Research Policy 38 (2009) 1–13

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Have university knowledge flows narrowed? Evidence from patent data

Carlos Rosell[a], Ajay Agrawal[b,*]

[a] University of Toronto, Canada
[b] University of Toronto and NBER, 105 St. George St., Toronto, Ontario, Canada M5S 3E6

ARTICLE INFO

Article history:
Received 18 October 2006
Received in revised form
10 September 2007
Accepted 29 July 2008
Available online 18 November 2008

JEL classification:
I23
O33
O34

Keywords:
University patents
Bayh-Dole Act
Knowledge flows
Anti-commons
Diffusion

ABSTRACT

The rate of university patenting increased dramatically during the 1980s. Did the manner in which knowledge embedded in university patents was managed change during this period of rapid patenting growth? Using a Herfindahl-type measure of knowledge flow concentration and employing a difference-in-differences estimation to compare university-to-firm patent citations across two time periods, we find that the university diffusion premium (the degree to which university knowledge outflows are more widely distributed than those of firms) declined by more than half during the 1980s. In addition, we find that the university diversity premium (the degree to which knowledge inflows used by universities are drawn from a more widely distributed set of prior art holders than those used by firms) also declined by more than half. However, these changes are mostly limited to a narrow set of technology fields (i.e., biotechnology and pharmaceuticals in the outflows case and sub-fields of electronics in the inflows case). The social welfare implications are ambiguous.

1. Introduction

Among the most striking developments on American university campuses over the past quarter century has been the rapid rise of patenting to lay claim to and collect rents from intellectual property associated with novel and practical inventions developed by university researchers. Indeed, in just 14 years, from 1980 to 1993, the number of patents issued annually to US universities increased by 316%, from 390 to 1622.[1] Social scientists have attributed this dramatic shift in academic behavior to many factors. Principal among these are developments in the fields of microbiology and computer science, an expansion in the range of patentable matter (e.g., genetically modified life forms, software), the creation of the Court of Appeals for the Federal Circuit, and, most commonly, the passage of the Bayh-Dole Act (1980), which granted universities extensive rights to patent and retain ownership of innovations produced with federal government funding.

Although many observers have characterized the dramatic rise of university patenting as a windfall for the American economy – *The Economist* described the Bayh-Dole Act as "possibly the most inspired piece of legislation to be enacted in America over the past half century"[2] – others have expressed a variety of concerns, most of which can be grouped into one of three categories: (1) a shift in focus from "basic" to "applied" university research,[3](2) a decline in the quality of university inventions, and (3) a decline in the dissemination of knowledge associated with university inventions.

Surprisingly, given the increasing level of concern over university patenting expressed in both policy circles and the popular press,[4] the evidence to date offers little support for the first two of these concerns. The first concern, that an increased focus on com-

* Corresponding author. Tel.: +1 416 946 0203.
E-mail addresses: carlos.rosell@utoronto.ca (C. Rosell), ajay.agrawal@rotman.utoronto.ca (A. Agrawal).

[1] By comparison, the number of patents issued to other US non-government organizations, namely firms, increased by only 48% over the same time period.

[2] "Innovation's Golden Goose," *The Economist*, December 14, 2002, vol. 365 (8303), p. 3.

[3] Notwithstanding Stokes' legitimate grievances with respect to the basic/applied taxonomy (Stokes, 1997), we reference it here since most of the discourse on this topic has characterized research this way.

[4] Lieberwitz (2005); "University Resolves Dispute On Stem Cell Patent License", *New York Times*, January 10, 2002, p. C11; "Bayhing for blood or Doling out cash?" *The Economist*, December 24, 2005, p. 115; "Lilly Loses Patent Case to Ariad," *New York Times*, May 5, 2006, p. C1.

0048-7333/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.respol.2008.07.014

mercialization may induce university researchers to divert their energies away from basic research (Cohen et al., 1998; Henderson et al., 1998), is predicated on the notion that universities should provide basic rather than applied research. This is because the market is more likely to under-provide basic than applied research due to greater appropriability problems with the former. Moreover, basic research is important since it provides knowledge for subsequent applied research and product development, which in turn is the basis for long-run productivity and economic growth.

However, empirical studies that examine whether professors substitute patenting for publishing, a rough proxy for changes in research focus, do not provide evidence of such substitution. Agrawal and Henderson (2002) examine the publishing and patenting output of electrical engineering, computer science, and mechanical engineering faculty at a major research institution (MIT) and present evidence suggesting that these two activities are complements rather than substitutes. Furthermore, Fabrizio and DiMinin (2008) and Goldfarb et al. (2006) examine the complement-substitute question more directly with data from a much broader sample of university researchers and find similar results. Moreover, these findings are not specific to US universities; several studies that examine the patenting-publishing relationship at various European institutions yield similar conclusions (Van Looy et al., 2006– K.U Leuven in Belgium; Buenstorf, 2005– Max Planck Institute in Germany; Carayol, 2007– University Louis Pasteur in France; Breschi et al., 2007; Calerini and Franzoni, 2004– various institutions in Italy).

The second concern is predicated on the notion that an increased focus on commercialization may induce researchers to shift resources toward the disclosure and patenting of lower quality inventions (Henderson et al., 1998). However, evidence presented by Mowery et al. (2004) shows that although the quality of inventions did decline after 1980, this was due to the entry of universities with little patenting experience, not to a general decline in quality of inventions patented by all universities. The implication of this finding is that the estimated decline is likely to be only temporary while inexperienced universities learn the patenting process and how to most effectively manage their intellectual property portfolio.

Thus, only the third concern, relating to how the anti-commons limits the flow of knowledge, has found traction in empirical evidence.[5] In a study employing a difference-in-differences identification based on patent-paper pairs, Murray and Stern (2007) report findings that although publications linked to patents are associated with a higher overall citation rate, after the patent is actually issued, the rate declines substantially (by 9–17%). The authors note that the decline is particularly salient for articles authored by researchers with public-sector affiliations, such as university professors. They interpret their findings as evidence of an anti-commons effect that results from moving intellectual property from the public into the private domain.

Our paper further addresses the third concern: restricting the widespread flow of knowledge associated with university inventions. However, where Murray and Stern focus on the decline in the *level* of knowledge flows, we focus on the *narrowing* of knowledge flows to a relatively more concentrated set of recipients. Specifically, we examine whether, over time and conditional on being patented, university inventions are more likely to be cited by a more concentrated set of subsequent patent owners. Such a finding could reflect the outcome of a change in the management objectives of university intellectual property from broad knowledge dissemination towards limiting access, perhaps to maximize private returns to licensors.

Although we describe above various changes in the environment that have plausibly led to the narrowing of university knowledge flows, these changes themselves are not causal mechanisms. Rather, behavioral changes by a variety of actors in the university knowledge commercialization system, most notably researchers, technology licensing offices, and licensees, are the most likely cause of this phenomenon.

For example, as the culture of commercialization (Owen-Smith and Powell, 2001; Bercovitz and Feldman, 2008) spread among university professors and other researchers, norms of openness and disclosure may have given way to secrecy in order to maximize value to the licensee and ultimately the researcher in the form of royalties. Such a cultural change could have dampened the ability to build on university inventions, despite disclosure of the new knowledge by way of claims in the patent document. This is because university inventions are usually early-stage (Thursby et al., 2005), and successful commercialization often requires face-to-face interaction for the purpose of tacit knowledge transfer (Jensen and Thursby, 2001; Agrawal, 2006). In other words, if inventors respond to royalty-maximizing incentives, as suggested by Lach and Schankerman (2008), their tendency to share tacit knowledge with others who are not licensees may diminish as the inventors become more commercialization-oriented. While such behavior might not reduce the number of subsequent inventions building on the focal university invention, it could reduce the number of distinct organizations that do so (e.g., closer and more exclusive relations between inventor and licensee could lead to more citations per citing organization but relatively fewer citing organizations).

Alternatively or perhaps additionally, technology licensing offices (TLOs) might have, on average, shifted their objective from dissemination-maximization (leading to predominantly non-exclusive, widely licensed patents) to profit-maximization (leading to predominantly narrowly licensed patents). Such a shift seems plausible given that performance metrics for the latter are easier to measure (Dranove et al., 2003) and that TLOs operate in an environment where they must balance the benefits from connectedness to industry for access to market information with potential industrial "capture" that limits patent impact (Owen-Smith and Powell, 2003). Perhaps in response to such a shift in researcher and TLO behavior, the National Institutes of Health (NIH), a major US government research funding agency, recently issued guidelines urging universities to increase the frequency with which they license genomic, NIH-funded, patented inventions on a non-exclusive, rather than exclusive, basis (National Institutes of Health, 2005).

Finally, it could be that licensees are increasingly applying pressure to be granted exclusive rights to commercialize university inventions. For example, the Geron Corporation, a small biotechnology firm, pressured the University of Wisconsin to gain a greater scope of exclusivity over the use of a human stem cell technology that was developed at and patented by the university.[6] The university ultimately sued Geron to facilitate licensing the patent more broadly (including to Geron's competitors) in order to "enable more academic researchers and private companies to join the research for

[5] In general, the anti-commons problem occurs when the right to exclude others from the use of a resource is bestowed on "too many" individuals or organizations; the transactions costs of coordinating the right to use the resource become inefficiently high. As a result, the resource is under-utilized (Heller and Eisenberg, 1998; Argyres and Liebskind, 1998; Shapiro, 2001). In the context of our setting, each upstream patent owner prices royalties without coordinating with owners of complementary patents. Without coordination, the marginal cost of utilizing complementary technologies is higher than if a single agent owned all patents.

