The Wandering Scholars: Understanding the Heterogeneity of University Commercialization*

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

April 27, 2025

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 about 15% to 30% 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, Scott Stern, and Reinhilde Veugelers (discussants), Isaac Kohlberg, Matt Marx, Chris Walters, and numerous seminar participants 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, hmanley@stanford.edu, 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 entailing the 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 fellows 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 its diffusion into the economy. Why might university-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 university-based characteristics in the commercialization of scientific 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, work such as Di Gregorio and Shane (2003) has correlated university-level commercialization activity with university- and area-level characteristics; see also Ho and Lee (2021), Siegel, Waldman, and Link (2003), Sine, Shane, and Gregorio (2003), and Stuart and Ding (2006).

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 then use observed changes in the e-mail domains authors choose to list on their published scientific papers to infer moves of

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

researchers across universities.[6]

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) 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 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 about 15% to 30% of geographic variation in commercial spillovers from university-based scientific research as measured by patent citations is attributable to university-specific factors.

[6] As we discuss in more detail in Section 2.4, some of the institutions in our are not necessarily universities, but rather research hospitals or university systems. However, the vast majority of these institutions are universities, and so we will typically refer to all of these institutions collectively as "universities."

The Bertozzi Lab / OliLux Biosciences example 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 chaired 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. We can use the estimates from these related papers in an effort to benchmark whether the place effects (or university effects, in our case) that we uncover here are comparatively large or small. Starting with Abowd, Kramarz, and Margolis (1999), much of the related work has focused on decomposing the variation in log wages into person-specific and firm-specific components. Our variance decomposition in Table 4 suggests that the variance in commercialization would fall by 88% if all researcher effects were made equal, and by 19% if all university effects were made equal. A study of West German workers from 1985 to 2009 (Card, Heining, and Kline, 2013) finds a larger role of firms: their results suggest that the variance in log wages would fall by about 80% if worker effects were equalized and by about 40% if firm effects were equalized. Song et al. (2018) find similar numbers for US workers from 1978 to 2013, estimating that the variance in log wages would fall by about 90% if worker effects were equalized and by about 50% if firm effects were equalized.[7]

Beyond wages, Finkelstein, Gentzkow, and Williams (2016) use a very similar mover design to investigate healthcare expenditures. The authors estimate that around 50% of the geographic variation in healthcare expenditure can be attributed to place-related factors, compared to the 15% to 30% that we estimate. Fewer papers have applied this methodology to innovation-related outcomes. However, there are a few notable exceptions. Bhaskarabhatla et al. (2021) studies the contributions of individuals versus firms

[7] We compute these numbers by using the variance decompositions in Table III in Card, Heining, and Kline (2013) and Table III in Song et al. (2018). We take the variance components and apply the formulas for $S^{var}_{researcher}$ and $S^{var}_{university}$ from Section 4.2. Due to covariance terms, these shares do not sum to one.

in patenting, and finds that only about 5% of the variance in cross-firm patenting can be explained by firm effects. Most closely related to this paper is contemporaneous work by Chandra and Xu (2024), which uses a mover design to decompose scientific productivity (measured using citation-weighted publications) into researcher and university effects. The authors find that 40% to 50% of the cross-university variation in productivity can be attributed to university-related factors. These results echo findings from a smaller-scale study of 179 mover-scientists by Allison and Long (1990). Overall, our estimated university effects appear to be comparatively small when compared to estimates of firm effects on wages. However, they appear to be somewhat in line with the range of estimates of place effects on innovation-related outcomes.[8]

It is important to note that because we define moves as moves across universities – which of course have fixed physical locations — our estimated university 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 university-specific factors are important. Higher-ranking schools, as well as schools with more active technology transfer offices, have larger estimated commercialization effects.

While our focus here is on how scientific papers spill over to commercial innovation, our work relates to spillovers from scientific research more generally. Such spillovers have been documented to operate in what Azoulay, Graff 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 (2024)), 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 con-

[8] They are also consistent with the results of Waldinger (2016), which finds that human capital matters more for scientific output than physical capital.

struct 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 cataloged in the Web of Science data from Clarivate Analytics, which compiles records of peer-reviewed scholarly journals, conference proceedings, and editorially selected books.[9] Our version of the Web of Science includes publications from 1900 to 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[10] 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 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

[9] 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 Graph, and Scopus. There are some differences across these datasets. For example, Microsoft Academic Graph (now known as OpenAlex) 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).

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

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.[11] As of 2022, Langer had 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. It is the most frequently observed commercialization-related measure in our sample, and is also in some sense the most general, as it does not focus on 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' published scientific research to be cited in patents. 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 dis-

[11] 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).

tinct 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). 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.[12] 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 cataloging 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 Marx and Fuegi data.[13] 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

[12] 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."

[13] 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.

of Science data, so we omit those papers from our analysis.[14]

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. Figure E.1 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. 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. However, because we must cut our sample off X years before 2020 in order to allow enough time for citations to accrue (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 $X = 5$ in our main specifications.

