policy. The second is the density of scientists and engineers per square mile (in 1995).[25] The third is a measure of the high-tech density in the city where the university is located (TechPole), which is constructed by the Milken Institute (Devol, 1999).

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.

[25] The data on the total number of scientists and engineers in each state are taken from the National Science Foundation, http://www.nsf.gov/statistics/pubseri.cfm?/seri_id=18.

In the patent citation equation, we expect positive coefficients on the interaction of the within-state dummy with strong local development objectives and S&E density, and a negative coefficient on the interaction with TechPole.

Local information: 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, which we call *'labor mobility'*, and 2) inventors working in a state are more likely to have been educated at the graduate level in that state, which we call *'local education'*.[26] We consider each in turn.

Local Education: 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 large-scale information we are aware of that links 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 (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 In-State Educated S&E.[27] The samples in this survey are relatively small, and the variable is certainly measured with substantial error. Attenuation bias will cause us to underestimate the true impact of local education on the border effect.

Labor Mobility: To examine the role of labor mobility, we use two complementary approaches. First, we use data from the March Current Population Survey (CPS) from the U.S. Bureau of the Census to construct a state-level measure of scientific labor mobility. The March CPS asks whether people have moved in the last year, the state they currently live in, and the state they lived it last year, as well as their occupation. We use the CPS data for the available years within our sample period to identify the level of both inward and outward migration for each state for the set of occupational categories that correspond to scientists and engineers (in our best judgment). We normalize this migration by the existing stock of scientists and engineers in each state, taken from the National Science Foundation. We construct a measure of gross migration

[26] 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. Agrawal, Cockburn and McHale (2006) present evidence using patent citations that support this argument.

[27] The information is taken from Table 8.2 in Sumell, Stephan and Adams (2008). The fraction of new hires educated in-state varies widely, from a low of 8.3 in the District of Columbia, and 19.5 percent in New Jersey, to a high of 57.4 percent in Utah and Iowa.

rates (inward + outward) for this purpose – which we call S&E Mobility – and interact it with the within-state dummy to test whether states with higher scientific labor turnover have weaker border effects.[28]

The second approach we use is to 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 increase the likelihood that employees who change jobs leave the state, creating less intrastate job mobility (Marx et. al. 2007; 2010). Our hypothesis is that, by inducing inventors with local information to leave the state, such laws should reduce the strength of the state border effect. To test this, we use the 'non-competition enforceability index' for each state constructed by Garmais (2009) and interact it with the within-state dummy.[29]

Table 5 presents the results.[30] In these regressions we include as an additional control the interaction of the within-state dummy with the level of economic activity (log of gross state product per capita). We begin by estimating a completely flexible specification that allows each state to have its own border effect, i.e., a different within-state dummy coefficient for each state (results not reported for brevity). Even with this general specification of the state border effect, our earlier findings on the impact of distance are robust. The citation probability declines sharply, and the effect of distance is exhausted after 150 miles.[31] However, the estimates indicate substantial variation across states in the magnitude of border effects. The average state border effect is large – at 0.194, it is equivalent is 40 percent of the sample mean citation probability – but the point estimates range from a low of 0.040 in New York to a high of 0.513 in Maine. We strongly reject the null hypothesis that the border effects are the same across states (p-value <.001). The map of the U.S. in Figure

[28] The mobility data are taken from http://www.bls.census.gov/cps/ads/adsmain.htm. Gross migration is the appropriate measure for this purpose because both inward and outward migration should reduce the presence of local information in the state. Scientists who leave the state with local knowledge and subsequently cite the university's patents or publications will represent out-of-state citations, and thus reduce the border effect. Scientists migrating to the state but educated elsewhere will have less local knowledge which again reduces the border effect.

[29] 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). We obtain a very similar point estimate with this measure, but it is not statistically significant.

[30] As before, standard errors are clustered at the patent level. In these regressions the interactions between university/state characteristics and the within-state citation dummy vary at the micro (citation) level. This makes it different from the case studied by Moulton (1990), where a micro regression includes an aggregate regressor with no variation over a subset of micro observations, and thus requires adjustment to standard errors.

[31] The estimates (s.e.) on the first 25-50, 50-100 and 100-150 mile distance intervals are -0.143 (.009), -0.208 (.011) and -0.238 (.011).

2 summarizes the cross-state variation by identifying the relevant quartile of the distribution into which each state falls. There is no obvious regional, size or other simple characterization of states in this figure. We turn next to an econometric analysis of the cross-state variation in the border effect.

Insert Table 5 and Figure 2 here

Column (1) presents the results with interactions of the within-state dummy and all of the factors discussed above, except the university local development dummy. We turn first to the local commercialization hypotheses. The coefficient on the S&E density interaction is positive (though marginally significant, p-value =.07), indicating that states with greater ability to exploit local knowledge spillovers have larger border effects. Controlling for this density, we find that the state border effect is smaller when high-tech activity is more concentrated at the university location. A one standard deviation increase in TechPole (corresponding roughly to a move from Phoenix to Boston) reduces the estimated border effect by 0.026, which is about 12 percent of the average state border effect.

We turn next to the role of local information. First, we find that non-compete laws reduce within-state knowledge spillovers, and the effect is large. The estimated coefficient of -0.018 implies that 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 0.162, which is 84 percent of the average border effect. Second, the results confirm that the border effect is stronger in states that have a higher fraction of locally educated scientists and engineers. Third, states with greater levels of labor mobility exhibit smaller within-state knowledge spillovers. These results show that policies that promote retention of local university graduates and reduce scientific labor turnover will increase the localization of knowledge spillovers. In addition, as we found earlier, the state border effect is smaller for private universities.

