The Wandering Scholars: Understanding the Heterogeneity of University Commercialization*

Josh Lerner, Harvard University and NBER
Henry Manley, NBER
Carolyn Stein, UC Berkeley
Heidi Williams, Dartmouth College and NBER

January 16, 2024

Abstract

University-based scientific research has long been argued to be a central source of commercial innovation and economic growth. Yet at the same time, there have been long-held concerns that many university-based discoveries never realize their potential social benefits. Looking across universities, research and commercialization activities such as start-up formation vary tremendously – variation that could reflect the composition and orientation of faculty research, university-level factors such as patenting and licensing efforts, or broader place-based factors such as location in a technology cluster. We take a first step towards unpacking this heterogeneity in university commercialization by analyzing how the propensity of academic research to spill over to commercial innovation changes when academics move across universities. Our estimates suggest that at least 15–25% of geographic variation in commercial spillovers from university-based research is attributable to place-specific factors.

*We are grateful to Ryan Broll, AnneMarie Bryson, Clara Chen, Gideon Moore, Deepika Prabhakar, Maya Roy, Ralph Skinner, Zahra Thabet, and Anthony Wang for excellent research assistance. This research was funded in part by Harvard Business School's Division of Research and Doctoral Programs, Harvard Economics' SUPER Program, the Alfred P. Sloan Foundation, and the Smith Richardson Foundation. We thank Ashish Arora and Scott Stern (discussants), Isaac Kohlberg, Matt Marx, and participants in seminars at Dartmouth, Harvard, Stanford, UC-Berkeley, UC-Santa Barbara, and UCL, as well as the Econometric Society's North American Winter Meeting, the NBER Productivity Lunch, and the NBER Organizational Economics and Science of Science Funding meetings, for helpful comments. Lerner has received compensation from advising institutional investors in venture capital funds, venture capital groups, and governments designing policies relevant to innovation and venture capital. All errors and omissions are our own. Contact: jlerner@hbs.edu, manleyh@nber.org, carolyn_stein@berkeley.edu, heidi.lie.williams@dartmouth.edu. First draft February 2022.

1 Introduction

University-based scientific research has long been argued to be a central source of commercial innovation and economic growth. For example, Nelson (1986) and Mansfield (1991) conclude from surveys that a large share of commercial firms' discoveries could not have been developed (without substantial delay) in the absence of academic research. Jaffe (1989) estimates that a 10% increase in university research expenditures is associated with a 6% increase in nearby corporate patents.[1]

Yet at the same time, there have been long-held concerns that many university-based discoveries never realize their potential social benefits. From a policy perspective, this concern surfaced during the debate leading up to the passage of the Bayh-Dole Act of 1980 in the U.S.[2] A common narrative at the time was that most federally funded inventions discovered at universities "languished" in the ivory tower, never diffusing out into the economy, due to a lack of clear title and property rights over those inventions. For example, a candidate drug discovered in a university lab might never be picked up by a pharmaceutical firm for further development due to a lack of clarity over title and patent rights. Bayh-Dole aimed to address that concern by granting universities the right to retain intellectual property derived from federally funded research.[3]

Academic research in this area has also highlighted a separate concern. As argued in the survey work of Jensen and Thursby (2001), most university inventions are "embryonic" when initially disclosed in academic publications and patents and require significant additional development — often including cooperation of the original inventor — before they can be commercially useful. Consistent with this idea, Zucker, Darby, and Brewer (1998) provide evidence that the timing and location of U.S. biotechnology enterprises are closely linked with the geographic location of the university-based scientists who undertook the relevant basic research; the survey work of Jensen and Thursby (2001) is suggestive that similar patterns hold outside of biotechnology as well. It is relatively rare for senior academics to directly take on executive roles in private firms, as most academics prefer to

[1] Of course, especially given that academic discoveries are routinely disclosed through publications, spillovers from academia to industry need not be geographically localized, but the existence of geographically localized spillovers as in Jaffe (1989) is nonetheless indicative of their existence.

[2] See, for example, the discussion in Chapter 5 of Mowery et al. (2004).

[3] The 1984 passage of Public Law 98-620 pushed further in the same direction, removing certain restrictions contained in Bayh-Dole; see Henderson, Jaffe, and Trajtenberg (1998) for a detailed discussion. A recent analysis of Bayh-Dole by Hausman (2022) documents that employment, payroll, and wages grew faster after Bayh-Dole in industries more closely related to the research focus of nearby universities.

keep their labs in operation and universities often express little enthusiasm for joint commitments. Senior academics will sometimes place one of their students or post-doctoral students with a start-up, but there are nonetheless concerns that there may be too few incentives for academics to engage in the diffusion process necessary to have their academic discoveries fully realize their potential private and social returns. Recognition of this potential concern has motivated research investigating the role of faculty incentives as a driver of university commercialization, as in Hvide and Jones (2018), Lach and Schankerman (2008), and Ouellette and Tutt (2020).

In this paper, we focus on the question of whether *where* an academic discovery is made affects whether it diffuses into the economy. Why might place-based factors matter? Bayh-Dole focused on facilitating university patenting and licensing, two natural university-specific efforts that may affect the likelihood that university-based discoveries spill over to commercial innovation. Beyond university patenting and licensing, a broader set of university policies and practices, such as responsiveness to demands from the local community for innovations and the training of students, have been argued to influence whether and when university discoveries spill over into commercial innovation. Peer effects may also matter: Marx and Hsu (2022) finds that having peers who are star commercializers is a strong predictor of whether researchers commercialize their own work.

Looking beyond the universities themselves, a broad range of local area characteristics have also been argued to influence academics' contribution to commercial innovation. For example, Kenney (1986) qualitatively argues that the local availability of venture capital played a significant role in the birth of U.S. biotechnology enterprises. More broadly, works such as Saxenian (1994) and Moretti (2021) have documented compelling evidence on the importance of place-based technology clusters such as Silicon Valley as a determinant of inventor productivity.

Quantitatively estimating the role of place-based characteristics in the commercialization of university-based discoveries is, however, empirically challenging. In a descriptive sense, we know that commercialization activity varies tremendously across universities, in ways that do not simply reflect differences in research activity across universities.[4] For example, in 2016 Duke and Stanford had similar expenditures on research ($910 million

[4] See e.g. Fisch et al. (2015), who tabulate counts of granted patents assigned to different universities, as well as the tabulations produced from Association of University Technology Managers (AUTM) surveys. In addition, past work such as Di Gregorio and Shane (2003) has correlated university-level commercialization activity with university- and area-level characteristics; see also Siegel, Waldman, and Link (2003), Sine, Shane, and Gregorio (2003), Stuart and Ding (2006), and Ho and Lee (2021).

versus $990 million), but Duke generated 9 start-ups while Stanford generated 32.[5] But such cross-university differences in commercialization activity are difficult to interpret. Field and sub-field differences may be central: e.g, medical schools produce translational biomedical research knowledge that may have inherently higher commercial relevance than scientific advances in more basic scientific fields like organic chemistry. Academics doing more applied or commercializable research may sort to universities that are better able to support such work, or those based in clusters such as Silicon Valley where their work may be more likely to be attract the attention of local entrepreneurs and venture capitalists.

We take a first step towards unpacking this heterogeneity in university commercialization by analyzing how the propensity of academic research to spill over to commercial innovation changes when academics move across universities. As an example, consider the case of Professor Carolyn Bertozzi, a renowned chemist who works at the intersection of chemistry, biology, and medicine. She began her academic career in 1996 at UC Berkeley. However, she has publicly discussed how the environment at Berkeley, while productive for basic science research, was not particularly conducive to translating basic scientific discoveries into commercializable technologies. In her words, "I kind of felt [that] Berkeley... was great for basic science and really great for chemistry because the graduate students are top-notch. But it was really hard for me to think about how to translate [ideas] to the clinic" (Jarvis, 2020). In 2015, she moved her lab to Stanford. The change in her lab's commercialization activity was stark: while two companies spun out of her lab during her 18-year tenure at Berkeley, she started four new companies within five years of moving to Stanford. She has attributed her "entrepreneurial fervor to the natural collaborations that bubble up at Stanford in a way they didn't at Berkeley" (Jarvis, 2020).

To construct a sample of academic movers like Professor Bertozzi, we start with the universe of academic discoveries, measured in the form of published scientific papers cataloged in the Web of Science data. For each published scientific paper in the Web of Science data, we extract — wherever available — the e-mail address listed for each author. For each e-mail address, we then extract the domain: for example, extracting `dartmouth` from `heidi.lie.williams@dartmouth.edu`. We stitch together linkages between e-mail domains and university technology transfer offices in order to use observed changes in the e-mail domains authors choose to list on their published scientific papers to infer moves of researchers across universities.

[5] Data on research expenditures and start-ups formed are drawn from the AUTM survey data.

With this panel data on authors' university affiliations and published scientific papers in hand, we then construct measures of these papers' propensity to spill over to commercial innovation. To understand the type of spillovers we are interested in, it is helpful to return to the example of Professor Bertozzi. In 2018, her lab published a paper documenting how a novel small molecule known as DMN-Tre could detect tuberculosis in a rapid, cheap, and highly accurate manner (Kamariza et al., 2018). Bertozzi's then-PhD student, Mireille Kamariza, was the lead author on the paper. Stanford also filed a patent application for the diagnostic, listing both Bertozzi and Kamariza as inventors. In 2019, Kamariza and Bertozzi co-founded a startup called OliLux Biosciences, dedicated to rapid and low-cost tuberculosis detection. Kamariza was the CEO, and Bertozzi the Chair of the Scientific Advisory Board (SAB).

More common than the spillover margins articulated in the example above is for an academic paper to be cited by a patent. Moreover, patent citations of academic papers are also a more comprehensive outcome measure, as they capture anyone (not just the academic authors) to be involved in the commercialization effort. For both reasons, our key outcome variable is the propensity of published scientific papers to spill over to commercial innovation, as measured via patent citations. Our reliance on patent citations to scientific papers as a measure of commercial spillovers builds closely on the recent work of Bryan, Ozcan, and Sampat (2020) and Marx and Fuegi (2020a). As has been documented in earlier work, patent citations to a given scientific paper accrue slowly over time, complicating application of a standard event study framework. To circumvent that challenge, our baseline analysis *predicts* the number of times a scientific paper will be cited in U.S. patents in the five years after its publication based on the journal in which the paper was published: for instance, scientific papers published in *Nature Biotechnology* are around three times more likely to be cited in a patent in the five years after publication than are those published in *Proceedings of the National Academy of Sciences*. As a robustness check, we also examine the counts of patent citations actually received by a given scientific paper as an outcome and obtain similar estimates. Taken at face value, our empirical estimates suggest that at least 15–25% of geographic variation in commercial spillovers from university-based research as measured by patent citations is attributable to place-specific factors.

The Bertozzi Lab / OliLux Biosciences example also suggests two additional outcomes. The first is so-called "patent-paper pairs," where the same idea is disclosed with near simultaneity by an inventor or lab in an academic paper and in a patent application. The second is the participation of scientists in startups formed to commercialize their ideas. Recall that in the Bertozzi Lab example, Kamariza served as the company's CEO and Bertozzi sat on the Scientific Advisory Board (SAB). In order to measure SAB participation, we hand-match records from Capital IQ to our Web of Science data in order to measure that dimension of academics' participation in commercial activities. Unfortunately, our ability to draw conclusions from these outcomes is hampered by a lack of statistical precision, due to the infrequency of these outcomes in our sample.

Methodologically, our empirical approach builds closely on related work using movers to estimate the relative importance of person-specific factors and place-specific factors in other markets (Abowd, Kramarz, and Margolis, 1999; Bertrand and Schoar, 2003; Finkelstein, Gentzkow, and Williams, 2016). A key difference in our setting is the long lags inherent in the process of academic research spilling over to commercial innovation, since most past empirical analyses of movers analyze person-year outcomes that can change at the point of move — such as worker wages in Abowd, Kramarz, and Margolis (1999) or person-level health expenditures as in Finkelstein, Gentzkow, and Williams (2016). These lags motivate our approach of predicting citations using characteristics of the publication, allowing us to sidestep waiting for citation accrual.

Because we define moves as moves across universities – which of course have fixed physical locations — our estimated place effects reflect both university-specific and geographic-specific factors. At a descriptive level, we can correlate our estimated place effects with university-level variables such as the volume of research activity at the school, and with local area-specific factors such as hubs of technical specialization in the commercial sector (Bikard and Marx, 2020; Moretti, 2021). Although we are hindered by statistical power, the evidence suggests that both university-specific and geographic-specific effects matter. Schools with more active licensing agreements have larger fixed effects. But we also see suggestive evidence that more local venture capital activity correlates with more commercialization, even though this venture capital activity occurs "outside" of the university.

While our focus here is on how scientific papers spill over to commercial innovation, our work of course relates to spillovers from scientific research more generally. Such spillovers have been documented to operate in what Azoulay, Zivin, and Wang (2010) refer to as "idea space," i.e., between colleagues working in similar areas, regardless of their location, Supporting evidence can be seen in the analyses of innovative movers by Agrawal, Cockburn, and McHale (2006) and Sharoni (2023). In other cases, knowledge spillovers are very localized, as the strand of research beginning with Jaffe (1989) suggests (e.g., Helmers and Overman (2017) and Andrews, Russell, and Yu (2023)), though this tendency may have fallen over time (Kantor and Whalley, 2019). Azoulay, Graff Zivin, and Sampat (2012) points to a more nuanced story, where article-to-article citations at a moving academic superstar's origin location are barely affected by their departure, but citations to their patents drop sharply.

The paper proceeds as follows. Section 2 describes our data, including how we construct our key outcomes and how we infer academics' moves across universities. Section 3 describes our sample construction and provides some descriptive statistics. Section 4 describes our empirical strategy alongside a presentation of our results. Section 5 concludes.

2 Data and measurement

2.1 Measuring academic research: Web of Science data

Before we can measure commercial spillovers from academic research, we first must measure the academic research itself. We measure academic research as publications catalogued in the Web of Science data from Clarivate Analytics, which compiles records of peer-reviewed scholarly journals, conference proceedings, and editorially selected books.[6] Our version of the Web of Science includes publications from 1900 through the end of 2020, and contains more than 80 million publication records.

For each author in our sample, we measure their publications in each year based on the Web of Science-assigned author ID. Clarivate Analytics generates Web of Science author IDs via a proprietary author name disambiguation algorithm; the only disclosed descriptions of which we are aware[7] clarify that the algorithm takes both ORCID and Publons (two opt-in systems for author disambiguation) as inputs, and that the algorithm analyzes data including author names, institution names, and citing/cited author relationships. The only validation study of which we are aware is Levin et al. (2012), which argued based on a comparison with hand-collected data from 200 authors that the Web of Science author ID erred (at least at the time of the analysis) on the side of over-counting publications: on average 8.8% of authors' Web of Science-linked articles did not appear in the hand-collected

[6] Analyses such as Martín-Martín et al. (2018) and Visser, van Eck, and Waltman (2021) have analyzed the degree of overlap between Web of Science and similar datasets such as Dimensions, Google Scholar, Microsoft Academic, and Scopus. There are some differences across these datasets. For example, Microsoft Academic aims to include documents such as blogs and news articles, whereas the Web of Science does not. Looking at Scopus and Web of Science, Scopus covers book chapters while Web of Science does not; Web of Science covers meeting abstracts and book reviews while Scopus does not. Visser, van Eck, and Waltman (2021) argue that, across data sources, documents included in one source but not another tend to have few to no citations — suggesting that in practice these differences may not be very consequential from the perspective of any given data set missing "important" publications. On the Web of Science data in particular, see also Birkle et al. (2020).

[7] See, e.g., https://clarivate.com/blog/author-data-made-better-together/.

publication lists. The accuracy of the Web of Science author ID has presumably improved since 2012, as evidenced by the fact that the quality of other author disambiguation algorithms are generally assessed against the (presumed-to-be-correct) Web of Science data. For example, Lerchenmueller and Sorenson (2016) notes that the developers of the Authority database assessed the quality of their author disambiguation through comparisons to Web of Science author IDs.

2.2 Measuring commercial linkages to academic research

Commercial spillovers from university-based research can take many forms, not all of which are possible to measure. Consider as an example the work of MIT professor Robert Langer.[8] Langer has authored over 1,500 scientific papers, and holds over 1,400 granted or pending patents. Langer's patents have been licensed or sublicensed to over 400 pharmaceutical, chemical, biotechnology and medical device companies. Unfortunately, such licensing agreements are (with rare exceptions) not publicly observed. Over 40 companies have been spun out of the Langer lab — including the mRNA therapeutics firm Moderna, which Langer co-founded. For one faculty member to be linked to so many start-ups is exceptional, and in practice looking at direct faculty involvement in start-ups as an outcome is too rare to be informative. However, as discussed below, we construct a closely related measure — faculty involvement on Scientific Advisory Boards — which arguably captures faculty involvement in start-ups at a broader level. In Langer's case, we observe him sitting on 45 Scientific Advisory Boards in our data. Our main analysis focuses on patent citations to academic papers as our primary measure of commercial spillover. This measure is by far the most commonly observed commercialization-related measure in our sample, and is also in some sense the most general, as it does not restrict to commercialization done solely by the academics. In addition, we also compute patent-paper pairs and membership on Scientific Advisory Boards (SABs) as additional measures. In the rest of this section, we detail how we construct these variables.

Actual and predicted patent citations to academic papers. Building on recent work by Bryan, Ozcan, and Sampat (2020) and Marx and Fuegi (2020a), our main outcome variables measure, in various ways described in more detail below, the propensity of authors'

[8] Most figures in this paragraph are drawn from Professor Langer's MIT website: https://langerlab.mit.edu/langer-bio/; the exceptions are the company spin-out figure that is drawn from Langer's citation for the 2019 Dreyfus Prize in the Chemical Sciences (https://www.dreyfus.org/robert-langer-2019/) and the Moderna example that is drawn from the Moderna website (https://www.modernatx.com/modernas-board-directors).

published scientific research to be cited in patents. In other words, for every academic publication, we start by counting the number of times (if any) the publication is cited by a patent. Patents may cite scientific research papers in two ways: on the front page of the application ("front-page citations") and in the body of the text ("in-text citations"). Bryan, Ozcan, and Sampat (2020) articulate why, from a legal perspective, front-page and in-text citations play distinct roles in a patent. Front-page patent citations are disclosed as "prior art" relevant to the patentability of inventions. Patent applicants submit front-page citations in applicant information disclosure statements (USPTO form 1449), and front-page citations are frequently added by the patent examiner over the course of the examiner's review of the application. In contrast, in-text citations appear in the specification of the patent, which is intended to teach someone "skilled in the art" how to make and use the invention. Consistent with the argument that front-page and in-text citations play distinct roles in a patent, Bryan, Ozcan, and Sampat (2020) document that these two types of citations have relatively little overlap: on average for a given patent, only 31% of in-text citations appear as front-page citations, and only 24% of front-page citations appear as in-text citations.

Dating back at least to the work of Narin (1994), Trajtenberg (1990), and Jaffe, Trajtenberg, and Henderson (1993), front-page patent citations have been the focus of academic research analyzing patent citation data. However, that choice largely reflected the fact that front-page patent citations are — from a practical perspective — much easier to extract than are in-text citations.[9] Our read of the nascent literature on this topic is that there is no clear theoretical reason to prefer front-page to in-text citations or vice versa in our application, so our baseline analysis analyzes the sum of front-page and in-text citations.

Until recently, there was no data systematically cataloguing in-text patent citations. This changed with the release of the Reliance on Science data (Marx and Fuegi, 2020a,b). The current version of the Reliance on Science data captures both front-page and in-text citations. Marx and Fuegi's data includes Microsoft Academic Graph (MAG) paper identifiers, but unfortunately to the best of our knowledge there does not exist a crosswalk between MAG paper identifiers and Web of Science paper identifiers. In order to link the publications in our Web of Science data to the Reliance on Science data, we construct a crosswalk as follows. Where available, we merge based on Distinct Object Identifiers (DOIs) and PubMed identifiers; together, these merges capture 69.4% of the papers in the

[9] Indeed, Bryan, Ozcan, and Sampat (2020) note that one of the first papers to empirically examine front-page citations — Narin and Noma (1985) — argued that in-text references "may be more related to the history, usefulness, and development of the invention" but that they instead analyze front-page citations since they are "far easier to extract."

Marx and Fuegi data.[10] Among the remaining unmatched papers in the Marx and Fuegi data, we attempt a merge based on ISSN number (a unique journal identifier), volume, issue, and first page number; this merge captures an additional 1.7% of papers. Hand-checks suggest there was not a clear way to match the remaining 28.9% of Marx-Fuegi papers to our Web of Science data, so we omit those papers from our analysis.[11]

One challenge that arises in constructing and using such a measure is that there are substantial time lags between when papers are published and when they accrue citations. Appendix Figure 3(a) documents the distribution of years from when a paper is published to when it receives its first citation, conditional on receiving at least one citation. The average lag is 3.7 years, and the median lag is 3 years.[12] For any X, counting the number of patent citations accrued over the first X years of a paper's life is more informative for a larger X — as longer time windows capture more across-paper variation — but because we must cut our sample off at year $2020-X$ (2020 is the last year for which we have patent citation data), choosing a larger X reduces our sample size. We resolve this trade-off by focusing on $X=5$.[13]

However, this observed citation measure poses a challenge in our mover design. To see why, consider an author who moves in 2010 but writes a paper in 2008. The paper was clearly written prior to the move. However, our five-year citation move will count citations that accrue in the years 2008, 2009, 2010, 2011, and 2012 — in other words, two years of citations accumulated in post-move years. Thus, for our five-year outcome, the distinction between pre and post becomes blurry.

To overcome this challenge, our baseline outcome *predicts* the number of times a scientific paper will be cited in U.S. patents in the five years after its publication based on the journal in which the paper was published. Intuitively, this measure captures the idea that scientific papers published in *Nature Biotechnology* are around three times more likely to

[10] In practice, PubMed identifiers do much less of the work here than do DOIs: of these matches, 59% match on both DOI and PubMed ID, 37% match on DOI with no PubMed ID, and only 4% match on PubMed ID alone.