However, this observed citation measure poses a challenge in our mover design. To understand 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, it will include 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 during our sample period (2000 to 2020), 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 scientific papers published in *Proceedings of the National Academy of Sciences*. Although patent citations and paper citations are positively correlated, journals in the right tail of the patent citation distribution include both high impact factor journals (such as *Nature Nanotechnology;* impact factor = 38.1) and low impact factor journals (such as *Vaccine;* impact factor = 4.5). See Figure E.2 for the distribution of five-year patent citations by journal and examples of top patent cited journals.

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

[14] 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.

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.

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, which includes data from over 3,800 public and private companies. 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. Research assistants hired on UpWork manually 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. 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, though one exception is Stuart and Ding (2006), who analyze data on 727 scientific advisors.

2.3 Measuring university affiliations

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 Figure E.3 — 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 university. Figure E.3(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. Figure E.3(b) shows how this publication is listed in the Web of Science data, and Figure E.3(c) provides a screenshot of the associated SQL tables from which we extract these data.

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. In practice, this definition includes primarily university-based TTOs (89%, e.g., Stanford University), but also some university system-wide TTOs (2%, e.g., University of California System), some affiliated research hospitals reporting separately from universities (3%, e.g., Brigham and Women's Hospital, which reports separately from Harvard University), and some standalone research hospitals (6%, e.g., St. Jude Children's Research Hospital). For some of our descriptive analyses, we geocode TTO locations to the Bureau of Economic Analysis (BEA) economic area codes used in Moretti (2021).[15]

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`. Note that e-mail addresses are reported much more frequently for first and last 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

[15] 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).

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. For each unique e-mail domain, we can therefore 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 subdomains 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 which should instead be kept separate because the hospital has its own independent TTO in the AUTM survey data.

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."[16]

2.4 Building the panel and inferring moves

Once we have organizational affiliations for each author-year, we can begin to build our panel.

Selecting author-years for inclusion. Table 1 walks through our sample construction. 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 contributor-publications if the contributor is not coded as an author (but rather as an editor, 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.[17] This process leaves us with about 7.4 million author-publications.

We then aggregate these data to the author-year level, taking the most common university within an author-year and randomly breaking ties where necessary. We further restrict attention to "biomedical authors" whom we define as authors with 50% or more of

[16] 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.

[17] 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 Figure E.6.

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	1,720,191
Author-years after including only non-movers and movers ± 10 years around move	1,709,137

[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 determine researchers that are eligible to be in our sample.

their publications 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. This leaves us with approximately 1.7 million author-year observations.

Measuring moves. With unique author-year affiliations in hand, we define movers as follows: movers are researchers that we observe at two 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. Using our email-based panel, we are able to identify 12,196 movers. However, as detailed in Figure A.1, we also use mailing address data provided in the Web of Science data to fill in affiliation gaps where possible in order to add some additional movers to our sample. We are able to increase the 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

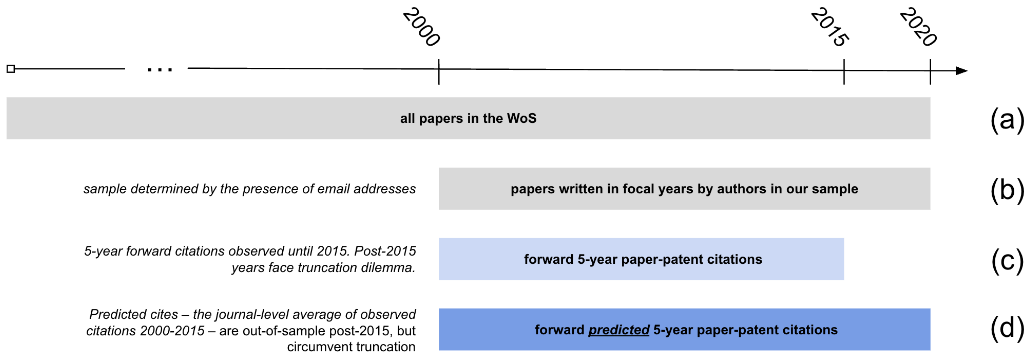


Figure 1: Timeline of coverage for key outcome variables

Notes: This figure shows the coverage over time of our key outcome variables: actual five-year patent citations and predicted five-year patent citations. We start with the universe of papers indexed by Web of Science (a). Due to the availability of email addresses, we then focus in on papers written from 2000 to 2020 (b). Actual five-year patent citations for papers written in (b) can be computed for the subset of years 2000 to 2015 (c). Predicted five-year patent citations can be computed for all years from 2000 to 2020 (d).

— 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 Figure E.4, around 25% of movers list more than one affiliation in the year they move. Figure E.5 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.

Final panel. As shown in Table 1, our final panel contains a total of 1,709,137 author-years, comprised of 14,195 movers and 499,784 non-movers. Figure 1 helps visualize the calendar years we have coverage for. While the Web of Science goes back to 1900, we restrict our attention to papers written in the 2000-2020 time frame, because very few email addresses were reported prior to that time (see Figure E.6). When using five-year forward citations as our outcome, we can only use papers published until 2015 (so that they have enough time to accrue citations). On the other hand, if we use predicted citations, we use papers written all the way until the end of our sample.