In column (2) we add the interaction between the dummy variable for strong local development objectives and the within-state dummy. This university policy 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 coefficient on the university policy interaction variable is positive and significant, confirming that universities that use technology transfer to promote local development create stronger within-state knowledge spillovers. Nevertheless, even after we control for university local development policy objectives, we still find that public universities are more effective at creating within-state knowledge spillovers than private ones. It remains an important task for future research to identify the reasons behind this difference.

The coefficients on the other interaction variables are generally robust when university policy is introduced. The point estimates of the impact of labor mobility and non-compete laws rise somewhat. The coefficient on S&E density is about twice as large as before, and in this case statistically significant. Finally, controlling for these other factors, we find that states with greater higher levels of per capita income have smaller state border effects.

5.3.1. The Michigan 'Natural' (Policy) Experiment

In the previous section we exploited the cross-state variation in characteristics and policy to identify the effects of interest. Of course, there is 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.

To do this, 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 strongly confirm this prediction. We observe a sharp, and statistically significant, drop in the coefficient after 1989, and essentially no change thereafter.[32] Moreover, the magnitude of this shift in the state border effect is consistent

[32] 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.

with the change implied by the parameter estimates obtained from columns (1) and (2), where we identify the effect from the cross-state variation. Using the estimate 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 0.162. This is very close to the estimate using the Michigan policy experiment, which yields a decline of 0.144 (= 0.222 - 0.078).

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.

In Table 6 we summarize the relative importance of the main factors that affect the state border effect, using the parameter estimates from column 2 in Table 5. We focus on the following three sets of determinants: 1. *university policy* (public/private status, and local development objectives), 2. *state policy* (non-compete laws), and 3. *scientific labor factors* (In-state educated S&E, S&E mobility and S&E density). We cannot meaningfully decompose the border effect into different components because the various factors above are not orthogonal, so any decomposition would be arbitrary. Instead, we compute the impact of a standard deviation change for continuous measures, and discrete impacts for the dummy variable measures. In addition, we show the impact of moving from the minimum to maximum values in the sample of each variable.

Insert Table 6 here

The results show that all of the above factors have large, and roughly similar, impacts on the state border effect. Turning first to policy variables, moving from public to private ownership is associated with a 29.2 percent decline in the state border effect, while a shift from no (or weak) to strong local development objectives raises it by 36 percent. A standard deviation increase in the strength of state non-compete laws reduces the border effect by 47.7 percent. Local information variables also have large impacts. A standard deviation increase in the fraction of locally educated scientists and engineers raises the border effect by 42.2 percent, while for S&E density the corresponding impact is 22.8 percent. A standard deviation rise in scientific labor mobility reduces the border effect by 36 percent.

5.4. Variation in localization across technology fields

The previous analysis was based on pooling data for different technology areas. In this section we disaggregate the data to examine whether our findings of localized knowledge spillovers is common to all fields, or driven by only a few technology areas. Table 7 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, Chemicals, Pharmaceuticals, Medical Instruments, Engineering, Electronics, Information Technology, and Telecommunications.[33]

Insert Table 7 here

We find substantial variation across fields in the degree to which knowledge diffusion is localized, both in terms of the distance gradient and the state border effect. While distance strongly mediates spillovers in all technology areas, the effects are less sharp in Biotechnology, Information Technology 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, the citation probability declines by an average of 13 percentage points (26 percent of the mean citation rate) after 100 miles for the newer fields, but by an average of 29 percentage points for the others. 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 border effect is statistically, and economically, significant only in engineering and the biomedical-related fields – Biotechnology, Chemicals, Pharmaceuticals and Medical Instruments. Since we control in these regressions for additive technology field-state fixed effects, this finding suggests that there may be some interaction between state or university policies and technological specialization –

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

e.g., university local development policies that target particular fields. We leave this for future investigation. In any event, 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.

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. The main question of interest is whether the geography of knowledge spillovers differs in an 'open science' (publication) regime as compared to a proprietary one (patents), 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 8 reports the results. In all regressions, we include university fixed effects, dummy variables for pairs of five high-technology clusters, and technology field-state interaction dummies. Column (1) reports the specification with the dummy variables for different distance intervals, but no within-state dummy. As with patents, the results show that geography sharply constrains citations to publications. Moving from 0-25 to 25-50 miles reduces the citation probability by 20.1 percentage points (40 percent of the mean citation rate), and moving out to 50-100 miles further reduces it by another 5.7 percentage points. However, citations continue to decline with distance up to about 1000 miles. This finding also holds in later specifications with the state border effect and university quality controls.

Column (2) reports a specification with the within-state dummy but no distance effects. While it appears that citation is much more likely from inventors located within the same state, the estimated state border effect falls dramatically (from 0.193 to 0.036) when we introduce both distance effects and the within-state dummy (column 3). This specification shows that both distance and the state border effect are statistically significant, and it is important to include both variables in the specification. However, after including distance dummies, the estimated state border effect is quite small, only 7.2 percent of the mean citation rate. This is only a third as large as we found for citations to patents.