[6] The university's intellectual property and technology transfer is technically managed by a foundation (the Wisconsin Alumni Research Foundation).

new therapies and cures for some of the world's most debilitating diseases."[7] Furthermore, licensee exclusivity, or even the threat of such, can have a limiting effect on the number of users of a particular invention. For example, in the Geron case, the firm decided to allow academic scientists to freely convert stem cells into derivative cells but to prevent them from commercializing their discoveries. As a result, some researchers explicitly avoided building on this prior art. For instance, a prominent Harvard-based scientist balked at working on the derivative cells under such terms. "Those conditions would mean that I am the ideal employee of Geron. They don't pay my salary, they don't pay my benefits, but anything I discover they own."[8]

To be clear, the objective of this paper is not to decompose the relative roles of changes in inventor, TLO, or licensee behaviors as causal mechanisms driving the change in knowledge flow patterns. Rather, our objective is to identify the existence of such a change and examine the scope of fields over which it applies. That is not to say that untangling the relative importance of mechanisms responsible for restricting knowledge flows is not important; it is, in fact, the obvious next step in this line of inquiry.

However, the basic question of whether the increased focus on commercialization (reflected in the sharp rise in university patenting) was accompanied by a narrowing of knowledge flows is important because such a trend would have significant implications for science policy and economic growth. Universities receive extensive government funding to produce basic knowledge that is intended to be widely disseminated.[9] Precisely because of their non-commercial focus and welfare-enhancing objectives, universities play a unique and important role in the national innovation system (Nelson, 1993; Rhoten and Powell, 2007). Since knowledge spillovers are central to economic growth (Romer, 1986, 1990), a finding that university flows are narrowing, even just those associated with patented inventions in certain fields, could throw into question the traditionally conceived arrangement between academia and society.

What do the data reveal? Using a Herfindahl-type measure of patent assignee concentration associated with forward citations as a dependent variable and employing a difference-in-differences estimation (taking the difference of the change in concentrations over time between university and firm patents), we estimate that the university diffusion premium (the degree to which knowledge flows from patented university inventions are more widely distributed across assignees than those of firms) declined by more than half during the 1980s. However, this phenomenon of narrowing knowledge outflows was mostly limited to just a few fields, namely biotechnology and pharmaceuticals.

In addition to examining the pattern of knowledge flowing out from university inventions, we also study the pattern of flows *into* these inventions. Although the approach we employ to examine inflows is similar to that for outflows, the phenomenon itself is distinct. Relative to firms, we expect universities to draw from a wider set of prior art holders since academia is largely shielded from the anti-commons problem.

While firms may consciously conduct R&D in a manner that minimizes exposure to the anti-commons,[10] university researchers are largely insulated for two reasons. First, universities have traditionally been shielded from patent infringement liability due to the "experimental use exemption" (Eisenberg, 2003). Under this doctrine, otherwise infringing activity is permitted if it occurs "for amusement, to satisfy idle curiosity, or for strictly philosophical inquiry."[11] Second, to the extent that university researchers choose their research projects to advance knowledge and only concern themselves with patenting *ex post*– after something they have discovered in the pursuit of knowledge also appears to offer commercial potential – their project selection and prior art decisions will not be influenced by concerns about potential hold-up during the subsequent product development phase.

However, as university patenting rose during the 1980s, we find that university researchers tended to draw from a more concentrated set of prior art holders. Specifically, our results suggest that the university diversity premium (the degree to which knowledge inflows used to develop patented university inventions are drawn from a more fragmented set of prior art holders than those used by firms) declined by over half between the early 1980s and early 1990s. However, this phenomenon was mostly limited to a few technology subfields within electronics.

This finding may reflect a change over time in the manner by which university researchers conduct research. Rather than merely worrying about patentability after an invention has been created, researchers may increasingly plan research projects with an eye towards commercialization.[12] If motivated by pecuniary gains, as evidence reported by Lach and Schankerman (2008) suggests, academic researchers may look forward in anticipating the burden of future licensees and reason back that the value of their intellectual property could be increased if they are able to plan their research approach so as to narrow the scope of prior art holders associated with complementary technologies.

Like Murray and Stern, our findings suggest caution with respect to the increasing tendency to patent university research. However, our findings are quite distinct. Their paper shows the impact of patenting on knowledge dissemination: an overall reduction in the level of knowledge outflows. Our results suggest that, controlling for any changes in overall flow levels, the rapid rise in patenting and the associated management of knowledge flows was associated with a narrowing of knowledge flows both to and from universities but only in specific technology fields. These fields include drugs and biotechnology (for knowledge produced by universities) and subfields of electronics (for knowledge consumed by universities).

[7] "Suit Seeks to Expand Access to Stem Cells," *New York Times*, August 14, 2001, p. C2.

[8] "Patent on Human Stem Cell Puts U.S. Officials in Bind," *New York Times*, August 17, 2001, p. A1.

[9] From 1980 to 1993, universities received approximately $144 billion (constant 1996 dollars) from all levels of government to fund basic R&D. This represents approximately 45% of all basic research undertaken in the US (National Science Board, 2004).

[10] Kodak offers a well-documented example of research caution with respect to prior art. From the outset of the firm's efforts to develop its instant photography technology, the firm employed its legal counsel to work along with its R&D engineers to minimize the likelihood that any new technology would infringe on existing Polaroid patents (Warshofsky, 1994; Rivette and Kline, 2000; Jaffe and Lerner, 2004). In addition, Hall and Ziedonis (2001) and Ziedonis (2004) present evidence suggesting that firms building on prior art that has more fragmented ownership patent more aggressively themselves in order to facilitate cross-licensing and mitigate against potential infringement costs.

[11] Walsh et al. (2005) present evidence suggesting that university researchers pay little attention to patents protecting research tools and are unlikely to modify their research due to impediments posed by existing patents. These findings are particularly interesting since they are based on data reflecting attitudes after the *Madey v. Duke* verdict of 2003.

[12] A patent does not grant the owner the right to practice an invention but rather the right to exclude others from practicing it without the owner's permission. Thus, the university is not able to grant a licensee the right to practise that university's invention since doing so may infringe on additional prior art owned by others.

Overall, this change in behavior may be counter to the stated mandate of most US universities, which is to maximize the dissemination of new knowledge. While the social welfare implications of our findings are non-obvious - limiting access to new knowledge can be welfare enhancing if the welfare lost to those denied access is less than the welfare gained by those granted exclusivity to invest in commercialization (Colyvas et al., 2002; Mazzoleni, 2005; Agrawal and Garlappi, 2007)—our results are consistent with the view that universities may be increasingly managing their intellectual property like profit-maximizing firms rather than as welfare-maximizing public institutions.[13]

The remainder of our paper proceeds as follows. In Section 2, we describe our empirical methodology, particularly the construction of our dependent variable, the "fragmentation index." In Section 3, we describe the patent citation data that we use to construct our measures. In Section 4, we present our empirical results for both knowledge outflows and inflows as well as provide some intuition to better understand the meaning of the estimated coefficients. Finally, in Section 5, we conclude by commenting on the welfare implications of our findings and offering directions for future research.

2. Methodology

Our empirical objective is to test whether university knowledge flows, in particular those associated with patented inventions, become more concentrated over time relative to those associated with firms. Thus, most importantly, we need to employ an estimation technique that facilitates the clean identification of a change in the concentration of knowledge flows over time that is university-specific. Furthermore, we require an appropriate measure of knowledge flow concentration. We describe each of these in turn.

2.1. Estimation

In order to estimate university-specific changes in the concentration of knowledge flows over time, we analyze data from two distinct periods.[14] We define these as Period 1 (1980–1983) and Period 2 (1989–1992). In order to identify changes in concentration that are university-specific as opposed to general changes in knowledge flow patterns, we employ a difference-in-differences estimation (taking the difference of the change in flow concentrations over time between university and firm patents). Thus, it is important to note that throughout the paper our focus is on the narrowing of university knowledge flows *relative* to that of firms. We adopt this focus because many elements of the intellectual property environment evolved over the 10-year period under study, such that controlling for each is not feasible. Instead, we assume that ambient changes influence both firm and university knowledge flows, such that the difference in the rate of change between the two allows us to isolate university-specific changes. In addition, we include control variables to address specific dimensions along which universities may systematically behave differently than firms that could otherwise bias our results (i.e., inventions that are more important, more basic, or more likely from a particular technology field).