[11] As discussed above, the intended coverage of the Web of Science data and the MAG data are not the same: for example, Visser, van Eck, and Waltman (2021) note that MAG covers documents not of a scientific nature (for example, news articles). In our data, the majority (52%) of unmatched MAG observations lack a journal ID, suggesting they are likely not academic journal articles.

[12] There are few differences across institutions of various research intensities (averaged over the entire sample, with institutions in the first quartile representing those with the lowest amount of research spending) in the timing of citations. Appendix Figure 3(b) shows the cumulative number of citations a paper receives over time (again, restricting to papers with at least one citation).

[13] We include the paper's publication year as one of the years a paper can accrue citations. So, for example, a paper published in 2000 will be allowed to accrue citations from 2000 to 2004.

be cited in a patent in the five years after publication than are scientific papers published in *Proceedings of the National Academy of Sciences* (see Appendix Figure 4 for more examples). Because both observed and predicted citations are highly skewed (and because the majority of papers receive zero patent citations), we apply the inverse hyperbolic sine transformation to each of these outcomes. Finally, we sum over all papers written in the same year, so that authors who write more papers in a given year will have more predicted patent citations, all else equal. This leaves us with a commercialization measure that is unique at the author-year level. The key benefit of this measure is that it is "instantaneous" — the outcome is computed using data from the year of publication. However, it is worth noting that this outcome also focuses our attention on a particular channel. Predicted patent cites will only increase if the author changes the content of their research and/or targets different journals. Notably, predicted patent citations *will not* increase if the author makes no changes to their research, even if the school increases efforts to publicize the work in such a way that increases patent citations. We will keep this caveat in mind when interpreting our results. We also present results where we use *actual* five-year citations. In these cases, we bear in mind that relative years -4 to 0 are difficult to interpret, because it is not clear whether they belong in the pre- or post-period.

Venture capital-backed patents. Patent citations originating from venture capital-backed firms are of particular interest, because these firms are thought to be particularly innovative (Kortum and Lerner, 2000). Records on all U.S. firms that raised venture capital funds from 1968-2020 are obtained via the Refinitiv Database. Following the methodology outlined by Bernstein, Giroud, and Townsend (2016) and Akcigit et al. (2024), we link USPTO patents to firms using the firm name, patent assignee name, and geocoded city information. Oftentimes there are exact one-to-one matches, but additional "fuzzy matches" are made by allowing some flexibility in firm name spelling or geography (e.g., non-exact candidate matches with the same city or state were ultimately considered a true match).

Given this linkage, we then want to restrict attention to the patents that should reasonably be coded as venture-backed. As examples of what we would want to exclude, many patents in our sample are awarded to firms like Microsoft and Intel, which had received venture financing decades before the grant date of the patents in our sample. One approach is to code patents as venture-backed if the application year falls between the first and last years of venture financing. However, this approach has two potentially undesirable features. First, it artificially deflates the number of measured venture-backed patents in time periods when firms went public very quickly, and artificially inflates the number of venture-backed patents during time periods where firms remained private for much longer periods. Second, a number of firms in the sample received their last round of financing many years after their initial round. In most cases, these instances are buyouts or other "take private" deals, private investments in public entities (PIPEs), or venture financings of companies that had been taken private after an initial public offering – none of which correspond to the traditional definition of venture activity. While it was feasible to purge a number of these financings, it is not possible to do so in all cases. Reflecting these concerns, our chosen approach is to define venture-backed patents as those that were applied for within five years of the firm's first round of venture financing. We use the actual count of five-year patent citations originating from these firms as an additional measure of commercial spillover.

Patent-paper pairs. Building on work by Ducor (2000), Murray (2002), and Murray and Stern (2007), we examine the propensity of authors' published research to be part of a "patent-paper" pair. This has become an increasingly popular metric for capturing the commercialization of academic science (see, for instance, Ahmadpoor and Jones (2017) and Fleming et al. (2019)). The idea behind these patent-paper pairs is that a given discovery can sometimes be disclosed by the same research team in both a scientific paper and a patent, but doing so requires careful orchestration of the relative timing of the two disclosures. Patenting is generally concomitant with publication, as there is only a one-year grace period during which patentability is not invalidated by prior publication (35 U.S.C. §102(b)(1)). One advantage of this outcome is a clean prediction of when they will occur: patent-paper pairs occur closely clustered in time, if they occur at all. We identify patent paper pairs through 2020, using the July 2022 release of the Reliance on Science data. We use the confidence score measure assigned by Marx and Fuegi, which estimates the quality of the match, to weight the observations in the analyses.

Scientific Advisory Board memberships. As a third measure of commercial spillover, we collect data on memberships on Scientific Advisory Boards (SABs) from Capital IQ. Capital IQ, a unit of S&P Global, covers both public and private firms; our version of the data includes 3,807 companies, many of which are backed by venture capitalists.[14]

[14] The Capital IQ data are sourced from Capital IQ's own researchers, S&P's Compustat, and a variety of additional data providers, including Preqin, Crunchbase, CreditSafe, and Dun & Bradstreet.

The database lists current and former SAB members and their biographies. We use these biographies to extract board members' current and past affiliations, based on a combination of text parsing and hand-coding. We then use the combination of names and affiliation(s) to match SAB members back to the Web of Science data. UpWorkers helped to eliminate cases where practitioners held adjunct positions at universities, served on university boards, or held other academic connections that do not constitute traditional faculty roles.

Through this process, we identify 20,969 board member-company pairs and 15,824 unique board members. Around 1,300 individuals serve on SABs for more than one company; 15 individuals serve for more than 10 companies. Faculty members often serve on the SABs of companies that seek to commercialize their knowledge, even in cases where they also are founders and even serve on traditional boards. Linking back to our example of Professor Bertozzi, she serves on five SABs in our data. The involvement of faculty on SABs has been much less scrutinized than patent-based metrics: one exception is Stuart and Ding (2006), who analyze data on 727 scientific advisors.

2.3 Measuring university affiliations and cross-university moves

In order to identify movers — defined as authors of research papers whose university affiliations change over time — we first need to determine each author's organizational affiliation in each year. In the Web of Science data, affiliations can be measured in a variety of ways: for example, each author of each publication in the Web of Science data is linked to one or more organization names; separately, the "corresponding author" of each paper in the Web of Science data has a listed address. However, in practice we found that measuring moves based on either of these variables has significant drawbacks. For example, Web of Science routinely lists multiple organization names for a given author on a given paper and does not offer any indication of which is the primary affiliation, leading to a "multiple affiliations problem." Given these challenges, our baseline approach — described below and illustrated by an example in Appendix Figure 1 — is to instead infer moves based on observed changes in the e-mail addresses that authors choose to list on their publications. We view the email address — and in particular, the email address domain — as a revealed preference measure of the author's true home institution. Appendix Figure 1(a) illustrates an example where several authors have multiple affiliations. In this case, the final author, Sangeeta Bhatia, is the designated corresponding author, and is the only author whose email address is reported. She has multiple affiliations (MIT, Broad Institute, Brigham and Women's Hospital, and Howard Hughes Medical Institute), but a single email address with an @mit.edu domain. Appendix Figure 1(b) shows how this publication is listed in the Web of Science data, and Appendix Figure 1(c) provides a screenshot of the associated SQL tables we extract this data from.

Defining universities. Given the central role of technology transfer offices in facilitating spillovers of academic research to industrial innovation, for the purpose of our analyses we define "universities" as technology transfer offices (TTOs). In particular, we focus attention on the approximately 300 U.S.-based technology transfer offices that participate in the Association of University Technology Managers (AUTM) surveys.[15] In practice, this definition includes primarily university-based TTOs (89%, e.g., Stanford University; AUTM ID 95165), but also some university system-wide TTOs (2%, e.g., University of California System; AUTM ID 95188), some affiliated research hospitals reporting separately from universities (3%, e.g., Brigham and Women's Hospital; AUTM ID 95017, which reports separately from Harvard University; AUTM ID 95070), and some standalone research hospitals (6%, e.g., St. Jude Children's Research Hospital; AUTM ID 95164). For some of our descriptive analyses, we geocode TTO locations to Bureau of Economic Analysis (BEA) economic area codes used in Moretti, 2021. Because there is no off-the-shelf mapping, this required us to construct two additional linkages: i) from TTO locations to ZIP codes, using data from Association of American Medical Colleges; and ii) from ZIP codes to county FIPS codes (Din and Wilson, 2018).

Extracting data on e-mail domains. For each author of each publication in the Web of Science data, we extract the listed e-mail address where available. Because no more than one e-mail address is listed for each author, this provides a solution to the multiple affiliations problem described above. For each e-mail address, we extract the domain — for example, extracting `dartmouth` from `heidi.lie.williams@dartmouth.edu`. Given our focus on U.S. universities, we restrict attention to e-mail addresses which contain an `@` symbol and end in `.edu`, and extract as the domain all text between the final `@` symbol and `.edu`.[16] Note that e-mail addresses are reported much more frequently for corresponding

15 We separately draw some TTO-level variables from the AUTM surveys between 1991 and 2019, such as measures of research expenditures, licenses, and licensing income. Across cross-sectional surveys, TTOs are identified by unique ID numbers, allowing us to follow individual TTOs over time even if the listed organization name changes.

16 We standardize and lightly clean the e-mail addresses prior to extracting this text, by converting all extracted e-mail addresses to lower case, filtering to only `a-z`, `0-9`, `.`, and `@` characters, and fixing commonly observed duplication errors (e.g. replacing `..` with `.`, replacing `.edu.edu` with `.edu`).

authors than non-corresponding authors (roughly 85% of the time versus 15% of the time in the years at the end of our sample), and are reported more often for first and last (and solo) authors than for middle authors.

Linking e-mail domains to organizations. We next need to link the domains extracted from e-mail addresses to universities, defined as AUTM TTOs. While conceptually straightforward, what we need is a systematic way of knowing that, for example, e-mail domains `harvard.edu` and `hbs.edu` both map to the same technology transfer office at Harvard University. We do so by constructing a crosswalk from e-mail domains to so-called "preferred organizations" in the Web of Science data to AUTM TTOs as follows.

As noted above, each author of each publication in the Web of Science data is linked to one or more organization names. Clarivate Analytics makes an attempt to standardize these organization names by creating unique organization identifiers referred to as "preferred organizations." For example, the preferred organization for Harvard University aggregates institutions such as the Harvard Stem Cell Institute, the Harvard School of Engineering and Applied Sciences, and the Harvard Society of Fellows.[17] We start by expanding the Web of Science data set to a table that is unique at the paper-, author-, preferred organization-level (that is, a given author-paper pair is listed more than once if more than one preferred organization is listed for that author-paper). In these data, we then restrict attention to observations reporting both an e-mail address and a preferred organization name. We drop observations linked to preferred organizations of the form "University of .* System" because — empirically — all Web of Science records with an affiliation of that form have at least one other, more granular organization; e.g., we drop "University of California System" in favor of keeping "University of California, Berkeley."

For each unique e-mail domain, we then select the Web of Science preferred organization that is most frequently associated with that domain. The only exceptions are that we manually correct some e-mail sub-domains for academic medical centers that have independent TTOs but our methodology would otherwise (incorrectly) pool with their affiliated university. For example, Brigham and Women's Hospital has the email domain `bws.harvard.edu`, which our method would default to pool with Harvard University, but

[17] In practice, Clarivate (understandably) appears to have only undertaken this standardization effort for relatively common organizations, so relying on these preferred organizations loses some information for rarely reported organization names. Clarivate has assigned just under 15,000 unique preferred organization identifiers, of which 2,040 are based in the United States. As one point of comparison, in 2019-2020 there were around 2,700 four-year degree-granting post-secondary institutions in the United States (National Center for Education Statistics, 2020).

which should instead be kept separate because the hospital has its own independent TTO in the AUTM survey data.[18]

Finally, we link Web of Science preferred organizations to AUTM TTOs by hand. Taken together, this method allows us to link e-mail domains — our measure of author affiliation — to AUTM TTOs, our measure of "universities."[19]

Constructing author-year panel data on university affiliations. For each author-article pair reporting an e-mail address in the Web of Science data, we can thus assign a unique AUTM organization. If within a given year, an author with multiple publications is linked to more than one AUTM organization, we assign him or her to the AUTM organization with which they are most frequently affiliated. In cases of ties (which occur for about 0.7% of author-years), we select an affiliation at random. This process generates a single affiliation for each author-year.

Measuring moves. With unique author-year affiliations in hand, we define movers as follows. We only identify movers that we can see are at distinct universities in two consecutive years. That is, we do not use cases where we need to impute the year of the move because there is a gap in publications, as imputing affiliation in years an author does not publish risks mis-coding the timing of their move.[20] As detailed in Figure Appendix Figure 2, we use mailing address data provided in the Web of Science data to attempt to fill in affili-

[18] The list of institutions for which we do this type of manual correction is Beth Israel Deaconess Medical Center, Boston Children's Hospital, Brigham and Women's Hospital, Children's Hospital Los Angeles, Dana-Farber Cancer Institute, Massachusetts General Hospital, Moffitt Cancer Center, Schepens Eye Research Institute, and the Whitehead Institute.

[19] E-mail domains that are not affiliated with an AUTM TTO are coded as missing and not included in our analysis. For each observation in the AUTM data that we fail to hand-match with a Web of Science preferred organization (30 of 309), we instead attempt to hand-match to raw (non-preferred) organization names in the Web of Science data; this generated an additional seven matched AUTM TTOs. Our remaining unmatched AUTM organizations are Albert Einstein Healthcare (AUTM ID 95003), Center for Innovative Technology (AUTM ID 95031), Competitive Technologies (AUTM ID 95046), Florida Institution of Technology (AUTM ID 95061), Hahnemann University (AUTM ID 95069), Hospital for Special Surgery (AUTM ID 95073), Houston Advanced Research Center (AUTM ID 95074), Mayo Foundation (AUTM ID 95098), Medical College of Pennsylvania (AUTM ID 95102), Medical University of South Carolina (AUTM ID 95103), Ramsey Foundation (AUTM ID 95141), Research Corporation Technologies (AUTM ID 95144), Children's Hospital Cincinnati (AUTM ID 96008), Louisiana State University Agricultural Center (AUTM ID 96016), University of Maryland Biotech (AUTM ID 96028), Florida Agricultural (AUTM ID 97003), Paper Science and Tech (AUTM ID 97023), Legacy Health System (AUTM ID 20000014), Center for Blood Research (AUTM ID 20000252), Tennessee Board of Regents (AUTM ID 20000254), NUtech Ventures (AUTM ID 20100289), WiSys Technology Foundation (AUTM ID 20120294), and Providence Health & Services (AUTM ID 20140302).

[20] For example, suppose we observe Josh Lerner at school A in 2012 and at school B in 2016, with no publication record in between. If we imputed moves based on the first publication where he was identified at a new school, we would code Josh to be at school A from 2013-2015, even if he moved in 2013.

Restriction	Observation Count
Contributor-publications	300,691,280
Contributor-publications where role=="author"[a]	280,429,056
Author-publications with DAIS IDs	280,323,438
Author-publications with Emails and DAIS IDs	47,739,862
Author-publications with Emails and numeric DAIS IDs	47,739,777
Author-publications with numeric DAIS IDs and valid domains[b]	7,417,802
Author-year-organization affiliations[c]	4,204,781
Author-years after randomly breaking ties[d]	4,138,812
Author-years after including only biomedical[e] authors	1,720,112
Author-years after including only non-movers and movers ± 10 years around move	1,709,102

[a] Other possible roles are "book," "book editor," "book_corp," "anon," and "corp"

[b] Note the vast majority ($> 99\%$) of observations discarded here are non-.edu emails

[c] Each organization here is either an AUTM TTO assigned via email or "missing"

[d] 60,872 author-years have a conflict. 59,624 are two-way ties, 1,212 are three-way ties, and 36 are four-way ties. This yields 30,225 author-years with at least two conflicting affiliations

[e] Defined as authors with 50% or more of their publications listing a PubMed ID

Table 1: Author-year sample construction

Notes: This table lists the sample restrictions we apply to the Web of Science data to code movers and non-movers. This process pertains to authors identified via our e-mail panel based approach, described in Section 2.3. The remainder of our sample construction process is illustrated in Appendix Figure 2 and detailed in Appendix Section A.

ation gaps where possible in order to add some additional movers to our sample. On its own, our e-mail-based panel identifies a total of 12,211 movers, and we are able to increase that sample by around 20% using geographic information and other methods described in Appendix Section A.

Once we code an individual as a mover, we check if — in the year prior to the move-year — there are a non-zero number of papers that list the destination as their affiliation. If yes, we code the move-year to be one year earlier. Our motivation here is two-fold: publication lags distort our ability to code move-years, and a researcher's change in affiliation might not occur cleanly at year's end. As shown in Appendix Figure 5, around 25% of movers list more than one affiliation in the year they move. Appendix Figure 6 illustrates how implementing this one-year correction improves the accuracy of move timing for a randomly selected hand-coded sample of 100 movers. Prior to the correction, we are more likely to code an author's move one year too late than we are to code it correctly. After implementing the correction, however, our accuracy improves.

3 Descriptive statistics

3.1 Sample construction

Table 1 walks through our sample construction. Ultimately, our data will be at the author-year level. However, we start with the Web of Science data at the contributor-publication level (contributor includes authors, but may also include book editors, etc.). We drop contributors who are not coded as authors (but rather as editors, etc.) and make a series of restrictions aimed at ensuring that the remaining author-publications have valid Web of Science author IDs and valid email domains.[21] We then aggregate these data to the author-year level, taking the most common institution and randomly breaking ties where necessary.

Finally, we restrict attention to "biomedical authors": i.e., authors for whom 50% or more of their publications are linked to a PubMed ID. Because biomedical research comprises such a large share of the commercial activity in our data — indeed, some of our outcomes such as patent-paper pairs and Scientific Advisory Board participation are largely relevant only to that sample — this sample restriction focuses our analysis on decomposing meaningful variation in a well-defined activity (biomedical research) across universities.

3.2 Descriptive statistics: Movers and their moves

Appendix Figure 7 provides a verification that our researcher move coding is meaningful, documenting the the share of publications among movers where the researcher's primary affiliation is the destination university, by year relative to the move.[22] Prior to the move, almost none of the publications come from the destination university, whereas post-move over 95% of publications come from the destination university. This share remains constant throughout the post period, suggesting we capture real, long-term moves by researchers.

Figure 1 provides an alternative proof of concept graph, documenting how the geography of patent citations to scientific papers changes before and after move. To construct this graph, we assign geographic locations to patents that cite papers in our sample by taking the address reported for the inventor on each patent and mapping that address to an economic area defined by the BEA. We then count the number of patent citations to a mover's papers originating from the mover's origin or destination economic area in each year surrounding their move in the five years after the article is published. Prior to a move, we observe more

[21] We lose the majority of author-publications here due to the email domain requirement; recall that Web of Science goes back as far as 1900, and email addresses were not commonly reported until the 2000s. The increasing count and share of emails with an associated email address by year is shown in Appendix Figure 8.

[22] We here restrict attention to publications where the researcher's primary affiliation is either the destination or the origin institution.

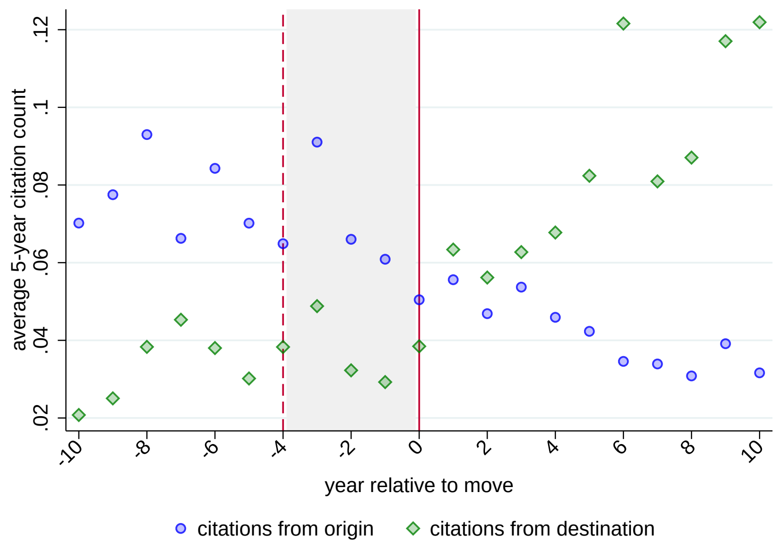


Figure 1: Citations from origin and destination by relative year

Notes: This figure shows the average count of five-year patent citations that a mover's papers accrue by relative year, plotted separately for citations from their geocoded origin (blue circles) vs. destination (green diamonds). When counting five-year citations, there is a five-year period where papers written at a mover's origin can accrue citations at their destination. This window, from relative year -4 to relative year 0, is shaded gray to emphasize the possibility for "contamination". A dashed red line marks the beginning of this potentially contaminated window, and a sold red line marks its end. This figure uses the publication record of our full sample of researcher movers (N = 14,213), which amounts to a total of 757,066 author-papers published ten years pre- or post-move.

citations from patents in movers' origin economic areas. In the five years immediately before the move, when the five-year citation counts include citations made both before and after the move, this remains true, but then this trend reverses in the year after the move, when we start to observe more patent citations from the mover's destination.

We also hand check 100 randomly sampled movers using faculty web pages, CVs, and LinkedIn profiles. Of these 100 movers, we are able to locate information about 75, while the remaining 25 are untraceable. Of the 75 we located, 68 appear to move from and to the universities that we had identified, while 7 appear not to actually move (a 91% success rate among the authors we could locate). Focusing on the 68 correctly identified movers, we code the move year exactly right nearly 40% of the time. We code the move as happening a year too late just over 50% of the time (which is not surprising given that we use publications — a measure that is likely to lag — to identify movers). The remaining 10% of authors have larger errors. We discuss how this measurement error affects the interpretation of our results in Section 4.