Insert Table 8 here

Moreover, the state border effect for publications depends strongly on whether the university is public or private, and on the quality of the institution. We report estimates for the model separately for public and private universities in columns (4) and (5). We find no statistically significant border effect for private universities. There is a small border effect for private institutions, which is equivalent to 8.6 percent of the mean citation rate. However, this average effect masks big differences across universities of different qualities. In columns (5) and (6) we report estimates where we allow for universities in different quality quartiles to have different state border effects (the reference group is the lowest quartile). Column (5) confirms that there is essentially no border effect for any of the quality groups for private universities. For public universities, there is no border effect for the top quartile, but the state border strongly constrains diffusion of information in publications for the lower quartiles, especially below the median quality level. We leave it for future research to identify the underlying mechanisms that cause this outcome.[34]

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 for patents, and constant thereafter. Distance also constrains the diffusion of knowledge in publications, but the effects are less sharp. 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 contrast, we find essentially no state border effect for patent citations to scientific publications except for lower quality public universities. Differences between the open science regime of scientific publications and the proprietary regime of patents seems to be important in shaping the geography of knowledge spillovers.

The state border effect is influenced by the characteristics and policies of the university and state.

[34] We checked whether the geographic profile of knowledge spillovers from scientific publications changed over time. To do this, we re-estimated the specification in column (3) for two sub-periods, 1976-1993 and 1994-2006, using either the date of the cited publication or the date at which the citation occurs. The point estimates of the state border effect decline somewhat, though are not statistically different, between the two periods: 0.047 (.017) versus 0.015 (.019) using the cited publication to date, 0.065 (.027) versus 0.026 (.014) using the citing patent to date. As with patents, the coefficients on the distance dummies show somewhat stronger localization for publications for the later period, using both dating methods, and in both periods the effects of distance persist up to about 1000 miles.

It is significantly larger for 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 density of scientists and engineers who can exploit the potential for knowledge spillovers, and state policy toward non-compete laws that affect intrastate and interstate labor mobility. Finally, we show that there are differences across technology areas in how distance and state borders affect knowledge diffusion.

The key challenge for future research is to identify the channels through which distance and state borders mediate knowledge spillovers – more specifically, why public and private universities differ and how open science and patent regimes affect diffusion. And, perhaps most important, what is the impact of intra- and interstate knowledge spillovers on economic growth at the state level? A promising way forward is to collect systematic information on university and state policies promoting spillovers, and firm strategies for exploiting them.

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 (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,* doi:10.1093/jleo/ewp033: 1-50

[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] Lach, Saul and Mark Schankerman (2008), "Incentives and Invention in Universities," *Rand Journal of Economics*, 39(2): 403-433

[21] Mansfield, Edwin (1995), "Academic Research Underlying Industrial Innovations: Sources, Characteristics, and Financing," *Review of Economics and Statistics*, 77: 1, 55-65

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

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

[24] 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

[25] Mowery, David and Arvids Ziedonis (2001), "The Geographic Reach of Market and Non-Market Channels of Technology Transfer: Comparing Citations and Licenses of University Patents", NBER Working Paper 8568

[26] National Research Council (1993), *National Survey of Graduate Faculty* (Washington D.C.: National Science Foundation)

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

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

[29] 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

[30] 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

[31] 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

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

[33] 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

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

[35]

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.[35] 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.

[35] 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).

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 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.[36] 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, C07 K4, C07K14, C07K16, C07K17, C07K19/00, C12M, C12N, C12P, C12Q, C12S, G01N27/327, G01N33/53, G01N33/54, G01N33/55,

[36] 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).

G01N33/57, G01N33/68, G01N33/74, G01N33/76, G01N33/78, G01N33/88, G01N 33/92

Chemicals: C0, C1, B01, D01F, A62D (excluding Biotechnology)

Pharmaceuticals: A61K, A61P

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

A.5. Other data sources

References

[1] *TechPole index:* http://milkeninstitute.org/pdf/hightech_metros.pdf

[2] *Stock of scientists and engineers (S&E)*: http://www.nsf.gov/statistics/pubseri.cfm?seri_id=18

[3] *Inward/outward migration of S&E*: Constructed from data available at http://www.bls.census.gov/cps/ads/adsmain.htm

[4] *Measures of university quality*: Constructed from data appendices *Research-Doctorate Programs in the United States: Continuity and Change*, available at http://www.nap.edu/html/researchdoc/intexp.html

[5] *State size (land area):* http://www.quickfacts.census.gov/qfd/states/04000.html

[6] *State population (in year 2000):* http://www.census.gov/compendia/statab/2007/population.html

[7] *Gross state product (in year 2000):* http:www.census.gov/compendia/statab/2007/ income_expenditures_wealth/gross_domestic_product_and_gross_state_product.html