Thus, we estimate the following relationship:

$$ForFrag_p = \alpha_o + \alpha_1 Uni_p + \alpha_2 Uni_p Period2_p + X_p\alpha_3 + X_p Period2_p \alpha_4 + \sum_{t=1}^{T} \beta_t Tech_{t,p} + \sum_{t=1}^{T} \lambda_t Tech_{t,p} Period2_p + \varepsilon_p \quad (1)$$

where $ForFrag_p$ measures the fragmentation of ownership dispersion of patents building upon patent p ("forward fragmentation" of knowledge outflows).[15] Uni_p is a university dummy variable that takes a value of one if p is assigned to a university and zero otherwise. Similarly, $Period2_p$ is a period dummy variable that takes a value of one if p is issued in Period 2 and zero otherwise.[16] X_p is a vector of variables that control for non-institutional factors that may also affect fragmentation. $Tech_{t,p}$ is a technology field fixed effect that equals one when patent p is assigned to technology field t and is zero otherwise.[17] Finally, ε_p is a mean zero random error.

We use Eq. (1) to test whether the university dummy explains some of the fragmentation of knowledge flows, $ForFrag_p$. The sign and significance of $\hat{\alpha}_1$ offers insight into the relationship between institution type and the patterns of related knowledge flows. We interpret a positive and statistically significant estimate of $\hat{\alpha}_1$ as suggestive evidence that university knowledge flows were more fragmented, or less concentrated, than those of firms, at least in Period 1. This finding is consistent with our prior belief about the differences between university and firm knowledge flows.

To identify how the initial difference in knowledge flows between universities and firms has changed over time, we focus on $\hat{\alpha}_2$, the coefficient on the interaction between Uni_p and $Period2_p$. We interpret a negative value of $\hat{\alpha}_2$ as indicating that knowledge flows from university patents narrowed disproportionately over time relative to those from firms.

2.2. Variables

We construct each of our variables using information found on the front page of the patents in our data, including the technology field, the assignee name (i.e., the patent's owner), and all prior patents on which the given innovation builds (i.e., backward citations). We use citations as a proxy measure for knowledge flows.[18]

2.2.1. Dependent variable

We construct our dependant variable, a measure of the concentration of knowledge flows, in the spirit of the "fragmentation index" as developed by Ziedonis (2004). Again, we describe only the knowledge outflows measure, or forward fragmentation, $ForFrag_{i,p}$, given that we define the backward measure, $BackFrag_{i,p}$ analogously using the citations a patent makes rather than receives.

Forward fragmentation measures the ownership dispersion of subsequent patents that cite a focal patent. Specifically, for a focal patent p issued to assignee i, the fragmentation measure $ForFrag_{i,p}$ is given by

$$ForFrag_{i,p} = \left[1 - \sum_{j \in J} \left(\frac{C_{j,i,p}}{C_{i,p}} \right)^2 \right] \frac{C_{i,p}}{C_{i,p} - 1}, \quad (2)$$

[13] Alternatively or perhaps additionally, firms may be increasingly managing their IP and associated knowledge like universities.

[14] As described in the Introduction, we are interested in university-specific changes in the concentration of both knowledge outflows and inflows. Since the estimation procedure is almost identical, we describe the outflows case only and comment in footnotes where the methodology differs for inflows.

[15] Similarly, for the case of knowledge inflows, $BackFrag_p$ measures the fragmentation of prior art holders upon which patent p builds (backward fragmentation).

[16] For the case of knowledge inflows, $Period2_p$ takes a value of one for patents that were *applied* for in Period 2.

[17] One might expect our difference-in-differences regression equation to include the *Period2* dummy variable on its own. We do this implicitly by interacting this variable with the technology field fixed effects.

[18] Although commonly used as a measure of knowledge flows, citations are a noisy proxy. We describe the limitations of this proxy in Section 3.2.

where J is the set of assignees whose patents cite the focal patent, $i \notin J$, and $C_{j,i,p}$ are all citations made to p by patents belonging to assignee $j \in J$. $C_{i,p}$ is the total number of patents citing patent p that do not belong to i.

Our fragmentation variable simply measures dispersion as the expected probability that two randomly selected citations made to a given patent refer to citing patents issued to two *different* assignees.[19,20] Consequently, the measure's range of possible values is the unit interval. For patents that have broader knowledge outflows (i.e., higher fragmentation), the probability that any two sampled citations belong to different assignees is closer to one. Conversely, the probability of this event is closer to zero when the citing patents are more concentrated.

To gain better intuition for this dispersion index, which is related to the familiar Herfindahl concentration measure, consider the following three examples of focal patents that are each cited by 10 patents (i.e., $C_{i,p} = 10$). First, suppose the focal patent is cited by 10 patents that are all issued to the same firm. In this case, citing patents are perfectly concentrated and thus make it impossible for any two citations to refer to different assignees, $ForFrag_{i,p} = 0$. Next, suppose the focal patent receives five citations each from two different assignees. This yields an intermediate measure of fragmentation; the probability that any two of the 10 citations are made by different assignees is approximately half, $ForFrag_{i,p} \simeq 0.556$.[21] Finally, suppose 10 different assignees cite the focal patent once each. In this case, it is certain that any two citations will come from different assignees, $ForFrag_{i,p} = 1$.

2.2.2. Control variables

Our identification of university specific fragmentation is based on a difference-in-differences estimation that compares differences in fragmentation over time between universities and firms. This approach is used to "difference out" overall changes in knowledge flow fragmentation that are not university specific. However, it may be the case that certain characteristics of university patents, rather than institutional characteristics of universities themselves, result in the identified changes in university knowledge flow fragmentation. For example, it may be the case that the probability of generating a biotechnology patent increased more over the time period for universities than firms and that biotechnology patents were more likely to generate concentrated knowledge outflows. This might appear as a university specific increase in knowledge flow concentration over time but is actually a biotechnology effect.

We control for three possible confounding effects: (1) generality, (2) technology field, and (3) importance.[22] First, we construct "generality" using the same citations used to calculate the dependant variable. However, rather than measuring the dispersion of citations received from different assignees, this variable measures dispersion of citations received from different technology fields defined by the US Patent and Trademark Office (USPTO) three-digit technology classification system.[23] Generality reflects the extent to which the knowledge embedded in a focal patent is applicable across other technology fields (Trajtenberg et al., 1997).

Second, we include technology field fixed effects using dummy variables coinciding with the NBER two-digit technology field classification.[24] Third, we control for invention importance using a simple count of total citations received by the focal patent.[25]

3. Data

We collect our data primarily from the NBER patent database as described by Hall et al. (2002).[26] This source provides all the raw citation data needed to construct the variables in our samples. In addition, we use two reports: (1) "US Colleges and Universities-Utility Patent Grants, Calendar Years 1969–2005"[27] to identify all US university patents granted from 1969 to 2002[28] and (2) "A Classification of Institutions of Higher Education—1987 Edition" to categorize universities by research ranking.

3.1. Sample construction

Since we ask two different but related questions concerning changes in the concentration of university knowledge (i.e., outflows and inflows), we require two distinct samples. Although the sample construction process used for each is similar, there are a few key differences. We describe each separately below.

3.1.1. Knowledge outflows sample

We collect all utility patents issued to US non-government organizations by the USPTO during the periods 1980–1983 and 1989–1992. We focus on the 1980s because this decade represents a

[19] This is a traditional interpretation for dispersion measures of the type defined by Eq. (2). See Easterly and Levine (1997) for an example of this interpretation in the context of measuring ethnic diversity.

[20] With this interpretation, one can easily understand the fragmentation measure defined by Eq. (2). Due to the count nature of citations (i.e., too few citations are typically made to make sampling with replacement an appropriate assumption), the conditional probability that two citing patents belong to different assignees, given that one of these two citations is known to belong to assignee j, is

$$1 - Pr(\text{Second citation belongs to assignee} j) = 1 - \frac{C_{j,i,p} - 1}{C_{i,p} - 1}.$$

Consequently, the expected probability that two randomly chosen citing patents belong to different assignees is

$$\sum_{j \in J} \frac{C_{j,i,p}}{C_{i,p}} \left(1 - \frac{C_{j,i,p} - 1}{C_{i,p} - 1}\right) = 1 - \sum_{j \in J} \frac{C_{j,i,p}}{C_{i,p}} \frac{C_{j,i,p} - 1}{C_{i,p} - 1}.$$

It then can be shown that

$$1 - \sum_{j \in J} \frac{C_{j,i,p}}{C_{i,p}} \frac{C_{j,i,p} - 1}{C_{i,p} - 1} = \left\{1 - \sum_{j \in J} \left(\frac{C_{j,i,p}}{C_{i,p}}\right)^2\right\} \frac{C_{i,p}}{C_{i,p} - 1}.$$

The term $(C_{i,p}/C_{i,p} - 1)$ in Eq. (2) corrects the empirical probability had we assumed that we could sample with replacement. Without this adjustment, our dispersion measure would be biased toward zero (Hall et al., 2002).

[21] $Forfrag_{i,p} = (1 - 2(5/10)^2)(10/9) \simeq 0.556$.

[22] In the case of inflows, we control for "originality" rather than generality and "citations made" rather than importance. These measures are similar in spirit.