Table Appendix Table 1 presents author-level descriptive statistics, separately, for non-movers (column 1) and movers (column 2). Our sample includes around 500,000 authors, roughly 3% of whom are coded as movers. Mechanically, movers have longer careers in our sample (8.6 versus 2.4 years), as we are more likely to code someone as a mover if we can observe him or her for more years. Movers similarly have more publications (61 versus 13) and more patent citations to papers (43 versus 8). Notably, the higher number of patent citations appears to be driven by the higher number of publications, rather than by a higher share of papers being cited (0.05 vs. 0.06) or more citations per paper (0.45 vs. 0.39).

Appendix Figure 9(a) provides additional summary statistics about the moves themselves.[23] The average move distance is 878 miles, equivalent to roughly one-third the width of the continental United States. However, the distribution of move distances is right-skewed, reflecting the density of universities on the coasts — the median move is 641 miles, and the standard deviation of the average move is 801 miles. Over the sample period, the median university received 20 movers.[24]

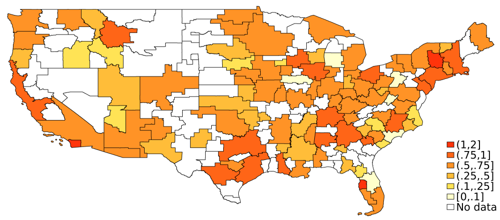
3.3 Descriptive statistics: Commercialization

How does commercialization vary across universities in our sample? Figure 2 illustrates our cross-sectional variation in our key outcome variable, the IHS of predicted patent citations. We collapse this author-year measure down to the institution level to document how commercialization varies across universities. Our collapsing procedure averages across author-years within a given university, with each author-year receiving equal weight. High values correspond to universities that receive more predicted patent citations per author-year, meaning that larger universities are not mechanically at an "advantage." Figure 2(a) documents this variation geographically across economic areas as defined by the BEA as of 2012. While this chart displays the familiar concentration in California and the Northeast, there are many other clusters, such as Texas and the St. Louis area. Figure 2(b) plots these institution-level measures as a histogram.[25] Figure 2(c) lists the top ten institutions by com-

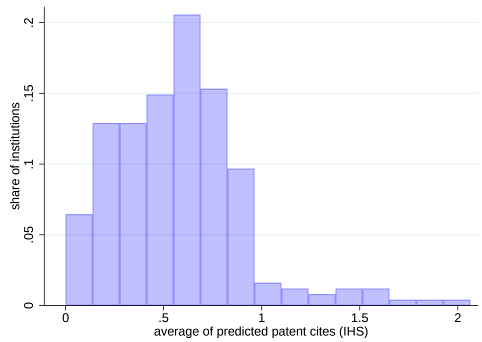
[23] Per the recommendation of Miller (2023), Appendix Figure 10 tabulates additional information about the nature of the movers in our sample. Panel (a) presents the number of researcher moves by the year in which they moved, highlighting the gradual increase over time. This reflects both the increase in the number of researchers, as well as the greater number of disclosed email addresses in articles, which allows us to discern moves. Panel (b) illustrates the number of movers who publish in each publication year— distinct from the number of moves that occur in each year– data our model uses to identify publication year fixed effects. Panel (c) shows the number of movers publishing articles in the years relative to the moves. The number of observations trails off in years far from the move year due both to truncation (we cannot observe publications 10 years after a 2015 move) and the exit of researchers from academe.

[24] On average, universities received 54 movers, though this statistic is inflated by our definition of university systems (e.g. State University of New York (SUNY) or the University of California System).

[25] We document in Appendix Figure 11(a) that if we divide universities into quintiles (in a base year) by commercialization rates, the mean commercialization rate (predicted patent citations) is roughly constant in each quintile. When we look at



(a) Map of commercialization propensities



(b) Commercialization propensities across institutions

rank	*Institution*
1	Whitehead Institute
2	City of Hope
3	Dana-Farber
4	Lee Moffitt
5	Sloan Kettering
6	Sanford Burnham
7	Cold Spring Harbor
8	Scripps Research
9	Rockefeller University
10	Salk Institute

(c) Top institutions by commercialization propensity

rank	*University*
9	Rockefeller University
13	Rensselaer Polytechnic
14	Massachusetts Institute of Technology
18	University of Texas Southwestern
19	Rice University
20	Case Western
21	Stanford University
22	Princeton University
23	Harvard University
24	Georgia State University

(d) Top universities by commercialization propensity

Figure 2: Variation in research commercialization

Notes: In all four panels, though aggregated at different levels, the outcome variable is the average of the inverse hyperbolic sine (IHS) of predicted five-year patent citations. Panel (a) shows geographic variation in average predicted patent citations across "economic areas," as defined by the U.S. Bureau of Economic Analysis. Areas with darker shading are those with higher rates of commercialization. Panel (b) is a histogram of the same outcome, but aggregated to the institution level instead of economic area. There are a total of 265 institutions included in this figure. Panel (c) lists the ten institutions with the highest average predicted citation rates and panel (d) lists the top ten *universities*, a subset of institutions, with the highest average predicted citation rates. Importantly, the commercialization rank computed for panels (c) and (d) are equivalent, and based on the institutional-level average– e.g., Rockefeller University is the ninth-ranked institution and the first-ranked university.

mercialization propensity, overall and separately for universities. Perhaps unsurprisingly, medically focused institutes (that have their own TTOs) tend to outperform their university counterparts on this measure. This is likely because the vast majority of their authors are focused on commercially relevant research, whereas universities have faculty in a broad array of fields that may be closer to the scientific frontier and less readily applicable.

3.4 Descriptive statistics: On-move changes in commercialization

We can now present a descriptive version of our key results in graphical form, to convey the basic patterns. Figure 3 plots the change in mover i's commercialization activity before and after the move against the destination-origin difference in mean commercialization propensity of the universities. Again, we use our IHS predicted patent cite measure to capture commercialization. We define the pre-move period as relative years -10 to -1. We define the post-move period as relative years 1 to 10. The "change on move" term is the difference between those averages for each mover. This focus slightly pares the number of author-years displayed in Appendix Figure 2, for instance, since the process of coding authors as movers/non-movers must precede a restriction on the number of pre- and post-move years we allow in our sample.

The positive correlation suggests that researchers who move to a higher-commercialization location commercialize more themselves, while those who move to a lower-commercialization location commercialize less themselves. If all the variation were due to individuals, we would expect this line to have a slope of zero. If all the variation were due to place, we would expect it to have a slope of one. Instead, we find a slope of 0.16, suggesting that roughly one-sixth of the heterogeneity in commercial activity by faculty is due to the place-specific factors.

4 Empirical strategy and results

4.1 Empirical framework

4.1.1 Additive decomposition

Our empirical approach closely follows that of Finkelstein, Gentzkow, and Williams (2016). Researcher i at university u in year t produces commercialized research (y) according to the following equation:

$$y_{iut} = \alpha_i + \gamma_u + \tau_t + \varepsilon_{iut} \tag{1}$$

a representative school from each quintile in Appendix Figure 11(b), we see that the ranking of schools stays relatively stable over time. This suggests that commercialization patterns across different universities have been roughly stable over time.

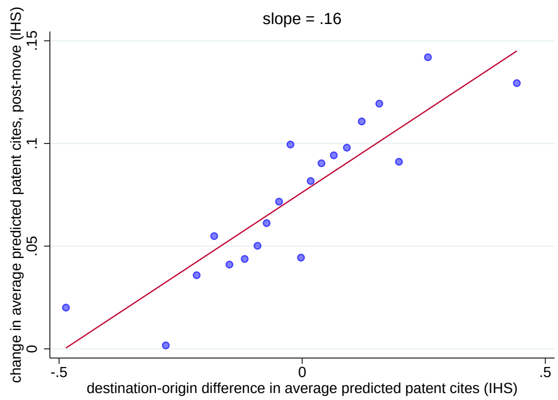


Figure 3: Change in commercialization propensity by size of move

Notes: This figure is a binned scatterplot that compares an individual's change upon moving in predicted five-year patent citations with the difference in average institution-level patent citations between their destination and origin institutions. For each mover, we compute two values. First, we separately calculate the average count of IHS predicted five-year citations for papers published pre- and post-move, and report the difference. We call this the change in an individual's commercialization level. Second, we generate $\hat{\delta}_i$ by taking the institution-level differences, as shown in Figure 2. We call this the change in the institution's commercialization level. The x-axis displays ventiles of this institution-level change, while the y-axis plots, for each ventile, the average change in individual commercialization. The slope of the line of best fit is reported above the graph. There are 119,338 author years in our sample, coming from the sample of 14,213 movers.

where y is some measure of commercialized research, α is an individual fixed effect, γ is a university fixed effect, and τ is a calendar year fixed effect.

In order to decompose variation in commercialization outcomes y into variation driven by the researcher versus variation driven by the researcher's environment (i.e., the university and/or geographic area), let $\bar{y}_{ut}$ be the average of y_{iut}'s within university u in year t, and let $\bar{y}_u$ be the average of $\bar{y}_{ut}$ across time. Similarly, let $\bar{\alpha}_{ut}$ be the average of the α_i's within university u in year t, and let $\bar{\alpha}_u$ be the average of $\bar{\alpha}_{ut}$ across time. If we consider two universities u and u', we have the following decomposition:

$$\bar{y}_u - \bar{y}_{u'} = (\gamma_u - \gamma_{u'}) + (\bar{\alpha}_u - \bar{\alpha}_{u'}). \tag{2}$$

Therefore, the share of the difference between u and u' attributable to the university is:

$$S_{university}(u,u') = \frac{\gamma_u - \gamma_{u'}}{\bar{y}_u - \bar{y}_{u'}} \tag{3}$$

while the share attributable to differing researchers is:

$$S_{researcher}(u,u') = \frac{\bar{\alpha}_u - \bar{\alpha}_{u'}}{\bar{y}_u - \bar{y}_{u'}}. \tag{4}$$

If we can construct sample analogs $\hat{y}_u$ of $\bar{y}_u$ and consistent estimates of $\hat{\gamma}_u$ of γ_u, then we are able to estimate $S_{university}(u,u')$. One minus this amount gives us the analogous $S_{researcher}(u,u')$.

Note that, of course, estimation of Equation (1) is only identified if the data include movers. The intuition here is identical to the classic AKM mover design (Abowd, Kramarz, and Margolis, 1999): if all researchers were non-movers, there would be no way to separate university differences in y due to differing researcher composition versus from fixed characteristics of the university.

Three assumptions are critical in our design. First, we assume that movers do not experience shocks to their commercialization y that are correlated with the timing and direction of the move. This will be the case if the individual fixed effects are time-invariant. This is important, because if movers from low commercialization institutions to high commercialization institutions systematically experience an increase in α_i at the same time as their move, we will overestimate the institution effect. Analyzing pre-trends in our event study analysis can help rule this out. If we believe that changes in an individual's commercialization propensity occur gradually over time, the absence of pre-trends can provide comfort that this assumption holds.

Second, we assume that α_i and γ_u are additively separable. Since our preferred outcomes are transformed with inverse hyperbolic sine transformation (which can be interpreted similarly to logs), this means we are assuming that individual and university effects impact the level of commercialization multiplicatively. Note that this assumption also rules out any type of match effects between researchers and universities. This implies that a chemist and biologist experience the same university effect from school u (though recall that we have limited our sample to biomedical researchers, which helps make this assumption more palatable).

Finally, we assume that the γ_u terms are the same for both movers and non-movers. In other words, both movers and non-movers experience the same university effects. Without this assumption, our decompositions are only relevant for the set of academics who move across institutions.

4.1.2 Event study

The researcher versus place decomposition gives us a static picture of how commercialization propensity is split between person and place effects. However, an event study allows us to observe the dynamics (and assess pre-trends). In this section, we outline our event study framework. For the sake of clarity, we start by ignoring calendar year effects in this discussion. Moreover, we assume that we have a balanced panel for the time being. We will relax both of these assumptions later in this section.

In this simplified setting, we can let $y_{iut} = \gamma_u + \varepsilon_{iut}$.[26] If all movers had the same origin and destination university (u and u'), then we could construct an event study by simply plotting the mean of the outcome y by the year relative to the move. If, however, the destinations and origins of the academics vary, this same plot would not be informative. In this example, if half of our movers went from u to u', while the other half went from u' to u, then the same graph described above would show no effect of the move, as the two moves would cancel each other out. This apparent non-result would appear even if the absolute value of the changes in either direction were quite large.

To solve this problem, we would want to scale y by the size and direction of the move. For a mover i who goes from origin university $o(i)$ to destination university $d(i)$, we can compute the size of the move (in terms of the difference in destination-origin propensity to commercialize) as follows:

$$\delta_i = \bar{y}_{d(i)} - \bar{y}_{o(i)} \tag{5}$$

where $\bar{y}_{d(i)}$ is computed by averaging over all author-years (including non-movers) at the destination university (and analogously for $\bar{y}_{o(i)}$ at the origin university). Figure 4 shows the distribution of these move sizes for all of the movers in our sample.

How should we scale the effect of the move across universities on the impact of propensity to commercialize? Following Bronnenberg, Dubé, and Gentzkow (2012), we define scaled commercialization as:

$$y_{it}^{scaled} = \frac{y_{it} - \bar{y}_{o(i)}}{\delta_i}. \tag{6}$$

This expression implies that y_{it}^{scaled} will equal zero if the mover's commercialization exactly equals the average commercialization at his origin university. Similarly, y_{it}^{scaled} will equal one if the mover's commercialization equals the average commercialization at his destination university. A value between zero and one implies the researcher adopts some (but not all) of the new university's propensity to commercialize. Estimating the equation

$$y_{it}^{scaled} = \theta_{r(i,t)} + \varepsilon_{it} \tag{7}$$

where $\theta_{r(i,t)}$ are the relative year coefficients would give us an easy-to-interpret event study,

[26] Given the balanced panel assumption, we ignore the individual fixed effects for now.

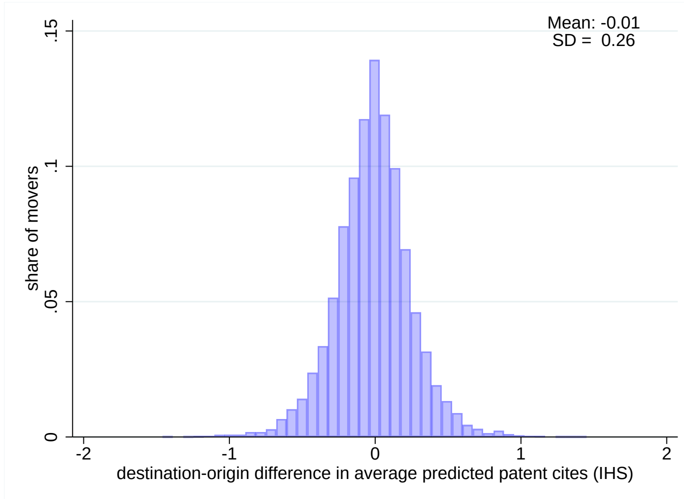


Figure 4: Distribution of destination-origin differences in commercialization propensities

Notes: This histogram shows the distribution of destination-origin differences in commercialization propensity ($\hat{\delta}_i$) across movers. For each author-publication (N = 16,380,070), we predict the number of citations from eventually granted patents filed within five years of publication by examining the journal it is published in. For each mover in our sample (N = 14,213), we define the difference between their destination and origin institution IHS five-year citation measure as $\hat{\delta}_i$, which is plotted in this histogram. The mean and standard deviation of this distribution is displayed in the top right corner.

where the size of the jump at the time of the move corresponds to the average value of $S_{university}$ across all movers. In other words, a larger jump would correspond to a larger share of cross-university heterogeneity being attributable to place-based characteristics. Moreover, if our model is correct, we would expect that the scaled outcome to be flat prior to the move — in other words, we should observe no pre-trends.

When taking this empirical framework to the data, we need to deal with a few additional complexities. First, if δ_i is close to zero (which Figure 4 suggests is a common occurrence), then y_{it}^{scaled} will behave badly as an outcome measure. Therefore, we want to avoid a regression specification that involves dividing by δ_i. Rearranging, we can instead re-write our event study as:

$$y_{it} = \bar{y}_{o(i)} + \theta_{r(i,t)}\delta_i + \varepsilon_{it}. \tag{8}$$

We also want to allow for calendar time controls. Moreover, because our panel is unbalanced, it is important to explicitly include individual level fixed effects. Adding these in (and combining $\bar{y}_{o(i)}$ and α_i into a single individual fixed effect $\tilde{\alpha}_i$) and replacing δ_i with its sample analogue $\hat{\delta}_i$ yields our final estimating equation:

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i + \tau_t + \varepsilon_{it} \tag{9}$$

	Top & Bottom 5%	*Top & Bottom 10%*	*Top & Bottom 25%*	*Above/Below Median*
Difference in IHS commercialization				
Overall	.85	.66	.45	.36
Researchers	.76	.54	.34	.31
Institutions	.09	.12	.11	.05
Share of difference attributable to				
Researchers	.89	.82	.77	.85
Institutions	.11	.18	.23	.15
	(±.10)	(±.11)	(±.07)	(±.07)

Table 2: Additive decomposition of IHS commercialization

Notes: This table is based on estimation of Equation (1), where the dependent variable y_{iut} is the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at institution u. Each column defines a set of universities R and R' based on commercialization rank $\hat{y}_u$. We rank institutions by averaging first across researchers within publication years to get $\hat{y}_{ut}$ and then averaging across publication years to obtain $\hat{y}_u$. The first row reports the difference in average commercialization ($\hat{y}_R - \hat{y}_{R'}$). The second row reports the difference in commercialization due to researchers ($\hat{\alpha}_R - \hat{\alpha}_{R'}$), and the third row reports the difference in commercialization due to universities ($\hat{\gamma}_R - \hat{\gamma}_{R'}$). The fourth row reports the share of the difference in commercialization due to researchers ($\hat{S}_{researcher}(R,R')$) and the fifth row reports the share of the difference in commercialization due to universities ($\hat{S}_{university}(R,R')$). The sixth row reports the 95% confidence interval constructed from bootstrapped standard errors on the share of geographic variation due to institutions. Standard errors are calculated using 50 repetitions of a resampled author-year panel; this resampling generates variation in both institution rank and levels. The sample is movers and nonmovers (N = 1,709,102 author-years).

where the relative-year coefficients $\theta_{r(i,t)}$ are the coefficients of interest. They measure the change in commercialization around the year of the move, *scaled by the size and direction of the move*.[27]

Figure 4 documents the distribution of the difference between the mean commercialization propensity of mover i's destination university and that of his or her origin university — essentially answering the question of how different origin and destination universities are in terms of their commercialization propensities. This distribution is symmetric around zero, with researchers making moves in both directions — from low to high commercialization universities and from high to low commercialization universities.[28]

[27] Finkelstein, Gentzkow, and Williams (2016) also include relative year fixed effects, which would allow for a common pre-trend among all movers (regardless of the size and direction of the move). We show in Section 4.5 that our results are robust to this alternative specification.

[28] It is worth keeping in mind that commercialization propensity is not collinear with university prestige. Appendix Figure 12 reports estimates from a regression of predicted patent citation rank over the period under study and the U.S. News and World Report school ranking in 2021, which estimates a coefficient of 0.47 and an R-2 of 0.2.

4.2 Additive decomposition estimates

Table 2 is based on our additive decomposition framework. The γ_u are estimated from a regression following Equation (1) using our sample of movers and non-movers, where y_{iut} is the predicted count of five-year patent citations to author i's papers in year t affiliated with university u. The estimates $\hat{\gamma}_u$ are consistent estimates of the true university effects as long as our identifying assumptions hold.

Table 2 presents statistics which make comparisons across *groups* of universities, rather than comparisons across individual universities. We define $\hat{y}_R$, $\hat{\gamma}_R$, and $\hat{\alpha}_R$ to be the simple average of all $\hat{y}_u$, $\hat{\gamma}_u$, and $\hat{\alpha}_u$ that fall within group R.

The first row of Table 2 presents estimates of $\hat{y}_R - \hat{y}_{R'}$: the difference in mean commercialization between the two groups of universities. The second row presents the component of that difference that is due to researcher characteristics ($\hat{\alpha}_R - \hat{\alpha}_{R'}$) and the third row presents the remaining difference, which is due to university effects ($\hat{\gamma}_R - \hat{\gamma}_{R'}$). Rows four and five convert these researcher and university components into shares. Finally, the last row presents the 95% confidence interval for these share estimates.

Columns in this table represent comparisons across different groups. The sizes of these comparison groups increases across columns, from the top and bottom 5% up to the top/bottom 50%. Thus, column (1) uses data from, and makes comparisons across, only institutions ranked in the top and bottom 5% of $\bar{y}_u$. The fourth and fifth rows of column (1) suggest that 79% of the variation in commercialization between these two groups of universities is due to differing researchers, while the remaining 21% is due to university effects. This division is fairly consistent across different comparison groups, with the university share ranging from 11 to 23%. Our confidence intervals let us rule out university shares smaller than 10% and larger than 33%.

Recall that the specific nature of our outcome — predicted citations based on journal of publication — implies that we are only investigating specific channels. Because the predicted citations are based on journal placement, we are observing an effect that operates through researchers changing the content of their research (or at least, the journals they are targeting). To the extent that different universities have an additional effect on patent citations, holding journal of publication fixed, we would not observe that in these results.