3 Descriptive statistics

3.1 Descriptive statistics: Movers and their moves

Figure E.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. 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.[18] This share remains constant throughout the post period, suggesting we capture real, long-term moves by researchers.

Figure 2 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 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 manually 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 seven 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 2 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 (23

[18] Note that the figure restricts attention to publications where the researcher's primary affiliation is either the destination or the origin university.

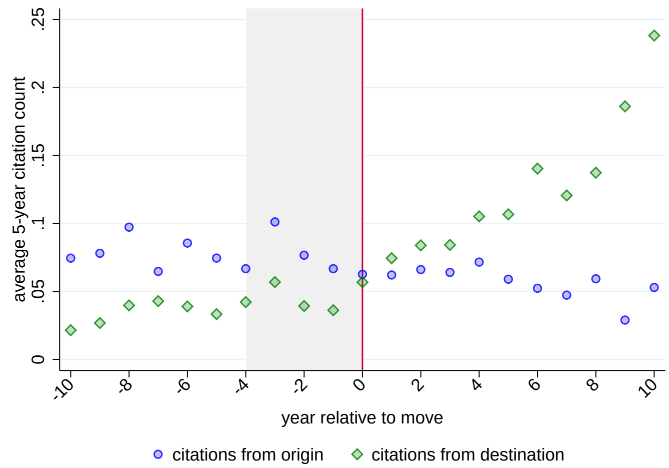


Figure 2: Citations from origin and destination by relative year

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." The event study is estimated using 11,441 movers and 81,725 mover-years: this sample is defined by dropping the years 2016-2020 from our main estimation sample with 14,195 movers and 119,214 mover-years.

versus 17 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 (60 versus 13) and more patent citations to papers (39 versus 7). 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 (0.48 vs. 0.38).

3.2 Descriptive statistics: Commercialization

How does commercialization vary across universities in our sample? Figure 3 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 university 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 3(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.

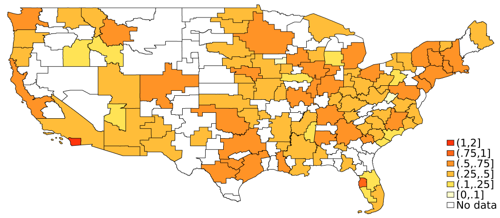
	(1) Non-movers mean	(2) Movers mean
Author metrics		
Total publications	12.51	60.16
Career length	17.05	22.91
Team size per publication	12.55	7.25
USPTO metrics		
Total USPTO patent citations	7.36	39.38
Total publications cited by 1+ USPTO patents	0.62	2.67
Share of authors cited by 1+ USPTO patents	0.22	0.52
Share of publications cited by 1+ USPTO patents	0.06	0.05
USPTO citations per publication	0.38	0.48
Outcome measures		
Annual 5-year patent cites	0.70	1.62
Annual 5-year predicted patent cites	0.47	1.01
Annual 5-year paper cites	21.98	60.98
Annual 5-year predicted paper cites	15.54	35.72
Annual patent-paper pairs	0.74	2.02
Annual 5-year VC patent cites	0.05	0.12
Annual SAB memberships	0.00	0.00
Observations	499,784	14,195

Table 2: Author-level summary statistics

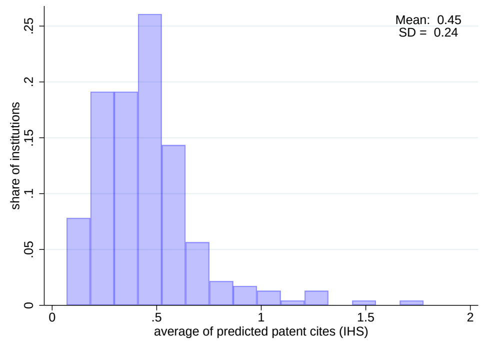
Notes: This table tabulates the means for movers vs. non-movers (in columns) for several measures of research productivity and commercialization (in rows). For total publications, total USPTO patent citations, and total number of publications cited by patents, we take the sum across all years for each 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.

Figure 3(b) plots these university-level measures as a histogram.[19] Figure 3(c) lists the top ten institutions by commercialization 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.

[19] We document in Figure E.8(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 representative school from each quintile in Figure E.8(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.



(a) Map of commercialization propensities



(b) Commercialization propensities across institutions

rank	*Institution*
1	Whitehead Inst
2	Dana-Farber
3	Scripps Research
4	Cedars Sinai
5	Cold Spring Harbor
6	Salk Institute
7	Rockefeller University
8	California Institute of Tech
9	Massachusetts General Hospital
10	Fox Chase

(c) Top institutions by commercialization propensity

rank	*University*
8	California Institute of Tech
11	Massachusetts Institute of Technology
14	University of Texas Southwestern
16	Harvard University
17	Rice University
18	WUSTL
19	University of Massachusetts
21	Stanford University
22	Rensselaer
23	Princeton University

(d) Top universities by commercialization propensity

Figure 3: 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 university level instead of economic area. There are a total of 230 institutions included in this figure. These institutions have at least 100 author-years affiliated with them. 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., California Institute of Technology is the eighth-ranked institution and the first-ranked university.