[23] To construct the "Generality" measure, Gen_p, we let $Tech_{t,p}$ denote the number of times patents from technology class t cite the focal patent p. Then we define Gen_p as

$$Gen_p = \left[1 - \sum_{t \in T} \left(\frac{Tech_{t,p}}{C_p}\right)^2\right] \frac{C_p}{C_p - 1},$$

where C_p reflects the total number of citations received by patent p and T reflects the USPTO three-digit technology field universe. We note that $C_p = \sum_{t \in T} Tech_{t,p}$. We construct "originality" in a similar manner but use backwards citations instead.

[24] Our conclusions are robust to using more disaggregated technology field fixed effects; dummy variables based on the USPTO three-digit technology classification codes do not change our conclusions.

[25] Scholars have widely used the generality and importance measures, as described by Hall et al. (2002), in the patent-based economics of innovation literature. The control for importance is particularly necessary since Sampat et al. (2003) show that the rate of university citations received slowed over time relative to that for firms.

[26] Specifically, we use the updated NBER patent dataset (1963–2002) available from Bronwyn Hall's website: http://elsa.berkeley.edu/ bhhall/bhdata.html.

[27] Information Products Division et al. (2005) produces this source.

[28] When referring to universities, we include universities, colleges, polytechnics, other post-secondary institutions, and university consortia.

period of dramatic growth in university patenting (i.e., over 300%). In our Introduction, we summarize a variety of developments that occurred during this time that likely contributed to the growth in university patenting, including key advances in the fields of microbiology and computer science, an expansion of patentable matter, the creation of the Court of Appeals for the Federal Circuit, and the passage of the Bayh-Dole Act. Furthermore, we note that some have argued these developments were accompanied by cultural changes on university campuses, which increased the propensity to patent inventions and more carefully manage the associated knowledge. So, by designing our study such that Period 1 is at the beginning of this time of change and Period 2 is approximately 10 years later, we aim to identify whether this duration of heightened intellectual property protection is accompanied by a narrowing of knowledge flows.

Furthermore, we are somewhat restricted in our choice of periods for two data-related reasons: (1) Since we base our analysis on citation data, we are limited in how early we can set Period 1 and how late we can set Period 2 due to truncation; our chosen periods allow 10 years of reasonably good quality citation data before Period 1 and after Period 2, and (2) We need sufficient time between Periods 1 and 2 to be able to observe any measurable changes in knowledge flow patterns; the approximately 10 years between our two periods is sufficient for this purpose. Moreover, our choice of periods is consistent with other related empirical research, such as Hall and Ziedonis (2001), who compare the periods 1979–1983 and 1989–1993 in part to study the determinants driving patenting in the US semiconductor industry over time.

Our sampling procedure results in 262,731 patents. From this set, we keep only those that receive at least two citations since our forward fragmentation and generality measures are otherwise undefined.[29,30]

Next, we define the specific citations we consider. We ignore assignee self-citations because we are interested in how knowledge flows across organizations in the economy. Furthermore, we do not consider citations received from patents applied for before the focal patent was issued since these citations are unlikely to represent knowledge flows due to the secrecy usually maintained during the patenting process. Finally, due to truncation issues, we remove citations that come from patents issued more than 10 years after the focal patent issue date.[31] Consequently, we are left with a final sample containing 194,500 focal patents that are, on average, referenced by 9.27 citing patents.

3.1.2. Knowledge inflows sample

We begin constructing this sample with the 281,963 focal patents issued to US non-government organizations that were applied for during 1980–1983 and 1989–1992. Similar to the outflows sample construction, we then remove patents that do not make at least two citations since our dependant variable, *BackFrag*, as well as our measure of originality are undefined for these patents. In addition, we only consider citations with particular characteristics. Since we are concerned about potential anti-commons effects on knowledge inflows, we only consider cited patents that can potentially hold up the utilization of follow-on inventions. Therefore, we focus on cited patents not owned by the focal assignee and that were issued before (but no more than 10 years before) the application of the focal patent. After this data processing, we generate a sample that includes 203,521 focal patents that, on average, cite 5.80 prior patents.

3.2. Data limitations

Though rich, our data have limitations. Most notably, some of the patents in the data do not include assignee information. This is problematic since we construct our dependent variable, the fragmentation index, using this information. As described in Hall et al. (2002), 18.4% of all patents in the NBER database have unidentified owners. We take a number of steps to minimize this problem. First, by construction, we only use focal patents for which we have assignee information. Recall that we draw our initial set of patents from those issued to US non-government organizations. Thus, only our citing patents may be missing assignee information.[32] Next, since we apply a 10-year window for constructing our backward fragmentation index and older patents are more likely to be missing assignee information, we further limit our exposure to this problem. As a result of these measures, only 14.1% of the citations made by our sampled patents are to unassigned patents and only 11.3% of citations received are from unassigned patents. Alternatively, each sampled patent, on average, cites 0.82 unassigned patents and receives 1.05 citations from unassigned patents. When calculating our fragmentation measure, we assume unassigned patents are not self-citations and that each belongs to a different assignee. We confirm the robustness of our result to the latter assumption by also examining the subset of focal patents for which we have complete information over the assignee identity of citations. The main result persists.

A second limitation of the data is the absence of ownership transfer information. We base our fragmentation measure on the assignee identified at the time each patent is issued. However, Serrano (2005) finds that the sale and purchase of patents is not uncommon. The uncorrected reassignment of patents would only bias our result if the likelihood of ownership transfer changed at a different rate for patents that cite university inventions than those that cite firms. The limited literature on this topic does not indicate any reason to believe this is the case.

A third limitation concerns noise in the assignee name data (e.g., spelling mistakes, abbreviations, changes in name for reasons such as acquisition). This is problematic since we base the construction of our fragmentation measure on the uniqueness of assignee names. We address this issue in Section 4.6.

A fourth limitation is the noisiness of citation data as a proxy for knowledge flows. Citations are a noisy measure since they are often added by patent examiners rather than inventors themselves (Alcacer and Gittelman, 2006; Hegde and Sampat, 2007). Furthermore, examiners do not add patents "randomly." For example, Cockburn et al. (2004) report that some examiners have "favorite" patents that they cite preferentially because they "teach the art" particularly well. Nonetheless, we believe that even examiner-added citations may reflect a knowledge flow. For example, Alcacer

[29] This follows from the definition of our forward fragmentation measure as defined in Eq. (2).

[30] We compare those patents that we remove from the sample due to too few citations with those that remain. The sets are similar on two key dimensions. First, 97.1% of the patents in the outflow sample were assigned to firms (rather than universities), compared to 97.9% of the patents removed from the sample. Similarly, the values for the inflow sample are 97.2% and 96.3%, respectively. Furthermore, these percentages do not change dramatically over time: 98.7% of excluded patents were assigned to firms in Period 1 and 97.0% in Period 2. Second, 8.6%, 19.6%, and 18.4% of the outflow sample are classified as drugs and medical, electronic, and mechanical, respectively, which is similar to the 7.6%, 15.0%, and 21.6% of the dropped outflow sample.

[31] We use a constant citation window to ensure our Period 1 and Period 2 measures are comparable. We are limited to a 10-year window since the updated NBER patent database contains citation data up to 2002.

[32] Similarly, for the knowledge inflows case, only our cited patents may be missing assignee information.

Table 1
Summary statistics, means, standard deviations and difference in means.

	Period 1: 1980–1983			Period 2: 1989–1992		
	University I	Firm II	Differences I – II	University III	Firm IV	Differences III – IV
Panel A: knowledge OUTFLOWS						
Forward fragmentation	0.910 (0.184)	0.887 (0.217)	0.023	0.858 (0.204)	0.867 (0.209)	−0.009
Generality	0.570 (0.332)	0.538 (0.361)	0.032	0.594 (0.301)	0.559 (0.326)	0.036
Citations received	8.651 (10.128)	6.382 (5.958)	2.269	13.619 (16.859)	10.839 (13.908)	2.779
Observations	1,127	70,709		4,301	118,363	
Panel B: knowledge INFLOWS						
Backward fragmentation	0.922 (0.195)	0.907 (0.211)	0.015	0.913 (0.199)	0.913 (0.190)	0.000
Originality	0.526 (0.368)	0.511 (0.383)	0.015	0.559 (0.354)	0.510 (0.362)	0.049
Citations made	5.553 (4.875)	4.821 (3.723)	0.732	6.531 (5.622)	6.390 (5.895)	0.141
Observations	1,302	76,546		4,198	121,475	

Note: Standard deviations in parentheses.

and Gittelman (2006) and Thompson (2006) find that examiners add a large portion of inventor and assignee self-citations, respectively, suggesting that examiners fill in reporting gaps (as opposed to knowledge gaps) arising from the "casualness" with which inventors complete their patent applications.

Moreover, Jaffe et al. (2002) survey cited and citing inventors to explore the "meaning of patent citations" and find that approximately one-quarter of the survey responses correspond to a "fairly clear spillover," approximately one-half indicate no spillover, and the remaining quarter indicate some possibility of a spillover. Based on their survey data, the authors conclude that "these results are consistent with the notion that citations are a noisy signal of the presence of spillovers. This implies that aggregate citation flows can be used as proxies for knowledge-spillover intensity, for example, between categories of organizations or between geographic regions" (p. 400). Subsequent studies that focus on examiner-added citations seem to support the conclusions reached by Jaffe et al. For example, Alcacer and Gittelman (2006) report results suggesting that, overall, examiner-added citations "do not change the presumption that patents trace out knowledge flows." Similarly, Sampat (2004) states that his findings concerning examiner-added citations are "broadly consistent with Jaffe et al."