Figure 1a. Distance and Patent Citation Probability

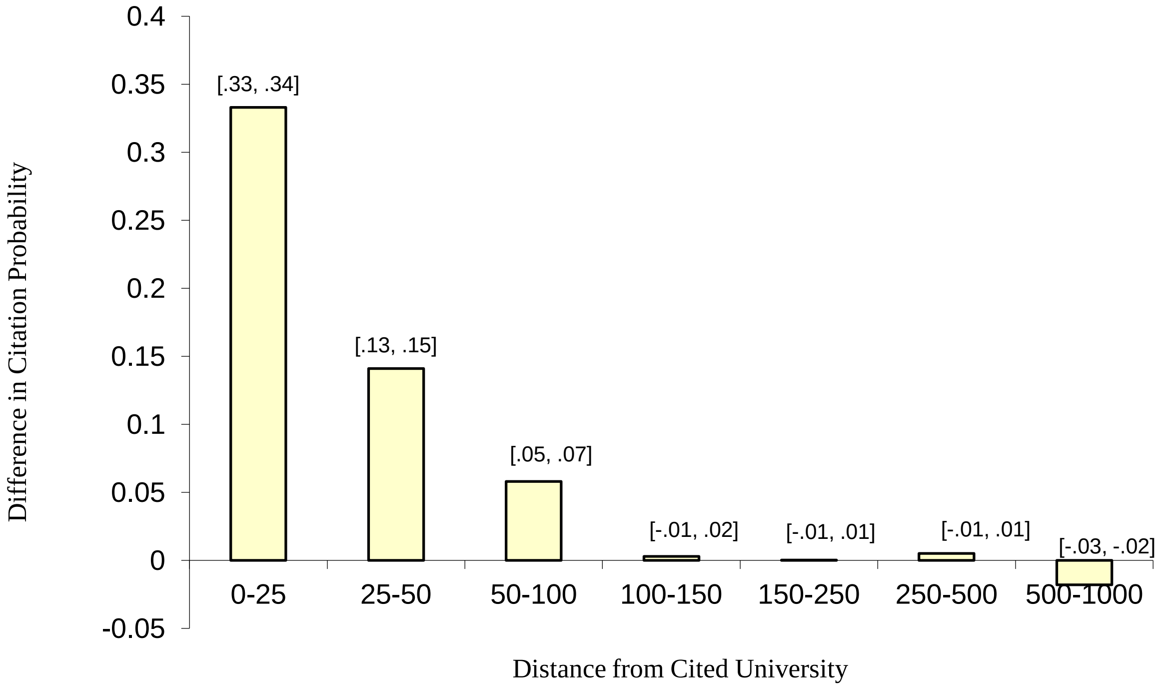
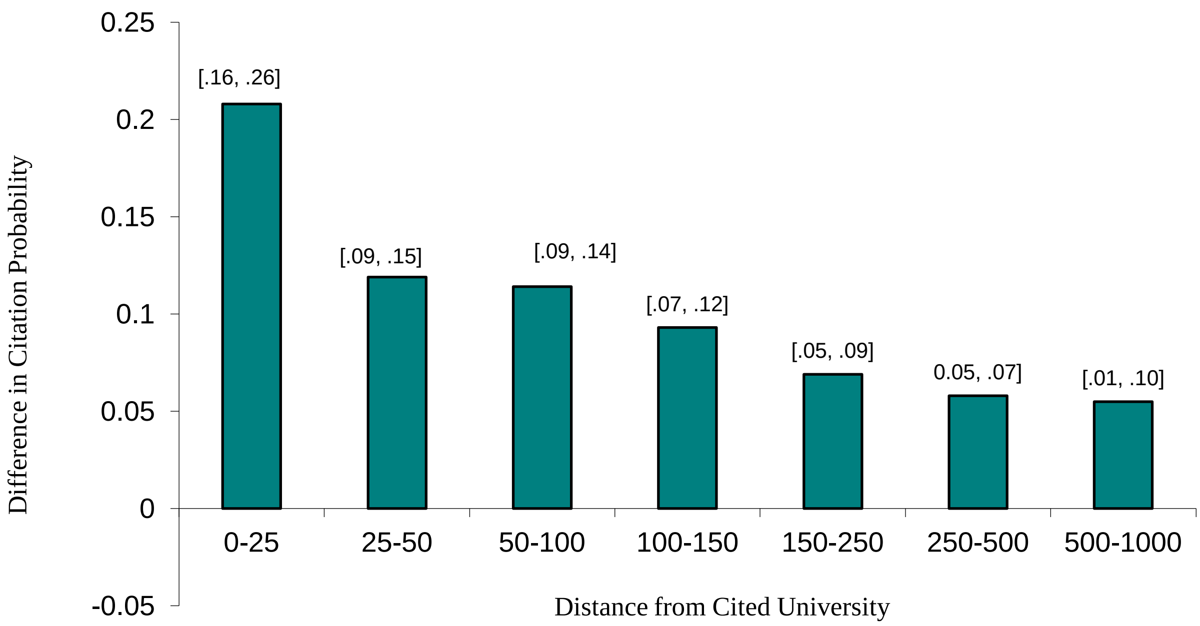
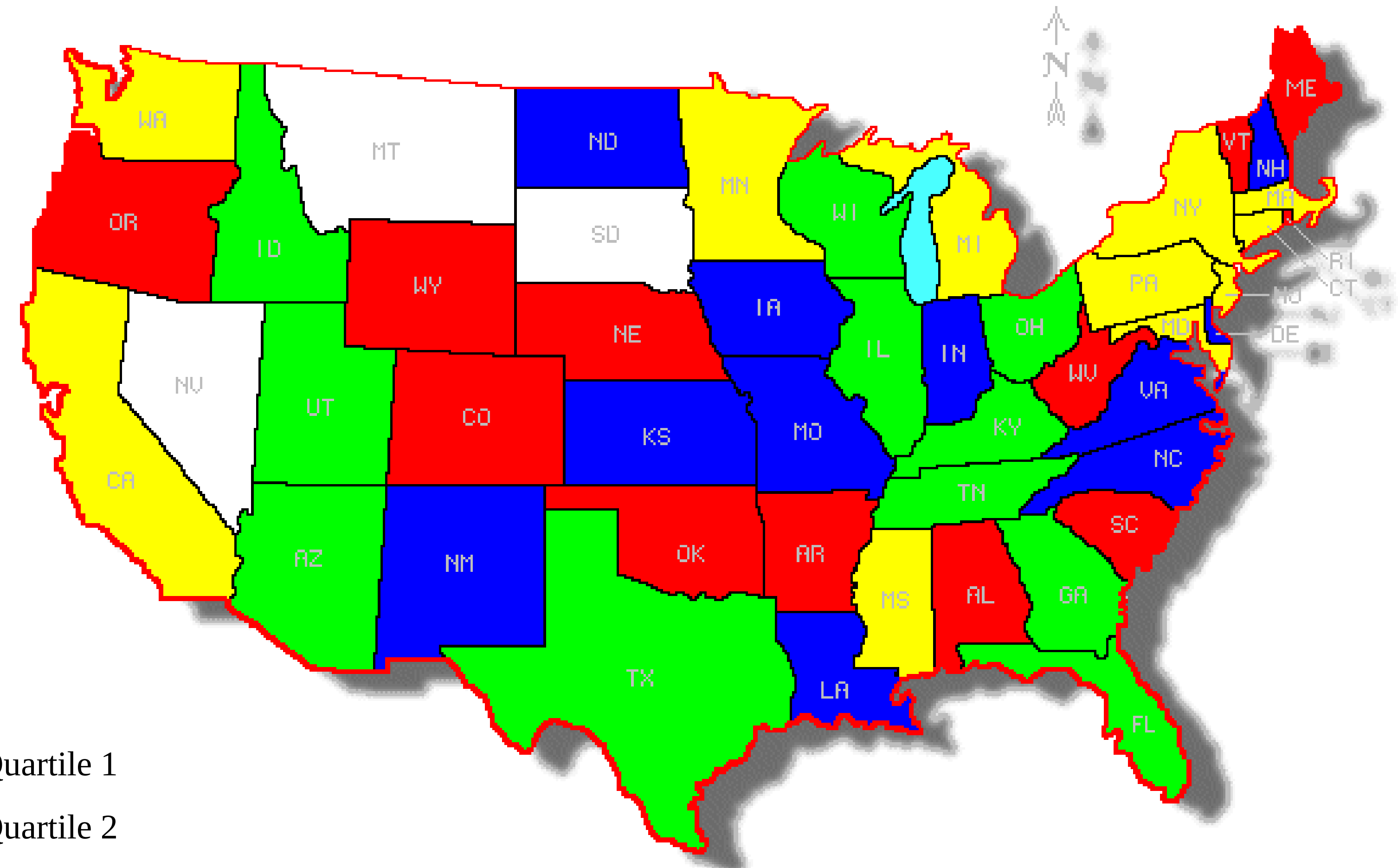