3.3 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 4 begins by describing the moves that our movers make. 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:

$$\hat{\delta}_i = \bar{y}_{d(i)} - \bar{y}_{o(i)} \tag{1}$$

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). We see that these moves are fairly centered around zero: movers go both up and down in the commercialization distribution.[20]

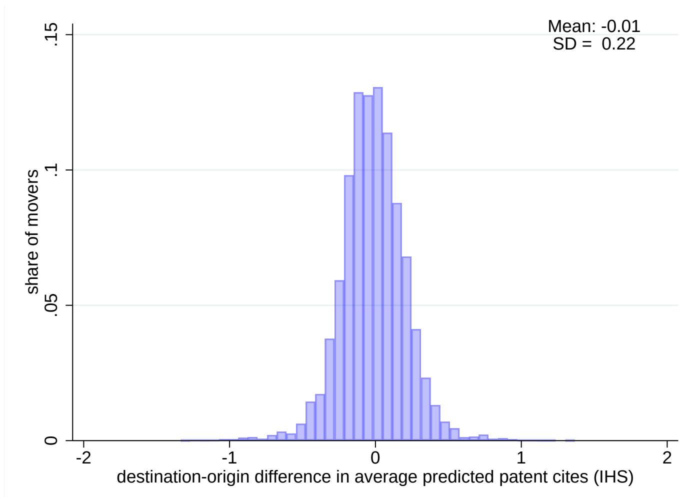


Figure 4: Distribution of destination-origin differences in commercialization propensities

Notes: This histogram shows the distribution of destination-origin differences in commercialization propensity ($\hat{\delta}_i$) across movers. For each mover in our sample (N = 14,195), we define the difference between their destination and origin university commercialization 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.

Figure 5 plots the change in mover i's commercialization activity before and after the move against these $\hat{\delta}_i$'s. 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

[20] It is worth keeping in mind that commercialization propensity is not collinear with university prestige. Figure E.9 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 0.51 and an R^2 of 0.24.

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 0.15, suggesting that roughly one-sixth of the heterogeneity in commercial activity by faculty is due to the place-specific factors.

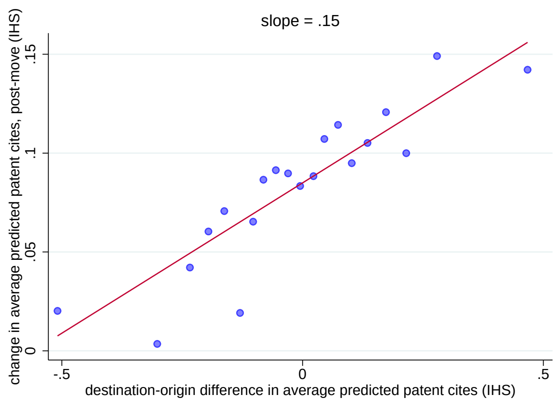


Figure 5: Change in commercialization propensity by size of move

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 university-level patent citations between their destination and origin universities. 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 university-level differences, as shown in Figure 3. We call this the change in the university's commercialization level. The x-axis displays ventiles of this university-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 13,954 movers in this graph: this sample comes from dropping 241 movers from our sample of 14,195 for whom we cannot calculate the difference between pre- and post-move commercialization.

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:

$$y_{iut} = \alpha_i + \gamma_u + \tau_t + \varepsilon_{iut} \tag{2}$$

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:

$$\bar{y}_u - \bar{y}_{u'} = (\gamma_u - \gamma_{u'}) + (\bar{\alpha}_u - \bar{\alpha}_{u'}). \tag{3}$$

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

$$S_{university}(u,u') = \frac{\gamma_u - \gamma_{u'}}{\bar{y}_u - \bar{y}_{u'}} \quad \text{and} \quad S_{researcher}(u,u') = \frac{\bar{\alpha}_u - \bar{\alpha}_{u'}}{\bar{y}_u - \bar{y}_{u'}}, \tag{4}$$

respectively.

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 (2) 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.

Two 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 universities to high commercialization universities systematically experience an increase in α_i at the same time as their move, we will overestimate the university 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 evidence 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. A corollary to this assumption is that the γ_u terms are the same for both movers and non-movers.

We perform two sets of tests in an effort to validate this assumption. First, in the spirit of Card, Heining, and Kline (2013), we compare the commercialization trajectories of researchers who make similar but opposite moves. We find approximately equal and opposite changes in commercialization for researchers making opposite moves, consistent with the additively separable model. In addition, following Bonhomme, Lamadon, and Manresa (2019), we directly estimate a model that allows for researcher-university complementarity. We find that the results of this analysis are very similar to the results from the additively separable estimation. More details on both of these validation exercises can be found in Appendix Section B.

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 $y_{iut} = \gamma_u + \varepsilon_{iut}$.[21] 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 want to scale y by the size and direction of the move, using the $\hat{\delta}$'s that are plotted in Figure 4.