Finally, it is important to note that to the extent that at least some fraction of examiner-added citations does not reflect knowledge flows, this could inflate our fragmentation index. However, this alone would not bias our main result, which is based on a comparison of universities and firms. In fact, even if examiners are systematically more or less likely to add citations to university patents, as compared to firm patents, this also would not bias our result since we focus on the extent to which the difference in fragmentation changes over time. Hence, for examiner-added citations to bias our main result, their relative roles with respect to university versus firm patents would have had to change between the two periods. We have no reason to believe this is the case.

4. Results

4.1. Summary statistics

We present summary statistics in Table 1 confirming the findings of Henderson et al. (1998) that university patents are more important, general, and original than firm patents. Beginning with Panel A, which presents data for the knowledge outflows sample, we see that university patents were more important (i.e., they received more citations) in both Periods 1 and 2. Similarly, university patents were more general in both periods. Turning to Panel B, we see university patents were also more original, and this difference seems to have increased over time.

Next, we consider our variable of interest, the fragmentation index.[33] Beginning with Panel A, we see that knowledge outflows from university patents were more fragmented than their private sector counterparts in Period 1. (We explain how to interpret the difference in index values in Section 4.5.) However, this difference seems to have disappeared by Period 2. Similarly, in Panel B, we see that knowledge inflows to university patents were more fragmented than those to firm patents in Period 1. Again, however, this difference seems to have disappeared by Period 2.

Although these statistics suggest a narrowing of university relative to firm knowledge flows, changes in institution-related fragmentation measures could be confounded with changes in non-institutional factors (such as technology field portfolio), as noted in the methodology section above. Thus, we turn next to regression analysis, which allows us to control for key invention characteristics.

4.2. Regression analysis: dispersion of outflows

We report the marginal effects of each variable specified in Eq. (1) based on coefficients estimated with fractional logit regressions (Table 2).[34] Recall that the dependent variable in this case is $ForFrag_{i,p}$. Referencing the fully specified model reported in Column II, we see from the estimated coefficient on the university dummy that university patents in Period 1 were more fragmented than their private-sector counterparts, even after controlling for the importance, generality, and technology field of the invention. On average, university fragmentation was 2.7% greater than firms. We refer to this difference – the degree to which university knowl-

[33] The forward fragmentation variable is positively correlated with the generality variable (Pearson correlation coefficient: 0.118). This is not surprising since we construct the two measures in a similar manner and base both on similar forward citation data. Whereas generality is a measure of concentration by technology field, fragmentation is a measure of concentration by assignee. Similarly, the correlation coefficient for backward fragmentation and originality is 0.114.

[34] Tables 2–4 report marginal effects rather than the actual estimated fractional logit coefficients so that we can interpret the results more easily (similar to OLS coefficient estimates). We use fractional logit estimation because the dependent variable is the fragmentation index, which is bound between zero and one. Papke and Wooldridge (1996) clearly describes fractional logit estimation. To implement this estimation technique, we assume a logistic functional form for the conditional mean of our fragmentation measure and estimate the parameters by quasi-maximum likelihood estimation. Using this procedure yields estimates that take values within the unit interval. The estimated coefficients do not directly provide the marginal effects; since we assume a non-linear functional form for the conditional mean of the dependant variable, we calculate the marginal effects as suggested by Ai and Norton (2003). We have confirmed that, at the sample mean, the marginal effect of each variable is very close in magnitude and significance to OLS estimates.

Table 2
FLOGIT, OUTFLOW fragmentation dependant variable regressions.

	Full sample		Subcat 31 & 33	Full sample except subcat 31 & 33	Full sample different uni. types
	I	II	III	IV	V
University dummy	0.032 ***(4.830)	0.027 ***(3.934)	0.046 **(2.327)	0.026 ***(3.863)	
University dummy × Period 2 dummy	−0.017 **(−2.248)	−0.016 **(−1.974)	−0.053 ***(−2.331)	−0.009 (−1.238)	
Citations received		0.000 (1.222)	0.004 ***(4.906)	0.000 (−0.106)	0.000 (1.226)
Citations received × Period 2 dummy		−0.001 ***(−3.603)	−0.004 ***(−4.400)	0.000 ***(−2.652)	−0.001 ***(−3.602)
Generality		0.071 ***(24.555)	0.118 ***(7.305)	0.069 ***(23.941)	0.071 ***(24.540)
Generality × Period 2 dummy		0.011 ***(3.143)	0.045 **(1.998)	0.009 ***(2.646)	0.011 ***(3.145)
Carnegie I					
High experience					0.026 ***(3.030)
High experience × Period 2 dummy					−0.009 (−0.997)
Low experience					0.008 (0.322)
Low experience × Period 2 dummy					−0.009 (−0.302)
Carnegie II					
High experience					0.049 ***(4.374)
High experience × Period 2 dummy					−0.041 ***(−3.195)
Low experience					0.052 ***(3.146)
Low experience × Period 2 dummy					−0.056 **(−2.459)
Other					
High experience					0.007 (0.420)
High experience × Period 2 dummy					0.024 (0.804)
Low experience					0.040 **(1.965)
Low experience × Period 2 dummy					−0.037 *(−1.647)
Tech. fixed effects	YES	YES	YES	YES	YES
Tech. fixed effects × Period 2 dummy	YES	YES	YES	YES	YES
Loglikelihood	−59,579	−58,878	−3,621	−55,240	−58,876
Observations	194,500	194,500	8,890	185,610	194,500
Clusters	29,589	29,589	1,365	29,022	29,589

Notes: The estimates reported here are marginal effects evaluated at the mean. The *Z*-statistics, shown in brackets, are clustered on assignee name and calculated using the delta method.
* Statistical significance: 10%.
** Statistical significance: 5%.
*** Statistical significance: 1%.

edge flows are more widely distributed across assignees than firm flows – as the university diffusion premium.

Turning to the coefficient on the interaction between the university and Period 2 dummies, we see that the university diffusion premium decreased by 0.016 in absolute terms. Thus, the premium was significantly diminished by the second period. In fact, by comparing the magnitudes of this coefficient with the coefficient on the university dummy (with no interaction), we see that the university diffusion premium measured in Period 1 was reduced by approximately 59% by Period 2. This is our main result. We elaborate on its interpretation in Section 4.5.

We next explore the degree to which the identified effect is concentrated in particular technology fields, noting that we have included field fixed effects to account for patenting activity shifting from one field to another. We find that two subfields mostly drive our main result: drugs and biotechnology (subcategories 31 and 33). We draw this conclusion from the result reported in Column III, which presents the estimated coefficients for the subsample that includes only these two fields where the effect is strong (the coefficients on the university dummy and the university–Period 2 interaction are 0.046 and −0.053, respectively, and both estimates are statistically significant). In contrast, when we exclude these two fields (Column IV), the estimated coefficients indicate that the narrowing effect is absent (the coefficient on the interaction term is not significantly different from zero).

4.3. Regression analysis: diversity of inflows

We turn next to examine the concentration of knowledge *inflows*. As described in the introduction, although the economic forces affecting the concentration of inflows were different than those affecting outflows, the econometric approach to identifying changes in concentration is much the same.

We report the estimated coefficients of Eq. (1) for the knowledge inflows sample in Table 3. Recall that the dependent variable in this case is $BackFrag_{i,p}$. Referencing the fully specified model reported in Column II, we see from the estimated coefficient on the university dummy that university patents in Period 1 were more fragmented than their private-sector counterparts, even after controlling for the originality, technology field, and overall number of citations made. On average, university fragmentation was 2.2% greater than firms. We refer to this difference – the degree to which university knowledge inflows are drawn from a more diverse set of prior art holders than firm inflows – as the university diversity premium.

Turning to the coefficient on the interaction between the university and Period 2 dummies, we see that the university diversity premium decreased by 0.013 in absolute terms. Thus, the premium was significantly diminished by the second period. In fact, by comparing the magnitudes of this coefficient with the coefficient on the university dummy (with no interaction), we see that the university diversity premium measured in Period 1 was reduced by 59% by Period 2.

We next explore the degree to which the identified effect is concentrated in particular technology fields, noting that we have included field fixed effects to account for patenting activity shifting from one field to another. We find that four subfields of electronics mostly drive our main result: measuring and testing, power systems, semiconductor devices, and miscellaneous electrical (subcategories 43, 45, 46, and 49). We report this result in Column III, which presents the estimated coefficients for the subsample that includes only these four fields where the effect is strong (the coefficients on the university dummy and the university–Period 2 interaction are 0.031 and −0.053, respectively, and both estimates

Table 3
FLOGIT, INFLOW fragmentation dependant variable regressions.