Table 3 presents an alternative decomposition, investigating what share of the cross-institution *variance* in commercialization is due to researchers versus place effects. The share of the cross-institution variance that would be eliminated if all institution effects were the same can be written as:

Cross-institution variance of mean:	
IHS commercialization	.043
Institution effects	.005
Researcher effects	.035
Correlation of average institution and researcher effects	.126
	(± .116)
Share variance would be reduced if:	
Institution effects were made equal	.193
	(± .081)
Researcher effects were made equal	.882
	(± .028)

Table 3: Variance decomposition of IHS commercialization

This table is based on estimation of Equation (1), where the dependent variable y_{iut} is the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at institution u. The results from a variance decomposition of y_{iut} are shown. This method is discussed in Appendix Section A. The first row reports the variance in $\hat{y}_u$, the second row reports the variance in $\hat{\gamma}_u$ (institution effect), and the third row reports the variance in $\hat{\alpha}_u$ (researcher effect). The correlation coefficient on $\hat{\gamma}_u$ and $\hat{\alpha}_i(u)$ is given by row four, accompanied by a 95% confidence interval constructed from the standard error obtained by an author-level re-sampling procedure with 50 bootstrap replications. The second half of this table displays the results from a counterfactual exercise that estimates the share of the variance in $\hat{y}_u$ that would be removed if the casual researcher or institution effects were equalized across institutions. The first row in the secind half reports the share of the variance that would be reduced if institution effects were equalized, again accompanied by a 95% confidence interval constructed by an empirical bootstrap. The third and final rows report the researcher effect analogs. The sample is movers and nonmovers (N = 1,709,102 author-years). Because $\hat{\gamma}_u$ can only be identified via movers, its precision relies on the number of movers per u. To address this, we restrict the sample to include only schools with at least 25 movers ($N = 161$ institutions). This choice is discussed further in Appendix Section C.

$$S^{var}_{university} = 1 - \frac{\mathrm{Var}(\bar{\alpha}_u)}{\mathrm{Var}(\bar{y}_u)}. \tag{10}$$

Similarly, the share of the variance that would be eliminated if all researcher effects were the same can be written as:

$$S^{var}_{researcher} = 1 - \frac{\mathrm{Var}(\gamma_u)}{\mathrm{Var}(\bar{y}_u)}. \tag{11}$$

Plugging in the empirical analogs of $\bar{y}_u$, $\bar{\alpha}_u$, and γ_u allows us to construct the estimates in Table 3. We find that 19% of the variance would be eliminated if university effects were equalized, and 88% would be eliminated if researcher effects were eliminated.[29] We also document a weak positive correlation between researcher and institution effects, with researchers who commercialize more sorting to universities that we estimate to have positive commercialization effects.[30]

[29] Note that these quantities do not sum to one because of the covariance term, which is positive.

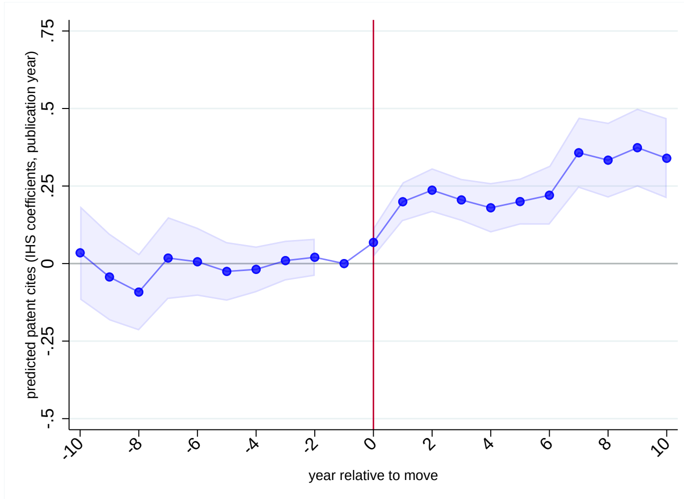


Figure 5: Event study: Predicted five-year patent citations

Notes: This figure is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9), where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the predicted count of citations received from patents filed within five years of publication for each of an author's papers in a given year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. The relative year -1 is omitted and normalized to zero. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution of the event study coefficients. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. There are N = 119,338 author years in our sample, coming from the sample of 14,213 movers. The red line denotes the researchers' move years.

4.3 Event study estimates

Figure 5 documents our main event study results. We plot the relative time coefficients (the $\theta_{r(i,t)}$ terms) from Equation (9). The dependent variable in this analysis is the predicted number of patent citations within five years to each of the author's papers published in a given year. We take the inverse hyperbolic sine transform of these predicted citations due to the skewness of this measure. Finally, we sum over all papers written in the same year (so that authors who are more prolific in a given year will have more predicted citations, all else equal). Since these coefficients are only identified up to a constant term, we normalize $r(i,t) = -1$ to be zero. The shaded area indicates the 95% confidence interval around these estimates, derived from bootstrapped standard errors clustered at the author level.

The figure shows a distinct pattern: after no clear trends in the decade prior to the move, there is a stark increase in the move year and the two years afterward. Academics

[30] Our institution effects are estimated with noise, which, as described by Andrews et al. (2008), can bias the correlation between person and institution effects downward, a phenomenon known as "limited mobility bias." To avoid this, we follow the advice of Andrews et al. (2012) and restrict to institutions estimated with at least 25 movers.

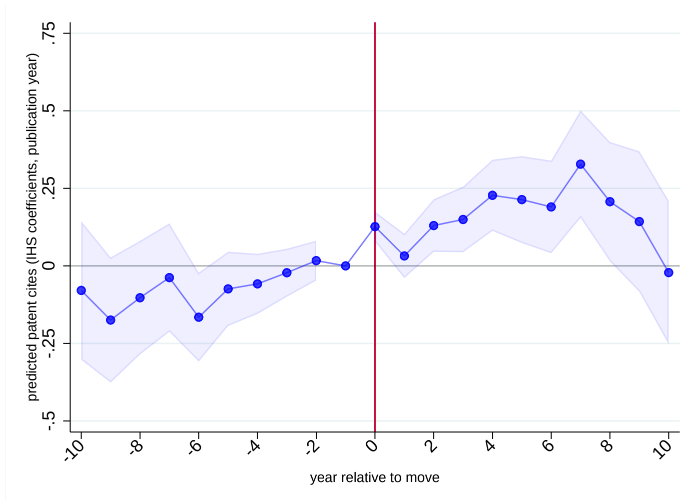
moving to universities with more of an orientation towards commercializing science end up producing research that, based on the publication outlet, is predicted to have more patent citations. After the initial climb, the coefficients remain fairly stable in the 0.25 to 0.35 range.

These magnitudes are similar, though not identical, to the additive decomposition results reported above. This is because the additive decomposition represents a slightly different thought experiment for several reasons. First, it analyzes the differences in commercialization rates across groups of universities, rather than averaging $S_{university}$ across all individuals who move. Second, it combines all pre- and post-move years, rather than separately estimating each relative year. And finally, the additive decomposition is estimated on both movers and non-movers, whereas the event study only uses movers.

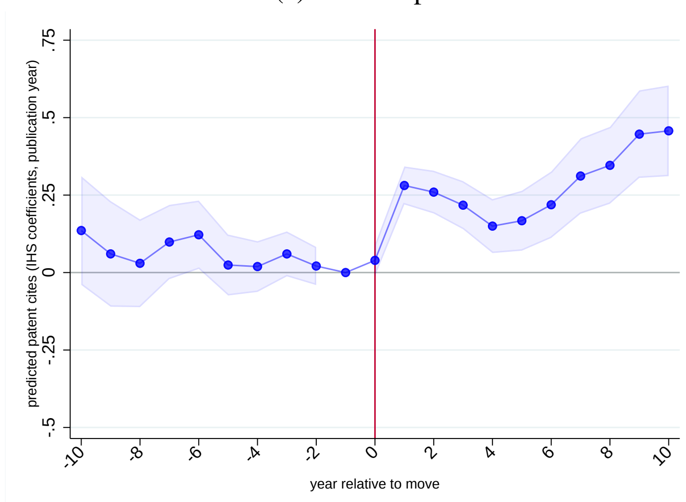
In Figure 6, we plot the results separately for researchers who move to a school that commercializes more (panel (a)) and researchers who move to a school that commercializes less (panel (b)). The degree of convergence in both panels is similar; if anything, we see slightly more convergence for downward moves. This suggests that movers pivot their research in both directions, depending on the type of move. Researchers who move to universities that commercialize more start to publish in journals that are more frequently cited by patents; the reverse is true for researchers who move to universities that commercialize less.

4.4 Additional outcomes

Actual five-year patent citations. We then repeat the analysis, looking at the *actual* patent citations received by the papers in the five years after their publication. As we discussed to in Section 3, this five-year cumulative measure poses some challenges in a traditional event study framework. To see why, we return to the example of a researcher who moves universities in the year 2000. A paper published in 1998 (before the move) will be allowed to accrue citations in the years 1998, 1999, 2000, 2001, 2002. At least two of those years (2001 and 2002) are post-move (2000 is ambiguous). Thus, while the paper was clearly written pre-move, the outcome of five-year citations contains some citations that were received post-move. This lag makes it challenging to classify five-year citations to a paper written in 1998 as a purely pre-move or post-move outcome. This challenge will exist for all papers written between 1996 and 1999. We refer to this four year window as the "contaminated period." More generally, for a move that occurs in year t, pre-move years $t-4$ to $t-1$ are contaminated.



(a) Moves up



(b) Moves down

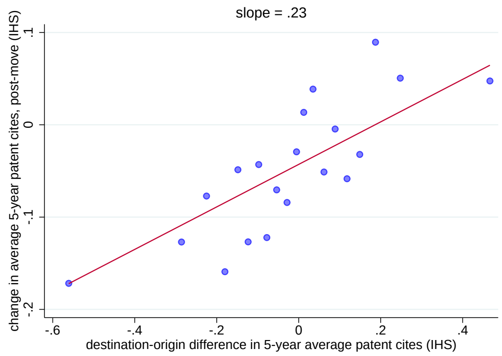
Figure 6: Comparing moves up and moves down

Notes: Each panel is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9), where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the predicted count of citations received from patents filed within five years of publication for each of an author's papers in a given year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. The relative year -1 is omitted and normalized to zero. Panel (a) is estimated using authors who move up (N =6,707), or those for which $\hat{\delta}_i > 0$. Panel (b) is estimated using authors who move down (N = 7,478), or those for which $\hat{\delta}_i < 0$. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution of the event study coefficients. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. The red line denotes the researchers' move years.

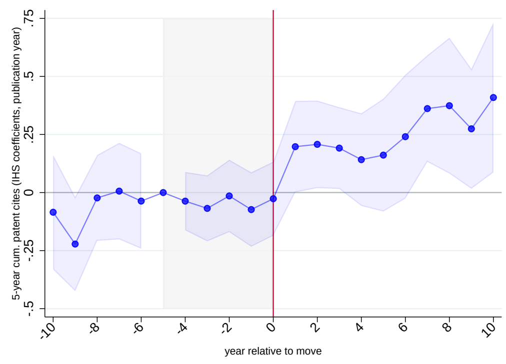
Despite this challenge, Figure 7 presents results for this outcome. Panel (a) again plots the change in mover i's commercialization activity (now measured by actual five-year patent citations) before and after the move against the destination-origin difference in commercialization activity (now also measured by actual five-year patent citations). We again define the pre-move period as relative years -10 to -1 and the post-move period as relative years 1 to 10. We find a slope of 0.23, which suggests that about one-fourth of heterogeneity in commercialization is due to place-specific factors.

Panel (b) presents the event study results. We see no evidence of a pre-trend, despite this "contamination" issue. This suggests that convergence does not begin until the author starts writing new papers at their destination university. The effect sizes are similar to those that we see in our predicted citation measure, with movers adopting around 25 to 50% of the commercialization propensity of their destination university. Again, this suggests that the key mechanism is the underlying research itself changing post-move. If the university were better at attracting commercial attention to the research, holding the research itself fixed, we would expect the effect to be larger for actual citations than for predicted citations, since the predicted citations outcome shuts this channel down. In addition, we might expect to see some effect starting in the "contaminated window" of the pre-period, because the academic's presence at the destination university could attract more attention to work written at the origin school. However, the lack of pre-trends suggests this is not occurring.

Weighted patent citations. When using patent citations to measure commercial spillovers, we may want to weigh patent citations from important patents more heavily. We compute each citing patent's importance by computing its "adjusted" patent citations as recommended by Lerner and Seru (2022). For each patent awarded in year t in technology class c, we calculate the ratio of that patent's forward citations after award divided by the average citations for all patents granted in year t and technology class c. We link this adjusted patent citation count to our five-year cumulative cites measure, where a publication accrues adjusted citations up to five years post-publication. Data on the count of forward patent citations and World Intellectual Property Organization (WIPO) patent-level technology classifications are downloaded from the USPTO PatentsView database and span the period 2000-2023. Patents with values above one have received above average citations compared to other patents in their same vintage and technology class. Appendix Figure 13 presents replicates Figure 7 using these weighted patent citations. The results are qualitatively very similar.



(a) Change in commercialization propensity by size of move



(b) Event study

Figure 7: Observed five-year patent citations

Notes: Panel (a) is a binned scatterplot that compares an individual's change upon moving in five-year patent citations with the difference in average institution-level patent citations between their destination and origin institutions. For each mover, we compute two values. First, we separately calculate the average count of IHS five-year citations for papers published pre- and post-move, and report the difference. We call this the change in an individual's commercialization level. Second, we generate $\hat{\hat{\delta}}_i$ by taking the institution-level differences, as shown in Figure 2. We call this the change in the institution's commercialization level. The x-axis displays ventiles of this institution-level change, while the y-axis plots, for each ventile, the average change in individual commercialization. The slope of the line of best fit is reported above the graph. Panel (b) is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9), where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the observed count of citations received from patents filed within five years of publication for each of an author's papers in a given year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. Because levels of event time are collinear when estimated together, we preemptively omit data from relative year = -5, forcing the value of its event study coefficient to zero, which can be seen on the graph without a standard error estimate. We normalize this particular relative year because it is the last year, pre move, that all five-year citations to papers written at a researcher's origin also accrue while the researcher is at their origin institution. The contaminated region is shaded in gray and a red line denotes the researchers' move years. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\hat{\delta}}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution on each coefficient. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. Each panel is estimated using 11,466 movers and 81,766 mover-years.

Venture capital-backed patent citations. One subset of patent citations of particular interest are those granted by venture capital-backed companies. Venture capital-backed firms have been long understood to be particularly innovative (Kortum and Lerner, 2000). In many cases, the commercialization of early-stage academic research has been due to the simultaneous provision of capital, mentorship, and governance by venture firms, as evidenced by successes such as Genentech, Netscape, and Moderna. We therefore recalculate our (actual) five-year patent citation measure, only counting patent citations that we are able to match with firms that received venture capital financing between 1968 and 2017. We link patents to firms based on company and city names and assignee IDs and call this set of patents "venture-backed." While the results are noisier (due to the smaller sample size), they are similar to our baseline five-year results. We observe a slope of 0.15 in Appendix Figure 14(a), suggesting that roughly one-fifth of commercialization heterogeneity on a move is due to place-specific factors; the implied jump in the event study in panel (b) ranges from 0.25 to 0.5, though with wider confidence intervals due to the sparser measure.

Patent-paper pairs. We also investigate the effect on patent-paper pairs. As discussed in Section 2, these represent papers that had an associated patent filed to protect the same discovery, by the same authors, at roughly the same time. The slope of 0.22 in Appendix Figure 15(a) suggests that nearly one-fifth of the commercialization heterogeneity as measured by patent-paper pairs is due to place specific factors. However, the event study results in panel (b) do show some pre-trends in the earlier years, warranting some caution in interpreting these results.

Scientific Advisory Boards. Lastly, we look at scientific advisory board (SAB) participation. We look at the propensity for researchers to join an SAB or an "SAB+" (which we define as an SAB *or* a regular firm board *or* a firm founder). Thus, the outcome is a 0/1 indicator for joining in a given year. Appendix Figure 16 and Appendix Figure 17 show the results. Given the sparsity of the outcome (we match only about 2,000 board join events to the 1,709,102 author-years in our sample), the results are unfortunately too noisy to draw many conclusions.

4.5 Robustness

In this section, we present several robustness checks bolstering our main results.

Measurement error in coding moves. We begin by exploring how measurement error affects our estimates. As first discussed in Section 2, we randomly drew 100 movers in our sample and attempted to use faculty web pages, CVs, and LinkedIn profiles to track whether we had coded their move years correctly. Of the 75 movers we were able to locate, 68 had indeed moved to and from the universities we identified (91%). Appendix Figure 18(a) then shows whether we correctly coded the move year correctly for those 68 movers. On average, we code movers as moving too late (we are most frequently off by one year). This is not surprising, given that we use publications (which may lag) to track moves.

To understand how this measurement error might affect our results, we simulate a very simple event study where movers experience a one-time convergence of 0.25 in relative year 1 (in other words, we set $\hat{\theta}_r = 0.25$ for all $r \geq 1$). That event study is shown in blue in Appendix Figure 18(b). Then, to understand how our timing errors affect the results, we re-assign move years so that the error in move years matches the empirical distribution in panel (a). The results from that event study are shown in red. Unsurprisingly, this leads to an increase that begins earlier, with a slight uptick in relative year -1 and a large uptick in relative year 0. This matches what we see in our event study results, and suggests that the fact that the effect manifests so early is likely due to measurement error.

Finally, we add in an additional 9% of movers who do not move (and thus have $\hat{\theta}_r = 0$ for all r) and re-run the event study. These results are shown in green. The effect size is dampened by about 10%, suggesting that our true effects could be about 10% bigger if we excluded false movers. To the extent that the 25% of authors that we could not locate are disproportionately like to be non-movers, then this dampening effect could be even larger.

Alternative specifications. An alternative to estimating equation (9) is to estimate:

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i + \rho_{r(i,t)} + \tau_t + \varepsilon_{it} \tag{12}$$

which allows for a common trend in commercialization for movers either pre- or post-move, as long as this trend is common across all movers regardless of the size and direction of their move. Including these additional covariates has little effect on our results, as shown in Appendix Figure 19.

Focusing on last authors. Next, we explore different "types of moves." Some moves are a professor moving from one position to another (as in the Carolyn Bertozzi example in the Introduction). However, other moves may represent different types of career steps;

for example, a PhD student moving to a post-doctoral position, or a post-doc moving to an assistant professor role. The results may be in some sense cleaner if we only include moves of principal investigators, whom are moving from one professorial role to another. One way to proxy for this would be to consider only those who are last authors before the move. It is common practice in scientific publications to have the senior author appear last, often regardless of his or her contribution. This step reduces the number of movers who are PhD students or postdocs at their origin institutions, and leaves us with 7,520 movers. The results are shown in Appendix Figure 20, and look very similar to our main results. In unreported analyses, we only include first and last authors, as well as only those above a threshold number of publications. The result are similar.

Alternative citation windows. In Appendix Figure 21, we replicate our event study for actual citations, looking at citation windows from one to ten years. The one and two-year citation windows are noisy, reflecting the fact that over one to two years, the vast majority of papers receive zero citations. The seven- and ten-year citation windows are also noisy, because we start to lose a larger share of our sample (when computing ten-year citation windows, any paper written after 2010 must be dropped, because our data runs through 2020). Still, the broad patterns look similar across the different citation windows.

Front page versus in-text citations. Next, we disaggregate the front-page citations from the in-text citations. Specifically, we replicate the event study in Figure 7(b), using five-year front-page citations and five-year in-text citations. The results are shown in Appendix Figure 22 and Appendix Figure 23, and look very similar to our main results.

4.6 Correlates of institution effects

Thus far, we have documented that universities matter for commercialization: institution effects explain 15-25% of the total heterogeneity that we observe in research commercialization. A natural next question is: why? What makes some universities "better" in this dimension than others?

We take a first step toward answering this question by correlating our estimated place effects $\hat{\gamma}_u$ with different university characteristics. We separate these characteristics into two broad groups. First, we have TTO features, sourced from the AUTM data, which describe characteristics of the university's TTO. These variables, among other things, include the number of licenses issued, the number of active licencses, and legal fee spending. Each TTO-level feature is calculated as an average of the 2000-2020 sample period, and then

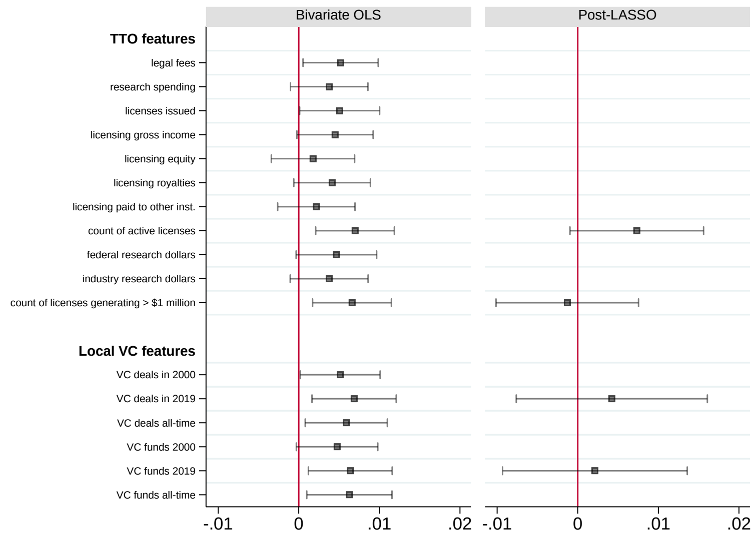


Figure 8: Correlates of average institution effects

Notes: This figure shows results from regressing the estimated institution effects – per Equation (2) – on various place-based factors. To minimize noise in the institution effect estimates, we first employ the empirical Bayes method described in Appendix Section C. All factors are standardized to have a mean of zero and standard deviation of one. The left panel shows the coefficients from bivariate regressions of IHS predicted five-year journal citations against each factor, individually. The right panel shows the coefficients from a multivariate regression, following a LASSO selection procedure. Specifically, a 10-fold cross-validation procedure that is designed to minimize mean squared error selects the optimal penalty weight to apply to the sum of the coefficients. The variables calculated to have non-zero coefficients are used in a separate OLS regression. Horizontal bars show the 95% confidence interval. The sample used in both panels are the 221 TTOs for which all covariates are available. Features of TTOs are retrieved from the AUTM data. For each TTO, AUTM variables are averaged across the years in our sample, from 2000 to 2021. Features of local venture capital funding are retrieved from Refinitiv, which compiles information on deals from securities filings, news stories, and anecdotal accounts. The venture capital covariates are summed at the county or county-year level.

transformed into a z-score. Second, we have local venture capital features. These are less specific to the university, and more related to the university's geography. We source these variables from Refinitiv, and aggregate the variables up to the county level. These variables are mostly focused on measure the amount of VC activity in the area, and include measures of transaction count and total funding in the years 2000 and 2022.