Figure 1b. State Borders and Patent Citation Probability



Notes: Figures 1a and 1b present the effect of distance and state-border on citations probability. For distance, each bracket compares the difference between the share of citations by inventors located in that bracket and citation by all other inventors located further away from the cited university. For state-border, each bracket compares the difference between the share of citations by in-state inventors and inventors located outside the state of the cited univesity, in the specified distance bracket.

Figure 2. Effects of State Borders on Patent Citations



Quartile 1

Quartile 2

Quartile 3

Quartile 4

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

	Panel A. Patents			*Panel B. Scientific Publications*		
	% of citations	# cited patents	# citing patents	% of citations	# cited articles	# citing patents
Dummy for within-state citation	16.3	11,732	27,942	17.3	4,207	5,266
Distance < 25	8.9	8,235	16,677	9.0	2,519	3,113
25 ≤ Distance < 50	2.8	2,600	5,185	2.8	717	957
50 ≤ Distance < 100	1.8	1,712	3,326	2.1	527	694
100 ≤ Distance < 150	1.4	1,547	2,689	1.6	446	527
150 ≤ Distance < 250	3.7	3,447	6,801	3.8	1,029	1,202
250 ≤ Distance < 500	10.7	8,011	18,583	10.8	2,730	3,252
500 ≤ Distance <1000	17.5	11,149	28,996	15.8	4,032	4,504
Distance ≥ 1000	53.2	19,732	73,948	54.1	11,518	12,333

Notes: Distance refers to the driving mileage between the locations 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

	Panel A: Patents			*Panel B: Scientific Publications*		
	(1)	(2)	(3)	(4)	(5)	(6)
Universities:	# Obs.	% Difference distance	% Difference Within-State	# Obs.	% Difference distance	% Difference Within-State
All	382,086	-6.9**	53.1**	69,428	-8.2**	44.2**
Private	176,174	-4.3**	49.0**	31,254	-6.2**	40.8**
Public	205,912	-9.2**	56.9**	38,174	-10.0**	47.1**
Cites received: lowest quartile						
All	98,319	-13.1**	62.5**	17,192	-6.6**	39.5**
Private	39,014	-10.8**	55.1**	9,390	-5.4**	35.0**
Public	59,305	-14.7**	67.6**	7,802	-8.0**	44.4**
Cites received: upper quartile						
All	95,357	-1.0*	45.0**	17,192	-6.6**	39.5**
Private	50,459	1.0	46.4**	9,390	-5.4**	35.0**
Public	44,898	-3.4**	43.1**	7,802	-8.0**	44.6**