	Full sample		Subcat 43, 45, 46 & 49	Full sample except subcat 43, 45, 46 & 49	Full sample different uni. types
	I	II	III	IV	V
University dummy	0.024 ***(3.813)	0.022 ***(3.656)	0.031 ***(4.265)	0.020 ***(3.064)	
University dummy × Period 2 dummy	−0.012* (−1.776)	−0.013 **(−2.049)	−0.053 ***(−4.303)	−0.007 (−0.957)	
Citations made		0.001 ***(2.759)	0.001 **(2.565)	0.000 **(2.215)	0.001 ***(2.764)
Citations made × Period 2 dummy		−0.000 (−0.809)	−0.002 **(−2.511)	−0.000 (−0.085)	−0.000 (−0.830)
Originality		0.057 ***(21.337)	0.047 ***(8.145)	0.058 ***(19.640)	0.057 ***(21.341)
Originality × Period 2 dummy		0.009 ***(2.612)	0.005 (0.691)	0.010 **(2.489)	0.009 ***(2.619)
Carnegie I					
High experience					0.021 ***(2.745)
High experience × Period 2 dummy					−0.011 (−1.462)
Low experience					0.026 (1.496)
Low experience × Period 2 dummy					−0.016 (−0.883)
Carnegie II					
High experience					0.036 ***(3.293)
High experience × Period 2 dummy					−0.028 *(−1.789)
Low experience					0.002 (0.084)
Low experience × Period 2 dummy					0.004 (0.136)
Other					
High experience					0.009 (0.386)
High experience × Period 2 dummy					−0.057 (−1.139)
Low experience					0.032 ***(2.958)
Low experience × Period 2 dummy					−0.013 (−0.912)
Tech. fixed effects	YES	YES	YES	YES	YES
Tech. fixed effects × Period 2 dummy	YES	YES	YES	YES	YES
Loglikelihood	−51,369	−50,568	−5,366	−45,195	−50,564
Observations	203,521	203,521	24,767	178,754	203,521
Clusters	30,827	30,827	4,614	28,907	30,827

Notes: The estimates reported here are marginal effects evaluated at the mean. The *Z*-statistics, shown in brackets, are clustered on assignee name and calculated using the delta method.

* Statistical significance: 10%.

** Statistical significance: 5%.

*** Statistical significance: 1%.

are statistically significant). In contrast, when we exclude these four fields (Column IV), the estimated coefficients indicate that the narrowing effect is absent (the coefficient on the interaction term is not significantly different from zero).

4.4. University patenting experience

The rapid rise in university patenting that occurred during the 1980s reflects a significant change in the overall landscape with respect to academia's approach to the management of intellectual property. During this period, many universities that did not have a formal technology transfer office established one and created standardized procedures for managing the disclosure, patenting, and licensing processes (Mowery et al., 2004). In addition, much of the increase in patent activity came from "inexperienced" institutions that had been issued few patents prior to 1980.

The increasing role of these inexperienced institutions in university patenting influenced the overall character of the "average" university patent. Indeed Mowery et al. (2004) show that the decrease in importance and generality of university patents over time identified by Henderson et al. (1998) was due to the entry of inexperienced schools. The implication of the Mowery et al finding is key; since the measured decrease in importance and generality was due to the entry of inexperienced universities, the effect was likely temporary while these schools learned to manage their intellectual property to become more like their experienced counterparts.

Since our study is similar in spirit to these papers, we check whether our effect is a result of entry by inexperienced universities. To accomplish this, we categorize our university patents in a similar way to Mowery et al. We divide universities into two categories based on their patenting experience prior to 1981. We define: (1) High Experience Universities as those obtaining at least 10 patents that were applied for after 1970 but before 1981 and (2) Low Experience Universities as those that obtained less than 10 during the same period. Based on this categorization, experienced universities account for 87% (983) and 65% (2,790) of the focal university patents in Periods 1 and 2, respectively. (For the knowledge inflows case, experienced universities account for 83% (1,084) and 66% (2,760) of the focal university patents in Periods 1 and 2, respectively.)

We also categorize universities by research resources as defined by the Carnegie Foundation.[35] The regression results in Column V (Table 2) show that, unlike the Mowery et al result, ours is not driven by the entry of inexperienced universities alone; rather, the narrowing phenomenon is also strong among experienced universities. Interestingly, however, even after controlling for experience, the narrowing seems to have been driven mostly by Carnegie II universities.[36] One might speculate that researchers at these institutions were willing to license their inventions more broadly during Period 1, after which cultural norms spread from Carnegie I to Carnegie II schools resulting in the reduced fragmentation.

In the knowledge inflows case, the narrowing result is at least partially driven by experienced universities and is again attributed

[35] We categorize universities as Carnegie I, Carnegie II, or Other, based on criteria such as level of research funding and breadth of doctoral programs, where Carnegie I universities are the most research oriented. The changing role of Carnegie II universities, relative to Carnegie I, is explored in Agrawal and Goldfarb (2008). The Carnegie Foundation for the Advancement of Teaching (1987) provides precise definitions of each category.

[36] It is important to note that the estimated coefficients here reflect marginal effects relative to the average fragmentation of firm patents. For example, we estimate that the Period 1 marginal effect of being a high experience, Carnegie I university is to increase fragmentation by 0.026 relative to the average firm patent in the same period.

to Carnegie II institutions (see Column V, Table 3). Overall, these results suggest that the narrowing of university knowledge flows, unlike the decline in importance and generality shown by Mowery et al, is not likely a temporary phenomenon that will diminish as universities gain experience although it will be bounded as the behavior at Carnegie II universities become more like that of their Carnegie I counterparts.

4.5. Interpretation of fragmentation index values

The meaning of the fragmentation index, our dependent variable, can be difficult to comprehend. Similar to the Herfindahl index, which, although often used in market concentration studies, is usually accompanied by the more intuitive "four firm concentration ratio," the fragmentation index is complex. This is because many states of the world (i.e., combinations of citation frequencies and assignee distributions) can generate similar values. Although the index is complicated, however, it is important to understand. Throughout most of the discussion so far, we have described changes in university knowledge flow concentration in relative terms. In other words, we have discussed the change in the university premium rather than the absolute change in the concentration of university knowledge flows. While the relative change in concentration between periods seems large (>50%), the absolute change seems small (<5%). Ultimately, we are interested in whether the change is economically important. To this end, we offer an example to help the reader develop intuition for comprehending the economic significance of the estimated changes in knowledge flow concentrations.[37]

Consider a patent that receives nine citations, roughly the mean number of citations received by university focal patents in Period 1. Further, suppose these citations are from six different assignees. If three different assignees each cite the patent twice while the other three assignees only cite the patent once, then the fragmentation measure equals 0.917.[38] To decrease the fragmentation measure by 0.053 (approximately the value of the coefficient on the university–Period 2 interaction term when we focus on drug and biotech patents) while holding constant the total number of citations, one less assignee would have to cite the patent. In this case, two assignees would each cite the patent once, two assignees would each cite the patent twice and the last assignee would cite the patent three times. With this distribution, the fragmentation index would decrease to about 0.861.[39,40]

4.6. Robustness to noise in assignee name data

The assignee name data in the NBER patent database, upon which we base our dependent variable, is notoriously noisy for a variety of reasons: misspellings, multiple names, subsidiaries, and mergers and acquisitions. The noise in these data could inflate the measurement of our dependent variable since assignee names that seem to represent different organizations may actually represent the same one. However, even though the noise in these data may inflate the measurement of our dependent variable, it does not bias our main result. Even if errors in name data are more likely for organizations that cite university patents than those that cite firm patents, although the estimated coefficient on the university dummy would be biased upwards, the main result (i.e., the coefficient on the university–Period 2 interaction) would not. Furthermore, we have no reason to believe such name data errors are biased in any case. Moreover, for the estimate of the coefficient on the interaction to be biased in the direction of our finding, the error in assignee name data associated with organizations that cite universities must decrease at a rate faster than that associated with organizations that cite firms. We are not aware of any reason that this might be the case.

We further minimize the possibility that noise in the assignee name data does not drive our results by performing a manual cleaning of a sub-sample of the data. We describe the manual cleaning process for the forward citation case below. The procedure for the backwards citation data is identical.

We focus the robustness check on our main results (narrowing of knowledge flows in particular fields: NBER Category 3 – Drugs and Medical, for forward citations; Category 4 – Electrical and Electronic, for backward citations). We randomly draw 20 patents with issue dates in Period 1 (1980–1983) and primary field classifications in Category 3 that have a university assignee. Next, we randomly match each university patent to a firm patent that: (1) was also issued in Period 1, (2) received an equal number of citations, and (3) was issued to the same two-digit technology field. We then repeat the process for Period 2 (1989–1992), randomly drawing 20 university patents that match the Period 1 university patents in terms of the two-digit technology class distribution; we then randomly draw 20 Period 2 firm patents that match the Period 2 university patents on technology class distribution and citations received. This process produces a total of 80 patents (20 university and 20 firm patents in Period 1 and 20 of each assignee type in Period 2).