For each feature, we begin by estimating a bivariate OLS regression, where the dependent variable is our estimated university fixed effect, and the independent variable is a single university characteristic. Because our university effects are estimated with noise, we employ an Empirical Bayes adjustment (Morris, 1983) to shrink the estimates toward the mean (see Appendix C and Appendix Table 2 for details). The results of these bivariate regressions are on the left-hand side of Figure 8. We then estimate a second regression, where we regress our fixed effects on *all* of the university characteristics, and let LASSO select the characteristics. We then take these selected covariates and estimate a post-LASSO OLS regression. These post-LASSO coefficients are shown in the right half of Figure 8.

Only three variables are selected by the LASSO procedure, and only one (count of active licenses) is marginally significant in the post-LASSO OLS results. We suffer from a small sample size (221 observations, after we restrict to TTOs for which we have all of the characteristics). Still, the results suggest that (a) universities with more active licenses do better, and (b) local availability of venture capital funding helps. Taken together, these results – although only correlational – suggest that institution-level factors and geography are both important for successful commercialization.

5 Conclusion

Given the potential importance of academic research for both regional development and overall economic growth, encouraging the diffusion of university-based discoveries into the economy is an important policy question. Looking across universities, research and commercialization activities such as start-up formation vary tremendously. We take a first step towards unpacking this heterogeneity by analyzing how the propensity of academic research to spill over to commercial innovation changes when academics move across universities. We see an abrupt discontinuity in orientation towards commercialization when individuals move institutions. Our baseline estimates suggest that at least 15-25% of geographic variation in commercial spillovers from university-based research is attributable to place-specific factors. This pattern remains robust when we use multiple measures of commercial orientation and measurement schemes.

The analysis raises many questions for future research. One relates to the drivers of shifts in the location of academics. Earlier literature has highlighted the importance of funding availability (Azoulay, Ganguli, and Zivin, 2017; Bernstein, 2014), religious persecution (Moser, Voena, and Waldinger, 2014; Waldinger, 2012), scientific productivity (Coupé, Smeets, and Warzynski, 2005; Ganguli, 2015; Lenzi, 2009; Zucker, Darby, and Torero, 2002), and personal circumstances.[31] Understanding the extent to which the benefits of moves to more commercially oriented schools are anticipated by academics seems ripe for further exploration.

[31] An early paper found that the presence of children in a household can constrain scientific mobility, particularly amongst women (Shauman and Xie, 1996). Building on these insights, Azoulay, Ganguli, and Zivin (2017) provides a more comprehensive study of personal and professional factors affecting mobility. Subsequent papers have continued to build on the literature (e.g. Liu and Hu (2021) on collaborations incentivizing moves; Laudel and Bielick (2019) on differences in mobility decisions across fields).

A second topic is motivated by the persistent differences in the propensity to commercialize across schools. Understanding the institutional features that lead to these differences remains challenging. As suggested in the introduction, much of the academic literature has focused on the formal rules governing ownership and royalty sharing. Our descriptive results suggest that both the university and the geography matter. Moreover, conversations with practitioners suggest the importance of other considerations, from the presence of faculty role models to the savvy of the technology transfer offices. Given the importance of academic science as a driver of economic growth, these questions deserve more attention.

References

Abowd, John M., Francis Kramarz, and David N. Margolis. 1999. "High Wage Workers and High Wage Firms." *Econometrica* 67 (2): 251–333.

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.

Ahmadpoor, Mohammad and Benjamin F. Jones. 2017. "The Dual Frontier: Patented Inventions and Prior Scientific Advance." *Science* 357 (6351): 583–587.

Akcigit, Ufuk, Sina T Ates, Josh Lerner, Richard R Townsend, and Yulia Zhestkova. 2024. "Fencing Off Silicon Valley: Cross-Border Venture Capital and Technology Spillovers." *Journal of Monetary Economics* 141: 14–39.

Andrews, Martyn, Len Gill, Thorsten Schank, and Richard Upward. 2008. "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?" *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 171 (3): 673–697.

— 2012. "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias." *Economics Letters* 117 (3): 824–827.

Andrews, Michael, Lauren Russell, and Lei Yu. 2023. "Higher Education and Local Educational Attainment: Evidence from the Establishment of U.S. Colleges." *Review of Economics and Statistics* forthcoming.

Azoulay, Pierre, Ina Ganguli, and Joshua Graff Zivin. 2017. "The Mobility of Elite Life Scientists: Professional and Personal Determinants." *Research Policy* 46 (3): 573–590.

Azoulay, Pierre, Joshua S. Graff Zivin, and Bhaven N. Sampat. 2012. "The Diffusion of Scientific Knowledge across Time and Space: Evidence from Professional Transitions for the Superstars of Medicine." In: *The Rate and Direction of Inventive Activity Revisited*. Ed. by Josh Lerner and Scott Stern. Chicago, IL: University of Chicago Press: 107–155.

Azoulay, Pierre, Joshua S. Graff Zivin, and Jialan Wang. 2010. "Superstar Extinction." *Quarterly Journal of Economics* 125 (2): 549–589.

Bernstein, Rachel. 2014. *Managing a Lab Move*. Science Careers. URL: http://www.science.org/content/article/managing-lab-move.

Bernstein, Shai, Xavier Giroud, and Richard R. Townsend. 2016. "The Impact of Venture Capital Monitoring." *Journal of Finance* 71 (4): 1591–1622.

Bertrand, Marianne and Antoinette Schoar. 2003. "Managing with Style: The Effect of Managers on Firm Policies." *Quarterly Journal of Economics* 118 (4): 1169–1208.

Bikard, Michaël and Matt Marx. 2020. "Bridging Academia and Industry: How Geographic Hubs Connect University Science and Corporate Technology." *Management Science* 66 (8): 3425–3443.

Birkle, Caroline, David A. Pendlebury, Joshua Schnell, and Jonathan Adams. 2020. "Web of Science as a Data Source for Research on Scientific and Scholarly Activity." *Quantitative Science Studies* 1 (1): 363–376.

Bronnenberg, Bart J., Jean-Pierre H. Dubé, and Matthew Gentzkow. 2012. "The Evolution of Brand Preferences: Evidence from Consumer Migration." *American Economic Review* 102 (6): 2472–2508.

Bryan, Kevin A., Yasin Ozcan, and Bhaven N. Sampat. 2020. "In-Text Patent Citations: A User's Guide." *Research Policy* 49 (4): Article 103946.

Coupé, Tom, Valérie Smeets, and Frédéric Warzynski. 2005. "Incentives, Sorting and Productivity along the Career: Evidence from a Sample of Top Economists." *Journal of Law, Economics, and Organization* 22 (1): 137–167.

Di Gregorio, Dante and Scott Shane. 2003. "Why Do Some Universities Generate More Start-Ups than Others?" *Research Policy* 32 (2): 209–227.

Din, Alexander and Ron Wilson. 2018. "Understanding and Enhancing the U.S. Department of Housing and Urban Development's ZIP Code Crosswalk Files." *Cityscape: A Journal of Policy Development and Research* 20 (2): 277–294.

Ducor, Philippe. 2000. "Coauthorship and Coinventorship." *Science* 289 (5481): 873–875.

Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence From Patient Migration." *Quarterly Journal of Economics* 131 (4): 1681–1726.

Fisch, Christian O., Tobias M. Hassel, Philipp G. Sandner, and Joern H. Block. 2015. "University Patenting: a Comparison of 300 Leading Universities Worldwide." *Journal of Technology Transfer* 40 (2): 318–345.

Fleming, Lee, Hillary Greene, Guan-Cheng Li, Matt Marx, and Dennis Yao. 2019. "Government-Funded Research Increasingly Fuels Innovation." *Science* 364 (6446): 1139–1141.

Ganguli, Ina. 2015. "Immigration and Ideas: What Did Russian Scientists "Bring" to the United States?" *Journal of Labor Economics* 33 (S1): S257–S288.

Hausman, Naomi. 2022. "University Innovation and Local Economic Growth." *Review of Economics and Statistics* 104 (4): 718–735.

Helmers, Christian and Henry G. Overman. 2017. "My Precious! The Location and Diffusion of Scientific Research: Evidence from the Synchrotron Diamond Light Source." *Economic Journal* 127 (604): 2006–2040.

Henderson, Rebecca, Adam B. 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* 80 (1): 119–127.

Ho, Jonathan C. and Demei Lee. 2021. "Research Commercialisation Performance in Different Types of Universities: Case from Taiwan." *Scientometrics* 126 (10): 8617–8634.

Hvide, Hans K. and Benjamin F. Jones. 2018. "University Innovation and the Professor's Privilege." *American Economic Review* 108 (7): 1860–1898.

Jaffe, Adam B. 1989. "Real Effects of Academic Research." *American Economic Review* 79 (5): 957–970.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *Quarterly Journal of Economics* 108 (3): 577–598.

Jarvis, Lisa M. 2020. "Carolyn Bertozzi's Glycorevolution." *Chemical and Engineering News* 98 (5): 18–23.

Jensen, Richard and Marie Thursby. 2001. "Proofs and Prototypes for Sale: The Licensing of University Inventions." *American Economic Review* 91 (1): 240–259.

Kamariza, Mireille, Peyton Shieh, Christopher S. Ealand, Julian S. Peters, Brian Chu, Frances P Rodriguez-Rivera, Mohammed R. Babu Sait, William V. Treuren, Neil Martinson, Rainer Kalscheuer, Bavesh D. Kana, and Carolyn R. Bertozzi. 2018. "Rapid Detection of Mycobacterium Tuberculosis in Sputum with a Solvatochromic Trehalose Probe." *Science Translational Medicine* 10 (430).

Kantor, Shawn and Alexander Whalley. 2019. "Research Proximity and Productivity: Long-Term Evidence from Agriculture." *Journal of Political Economy* 127 (3): 819–854.

Kenney, Martin. 1986. *Biotechnology: The University-Industrial Complex*. New Haven, CT: Yale University Press.

Kortum, Samuel and Josh Lerner. 2000. "Assessing the Contribution of Venture Capital to Innovation." *RAND Journal of Economics* 31 (4): 674–692.

Lach, Saul and Mark Schankerman. 2008. "Incentives and Invention in Universities." *RAND Journal of Economics* 39 (2): 403–433.

Laudel, Grit and Jana Bielick. 2019. "How Do Field-Specific Research Practices Affect Mobility Decisions of Early Career Researchers?" *Research Policy* 48 (9): Article 103800.

Lenzi, Camilla. 2009. "Patterns and Determinants of Skilled Workers' Mobility: Evidence from a Survey of Italian Inventors." *Economics of Innovation and New Technology* 18 (2): 161–179.

Lerchenmueller, Marc J. and Olav Sorenson. 2016. "Author Disambiguation in PubMed: Evidence on the Precision and Recall of Author-ity among NIH-Funded Scientists." *PLoS ONE* 11 (7): e0158731.

Lerner, Josh and Amit Seru. 2022. "The Use and Misuse of Patent Data: Issues for Finance and Beyond." *Review of Financial Studies* 35 (7): 2667–2704.

Levin, Michael, Stefan Krawczyk, Steven Bethard, and Dan Jurafsky. 2012. "Citation-Based Bootstrapping for Large-Scale Author Disambiguation." *Journal of the American Society for Information Science and Technology* 63 (5): 1030–1047.

Liu, Meijun and Xiao Hu. 2021. "Will Collaborators Make Scientists Move? A Generalized Propensity Score Analysis." *Journal of Informetrics* 15 (1): 101113.

Mansfield, Edwin. 1991. "Academic Research and Industrial Innovation." *Research Policy* 20 (1): 1–12.

Martín-Martín, Alberto, Enrique Orduna-Malea, Mike Thelwall, and Emilio Delgado López-Cózar. 2018. "Google Scholar, Web of Science, and Scopus: A Systematic Comparison of Citations in 252 Subject Categories." *Journal of Informetrics* 12 (4): 1160–1177.

Marx, Matt and Aaron Fuegi. 2020a. "Reliance on Science by Inventors: Hybrid Extraction of In-text Patent-to-Article Citations." Working Paper 27987. National Bureau of Economic Research.

— 2020b. "Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles." *Strategic Management Journal* 41 (9): 1572–1594.

Marx, Matt and David H. Hsu. 2022. "Revisiting the Entrepreneurial Commercialization of Academic Science: Evidence from "Twin" Discoveries." *Management Science* 68 (2): 1330–1352.

Miller, Douglas L. 2023. "An Introductory Guide to Event Study Models." *Journal of Economic Perspectives* 37 (2): 203–30.

Moretti, Enrico. 2021. "The Effect of High-Tech Clusters on the Productivity of Top Inventors." *American Economic Review* 111 (10): 3328–3375.

Morris, Carl N. 1983. "Parametric Empirical Bayes Inference: Theory and Applications." *Journal of the American Statistical Association* 78 (381): 47–55.

Moser, Petra, Alessandra Voena, and Fabian Waldinger. 2014. "German Jewish Émigrés and U.S. Invention." *American Economic Review* 104 (10): 3222–3255.

Mowery, David C., Bhaven N. Sampat, Arvids A. Ziedonis, and Richard R. Nelson. 2004. *Ivory Tower and Industrial Innovation: University-Industry Technology Transfer Before and After the Bayh-Dole Act*. Redwood City, CA: Stanford University Press.

Murray, Fiona. 2002. "Innovation as Co-Evolution of Scientific and Technological Networks: Exploring Tissue Engineering." *Research Policy* 31 (8-9): 1389–1403.

Murray, Fiona and Scott Stern. 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 (4): 648–687.

Narin, Francis. 1994. "Patent Bibliometrics." *Scientometrics* 30 (1): 147–155.

Narin, Francis and Elliot Noma. 1985. "Is Technology Becoming Science?" *Scientometrics* 7 (3): 369–381.

National Center for Education Statistics. 2020. "Digest of Education Statistics."

Nelson, Richard. 1986. “Institutions Supporting Technical Advance in Industry.” *American Economic Review* 76 (2): 186–189.

Ouellette, Lisa Larrimore and Andrew Tutt. 2020. “How Do Patent Incentives Affect University Researchers?” *International Review of Law and Economics* 61: Article 105883.

Saxenian, Annalee. 1994. *Regional Networks: Industrial Adaptation in Silicon Valley and Route 128*. Cambridge, MA: Harvard University Press.

Sharoni, Brit. 2023. “The Effect of Inventor Mobility on Network Productivity.” Working Paper. Harvard University.

Shauman, Kimberlee and Yu Xie. 1996. “Geographic Mobility of Scientists: Sex Differences and Family Constraints.” *Demography* 33 (4): 455–468.

Siegel, Donald S., David Waldman, and Albert Link. 2003. “Assessing the Impact of Organizational Practices on the Relative Productivity of University Technology Transfer Offices: An Exploratory Study.” *Research Policy* 32 (1): 27–48.

Sine, Wesley David, Scott Shane, and Dante Di Gregorio. 2003. “The Halo Effect and Technology Licensing: The Influence of Institutional Prestige on the Licensing of University Inventions.” *Management Science* 49 (4): 478–496.

Stuart, Toby E. and Waverly W. Ding. 2006. “When Do Scientists Become Entrepreneurs? The Social Structural Antecedents of Commercial Activity in the Academic Life Sciences.” *American Journal of Sociology* 112 (1): 97–144.

Trajtenberg, Manuel. 1990. “A Penny for Your Quotes: Patent Citations and the Value of Innovations.” *RAND Journal of Economics* 21 (1): 172–187.

Visser, Martijn, Nees Jan van Eck, and Ludo Waltman. 2021. “Large-Scale Comparison of Bibliographic Data Sources: Scopus, Web of Science, Dimensions, Crossref, and Microsoft Academic.” *Quantitative Science Studies* 2 (1): 20–41.

Waldinger, Fabian. 2012. “Peer Effects in Science: Evidence from the Dismissal of Scientists in Nazi Germany.” *Review of Economic Studies* 79 (2): 838–861.

Zucker, Lynne G., Michael R. Darby, and Marilynn B. Brewer. 1998. “Intellectual Human Capital and the Birth of U.S. Biotechnology Enterprises.” *American Economic Review* 88 (1): 290–306.

Zucker, Lynne G., Michael R. Darby, and Maximo Torero. 2002. “Labor Mobility from Academe to Commerce.” *Journal of Labor Economics* 20 (3): 629–660.

Online Appendix for *The Wandering Scholars: Understanding the Heterogeneity of University Commercialization*

By Josh Lerner, Henry Manley, Carolyn Stein, and Heidi Williams

A Details on sample construction

To increase the count of movers in our sample, we follow a multistep procedure illustrated in Figure Appendix Figure 2. First, we code 12,211 movers using an author-year panel of corresponding e-mail address stems. Described in Section 2.3, it is a requirement that we observe one's move in the data to code them as a mover. That is, there cannot be any "hole" in their e-mail-based affiliation between the move. But in cases where there are only one-to-three years to fill between the move, we regard these authors as *movers* that we would like to add to our sample.[1] We incorporate additional information from the Web of Science (the preferred organization field, described in Section 2.3) to attempt to fill the affiliation gaps for these candidate movers. In some cases, this helps to tell us the exact year in which they leave their origin. If using organization data is not enough to close[2] the gap, we set the candidate mover aside for an additional imputation step described later.

Until this point, an author could only make it into our sample if they were ever listed as the corresponding author on a publication, and that publication shared their e-mail address. This approach to coding movers is helpful in bypassing the noise in the Web of Science organization field and the "multiple affiliations problem" described in Section 2.3. However, it is also quite restrictive and potentially misses out on movers that were never corresponding authors. More generally, one might also worry that using e-mail addresses to code movers is too narrow, since Appendix Figure 8 shows that even in 2020 less than a third of all publications link a corresponding email. To address this, and in hopes of adding to our sample, we examine 5,324,663 authors who never appear in our e-mail-based panel. We start by seeing if any of these authors could be added to our sample as-is; meaning, their existent organization record is enough to determine where they sat in each year. There are 1,347 movers and 165,812 non-movers that are added to our sample with this approach. The top right panel of Figure Appendix Figure 2 shows this step.

[1] We might worry that an author moves more than once if there are many affiliation-years to impute.

[2] Here, "close" does not necessarily mean every missing affiliation year between the move is filled. Instead, it means the move year is observed in the data. Suppose we observe Josh Lerner at school A in 2012 and at school B in 2016, and so we need to back out where he was 2013-2015. Learning that Josh was at school A in 2015 or that he was at school B in 2013 is enough to know the year he moved in.

As before with the e-mail panel, there are some authors that we know are movers but cannot add to our sample because there is a gap in their organization record. As a final attempt to include these movers in our sample, we pool them with the candidate movers identified via the e-mail approach who also need additional imputation. Figure Appendix Figure 2 shows this as two arrows (one from the top left e-mail based panel and one from the top right "email-less" panel) pointing into the same node in the bottom center panel. There are 20,729 candidate movers that we could add, at maximum.

The final step in our procedure is to use organization ZIP codes to fill affiliation gaps for the remaining 20,729 candidate movers. The idea is that if we know a mover's origin and destination ZIP codes, we can use a *change* in ZIP codes to code the move-year and solve the same imputation problem. In general, the Web of Science ZIP code data is noisy, often identifying multiple codes for an author within the same year. This is especially true around the time of the move, which makes it challenging to confidently discern which ZIP code change is the one that encodes a move. To get around this, we let the first year with non-missing ZIP code data correspond to the mover's origin and the last year of non-missing data correspond to their destination. Then, for each year we need to fill an affiliation for, we count the number of times there is a match between the origin ZIP codes and the ones listed between the move. We repeat this process for the destination ZIP codes. We then compare these counts. For example, if there are two ZIP code matches to one's origin and zero to one's destination, we code them to still be at their origin. In case of ties, and in accordance with move-timing work described in Section 2.3, we code them to be at their destination. Only 3,349 of the 20,729 candidate movers ever have a ZIP code, of which 655 are coded as movers. Taken together, this procedure codes $12,211+1,347+655=14,213$ movers and $337,178+165,812=499,705$ non-movers, corresponding to a total of 1,709,102 author-years.

B Variance decomposition

Like in Finkelstein, Gentzkow, and Williams (2016), we provide results from a variance decomposition of our main outcome: predicted five-year patent citations. To begin with, we use the estimates generated by Equation (1)— $\hat{y}_{it}$, $\hat{\gamma}_u$, and $\hat{\alpha}_i$— and average each to the institution-year level, and then again to the institution level. This produces a dataset with $\hat{y}_u$, $\hat{\gamma}_u$, and $\hat{\alpha}_u$ for each institution with a non-zero count of movers ($N=248$). Then, the share of cross-institution variance that would be reduced by equalizing researchers across universities is written as

$$S^{var}_{researcher} = 1 - \frac{Var(\hat{\gamma}_u)}{Var(\hat{y}_u)} \tag{13}$$

where the *reduction* in variance is modeled to depend on only differences in institution effects, scaled by the overall variance across institutions. This means that any variance that remains is attributed to differences in the stock of researchers across universities. The analog for variance reduction as a function of equalizing place-specific factors is written as

$$S^{var}_{university} = 1 - \frac{Var(\hat{\alpha}_u)}{Var(\hat{y}_u)}. \tag{14}$$

Note, unlike $S_{researcher}$ and $S_{university}$, $S^{var}_{researcher}$ and $S^{var}_{university}$ are not additive, meaning they will not sum to one so long as $Cov(\hat{\alpha}_u, \hat{\gamma}_u) > 0$. One additional caveat is that because Equation (1) is identified via movers, our ability to precisely estimate γ_u depends, somewhat, on the number of movers observed at u. We want to minimize the extent to which cross-institution variance in $\hat{y}_u$ is driven by noise resulting from few movers, or "limited mobility bias." In particular, we know from Equation (13) that as $Var(\hat{\gamma}_u)$ increases, $S^{var}_{researcher}$ decreases. The same is true for $Cov(\hat{\alpha}_u, \hat{\gamma}_u)$, where for a fixed $Var(\hat{y}_u)$, noise generated by few movers will make $\hat{\gamma}_u$ either an over- or under-estimate of γ_u. And because we take institution and researcher effects as additive, if $\hat{\gamma}_u > \gamma_u$ we know $\hat{\alpha}_u < \alpha_u$.

Empirically, Appendix Figure 24 shows that the number of movers varies greatly across schools. To minimize noise-driven mistakes in $\hat{\gamma}_u$, and in accordance with Andrews et al. (2012), we restrict the variance decomposition to include only institutions with 25 movers. Analysis in Appendix Section C explores the sensitivity of this choice.