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:

$$y_{it}^{scaled} = \frac{y_{it} - \bar{y}_{o(i)}}{\delta_i}. \tag{5}$$

This expression implies that y_{it}^{scaled} will equal zero if the mover's commercialization ex-

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

actly 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

$$y_{it}^{scaled} = \theta_{r(i,t)} + \varepsilon_{it} \tag{6}$$

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

$$y_{it} = \bar{y}_{o(i)} + \theta_{r(i,t)}\delta_i + \varepsilon_{it}. \tag{7}$$

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:

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i + \tau_t + \varepsilon_{it} \tag{8}$$

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*.[22]

[22] 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.

	Above/Below Median	*Top & Bottom 25%*	*Top & Bottom 10%*	*Top & Bottom 5%*
Difference in IHS commercialization				
Overall	.37	.46	.69	.89
Researchers	.31	.36	.57	.83
Universities	.06	.10	.12	.06
Share of difference attributable to				
Researchers	.84	.78	.83	.93
Universities	.16	.22	.17	.07
	($\pm$.07)	($\pm$.08)	($\pm$.11)	($\pm$.10)

Table 3: Additive decomposition of IHS commercialization

Notes: This table is based on estimation of Equation (2), where the dependent variable y_{iut} is the IHS of the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at university u. Each column defines a set of universities R and R' based on commercialization rank $\hat{y}_u$. We rank universities 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 universities. Standard errors are calculated using 50 repetitions of a resampled author-year panel; this resampling generates variation in both university rank and levels. The sample is movers and nonmovers (N = 1,720,191 author-years).

4.2 Additive decomposition estimates

Table 3 is based on our additive decomposition framework. The γ_u's are estimated from a regression following Equation (2) using our sample of movers and non-movers, where y_{iut} is the IHS of 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 3 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 3 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 estimates of the shares.

Columns in this table represent comparisons across different groups. The sizes of these comparison groups decrease across columns, from the top and bottom 50% up to the top/bottom 5%. Thus, column (1) uses data from, and makes comparisons across, universities ranked in the top and bottom 50% of $\bar{y}_u$. The fourth and fifth rows of column (1) suggest that 84% of the variation in commercialization between these two groups of universities is due to differing researchers, while the remaining 16% is due to university effects. This division is fairly consistent across different comparison groups, with the university share ranging from 7% to 22%.

Table 4 presents an alternative decomposition, investigating what share of the cross-university *variance* in commercialization is due to researchers versus place effects. The share of the cross-university variance that would be eliminated if all university effects were the same can be written as:

$$S^{var}_{university} = 1 - \frac{\mathrm{Var}(\bar{\alpha}_u)}{\mathrm{Var}(\bar{y}_u)}. \tag{9}$$

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

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

Plugging in the empirical analogs of $\bar{y}_u$, $\bar{\alpha}_u$, and γ_u allows us to construct the estimates in Table 4. 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.[23] We also document a weak positive correlation between researcher and university effects, with researchers who commercialize more sorting to universities that we estimate to have positive commercialization effects.[24]

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.

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

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

Cross-institution variance of mean:	
IHS commercialization	.043
University effects	.005
Researcher effects	.035
Correlation of average researcher and university effects	.106
	(± .107)
Share variance would be reduced if:	
University effects were made equal	.186
	(± .069)
Researcher effects were made equal	.881
	(± .029)

Table 4: Variance decomposition of IHS commercialization

This table is based on estimation of Equation (2), where the dependent variable y_{iut} is the IHS of the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at university u. The results from a variance decomposition of y_{iut} are shown. This method is discussed in Appendix Section B.2. The first row reports the variance in $\hat{y}_u$, the second row reports the variance in $\hat{\gamma}_u$ (university effect), and the third row reports the variance in $\hat{\alpha}_u$ (average researcher effect). The correlation coefficient on $\hat{\gamma}_u$ and $\hat{\alpha}_i(u)$ is given in 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 causal researcher or university effects were equalized across institutions. The first row in the second half reports the share of the variance that would be reduced if university 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. 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 ($N = 164$ universities). This leaves us with a sample of 1,663,147author-years. This choice is discussed further in Appendix Section D.

4.3 Event study estimates

Figure 6 documents our main event study results estimated using our sample of movers. We plot the relative time coefficients (the $\theta_{r(i,t)}$ terms) from Equation (8). Similar to the results presented thus far, 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. 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.

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 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 0.2 to 0.3 range.

The dynamics of the event study provide some clues about the potential mechanisms. If the effects were mostly about a startup package of resources offered to the academics at commercialization-intensive schools – for example, subsidized space or human resources for a fledgling venture – we might expect the effect to fade over time after the move. On the other hand, if the effects were more about culture, peers, and exposure to nearby investors, we might expect them to increase with exposure. The fact that the effects grow over time is consistent with the latter exploration.