We then manually inspect the citation data associated with each of the 80 patents. Specifically, we review the assignee names associated with each individual patent to identify and correct any misspelling or multiple names that would result in assignees of citing patents appearing as distinct organizations when in fact they are the same (and thus erroneously inflating the fragmentation measure). Next, using the *Who Owns Whom* directory published by Dun & Bradstreet, we map the existing assignee names to the name of the parent company to account for subsidiaries and mergers and acquisitions (corporate ownership as of the application year of the focal patent's latest citing patent).[41] We run the base regression on this sub-sample and report the raw and manually cleaned results in Table 4. The coefficient estimates for the university dummy and

[37] Rosell and Agrawal (2006) provide additional intuition concerning the interpretation of fragmentation values.

[38] $ForFrag_p = \{1 - 3(1/9^2) - 3(2^2/9^2)\}(9/9 - 1) \simeq 0.917$.

[39] $ForFrag_p = \{1 - 2(1/9^2) - 2(2^2/9^2) - (3^2/9^2)\}(9/9 - 1) \simeq 0.861$.

[40] It is important to note that for illustrative purposes we hold the number of citations constant in the above example. However, if the number of citations increases, it is possible that the number of unique citing assignees also increases, even though the fragmentation measure decreases. Based on our data, this is in fact the case. The number of unique organizations citing university patents increased over the period under investigation, from an average of 6.53–7.73. However, the relative increase in the number of organizations citing university patents is less than that for firms. In other words, although the number of organizations citing university patents increased, it did not increase as fast as we would expect. Since the overall level of patenting was increasing over time, it is not surprising that on average more organizations were citing more recent patents given the constant citation window. We thus measure the increase in the number of organizations citing firm patents to control for the ambient level of increasing citations. So, while university patents received citations from 34% more organizations than firm patents on average in Period 1, they only received citations from 15% more in Period 2. Thus, we conclude that university knowledge flows were narrowing to a *relatively* smaller set of recipients. This point is particularly salient in the key technology fields where our main effect is strongest. In subcategories 31 and 33 (drugs and biotechnology, respectively), university patents received citations from 40% more organizations than firm patents on average in Period 1 but only from 5% more in Period 2.

[41] Manually cleaning the name data to address misspellings, subsidiaries, and mergers/acquisitions results in a small reduction in the number of unique citing/cited assignees. In the outflows case, for example, the number of unique assignees citing university patents drops from 92 to 85 and 97 to 94 for Periods 1 and 2, respectively. Similarly, the number of unique assignees citing firm patents drops from 85 to 82 and 96 to 94 for Periods 1 and 2, respectively.

Table 4
Effects of fragmentation errors on FLOGIT regressions.

	Outflow fragmentation		Inflow fragmentation	
	Sample data	Manually cleaned data	Sample data	Manually cleaned data
University dummy	0.084 ***(3.329)	0.088 ***(3.444)	0.033 **(2.102)	0.040 **(2.424)
University dummy × Period 2 dummy	−0.114 *(−1.926)	−0.112 *(−1.664)	−0.067 ***(−3.548)	−0.076 ***(−3.553)
Citations received	0.011 **(2.444)	0.011 ***(2.193)		
Citations received × Period 2 dummy	−0.015 **(−2.209)	−0.017 **(−2.167)		
Generality	0.024 (0.758)	0.023 (0.699)		
Generality × Period 2 dummy	0.240 **(2.353)	0.191 *(1.648)		
Citations made			0.002 (0.699)	0.002 (1.047)
Citations made × Period 2 dummy			−0.002 (−0.442)	−0.001 (−0.549)
Originality			0.024 (1.241)	0.011 (0.866)
Originality × Period 2 dummy			−0.019 (−0.738)	−0.004 (−0.202)
Tech. fixed effects	YES	YES	YES	YES
Tech. fixed effects × Period 2 dummy	YES	YES	YES	YES
Loglikelihood	−22.820	−23.689	−18.054	−16.396
Observations	80	80	80	80
Number of clusters	70	70	51	51

Notes: The estimates reported here are marginal effects evaluated at the mean. The *Z*-statistics, shown in brackets, are clustered on assignee name and calculated using the delta method.
* Statistical significance: 10%.
** Statistical significance: 5%.
*** Statistical significance: 1%.

the university–Period 2 interaction are virtually identical across the two samples. In other words, the main result persists.

5. Conclusion

Researchers have examined the dramatic rise in university patenting over the 1980s along a variety of dimensions. To our knowledge, this paper is the first study to determine that the increasing trend towards formal intellectual property protection was accompanied by a contraction in the breadth of knowledge flows. However, although the degree of narrowing was large relative to the magnitude of the premium universities exhibited over firms prior to the surge in patenting, the absolute change was modest.

Furthermore, the narrowing of knowledge flows was not uniform across fields. In fact, in the case of outflows, two related fields primarily drive the identified effect: biotechnology and pharmaceuticals. Prior studies on the commercialization of university research have identified these fields as outliers on a variety of dimensions (Powell and Owen-Smith, 1998; Cohen and Walsh, 2001; Walsh et al., 2003). It is well known that patent protection works particularly well in biotechnology and pharmaceuticals since: (1) enforcement is feasible (infringement is detectable in contrast to, for example, software), (2) the time required for patent examination and granting as well as the duration of protection (17 years at that time) is commensurate with the time required for drug development, FDA approval, and the product lifecycle, and (3) the importance of exclusivity is heightened due to the high upfront costs and risks associated with the drug discovery and approval processes. As such, the narrowing of knowledge outflows may be viewed as a somewhat isolated phenomenon that was at least partially justified given the economics and evolution of these particular industries. Importantly, it does not seem to have been an overall trend influencing all fields of university research. However, experienced universities at least partly drove the changes, suggesting that the identified effect is likely not temporary.

The explanation for the narrowing of knowledge inflows used by researchers in particular subfields of electronics is likely similar to that as described by Ziedonis (2004), whose study is set in the specific context of the semiconductor industry. To the extent that the ownership of related intellectual property in certain sub-fields in electronics is particularly fragmented, inventors in these fields might look forward and reason back in a manner that leads them to draw upon prior art in a manner that minimizes the potential for hold-up that future licensees will face. Our data suggest the salience of this issue has increased more in electronics than other fields in the university setting, although we have found no other evidence of this in the literature.

Earlier, we stated that our objective is to identify the existence of the narrowing phenomenon and examine the scope of fields over which it applies. We have done this. Furthermore, we speculate on several possible causal mechanisms involving researchers, technology licensing offices, and licensees. The next step in this line of inquiry is to unpack the relative importance of these mechanisms. However, this empirical identification exercise will be nontrivial given the lack of exogenous shocks leading to measurable variation at appropriate levels of aggregation (perhaps explaining the abundance of research focusing on the Bayh-Dole Act). Still, this does not diminish the importance of answering this question.

Finally, what are the broader social welfare implications of our findings? Although it is tempting to assume that the narrowing of knowledge flows is welfare reducing, this is not necessarily the case. Knowing that knowledge spillovers contribute to economic growth but also recognizing the importance of exclusivity for creating incentives to develop and commercialize, scholars are unclear as to how the narrowing of university knowledge flows affects welfare. That said, the striking rise in university patenting and the associated narrowing of knowledge flows in particular fields surely do influence welfare in some way. Thus, given the importance of university knowledge to both industrial competitiveness and social welfare, this topic warrants further research.

Acknowledgements

We are grateful to Stefano Breschi, Azim Essaji, Maryann Feldman, Avi Goldfarb, Brent Goldfarb, Aamir Hashmi, Ignatius Horstman, Rhys Mendes, Diego Puga, Frank Rothaermel, Bhaven Sampat, Mark Schankerman, Tim Simcoe, Nadia Soboleva, Marie Thursby, Daniel Trefler, Rosemarie Ziedonis, and participants at the 2005 Canadian Economic Association Meeting, 2005 Roundtable for Engineering Entrepreneurship Research Conference (Georgia Tech), and 2006 ExTra Workshop (EPFL) for helpful comments. We

also thank the editor, Woody Powell, and three anonymous referees for many thoughtful suggestions. Errors and omissions are our own. This research was partially funded by the Social Sciences and Humanities Research Council of Canada (Grant No. 410-2004-1770) and the Martin Prosperity Institute. Their support is gratefully acknowledged.

References

Agrawal, A., 2006. Engaging the inventor: exploring licensing strategies for university inventions and the role of latent knowledge. Strategic Management Journal 27 (1), 63–79.

Agrawal, A., Garlappi, L., 2007. Public sector science and the strategy of the commons. Economics of Innovation and New Technology 16 (7–8), 517–539.

Agrawal, A., Goldfarb, A., 2008. Restructuring research: communication costs and the democratization of university innovation. The American Economic Review 98 (4), 1578–1590.

Agrawal, A., Henderson, R., 2002. Putting patents in context: exploring knowledge transfer from MIT. Management Science 48 (1), 44–60.

Ai, C., Norton, E., 2003. Interaction terms in logit and probit models. Economic Letters 80 (1), 123–129.

Alcacer, J., Gittelman, M., 2006. How do i know what you know? Patent examiners and the generation of patent citations. Review of Economics and Statistics 88 (4), 774–779.

Argyres, N., Liebskind, J., 1998. Privatizing the intellectual commons: universities and the commercialization of biotechnology research. Journal of Economic Behavior and Organization 35 (4), 427–454.

Bercovitz, J., Feldman, M., 2008. Academic entrepreneurs: organizational change at the individual level. Organization Science 19 (1), 69–89.