The results from this exercise are shown in Table 3 and are generally consistent with estimates presented by the event study (Figure 5) and additive (Table 2) analogs. Specifically, it is estimated that about 20% of the cross-institution variance in predicted citations is attributable to variance in institution effects.

C Empirical Bayes adjustment

One additional step we take to minimize noise in our estimate of $\hat{\gamma}_u$ is to employ the empirical Bayes method (Morris, 1983). Broadly, this approach trades off sampling variance in the distribution of the estimated place effects with potential bias induced by their posterior mean.

To begin, we assume that $\hat{\gamma}_u$ is an unbiased estimate of γ_u with known variance s^2_u. Next, we assume that the latent parameters γ_u are randomly drawn from the distribution G_γ

which is defined in the population of institutions, is normally distributed, and is invariant to s_u^2 across institutions. Together, these assumptions yield the hierarchical model:

$$\begin{aligned}\hat{\gamma}_u|\gamma_j u, s_u^2 &\sim N(\gamma_u, s_u^2)\\ \gamma_u|s_u^2 &\sim N(\mu_\gamma, \sigma_\gamma^2).\end{aligned} \tag{15}$$

This model has two hyperparameters, μ_γ and σ_γ^2, which we estimate by method of moments as:

$$\hat{\mu}_\gamma = \frac{1}{U}\sum_{u=1}^{U}\hat{\gamma}_u, \quad \hat{\sigma}^2 = \frac{1}{U}\sum_{u=1}^{U}\left((\hat{\gamma}_u - \hat{\mu}_\gamma)^2 - \hat{s}_u^2\right). \tag{16}$$

The subtraction of $\hat{s}_u^2$ in Equation (16) is a bias correction for the sampling variance of γ_u. Once these two hyperparameters are estimated, the empirical Bayes method produces posteriors as:

$$\hat{\gamma}_u^{EB} = \hat{\mu}_\gamma\left(\frac{s_u^2}{s_u^2+\sigma_\gamma^2}\right) + \hat{\gamma}_u\left(\frac{\sigma_\gamma^2}{s_u^2+\sigma_\gamma^2}\right). \tag{17}$$

Appendix Figure 25 shows the result of this adjustment on the distribution of institution effects. Appendix Table 2 explicitly reports the value of select institution effects, pre- and post-adjustment.

Additionally, Appendix Figure 26 runs a horse race between this empirical Bayes method and the recommendation of Andrews et al. (2012) to restrict the sample to include institutions with more than 25 movers. First, panel (a) plots the estimated $Cov(\hat{\alpha}_u, \hat{\gamma}_u)$ and $S_{university}$ generated from increasingly strict restrictions on the number of movers per institution. The leftmost point on each series is the estimate from the unrestricted sample, and shows the adverse effects of limited mobility bias. From around $25-75$ movers per school, each estimate stabilizes. A green line shows the number of institutions that remain in each sample, according to the restriction imposed by the x-axis. Panel (b) takes the blue series from Panel (a) and plots it with an analog computed with institution effects that have been adjusted by the empirical Bayes method. That is, for every level of the x-axis, each series uses the same sample of institutions and only differ by whether shrinkage occurs. This exercise teaches us that implementing shrinkage on the full sample does not arrive the practitioner at the same result as restricting to institutions with > 25 movers. It is not until each institution in the sample has at least about 70 movers that the results from each of these approaches converge.

D Supplemental figures and tables

(a) Original journal article

Programmable probiotics for detection of cancer in urine

Tal Danino,[1*] Arthur Prindle,[2*] Gabriel A. Kwong,[1†] Matthew Skalak,[1] Howard Li,[2] Kaitlin Allen,[1] Jeff Hasty,[2,3,4‡] Sangeeta N. Bhatia[1,5,6,7,8‡§]

[1]Institute for Medical Engineering and Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139, USA. [2]Department of Bioengineering, University of California, San Diego, La Jolla, CA 92093, USA. [3]BioCircuits Institute, University of California, San Diego, La Jolla, CA 92093, USA. [4]Molecular Biology Section, Division of Biological Science, University of California, San Diego, La Jolla, CA 92093, USA. [5]Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA. [6]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA. [7]Electrical Engineering and Computer Science and David H. Koch Institute for Integrative Cancer Research, MIT, Cambridge, MA 02139, USA. [8]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA.

*Equally contributing lead authors.

†Present address: Wallace H. Coulter Department of Biomedical Engineering, Georgia Tech and Emory School of Medicine, Atlanta, GA 30332, USA.

‡Equally contributing senior authors.

§Corresponding author. E-mail: sbhatia@mit.edu

(b) Web of Science XML file


```
<name seq_no="8" role="author" reprint="Y" addr_no="1 5 6 7 8 9">
<display_name>Bhatia, Sangeeta N.</display_name>
<full_name>Bhatia, Sangeeta N.</full_name>
<wos_standard>Bhatia, SN</wos_standard>
<first_name>Sangeeta N.</first_name>
<last_name>Bhatia</last_name>
<email_addr>sbhatia@mit.edu</email_addr>
```


(c) SQL tables

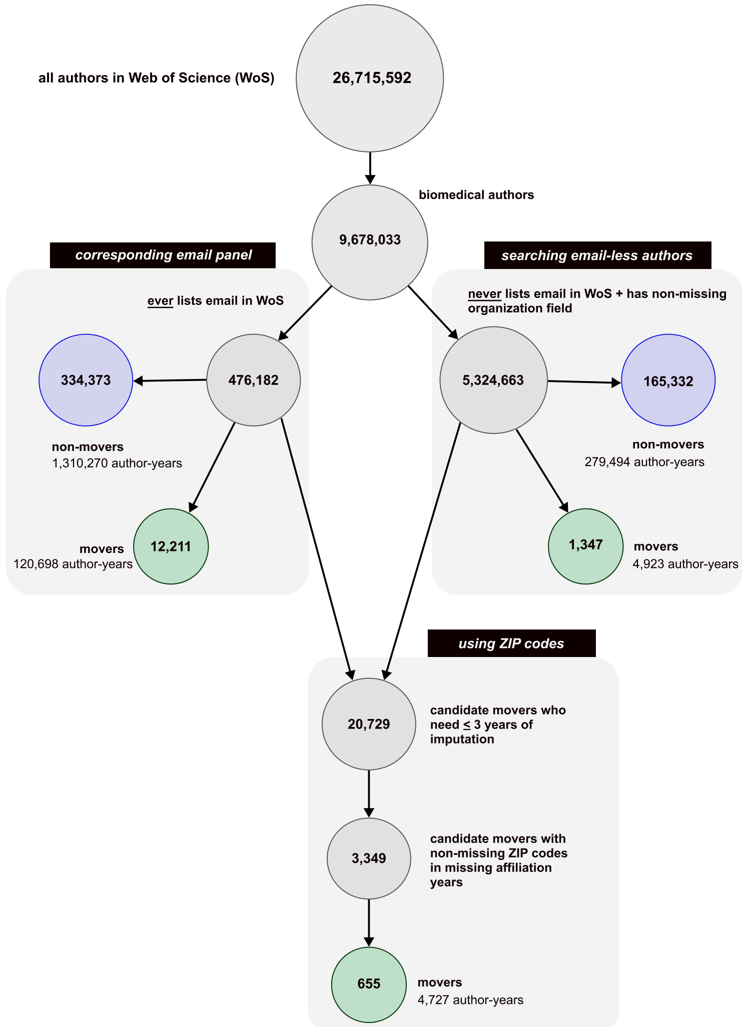
wos_uid	address_spec_addr_no	organization
WOS:000355179800007	1	Massachusetts Institute of Technology (MIT)
WOS:000355179800007	2	University of California System
WOS:000355179800007	2	University of California San Diego
WOS:000355179800007	3	University of California System
WOS:000355179800007	3	University of California San Diego
WOS:000355179800007	4	University of California System
WOS:000355179800007	4	University of California San Diego
WOS:000355179800007	5	Harvard University
WOS:000355179800007	5	Massachusetts Institute of Technology (MIT)
WOS:000355179800007	5	Broad Institute
WOS:000355179800007	6	Harvard University
WOS:000355179800007	6	Brigham & Women's Hospital
WOS:000355179800007	7	Massachusetts Institute of Technology (MIT)
WOS:000355179800007	8	Massachusetts Institute of Technology (MIT)
WOS:000355179800007	9	Howard Hughes Medical Institute

wos_uid	record_order	reprint	email_addr	full_name
WOS:000355179800007	1			Danino, Tal
WOS:000355179800007	2			Prindle, Arthur
WOS:000355179800007	3			Kwong, Gabriel A.
WOS:000355179800007	4			Skalak, Matthew
WOS:000355179800007	5			Li, Howard
WOS:000355179800007	6			Allen, Kaitlin
WOS:000355179800007	7			Hasty, Jeff
WOS:000355179800007	8	Y	sbhatia@mit.edu	Bhatia, Sangeeta N.

Appendix Figure 1: Measuring university affiliations of authors: An example

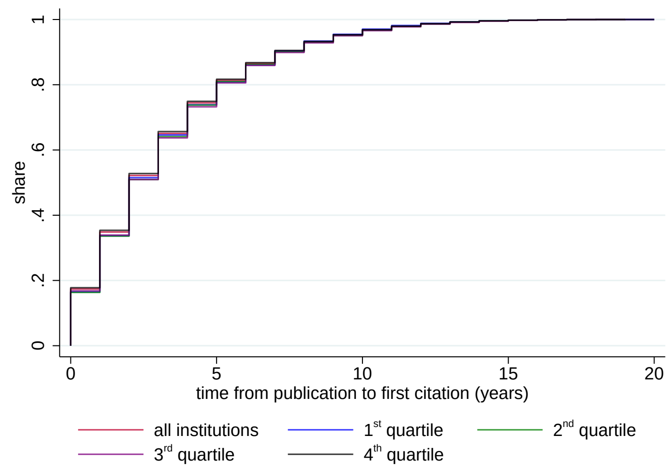
Notes: Example taken from Danino et al. (2015), Programmable Probiotics for Detection of Cancer in Urine

.

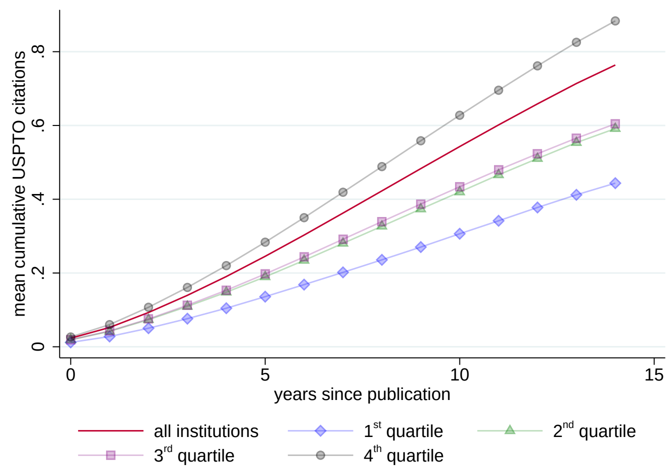


Appendix Figure 2: Author-year panel construction

Notes: This diagram tracks the "flow" of authors through the various sample restrictions we impose to isolate biomedical movers (green nodes) and non-movers (blue nodes). Summing across the values in the green nodes generates the total count of movers in our sample, 14,213. Summing across the values in the blue nodes equates the number of non-movers, 499,705. The same statements are true for the count of author-years, which are listed in the periphery of each colored node. The flowchart is partitioned into three gray regions, and is labeled as such. However, the first step in our sample construction is to include only biomedical researchers. The leftmost region shows the authors we capture from the basic email panel approach described in Section 2.3. This method begins with the author-year email skeleton shown in Table 1. The rightmost gray region displays the authors we add to our sample by examining authors that never list corresponding email addresses, and so would never appear in our email panel skeleton. The central gray region takes the "scraps," or authors difficult to code from each preceding procedure, and uses changes in ZIP codes to add additional movers to our sample. These approaches are described in Appendix Section A.



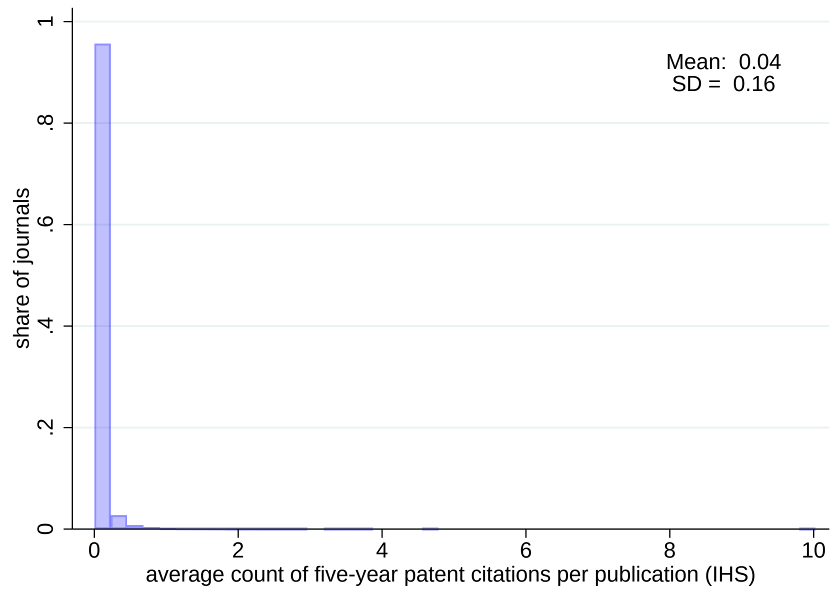
(a) Lag between paper publication and first patent citation



(b) Cumulative citations by year after publication

Appendix Figure 3: Lags and citation counts

Notes: Panel (a) shows the CDF of the time between a paper's publication and its first citation by a patent conditional on the paper being cited by at least one patent. Patent citations are dated to the year that the patent application was filed. The figure contains the CDF for publications from all institutions, as well as separate CDFs for papers from each quartile of institutions ranked by total research expenditure (averaged over years in our data). Panel (b) shows the average cumulative number of patent citations to papers in each year after paper publication, conditional on the paper being cited by at least one patent. Like Panel (a), this figure presents the cumulative citation count for all publications, as well as the cumulative citations counts for publications from each quartile of institutions ranked by total research expenditure (averaged over years in our data). The sample for each panel is all cited publications for which we have lag data and affiliation information from publication years between 2000 and 2020 (N = 16,380,070).



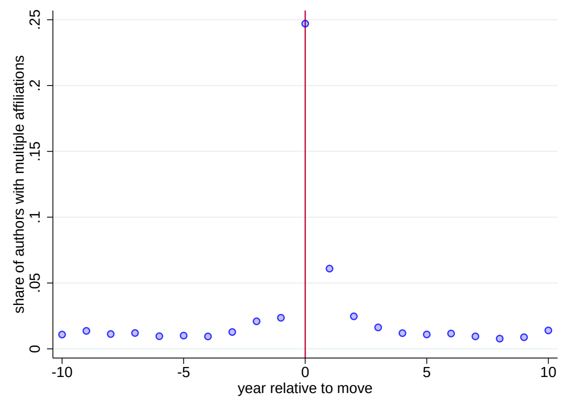
(a) Distribution of commercialization across journals

rank	*journal*	*publication count*	*mean patent citations*
1	2008 SID International Symposium, Technical Digest	545	10.03
2	2007 SID International Symposium, Technical Digest	478	4.78
3	International Electron Devices Meeting 2000, Technical Digest	201	4.71
4	Annual Review Of Immunology	705	4.68
5	Nature Biotechnology	10257	3.65
6	mAbs	1220	3.60
7	Annual Review Of Biomedical Engineering	201	3.60
8	IEEE International Electron Devices Meeting 2004, Technical Digest	251	3.53
9	2007 IEEE International Electron Devices Meeting, Technical Digest	240	3.31
10	International Electron Devices Meeting - 1997, Technical Digest	216	2.90

(b) Ranked fields by commercialization propensity

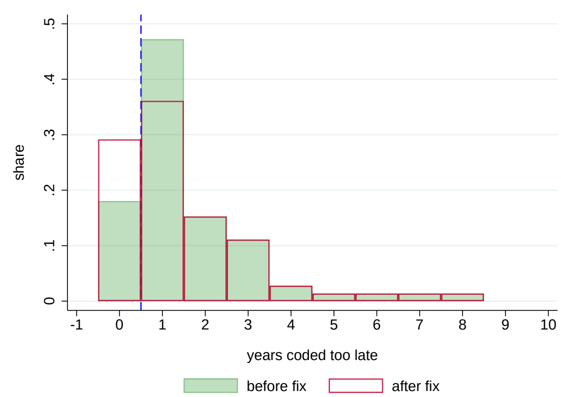
Appendix Figure 4: Commercialization among the most frequently cited journals

Notes: Panel (a) plots the distribution of average citations within five years of a paper's publication, across scientific journals. The sample used to construct this distribution is the universe of papers in the Web of Science database with journal information provided. The mean and standard deviation of journal-level count of five-year patent citations are denoted in the top right corner of the panel. Panel (b) lists the ten journals whose publications have the highest average annual count of five-year patent citations (Column 4), by name (Column 2). Annual average count of publications is included in Column 3.



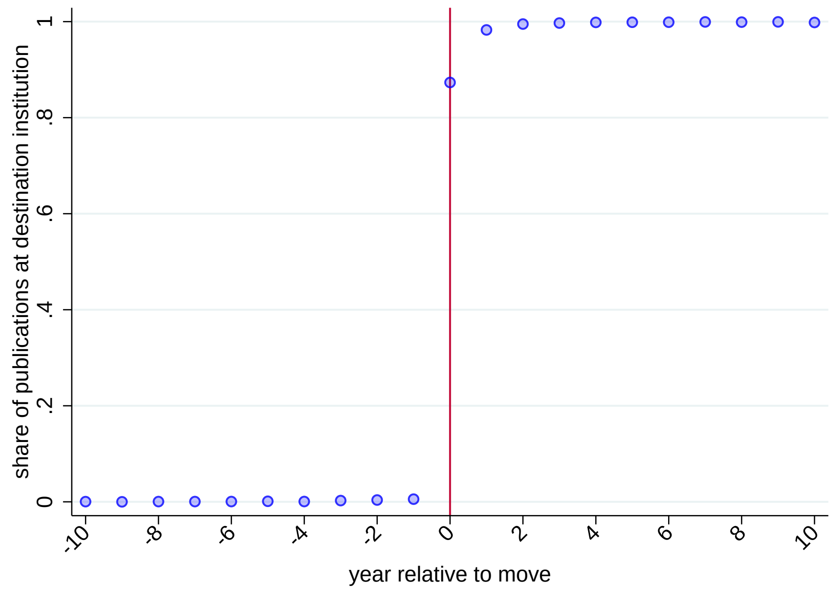
Appendix Figure 5: Multiple affiliations in the WoS by relative year

Notes: This figure plots the share of authors whose publications, within a single publication year, list more than one affiliation. This share is aggregated across all authors by year relative to one's move. A red vertical line denotes the move year. The sample of all mover-years, which includes the move year and ten years before and after the move, is used to calculate this share (N =119,338).



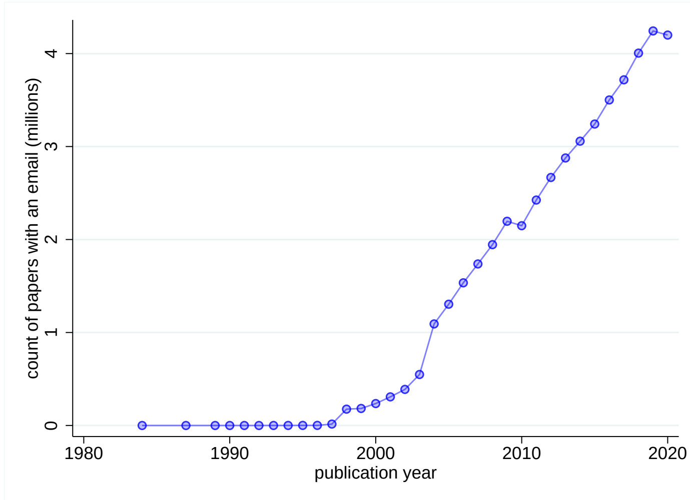
Appendix Figure 6: Move timing correction

Notes: Shown is the the distribution of move-timing errors encoded by a hand-coded, randomly selected sample of 100 movers from our full sample of 14,213. By collecting information from curriculum vitae, researcher LinkedIn profiles, and university lab webpages, we can observe the "ground truth" for the year in which a researcher moved. We use this ground truth and compare it against what we are able to code from the Web of Science publication records. The x-axis is the difference between the year in which a researcher actually moves and the year in which we have coded them to move. Per the green bars, the modal move timing error is one year, which informs an ex-post move-year correction procedure that takes into account the share of publications that list a mover's destination, one year before we originally coded them to move. The red bars show the result of implementing our correction procedure. The y-axis in each plot is the count of the hand-coded movers whose difference in actual and observed move appears in each bin.

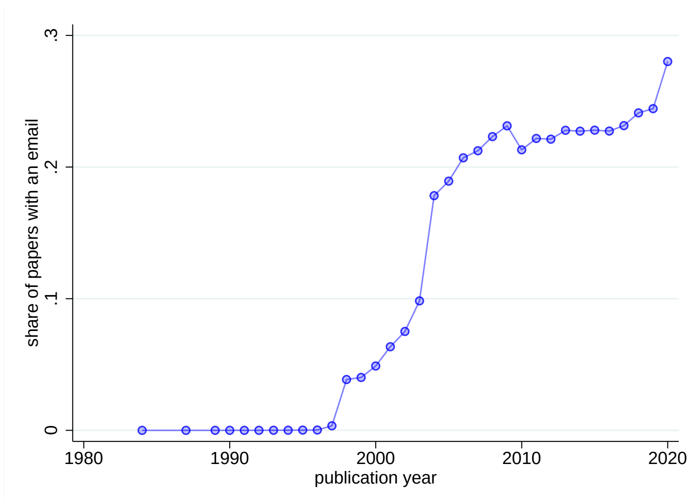


Appendix Figure 7: Share of publications associated with destination institution by relative year

Notes: The full sample of biomedical researcher movers (N = 14,213) is used to plot the average share of a mover's publications per relative year that list their destination institution as their academic affiliation. Author-year affiliations are drawn from corresponding email addresses disclosed on publications; this graph considers all the affiliations that an author might state within a given year, and calculates the share of these that match their identified destination institution. This value is then averaged over all movers, by relative year. A red line is drawn for relative year zero, the move year, to indicate the year when there is expected to be a discontinuous increase in the share of publications associated with a mover's destination.