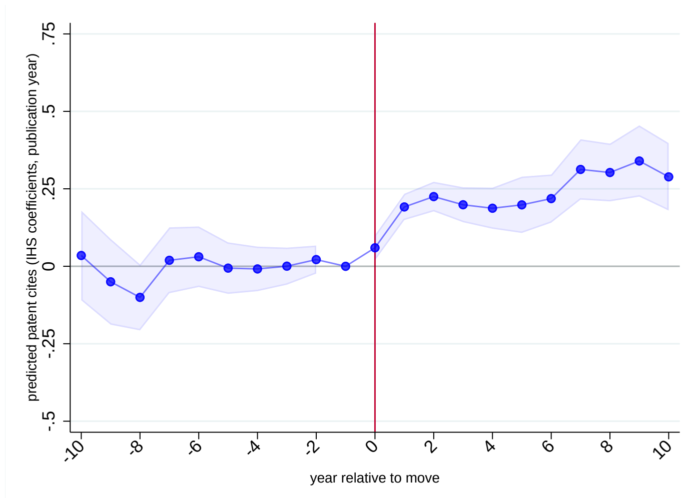
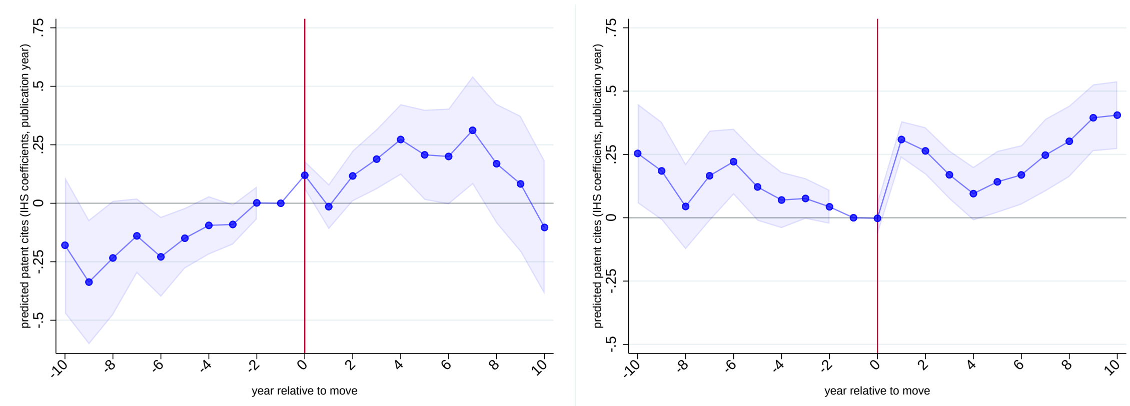


Figure 6: Event study: Predicted five-year patent citations

Notes: This figure is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (8),where the dependent variable y_{it} is the IHS of the predicted count of five-year patent citations to author i's papers in year t. 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. The 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. There are N = 119,214 author years in our sample, coming from the sample of 14,195 movers.

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 7, 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.



(a) Moves up

(b) Moves down

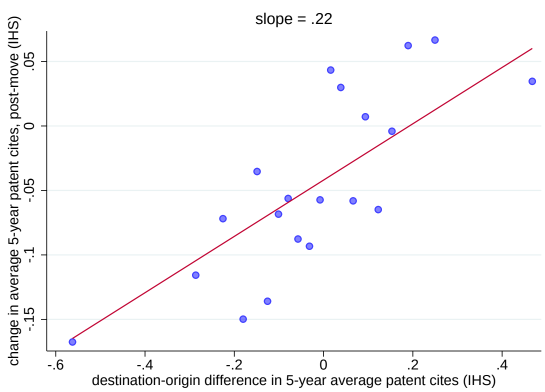
Figure 7: Comparing moves up and moves down

Notes: Each panel is an event study plot that shows the coefficients $\tilde{\theta}_{r(i,t)}$ estimated from Equation (8) and analogous to Figure 6. Panel (a) is estimated using authors who move up (those for whom $\hat{\delta}_i > 0$; N = 6,605). Panel (b) is estimated using authors who move down (those for whom $\hat{\delta}_i < 0$; N = 7,590).

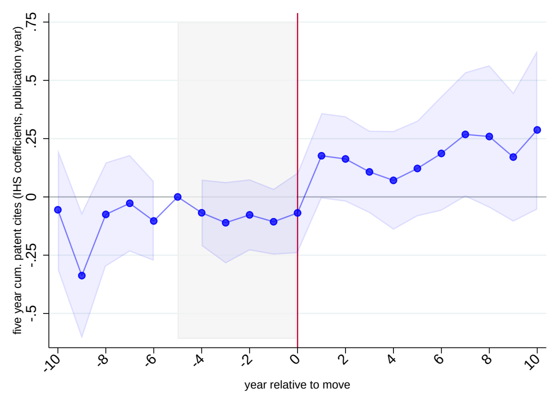
4.4 Additional outcomes

Actual five-year patent citations. We then repeat this 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.

Despite this challenge, Figure 8 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 0.22, which suggests that about one-fourth of heterogeneity in commercialization is due to place-specific factors.



(a) Change in commercialization propensity by size of move



(b) Event study

Figure 8: Observed five-year patent citations

Notes: Panel (a) is analogous to Figure 5, but uses actual IHS five-year patent cites rather than predicted values. Similarly, panel (b) is analogous to Figure 6, but uses actual IHS five-year patent cites rather than predicted values. In the event study, we omit data from relative year = -5. 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. These graphs are based on 11,441 movers and 81,725 mover-years: this sample is defined by dropping the years 2016-2020 from our main estimation sample with 14,195 movers and 119,214 mover-years.