Notes: Panel A reports mean comparison tests between cited and control patents for distance from, and fraction 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					
	(1)	(2)	(3)	(4)	(5)
University cited patents:	All	All	All	All	Exc. Top Patenting Universities
Intra-State Citation		0.196**	0.084**	0.089**	0.095**
		(0.004)	(0.007)	(0.007)	(0.009)
Match on 6-digit IPC	0.296**	0.300**	0.295**	0.319**	0.321**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
25 ≤ Distance < 50	-0.165**		-0.155**	-0.155**	-0.182**
	(0.008)		(0.008)	(0.008)	(0.011)
50 ≤ Distance <100	-0.252**		-0.217**	-0.218**	-0.246**
	(0.010)		(0.010)	(0.011)	(0.012)
100 ≤ Distance <150	-0.302**		-0.249**	-0.250**	-0.271**
	(0.010)		(0.011)	(0.011)	(0.013)
150 ≤ Distance <250	-0.306**		-0.239**	-0.239**	-0.263**
	(0.007)		(0.009)	(0.009)	(0.011)
250 ≤ Distance <500	-0.302**		-0.238**	-0.239**	-0.258**
	(0.005)		(0.008)	(0.008)	(0.010)
500 ≤ Distance <1000	-0.321**		-0.241**	-0.241**	-0.264**
	(0.005)		(0.008)	(0.008)	(0.010)
1000 ≤ Distance <1500	-0.318**		-0.236**	-0.235**	-0.256**
	(0.005)		(0.008)	(0.009)	(0.011)
1500 ≤ Distance < 2500	-0.312**		-0.229**	-0.227**	-0.244**
	(0.005)		(0.009)	(0.009)	(0.011)
Distance ≥ 2500	-0.280**		-0.197**	-0.194**	-0.214**
	(0.005)		(0.009)	(0.009)	(0.011)
Technology × State Fixed Effects	No	No	No	Yes	Yes
R^2	0.079	0.073	0.080	0.085	0.086
Observations	382,086	382,086	382,086	382,086	282,466

Notes: This table reports parameter estimates for a linear probability model for citation to university patents. Intra-State Citation is a dummy variable equal to one when the citing (or control patent) inventor resides in the same state as the cited university. Match on 6-digit IPC is a dummy variable equal to one when the citing (or control) and cited patents share the same six-digit IPC code. All columns include complete sets of university and high-tech cluster pair dummies. 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

	(1)	(2)	(3)	(4)	(5)	(6)
University cited patents:	Private	Public	All	Cites received $\leq 25^{th}$	Cites received $>75^{th}$	All
Intra-State Citation	0.065**	0.102**	0.125**	0.113**	0.158**	0.171**
	(0.010)	(0.011)	(0.009)	(0.012)	(0.023)	(0.016)
Intra-State Citation × Private			-0.080**	-0.096**	-0.056**	-0.069**
			(0.008)	(0.012)	(0.023)	(0.009)
Intra-State Citation × QualityQ2						-0.027
						(0.018)
Intra-State Citation × QualityQ3						-0.031
						(0.016)
Intra-State Citation × QualityQ4						-0.061**
						(0.016)
Match on 6-digit IPC	0.305**	0.335**	0.318**	0.336**	0.312**	0.318**
	(0.004)	(0.005)	(0.003)	(0.004)	(0.009)	(0.003)
25 ≤ Distance < 50	-0.139**	-0.191**	-0.176**	-0.207**	-0.114**	-0.171**
	(0.011)	(0.013)	(0.013)	(0.018)	(0.031)	(0.013)
50 ≤ Distance < 100	-0.233**	-0.230**	-0.211**	-0.261**	-0.160**	-0.209**
	(0.015)	(0.015)	(0.014)	(0.020)	(0.040)	(0.014)
100 ≤ Distance < 150	-0.236**	-0.288**	-0.264**	-0.269**	-0.316**	-0.261**
	(0.015)	(0.016)	(0.015)	(0.022)	(0.040)	(0.015)
150 ≤ Distance < 250	-0.237**	-0.257**	-0.228**	-0.271**	-0.175**	-0.226**
	(0.012)	(0.014)	(0.012)	(0.018)	(0.040)	(0.012)
250 ≤ Distance < 500	-0.220**	-0.281**	-0.252**	-0.284**	-0.207**	-0.247**
	(0.011)	(0.012)	(0.009)	(0.013)	(0.026)	(0.009)
500 ≤ Distance < 1000	-0.241**	-0.268**	-0.234**	-0.297**	-0.163**	-0.231**
	(0.012)	(0.012)	(0.010)	(0.014)	(0.026)	(0.010)
1000 ≤ Distance < 1500	-0.223**	-0.267**	-0.233**	-0.296**	-0.125**	-0.230**
	(0.012)	(0.013)	(0.010)	(0.014)	(0.027)	(0.010)
1500 ≤ Distance < 2500	-0.219**	-0.258**	-0.223**	-0.289**	-0.131**	-0.221**
	(0.012)	(0.013)	(0.010)	(0.014)	(0.027)	(0.010)
Distance ≥ 2500	-0.174**	-0.246**	-0.205**	-0.269**	-0.107**	-0.204**
	(0.012)	(0.014)	(0.009)	(0.013)	(0.024)	(0.009)
Distance Dummies × Private	No	No	Yes	Yes	Yes	Yes
R^2	0.079	0.092	0.085	0.105	0.076	0.085
Observations	176,174	205,912	382,086	98,319	95,357	382,086