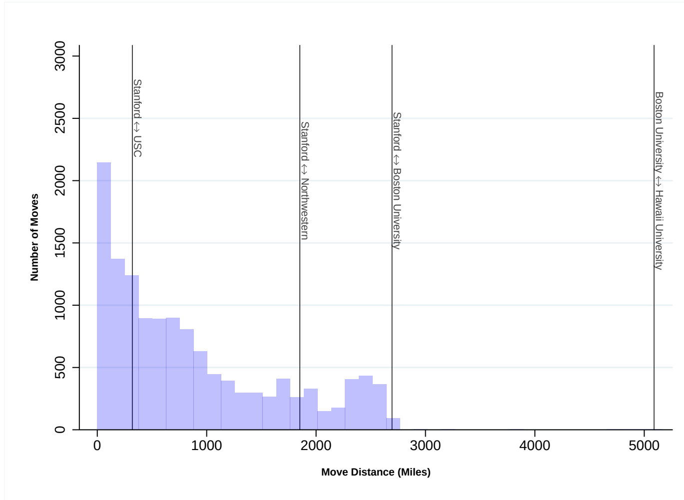
(a) Count



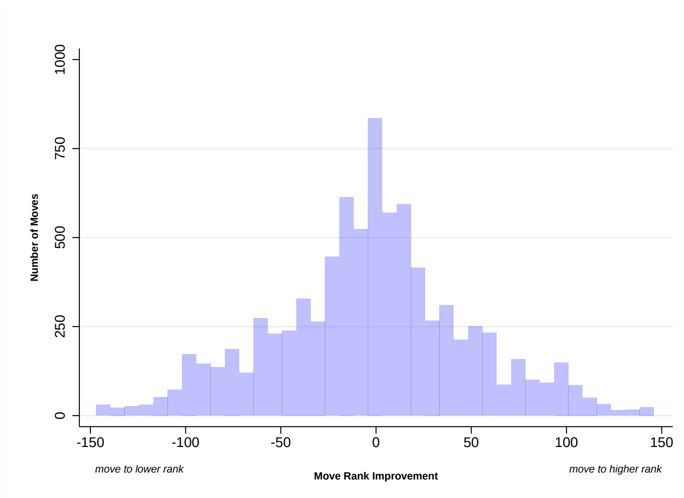
(b) Share

Appendix Figure 8: Email prevalence over time

Notes: This figure illustrates the change in the propensity for publications recorded in the Web of Science to report a corresponding author e-mail address. Panel (a) plots the count (in millions) of publications with a corresponding email address by publication year (1984-2020). Panel (b) plots the share of publications with a corresponding email address over the same period. The sample used to construct these tabulations is the universe of publications in the Web of Science database from 1984-2020.



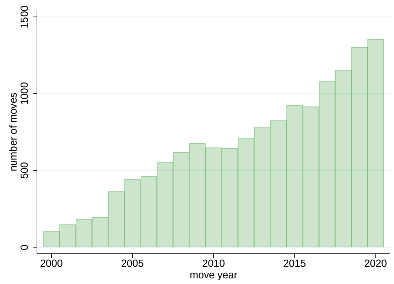
(a) Distance moved



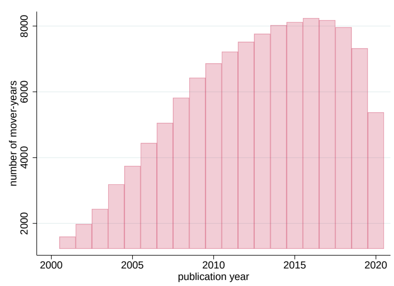
(b) Share

Appendix Figure 9: Descriptives on researcher moves

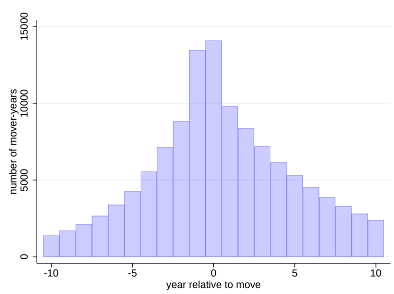
Notes: Panel (a) shows the distribution of miles traveled for the researcher moves in our sample. There are a total of 14,213 movers in our sample, of which 13,266 have origin and destination institutions that each can be geocoded using the python software package `geopy`. The geodesic distance between origin and destination institution for these movers is plotted on the x-axis, and the count of movers that move that distance is plotted on the y-axis. There are four black vertical lines that exemplify the "distance traveled" for four moves. Panel (b) plots the distribution of move-induced rank change. We link U.S. World News 2021 rankings of institutions to the origin and destination of each mover and calculate the difference in rank, post-move; this difference is plotted on the x-axis. The count of movers in each rank-change bin is plotted on the y-axis. Researchers that increase their rank post-move appear to the right of zero. Note, only colleges and universities are ranked by U.S. World News, and so we do not include any researchers who move to or from research institutions or hospitals (e.g., The Broad Institute), which leaves a sample of 6,387 movers.



(a) Number of researcher moves per publication year



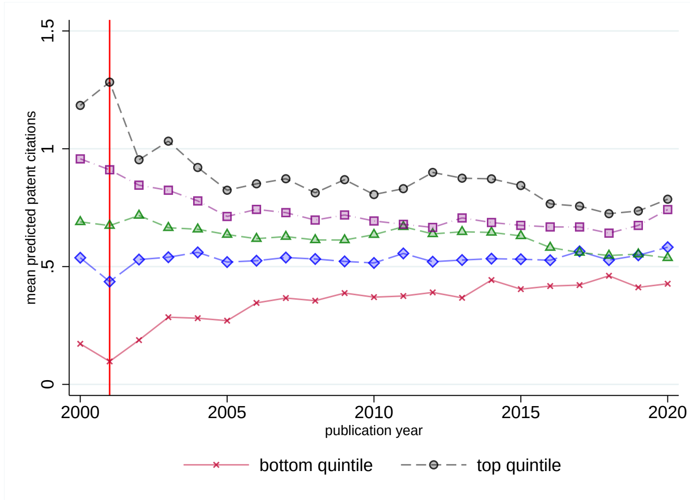
(b) Number of movers per calendar year



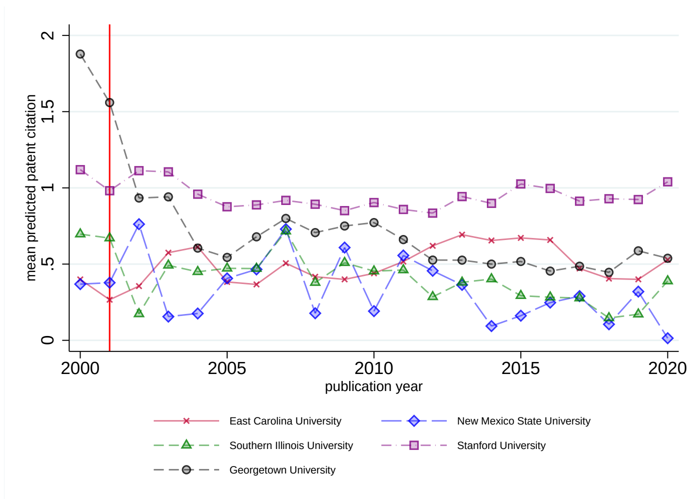
(c) Number of movers per relative year

Appendix Figure 10: Tabulating move time, calendar time, and event time

Notes: This figure tabulates the three "types" of time used in this paper to identify the effect of place on research commercialization propensity– move years, publication years, and relative years. The three are collinear by the following relationship: relative year = calendar year - move year. Using our full sample of (N = 14,213) movers, Panel (a) shows the number of researcher moves by publication year (move time), Panel (b) shows the number of movers who publish in each publication year (calendar time), and Panel (c) shows the number of movers who publish in each relative year (event time). Not every mover appears in each calendar year in our sample, and therefore our panel is imbalanced. From Equation (9), data from Panel (b) helps identify publication year fixed effects τ_t, while data from Panel (c) identifies event study coefficients $\theta_{r(i,t)}$.



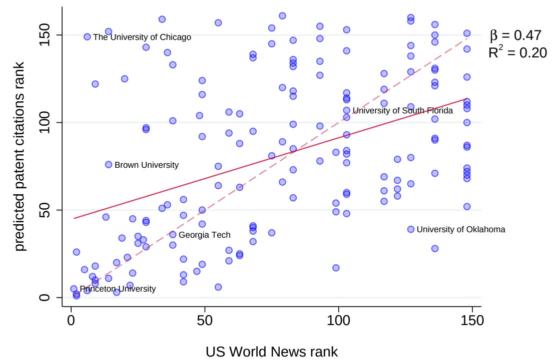
(a) Quintiles over time



(b) Sampled institution within quintile over time

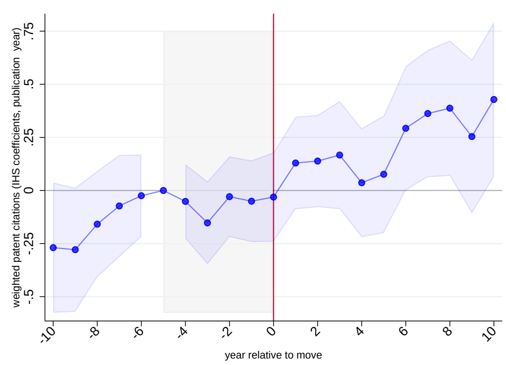
Appendix Figure 11: Stability in commercialization quintiles over time

Notes: Panel (a) plots quintiles of institution-level predicted patent citations over publication years from 2000 to 2020. These quintiles are calculated by estimating the distribution of the 2001 institution-level average count of predicted patent citations per publication written by authors at a given TTO. Then, mean counts of predicted patent citations are calculated over the author-years assigned to each quintile, by publication year. These quintile means are calculated using both movers and non-movers, totaling 513,918 authors. Panel (b) shows the average count of predicted citations for a single, randomly sampled TTO within each quintile over time. Only TTOs that appear in every year are eligible to be sampled. The quintile for which a TTO is sampled from is denoted by the pattern and color of its time series– for example, East Carolina University is in the lowest quintile and Georgetown University is in the highest. In each panel, a red vertical line denotes 2001 as the publication year used to construct these quintiles.



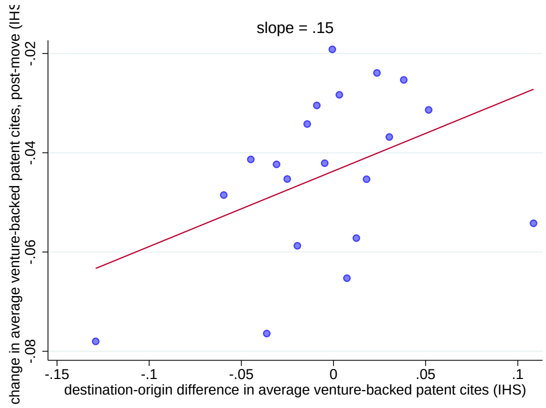
Appendix Figure 12: U.S. World News ranking vs. commercialization rank

Notes: This figure plots institution-level commercialization rank against U.S. World News' 2021 institution rank. Commercialization rank is calculated as the average count of predicted citations that researchers affiliated with a given school receive, across all publication years in our sample (from 2000 to 2020). A solid red line displays the slope coefficient from a regression of commercialization rank on U.S. World News rank; the value of the slope and R^2 are recorded in the top right of the figure. A dashed red line shows the 45^o line along these two axes. There are five schools labeled by name of the graph, each with various ranks. While our full sample spans researchers from 265 institutions, we are limited in scope by the institutions that have U.S. World News rankings. Because only colleges and universities are ranked by U.S. World News, we do not include any research institutions or hospitals (e.g., the Broad Institute), leaving a total of 161 institutions (blue dots).

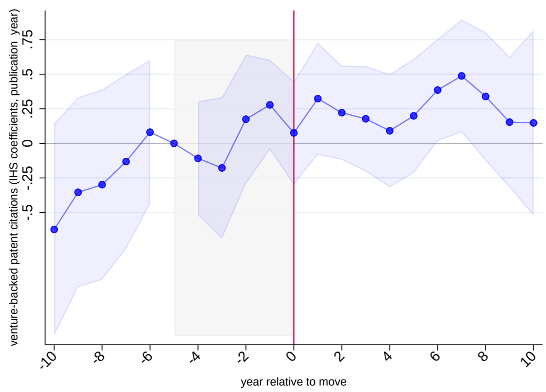


Appendix Figure 13: Event study: Weighted observed five-year patent citations

Notes: This figure is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9) where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the observed count of normalized citations received from patents filed within five years of publication for each of an author's papers in a given year. Patent citations are weighted by the cumulative count of patent citations they receive following award. In particular, and in accordance with (Lerner and Seru, 2022), we scale a cited patent's count by the average count of citing patents by WIPO technology class and award year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. The relative year -1 is omitted and normalized to zero. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution of the event study coefficients. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. There are 119,338 author years in our sample, coming from the sample of 14,213 movers. The red line denotes the researchers' move years.



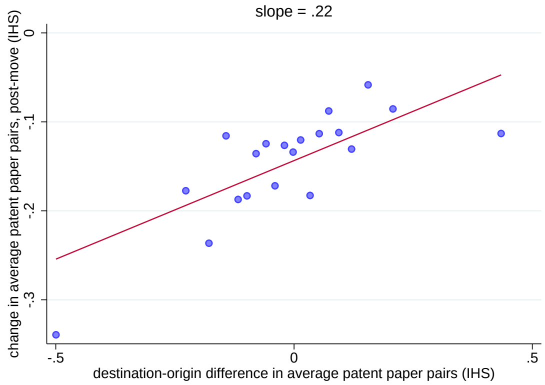
(a) Change in commercialization propensity by size of move



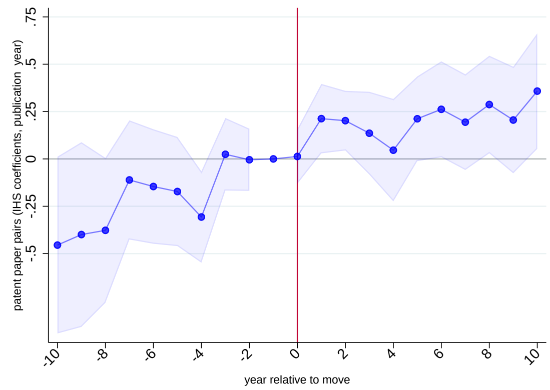
(b) Event study

Appendix Figure 14: Venture capital-backed patents

Notes: Panel (a) is a binned scatterplot that compares an individual's change upon moving in five-year venture-backed patent citations with the difference in average institution-level five-year venture-backed patent citations between their destination and origin institutions. For each mover, we compute two values. First, we separately calculate the average count of IHS five-year venture-backed citations for papers published pre- and post-move, and report the difference. We call this the change in an individual's commercialization level. Second, we generate $\hat{\delta}_i$ by taking the institution-level differences, as shown in Figure 2. We call this the change in the institution's commercialization level. The x-axis displays ventiles of this institution-level change, while the y-axis plots, for each ventile, the average change in individual commercialization. The slope of the line of best fit is reported above the graph. Panel (b) is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9), where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the observed count of citations received from patents filed within five years of publication for each of an author's papers in a given year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. Because levels of event time are collinear when estimated together, we preemptively omit data from relative year = -5, forcing the value of its event study coefficient to zero, which can be seen on the graph without a standard error estimate. We normalize this particular relative year because it is the last year, pre move, that all five-year citations to papers written at a researcher's origin also accrue while the researcher is at their origin institution. The contaminated region is shaded in gray and a red line denotes the researchers' move years. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution on each coefficient. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. Each panel is estimated using 11,466 movers and 81,766 mover-years.



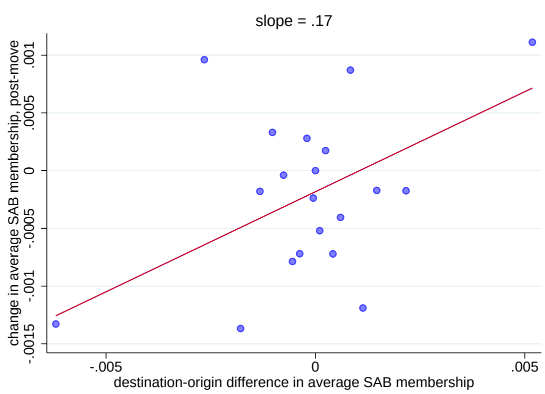
(a) Change in commercialization propensity by size of move



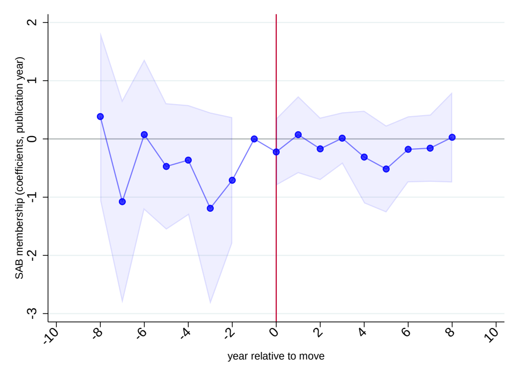
(b) Event study

Appendix Figure 15: Patent-paper pairs

Notes: Same as Appendix Figure 14. There are 119,338 author years in our sample, coming from the sample of 14,213 movers.



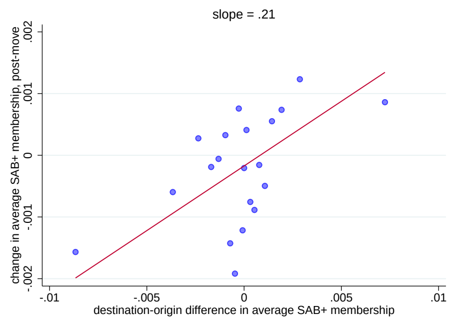
(a) Change in commercialization propensity by size of move



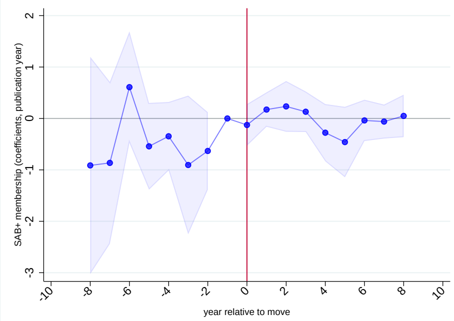
(b) Event study

Appendix Figure 16: Scientific advisory board membership

Notes: Same as Appendix Figure 14. We restrict the binscatter and event study to ± 8 years around the move, instead of ± 10, since the sparsity of the outcome leads to noisy estimates in relative years with few authors joining SABs. As a result, there are 111,042 author years in our sample, coming from the sample of 14,213 movers.



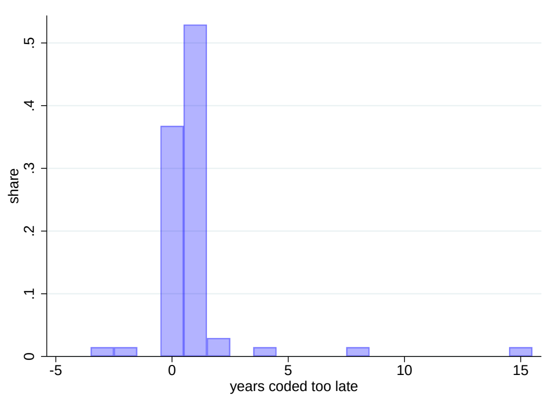
(a) Change in commercialization propensity by size of move



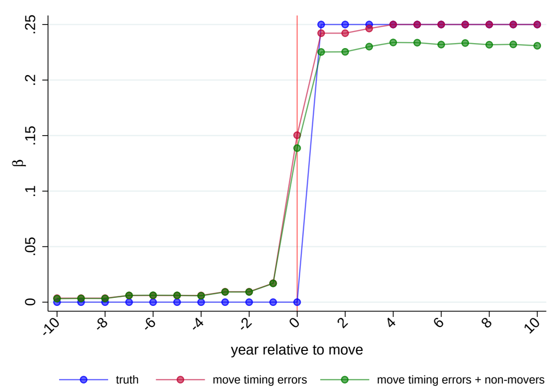
(b) Event study

Appendix Figure 17: Scientific advisory board membership+

Notes: Same as Appendix Figure 14. We restrict the binscatter and event study to $\pm$ 8 years around the move, instead of $\pm$ 10, since the sparsity of the outcome leads to noisy estimates in relative years with few authors joining SABs. As a result, there are 111,042 author years in our sample, coming from the sample of 14,213 movers.