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 15% to 25% 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 that 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 patent classifications (a simplified version of the International Patent Classification scheme with 35 subgroups) are downloaded from the USPTO PatentsView database and span the period from 2000 to 2023. Patents with values above one have received above average citations compared to other patents in their same vintage and technology class. Figure E.10 presents replicates Figure 8 using these weighted patent citations. The results, though noisy, are qualitatively similar.

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 0.19 in Figure E.11(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 and thus, the outcome is a 0/1 indicator for joining in a given year. Figure E.12 shows the results. Given the sparsity of the outcome (we match only about 2,000 board join events to the approximately 1.7 million 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.

Motivation for moves. In our analysis, we are assuming moves are orthogonal to considerations such as how well a university might assist in the commercialization of a scientists' work. We sought to investigate this more systematically by directly asking researchers why they move. More specifically, we surveyed 3,000 movers via email from our sample who had moved in the latter five years of our sample period (2016-2020). 311 respondents completed the survey, consistent with response rates that other recent papers surveying scientists have received (Carson, Zivin, and Shrader, 2023; Hill and Stein, 2025). We asked researchers about the most important reasons for their move, asking them to rank at least one choice, and no more than three choices, of a randomly ordered list of twelve choices. As shown in Figure E.13, better salary and better resources for research were the top two reasons selected for moving, while better commercialization was the *least* likely reason that respondents cited for moving. These results are consistent with the narratives we heard when we interviewed Harvard University officials responsible for coordinating faculty moves in the life sciences. Overall, this gives us additional confidence that moves are happening for reasons that are largely uncorrelated with a school's commercialization prowess.

Pre-trends. We implement the sensitivity tests developed by Rambachan and Roth (2023) to probe whether our results are robust to different pre-trends that are possible given pre-period coefficients that we estimate. We focus on the robustness of the $r = 1$ relative year coefficient. We find that coefficient remains statistically significant for most of the pre-trends that are possible to construct with the data that we observe. See Figure E.14.

Measurement error in coding moves. Our coding of researcher moves likely suffers from some degree of measurement error. 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%). Figure E.15(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 Figure E.15(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 (8) is to estimate:

$$y_{it} = \tilde{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i + \rho_{r(i,t)} + \tau_t + \varepsilon_{it} \tag{11}$$

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 Figure E.16.

Focusing on last authors. Next, we explore variation in seniority at the time of moving. 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 postdoc 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 universities, and leaves us with 7,552 movers. The results are shown in Figure E.17, 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.

Moves across different institution types. While we refer to TTOs as "universities" throughout the paper, a small minority of these institutions are not actually universities. AUTM classifies around 10% of our TTOs as either university-affiliated research hospitals or standalone academic medical centers. To check whether moves across different institution types are driving our results, we split our sample of movers into four mutually exclusive types of moves: university to university, university to non-university, etc. We then replicate our event study analysis. The vast majority of moves (90%) are university to university, and the results (shown in panel (a) of Figure E.18) are virtually identical to our baseline results. The remaining move types (shown in panels (b) through (d)) have too few movers to draw significant conclusions.

Alternative citation windows. In Figure E.19, we replicate our event study for actual citations, looking at citation windows ranging 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 results using five-year front-page citations and five-year in-text citations. The results are shown in Figure E.20 and Figure E.21, and look very similar to the results in Figure 8(b).

4.6 Correlates of university effects

Thus far, we have documented that universities matter for commercialization: across multiple analyses and subgroups, university effects explain 15% to 30% 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 three broad groups. First, we have university-level features. The majority of these are sourced directly from the AUTM data, which describe characteristics of the university's TTO. These variables include, among other things, the number of licenses issued, the number of active licenses, and legal fee spending. Each TTO-level feature is calculated as an average of the 2000-2020 sample period, and then transformed into a z-score. We also include measurs of univerisity prestige, such as its *U.S. News and World Report* ranking and its count of Nobel laureates.

The second group of covariates are 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, 2020, and the total across the time period. The final group of covariates are local labor market features. We assemble data on local labor market characteristics, such as employment, population, and educational attainment. These come from a variety of sources, including the Decennial Census, Current Population Survey, and American Community Survey.

For each feature, we begin by estimating a bivariate OLS regression, where the dependent variable is our estimated, standardized university fixed effect, and the independent variable is a single, standardized university characteristic. Because our university effects are estimated with noise, we employ an Empirical Bayes adjustment (Morris, 1983) prior to standardizing to shrink the estimates toward the mean (see Appendix D for details).

The results of these bivariate regressions are on the left-hand side of Figure 9. We then estimate a second regression, where we regress our fixed effects on *all* of the university characteristics, and let LASSO select the relevant covariates. 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 9.