Breschi, S., Lissoni, F., Montobbio, F., 2007. The scientific productivity of academic inventors: new evidence from Italian data. Economics of Innovation and New Technology 16 (2), 101–118.

Buenstorf, G., 2005. Commercializing basic science as a competitor or complement of academic accomplishment? The case of Max Planck directors. Max Planck Institute of Economics, mimeo.

Calerini, M., Franzoni, C., 2004. Is academic patenting detrimental to high quality research? An empirical analysis of the relationship between scientific careers and patent applications. CESPRI Working Paper No. 162.

Carayol, N., 2007. Academic incentives, research organization, and patenting at a large French university. Economics of Innovation and New Technology 16 (1–2), 119–138.

Cockburn, I., Kortum, S., Stern, S., 2004. Are all patent examiners equal? The impact of examiner characteristics on patent statistics and litigation outcomes. In: Cohen, W., Merrill, S. (Eds.), Patents in the Knowledge-based Economy. National Academy Press.

Cohen, W., Florida, R., Randazzese, L., Walsh, J., 1998. Industry and the academy: uneasy partners in the cause of technological advance. In: Noll, R. (Ed.), Challenges to the Research University. Brookings Institution, pp. 171–200.

Cohen, W., Walsh, J., 2001. Public research, patents, and implications for industrial R&D in the drug, biotechnology, semiconductor and computer industries. In: Wessner, C. (Ed.), Capitalizing on New Needs and New Opportunities: Government-Industry Partnerships in Biotechnology and Information Technologies. National Academy Press.

Colyvas, J., Crow, M., Gelijns, A., Mazzoleni, R., Nelson, R., Rosenberg, N., Sampat, B., 2002. How do university inventions get into practice? Management Science 48 (1), 61–72.

Dranove, D., Kessler, D., McClellan, M., Satterthwaite, M., 2003. Is more information better? The effects of report cards on health care providers. Journal of Political Economy 111 (3), 555–588.

Easterly, W., Levine, R., 1997. Africa's growth tragedy: policies and ethnic division. Quarterly Journal of Economics 112 (4), 1203–1250.

Eisenberg, R., 2003. Patent swords and shields. Science 299 (5609), 1018.

Fabrizio, K., DiMinin, A., 2008. Commercializing the laboratory: faculty patenting and the open science environment. Research Policy 37, 914–931.

Goldfarb, B., Marshke, G., Smith, A., 2006. Scholarship and Inventive Activity in the University: Complements or Substitutes? University of Maryland, Mimeo.

Hall, B., Jaffe, A., Trajtenberg, M., 2002. The NBER patent citations data file: lessons, insights and methodological tools. In: Jaffe, A., Trajtenberg, M. (Eds.), Patents, Citations, & Innovations: A Window on the Knowledge Economy. The MIT Press, pp. 403–459.

Hall, B., Ziedonis, R.H., 2001. The patent paradox revisited: an empirical study of patenting in the U.S. semiconductor industry, 1979–1995. The Rand Journal of Economics 32 (1), 101–128.

Hegde, D., Sampat, B., 2007. Examiner citations, applicant citations, and the private value of patents. Working paper available at: http://faculty.haas.berkeley.edu/hegde/research.html.

Heller, M., Eisenberg, R., 1998. Can patents deter innovation? The anticommons in biomedical research. Science 280 (5364), 698–701.

Henderson, R., Jaffe, A., Trajtenberg, M., 1965–1988. Universities as a source of commercial technology: a detailed analysis of university patenting. The Review of Economics and Statistics 80 (1), 119–127.

Information Products Division, Technology Assessment and Forecast Branch, 2005. U.S. colleges and universities—utility patent grants, calendar years 1969–2005. Tech. Rep. U.S. Patent and Trademark Office. http://www.uspto.gov/go/oeip/taf/univ/univtoc.htm.

Jaffe, A., Lerner, J., 2004. Innovation and Its Discontents: How Our Broken Patent System is Endangering Innovation and Progress and What to Do About It. Princeton University Press.

Jaffe, A., Trajtenberg, M., Fogarty, M., 2002. The meaning of patent citations: report on the NBER/case western reserve survey of patentees. In: Jaffe, A., Trajtenberg, M. (Eds.), Patents, Citations, & Innovations: A Window on the Knowledge Economy. The MIT Press, pp. 379–402.

Jensen, R., Thursby, M., 2001. Proofs and prototypes for sale: the licensing of university inventions. American Economic Review 91 (1), 240–259.

Lach, S., Schankerman, M., 2008. Incentives and invention in universities. The Rand Journal of Economics 39 (2), 403–433.

Lieberwitz, R., 2005. Confronting the privatization and the commercialization of academic research: an analysis of social implications at the local, national, and global levels. Indiana Journal of Global Legal Studies 12 (1), 109–152.

Mazzoleni, R., 2005. University patents, R&D competition, and social welfare. Economics of Innovation and New Technology 14 (6), 499–515.

Mowery, D., Nelson, R., Sampat, B., Ziedonis, A., 2004. Ivory Tower and Industrial Innovation University-Industry Technology Transfer Before and After the Bayh-Dole Act in the United States. Stanford Business Books.

Murray, F., Stern, S., 2007. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. Journal of Economic Behavior & Organization 63, 648–687.

National Institutes of Health, 2005. Best practices for the licensing of genomic inventions. Tech. Rep. National Institutes of Health, Federal Register.

National Science Board, 2004. Science and engineering indicators 2004. Tech. Rep. National Science Board, Appendix, Table 4.8. http://www.nsf.gov/statistics/seind04.

Nelson, R., 1993. National Innovation Systems: A Comparative Analysis. Oxford University Press.

Owen-Smith, J., Powell, W., 2001. To patent or not: faculty decisions and institutional success at technology transfer. Journal of Technology Transfer 26 (1), 99–114.

Owen-Smith, J., Powell, W., 2003. The expanding role of university patenting in the life sciences: assessing the importance of experience and connectivity. Research Policy 32 (9), 1695–1711.

Papke, L., Wooldridge, J., 1996. Econometric methods for fractional response variables with an application to 401 (K) plan participation rates. Journal of Applied Econometrics 11 (6), 619–632.

Powell, W., Owen-Smith, J., 1998. Universities and the market for intellectual property in the life sciences. Journal of Policy Analysis and Management 17 (2), 253–277.

Rhoten, D., Powell, W., 2007. The frontiers of intellectual property: expanded protection vs. new models of open science. Annual Review of Law and Social Science 3, 345–373.

Rivette, K., Kline, D., 2000. Rembrandts in the Attic: Unlocking the Hidden Value of Patents. Harvard Business School Press.

Romer, P., 1986. Increasing returns and long-run growth. Journal of Political Economy 94 (5), 1002–1037.

Romer, P., 1990. Endogenous technological change. Journal of Political Economy 5, 98.

Rosell, C., Agrawal, A., 2006. University patenting: estimating the diminishing breadth of knowledge diffusion and consumption. National Bureau of Economic Research Working Paper 12,640.

Sampat, B., 2004. Examining patent examination: an analysis of examiner and applicant generated prior art. Working Paper. Georgia Institute of Technology.

Sampat, B., Mowery, D., Ziedonis, A., 2003. Changes in university patent quality after Bayh-Dole: a re-examination. International Journal of Industrial Organization 21, 1371–1390.

Serrano, C.J., 2005. The market for intellectual property: evidence from the transfer of patents. Working Paper. University of Toronto.

Shapiro, C., 2001. Navigating the patent thicket: cross licenses, patent pools and standard setting. Innovation Policy and the Economy 1, 119–150.

Stokes, D.E., 1997. Pasteur's Quadrant: Basic Science and Technological Innovation. Brookings Institution Press.

The Carnegie Foundation for the Advancement of Teaching, 1987. A classification of institutions of higher education, 1987 edition. Tech. Rep. The Carnegie Foundation for the Advancement of Teaching.

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

Thursby, M., Thursby, J., Dechenaux, E., 2005. Shirking, sharing risk, and shelving: the role of university contracts. National Bureau of Economic Research Working Paper 11, p. 128.

Trajtenberg, M., Henderson, R., Jaffe, A., 1997. University versus corporate patents: a window on the basicness of invention. Economics of Innovation and New Technology 5 (1), 19–50.

Van Looy, B., Callaert, J., Debackere, K., 2006. Publication and patent behavior of academic researchers: conflicting, reinforcing or Merely Co-existing? Research Policy 35 (4), 596–598.

Walsh, J., Arora, A., Cohen, W., 2003. Effects of research tool patents and licensing on biomedical innovation. In: Cohen, W., Merrill, S. (Eds.), Patents in the Knowledge-Based Economy. National Academy Press.
Walsh, J., Cho, C., Cohen, W., 2002. View from the bench: patents and material transfer. Science 309, 3.
Warshofsky, F., 1994. The Patent Wars: The Battle to Own the World's Technology. John Wiley & Sons, Inc.
Ziedonis, R.H., 2004. Don't fence me in: fragmented markets for technology and the patent acquisition strategies of firms. Management Science 50 (6), 804–820.