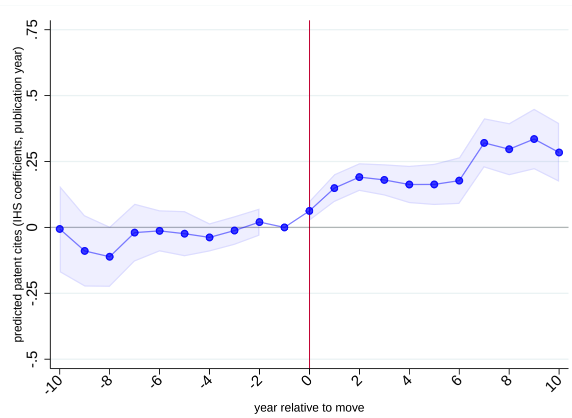
(a) Move timing errors



(b) Simulated event study

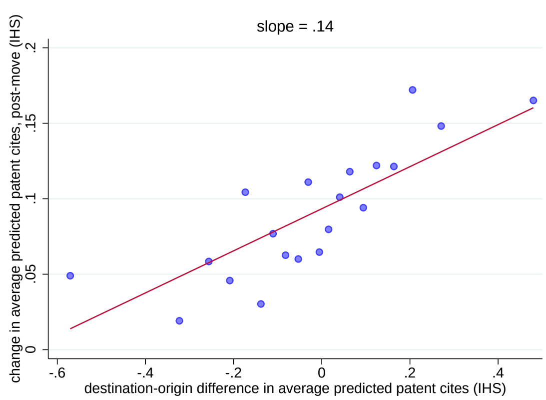
Appendix Figure 18: Simulated event study

Notes: For a sample of 100 randomly selected movers, Panel (a) shows the distribution of move years coded too late. Using faculty webpages, CV, and LinkedIn profiles, we are able to determine the actual move year for 68 movers in our sample. The remaining 32 are either untraceable using name and email information (N = 25), or we determine them to have not actually moved (N = 7). Panel (b) shows the results from a simulated event study. The event study uses a panel parametrized to have 14,213 movers observed in the same calendar years ten years pre- and post-move, a constant treatment effect of $\hat{\theta}_r = 0.25$ for all $r \geq 1$, and a distribution of δ that matches that shown by Figure 4. Because the treatment effect is constant across units and we do not add in a noise term, there are no standard errors. The "truth" coefficient path (blue) plots the estimates from an event study estimation using this panel as-is. The "move timing errors" path (red) shows the results from an estimation where the panel is simulated to have a distribution of move timing errors that matches Panel (a). The "move timing errors + non-movers" path (green) shows the results from a final estimation that incorporates the distribution of timing errors and also recodes 7% of the movers as non-movers. These recoded non-movers are not exposed to treatment, but remain coded as movers to match the mistake we make in our sample construction. A vertical red line denotes the move year.

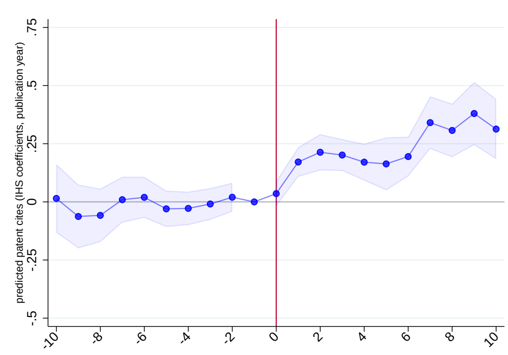


Appendix Figure 19: Event study: Predicted five-year patent citations with relative year fixed effects

Notes: This figure is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9) with the addition of uninteracted relative year fixed effects, where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the predicted count of citations received from patents filed within five years of publication for each of an author's papers in a given year. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. The relative year -1 is omitted and normalized to zero. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution of the event study coefficients. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above. There are 119,338 author years in our sample, coming from the sample of 14,213 movers. The red line denotes the researchers' move years.



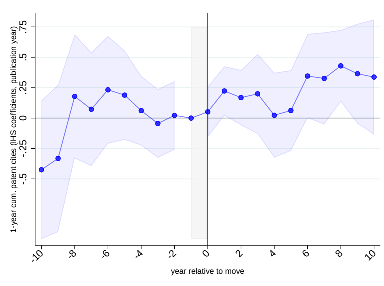
(a) Change in commercialization propensity by size of move



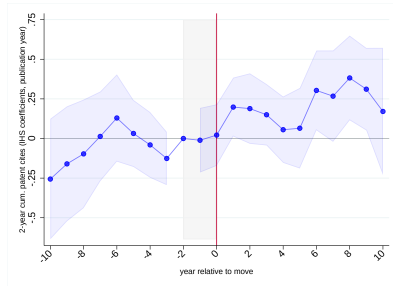
(b) Event study

Appendix Figure 20: Last authors

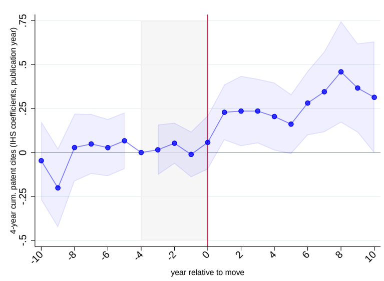
Notes: Same as Appendix Figure 14. There are 7,520 last author movers in our sample, which corresponds to a total of 78,897 author-years. There are 193,706 last author non-movers, which we define as ever being the last author on a publication. A red line denotes a researcher's move year.



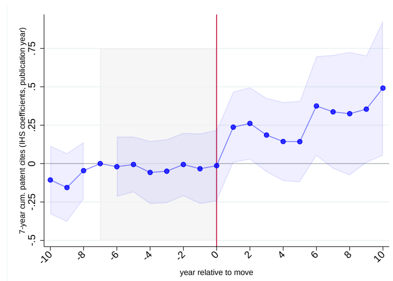
(a) 1-year cites



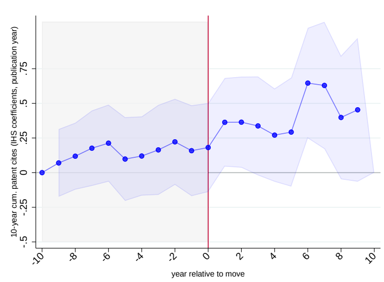
(b) 2-year cites



(c) 4-year cites



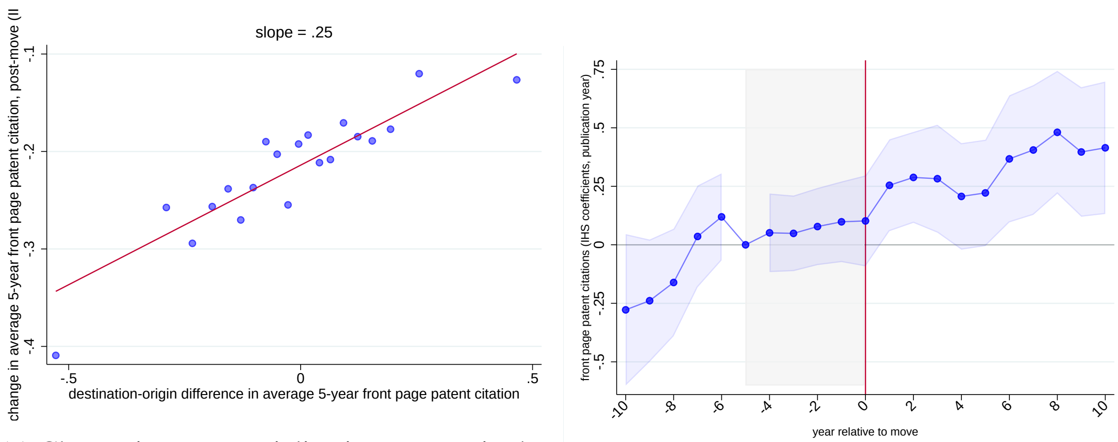
(d) 7-year cites



(e) 10-year cites

Appendix Figure 21: Citation windows

Notes: These event study plots show the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (9), where the dependent variable y_{it} is the inverse hyperbolic sine (IHS) of the observed count of citations received from patents filed within X years of publication for each of an author's papers in a given year. The value of X changes across the panel. Event time is plotted on the x-axis and the rate of convergence to a mover's destination mean commercialization is plotted on the y-axis. Because levels of event time are collinear when estimated together, we preemptively omit data from relative year $= -X$, forcing the value of its event study coefficient to zero, which can be seen on the graph without a standard error estimate. We normalize this particular relative year because it is the last year, pre move, that all lagged citations to papers written at a researcher's origin also accrue while the researcher is at their origin institution. Note, the number of movers, and thus observations, per event study decreases as X increases. This is because only publication years $2020 - X$ will have consistent measurements of y_{it}, due to truncation. For the ten-year lagged citation outcome, the sample gets so small that it cannot sufficiently identify the $\tilde{\theta}_{10}$ coefficient. Resultantly, we do not report it. For each, the contaminated region is shaded in gray and a red line denotes the researchers' move years. Standard error estimates are denoted by the shaded region around the plotted coefficients. Standard errors are computed by a bootstrapping procedure that randomly samples authors, with replacement, to construct estimates of $\hat{\delta}_i$ and $\tilde{\theta}_{r(i,t)}$. We perform 50 bootstrap iterations to form an empirical distribution of the event study coefficients. The standard deviation of this distribution, for each relative year, is the standard error used to calculate the 95% confidence interval plotted above.

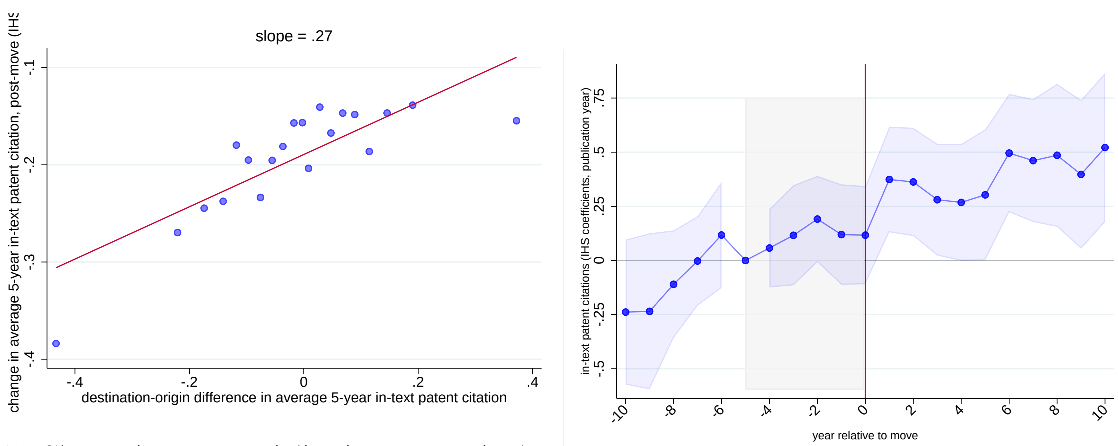


(a) Change in commercialization propensity by size of move

(b) Event study

Appendix Figure 22: Five-year front page patent citations

Same as Appendix Figure 14. Each panel is estimated using 11,466 movers and 81,766 mover-years.

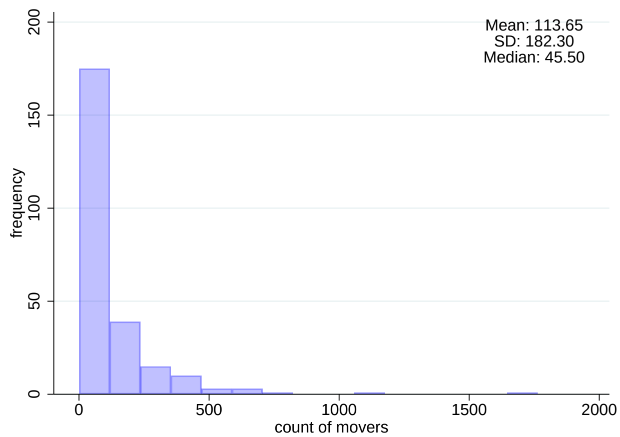


(a) Change in commercialization propensity by size of move

(b) Event study

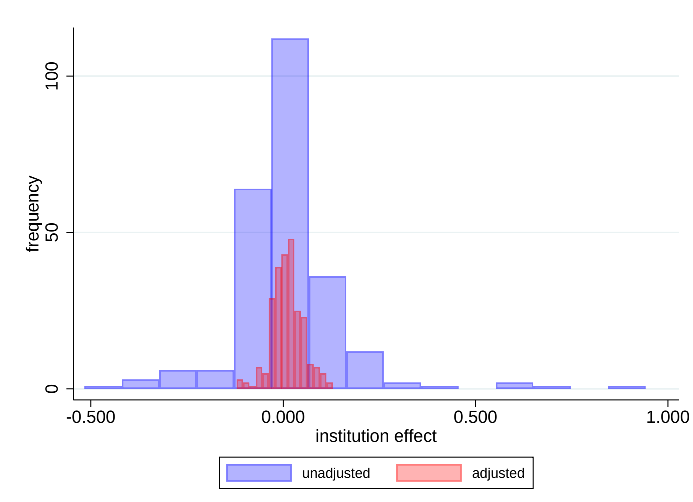
Appendix Figure 23: Five-year in-text patent citations

Same as Appendix Figure 14.. Each panel is estimated using 11,466 movers and 81,766 mover-years.



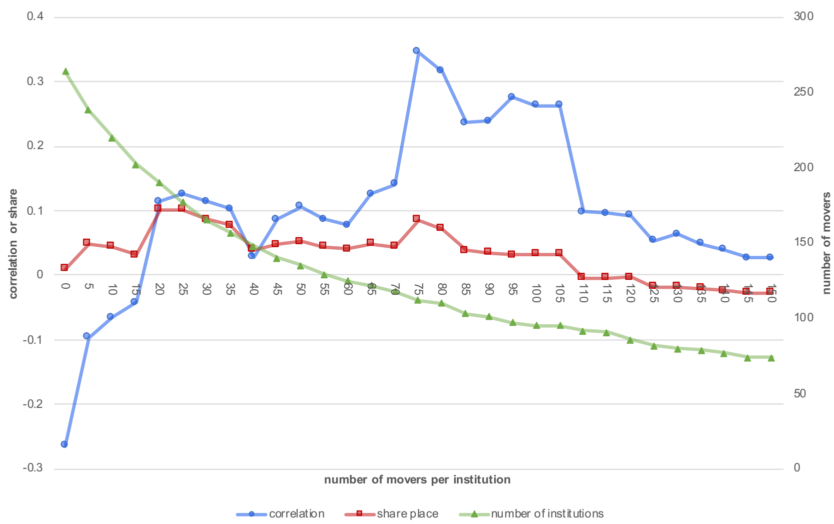
Appendix Figure 24: Movers by institution

Notes: This figure shows the distribution of the number of movers per institution, conditional on observing at least one move. Here, a mover is counted once for their origin and once for their destination. This definition is helpful to isolate the author-years that generate the variation necessary to identify the casual institution effects, as delineated in Equation (1). The count of movers is plotted on the x-axis. The number of schools with each mover count is plotted on the y-axis. The mean, median, and standard deviation are shown in the top right corner. Our sample of 14,213 movers is used; they span 248 institutions.

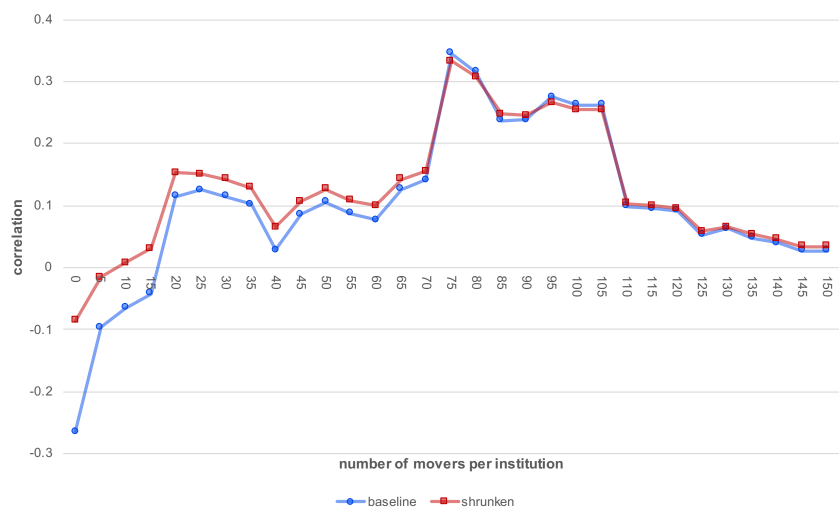


Appendix Figure 25: Empirical Bayes: Distribution of institution effects

Notes: This figure is based on estimation of Equation (1), where the dependent variable y_{iut} is the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at institution u. In particular, this figure displays the results of an empirical Bayesian shrinkage of the casual institution effects $(\hat{\gamma}_u)$. The x-axis is the estimated value of the institution effect for a given school, and the y-axis is the frequency of schools within an institution effect bin. The blue bars show the distribution of the institution effects estimated from Equation (1). The red bars show the distribution of the adjusted institution effects, following the Bayesian shrinkage $(\hat{\gamma}_u^{EB})$. To center the mean of this distribution at zero, we pre-omit Syracuse University when estimating Equation (1). While our full sample includes 265 institutions, only 248 have movers. This exercise uses our full sample of movers ($N = 14,213$).



(a) Point estimates at varying sample sizes



(b) Comparing empirical Bayes and sample restriction

Appendix Figure 26: Sensitivity of estimates to choice of variance reduction

Notes: This figure is based on estimation of Equation (1), where the dependent variable y_{iut} is the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at institution u. Particularly, it shows the sensitivity of select variance decomposition estimates with respect to the choice of variance reduction method. Panel (a) shows how the estimate of $Cov(\hat{\alpha}_u, \hat{\gamma}_u)$ (blue line) and $S_{university}$ (red line) vary when the sample is restricted to include only those with at least a certain number of movers (x-axis). The number of schools that satisfy this restriction is plotted by the green line, and is indexed by the secondary y-axis. Panel (b) borrows the blue line from Panel (a) and plots the estimate of $Cov(\hat{\alpha}_u, \hat{\gamma}_u)$ for each sample restriction, additionally implementing empirical Bayes (red line). This exercise uses our full sample of movers $(N = 14,213)$.

	(1) Non-movers mean	(2) Movers mean
Total publications	12.52	59.91
Total USPTO patent citations	7.38	38.55
Total publications cited by 1+ USPTO patents	0.62	2.65
Share of authors cited by 1+ USPTO patents	0.22	0.52
Share of publications cited by 1+ USPTO patents	0.06	0.05
USPTO citations per publication	0.37	0.43
Career length (years)	2.36	8.35
Outcome measures		
Average annual count of patent cites	0.70	1.59
Average annual count of predicted patent cites	0.41	0.88
Average annual count of papers paired with patents	2.38	5.94
Number of citations from VC-backed patents	0.05	0.12
Average annual SAB memberships	0.00	0.00
Observations	499,705	14,213

Appendix Table 1: Author-level summary statistics

Notes: This table tabulates the means for movers vs. non-movers (in columns) for several measures of research commercialization (in rows). For total publications, total USPTO patent citations, and total number of publications cited by patents, we take the sum across all years for each author. The share of publications cited by 1+ USPTO patents is the author's total number of cited papers divided by their total number of publications over all years. USPTO citations per publication is the number of citations received by all papers divided by the total number of papers over all years. Career length is calculated as the difference between the last and the first calendar years in which an author appears in the Web of Science publication data.

rank	*Institution*	*Average Commercialization*	*Adjustment Factor*	*Adjusted Average Commercialization*
Top 15				
1	Scripps Research	0.224	0.706	0.158
2	Northeastern University	0.206	0.660	0.136
3	Cedars Sinai	0.286	0.414	0.118
4	Rensselaer	0.196	0.584	0.114
5	University of Texas Medical	0.158	0.704	0.111
6	Rockefeller University	0.163	0.668	0.109
7	Blood Center of Wisconsin	0.587	0.180	0.106
8	Baylor College of Medicine	0.139	0.746	0.103
9	University of Memphis	0.168	0.609	0.102
10	Salk Institute	0.158	0.634	0.101
11	Children's Hospital Los Angeles	0.151	0.642	0.097
12	Lee Moffitt	0.269	0.360	0.097
13	Dana-Farber	0.139	0.687	0.095
14	Massachusetts General Hospital	0.125	0.749	0.093
15	New Jersey Institute of Technology	0.254	0.368	0.093
Middle 15				
117	Pennsylvania State University	0.016	1.497	0.024
118	Lehigh University	0.024	1.016	0.024
119	Arizona State University	0.023	1.016	0.024
120	Oklahoma State University	0.022	1.042	0.023
121	University of Utah	0.023	1.029	0.023
122	Kansas State University	0.022	1.042	0.023
123	University of Virginia	0.021	1.045	0.022
124	University of Dayton	0.009	2.515	0.022
125	University System of Maryland	0.021	1.049	0.022
126	San Diego State University	-0.065	-0.339	0.022
127	University of Idaho	0.018	1.183	0.021
128	Tulane University	0.019	1.113	0.021
129	Brigham & Women's	0.008	2.587	0.021
130	Tufts Medical Center	0.002	8.436	0.020
131	Harvard University	0.018	1.099	0.020
Bottom 15				
233	University of Toledo	-0.077	0.482	-0.037
234	SUNY	-0.065	0.600	-0.039
235	Miami University	-0.111	0.361	-0.040
236	Saint Louis University	-0.090	0.515	-0.046
237	Ohio University	-0.117	0.420	-0.049
238	Schepens Eye	-0.317	0.155	-0.049
239	Utah State University	-0.201	0.249	-0.050
240	University of Oklahoma Health Sciences	-0.105	0.510	-0.054
241	Iowa State University	-0.101	0.541	-0.055
242	Portland State University	-0.139	0.424	-0.059
243	Eastern Virginia Med	-0.291	0.283	-0.082
244	Wichita State University	-0.376	0.225	-0.084
245	California Institute of Tech	-0.159	0.558	-0.089
246	University of New Orleans	-0.280	0.330	-0.092
247	Whitehead Institute	-0.258	0.407	-0.105

Appendix Table 2: Causal institution effects

Notes: This table is based on estimation of Equation (1), where the dependent variable y_{iut} is the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at institution u. As described in Section 4.1.1, predictions and fixed effects are collapsed to the institution-year level, and then once more to the institution level. Institutions are sorted and ranked (column one) on the basis of their $\hat{\gamma}_u$, the causal institution effect. This value is reported in column three, for the top, middle, and bottom fifteen institutions. To minimize noise in the estimation of $\hat{\gamma}_u$ driven by few movers in a given u, we employ an empirical Bayesian shrinkage procedure. This procedure is discussed in Onine Appendix Section C. The adjusted effect, $\hat{\gamma}_u^{EB}$ is displayed in column five. Column four is the scalar by which the raw institution effect is multiplied by to produce the adjusted effect, or $\frac{\hat{\gamma}_u^{EB}}{\hat{\gamma}_u}$. While there are 265 institutions in our sample, only 248 have movers— institution effects are only estimable for those with movers. We choose Syracuse University as the omitted institution in this regression, leaving 247 institutions for which institution effects are produced. This exercise uses our full sample of movers ($N = 14,213$).