Notes: This table reports parameter estimates for a linear probability model of citations to university patents, focusing on differences between private and public institutions, university quality, and patent quality (as measured by total number of citations received). University quality is constructed on the basis of quality scores of academic quality for individual departments in the hard sciences (aggregated using faculty size weights) produced by the National Research Council of the U.S. Academy of Sciences. See Lach and Schankerman (2008) for details. Univeristies are grouped into quality quartiles in the table, with the lowest quartile as the reference group. All columns include complete sets of university, high-tech cluster pair, and state-technology interaction dummies. 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

Dependent variable: Citation Dummy						
	(1)	(2)	(3)	(4)	(5)	(6)
	State effects		*Michigan "experiment"*			
	All	Non-missing LocDev	No controls	Controls: bordering states	Controls: similar border effect	Controls: All non-Michigan states
Intra-State Citation	1.123** (0.168)	0.999** (0.254)				
Intra-State Citation × Private	-0.027** (0.010)	-0.026* (0.012)				
Intra-State Citation × Strong Local Development Objectives		0.032** (0.013)				
Intra-State Citation × Techpole ($\times 10^{-1}$)	-0.031** (0.008)	-0.043** (0.009)				
Intra-State Citation × Index of Non-Compete Laws	-0.018** (0.003)	-0.024** (0.004)				
Intra-State Citation × In-state educated S&E ($\times 10^{-1}$)	0.014** (0.005)	0.032** (0.007)				
Intra-State Citation × S&E Mobility ($\times 10^{-1}$)	-0.018** (0.004)	-0.023** (0.005)				
S&E Density	0.009 (0.005)	0.023** (0.007)				
ln(GSP Per Capita)	-0.252** (0.058)	-0.176* (0.085)				
Dummy for Michigan × :						
Dummy for Pre-1986			0.159 (0.085)	0.159 (0.085)	0.141 (0.088)	0.141 (0.083)
Dummy for 1986-1989			0.222** (0.072)	0.222** (0.072)	0.156* (0.073)	0.209** (0.072)
Dummy for 1990-1995			0.078* (0.034)	0.078* (0.034)	0.019 (0.035)	0.064** (0.034)
Dummy for Post-1995			0.082** (0.020)	0.081** (0.020)	-0.008 (0.019)	0.067** (0.019)
Dummy for Control States × :						
Dummy for Pre-1985				0.070 (0.059)	0.017 (0.024)	0.045* (0.020)
Dummy for 1986-1989				0.078 (0.078)	0.065** (0.016)	0.069** (0.014)
Dummy for 1990-1995				0.133** (0.037)	0.059** (0.012)	0.075** (0.009)
Dummy for Post-1995				0.164** (0.021)	0.089** (0.009)	0.096** (0.007)
R^2	0.085	0.084	0.085	0.085	0.085	0.084
Observations	381,994	259,810	382,086	382,086	382,086	382,086

Notes: This table reports parameter estimates for a linear probability model focusing on the determinants of the state border effect for citations to university patents. Local Objectives measures the weight the university technology licensing office attaches to local/regional development in its licensing policies. TechPole is a measure of high-tech density constructed by the Milken Institute (Devol and Wong, 1999). All columns include a set of dummies for distance, a dummy for matching on the same IPC, and complete sets of university, high-tech cluster pairs, and state-technology interaction dummies. 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. Impact of University and State Policies/Characteristics on Within-State Spillovers

	% change in state border effect					
	(1)	(2)	(3)	(4)	(5)	(6)
	Private Ownership	Strong Local Development Objectives	State Non-Compete Laws	In-state Educated S&E	S&E Mobility	S&E Density
Standard deviation increase	-29.2	36.0	-47.7	42.4	-36.0	22.8
Minimum to maximum value	-29.2	36.0	-188.8	176.5	-170.6	93.6

Notes: This table shows the impact of different university and state policies/characteristics on the state border effect for citations to university patents. Estimates are based on column 2 in Table 5. Percentage changes are computed with respect to the pooled state-border effect from column 4 in Table 3.

Table 7. The Effects of Distance and State Border on Patent Citations, by Technology Field

	Dependent variable: Citation Dummy							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Technology area:	*Biotechnology*	*Chemicals*	*Pharma*	*Medical Equipment*	*Engineering*	*Electronics*	*Information Technology*	*Telecommuni-cations*
Intra-State Citation	0.161**	0.153**	0.116**	0.099**	0.056**	0.015	0.034	0.032
	(0.024)	(0.023)	(0.022)	(0.021)	(0.016)	(0.020)	(0.037)	(0.045)
Matched on six-digit IPC	0.346**	0.273**	0.341**	0.310**	0.388**	0.284**	0.228**	0.262**
	(0.011)	(0.009)	(0.009)	(0.009)	(0.007)	(0.008)	(0.013)	(0.018)
25 ≤ Distance < 50	-0.075**	-0.215**	-0.161**	-0.136**	-0.143**	-0.171**	-0.130**	-0.100*
	(0.028)	(0.024)	(0.024)	(0.025)	(0.019)	(0.024)	(0.050)	(0.049)
50 ≤ Distance < 100	-0.104**	-0.248**	-0.252**	-0.200**	-0.227**	-0.255**	-0.159**	-0.120
	(0.046)	(0.030)	(0.031)	(0.026)	(0.022)	(0.031)	(0.049)	(0.089)
100 ≤ Distance < 150	-0.146**	-0.276**	-0.280**	-0.221**	-0.251**	-0.330**	-0.138*	-0.159*
	(0.035)	(0.031)	(0.032)	(0.030)	(0.025)	(0.032)	(0.065)	(0.072)
150 ≤ Distance < 250	-0.172**	-0.269**	-0.258**	-0.210**	-0.272**	-0.265**	-0.121*	0.008
	(0.030)	(0.026)	(0.028)	(0.026)	(0.020)	(0.028)	(0.053)	(0.059)
250 ≤ Distance < 500	-0.138**	-0.262**	-0.253**	-0.209**	-0.245**	-0.290**	-0.152**	-0.114**
	(0.027)	(0.023)	(0.024)	(0.024)	(0.017)	(0.021)	(0.040)	(0.047)
500 ≤ Distance < 1000	-0.137**	-0.291**	-0.265**	-0.191**	-0.258**	-0.300**	-0.164**	-0.105
	(0.029)	(0.025)	(0.025)	(0.024)	(0.019)	(0.024)	(0.045)	(0.059)
1000 ≤ Distance < 1500	-0.122**	-0.265**	-0.208**	-0.181**	-0.255**	-0.322**	-0.187**	-0.078
	(0.032)	(0.026)	(0.027)	(0.025)	(0.020)	(0.025)	(0.045)	(0.063)
1500 ≤ Distance < 2500	-0.132**	-0.242**	-0.206**	-0.161**	-0.250**	-0.318**	-0.159**	-0.120*
	(0.030)	(0.026)	(0.030)	(0.025)	(0.020)	(0.025)	(0.046)	(0.062)
Distance ≥ 2500	-0.112**	-0.242**	-0.171**	-0.144**	-0.228**	-0.274**	-0.120**	-0.061
	(0.035)	(0.026)	(0.029)	(0.026)	(0.020)	(0.025)	(0.044)	(0.062)
R^2	0.104	0.096	0.101	0.084	0.106	0.072	0.040	0.049
Observations	25,804	45,704	35,245	70,005	62,701	43,019	16,118	8,044

Notes: This table reports parameter estimates for a linear probability model of citation to university patents, focusing on differences across technology fields in the effects of distance and state borders. For definitions of the technology fields, see the Data Appendix. All columns include complete sets of university, high-tech cluster pairs, and state-technology interaction dummies. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.

Table 8. Citations to University Scientific Publications

	Dependent variable: Citation Dummy					
	(1)	(2)	(3)	(4)	(5)	(6)
University cited patents:	All	Private	Public	Cites received ≤25th	Cites received >75th	All
Intra-State Citation	0.036** (0.014)	0.021 (0.019)	0.043* (0.020)	0.149** (0.037)	0.089* (0.038)	0.266** (0.040)
Intra-State Citation × Private				-0.152** (0.036)	-0.137** (0.035)	-0.091** (0.016)
Intra-State Citation × QualityQ2						-0.050** (0.043)
Intra-State Citation × QualityQ3						-0.141** (0.041)
Intra-State Citation × QualityQ4						-0.211** (0.039)
25 ≤ Distance < 50	-0.205** (0.017)	-0.164** (0.024)	-0.283** (0.023)	-0.299** (0.046)	-0.230** (0.056)	-0.228** (0.023)
50 ≤ Distance < 100	-0.242** (0.022)	-0.200** (0.030)	-0.325** (0.032)	-0.327** (0.055)	-0.235* (0.112)	-0.274** (0.029)
100 ≤ Distance < 150	-0.291** (0.022)	-0.278** (0.031)	-0.351** (0.030)	-0.315** (0.066)	-0.277** (0.067)	-0.282** (0.029)
150 ≤ Distance < 250	-0.278** (0.018)	-0.235** (0.024)	-0.368** (0.027)	-0.310** (0.053)	-0.298** (0.061)	-0.283** (0.024)
250 ≤ Distance < 500	-0.302** (0.015)	-0.251** (0.021)	-0.398** (0.020)	-0.322** (0.040)	-0.392** (0.039)	-0.313** (0.017)
500 ≤ Distance < 1000	-0.315** (0.017)	-0.276** (0.024)	-0.396** (0.023)	-0.290** (0.044)	-0.316** (0.039)	-0.304** (0.018)
1000 ≤ Distance < 1500	-0.310** (0.017)	-0.289** (0.025)	-0.382** (0.024)	-0.287** (0.044)	-0.309** (0.041)	-0.287** (0.019)
1500 ≤ Distance < 2500	-0.314** (0.017)	-0.270** (0.025)	-0.397** (0.024)	-0.289** (0.043)	-0.324** (0.040)	-0.305** (0.018)
Distance ≥ 2500	-0.271** (0.017)	-0.217** (0.023)	-0.377** (0.025)	-0.256** (0.041)	-0.280** (0.037)	-0.278** (0.017)
R^2	0.025	0.022	0.032	0.030	0.023	0.027
Observations	69,428	31,290	38,796	17,632	17,192	69,428

Notes: This table reports parameter estimates for a linear probability model relating citations by patents to university scientific publications to the distance of citing inventors from the cited university, state borders, university quality and the quality of the publications (as measured by total number of citations received). All columns include complete sets of university, high-tech cluster pairs, and state-technology interaction dummies. Standard errors (in brackets) are clustered by cited patent. * and ** denote statistical significance at the 5 and 1 percent levels, respectively.