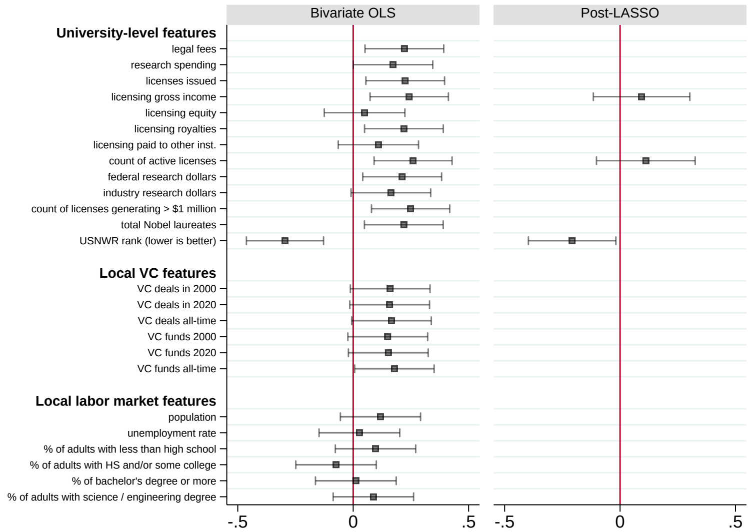


Figure 9: Correlates of average university effects

Notes: This figure shows results from regressing the estimated university effects – per Equation (2) – on various place-based factors. To minimize noise in the university effect estimates, we first employ the Empirical Bayes method described in Appendix Section D. All factors are standardized to have a mean of zero and standard deviation of one. In addition, the university effects are *also* standardized to have a mean of zero and a standard deviation of one. The left panel shows the coefficients from bivariate regressions of the standardized university effects against each factor, individually. The right panel shows the coefficients from a multivariate regression, following a LASSO selection procedure. Specifically, a ten-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 133 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 (2000-2020). Nobel laureates are all laureates at each TTO over the 2000-2020 time period. USNWR rank is the 2025 *U.S. News and World Report* ranking. 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. Population counts are obtained from the 2010 Census tabulations. Unemployment figures are obtained from the USDA Economic Research Service. Educational attainment figures are obtained from the 2012 five-year American Community Survey sample.

While many of the variables are statistically significant individually, only three variables are selected by the LASSO procedure. A lower (better) *U.S. News and World Report* ranking is associated with a higher commercialization effect, suggesting that overall university prestige is an important channel through which commercialization operates. If prestige is correlated with resources (research money, student quality, etc.), then it makes sense that this "absolute advantage" in research productivity is correlated with more research commercialization. The estimates in Figure 9 suggest that a one standard deviation improvement in ranking is correlated with a 0.2 standard deviation increase in the commercialization effect. In addition, total licensing income and count of active licenses are both (weakly) positively correlated with commercialization, suggesting that TTO experience may also help.

We emphasize that these results are only suggestive. We are simply showing correlations, and cannot make any causal claims. Moreover, we suffer from a small sample size (131 observations, after we restrict the sample to TTOs for which we have all of the characteristics), making us cautious to over-interpret these results.

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 universities. Our baseline estimates suggest that between 15% to 30% 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 Graff 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.[25] Understanding the extent to which the benefits of moves to more commercially oriented schools are anticipated by academics seems ripe for further exploration.

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

[25] An early paper found that the presence of children in a household can constrain scientific mobility, particularly among women (Shauman and Xie, 1996). Building on these insights, Azoulay, Ganguli, and Graff 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).

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 savviness 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.

Allison, Paul D. and J. Scott Long. 1990. "Departmental Effects on Scientific Productivity." *Americal Sociological Review* 55 (4): 469–478.

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. 2024. "Higher Education and Local Educational Attainment: Evidence from the Establishment of U.S. Colleges." *Review of Economics and Statistics* 106 (4): 1146–1156.

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.

Bhaskarabhatla, Ajay, Luis Cabral, Deepak Hegde, and Thomas Peeters. 2021. "Are Inventors or Firms the Engines of Innovation?" *Management Science* 67 (6): 3899–3920.

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.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica* 87 (3): 699–739.

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): 103946.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128 (3): 967–1015.

Carson, Richard T., Joshua S. Graff Zivin, and Jeffrey G. Shrader. 2023. "Choose Your Moments: Peer Review and Scientific Risk Taking." 31409. National Bureau of Economic Research.

Chandra, Amitabh and Connie Xu. 2024. "Productivity Differences in Fundamental Life-Sciences Discovery." Unpublished Working Paper. Harvard University.

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.

Hill, Ryan and Carolyn Stein. 2025. "Race to the Bottom: Competition and Quality in Science." *Quarterly Journal of Economics*: forthcoming.

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): eaam6310.

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.

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.

Marx, Matt and Aaron Fuegi. 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.

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.

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.

Rambachan, Anesh and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *Review of Economic Studies* 90 (5): 2555–2591.

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." Unpublished 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.

Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter. 2018. "Firming Up Inequality." *Quarterly Journal of Economics* 134 (1): 1–50.

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.

— 2016. "Bombs, Brains, and Science: The Role of Human and Physical Capital for the Creation of Scientific Knowledge." *Review of Economics and Statistics* 133 (2): 927–991.

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.

Supplement to *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 A.1. First, we code 12,196 movers using an author-year panel of corresponding e-mail address stems. Described in Section 2.4, 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 gap 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 2.4) 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 2.4. 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 Figure E.6 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 resided 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 A.1 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.

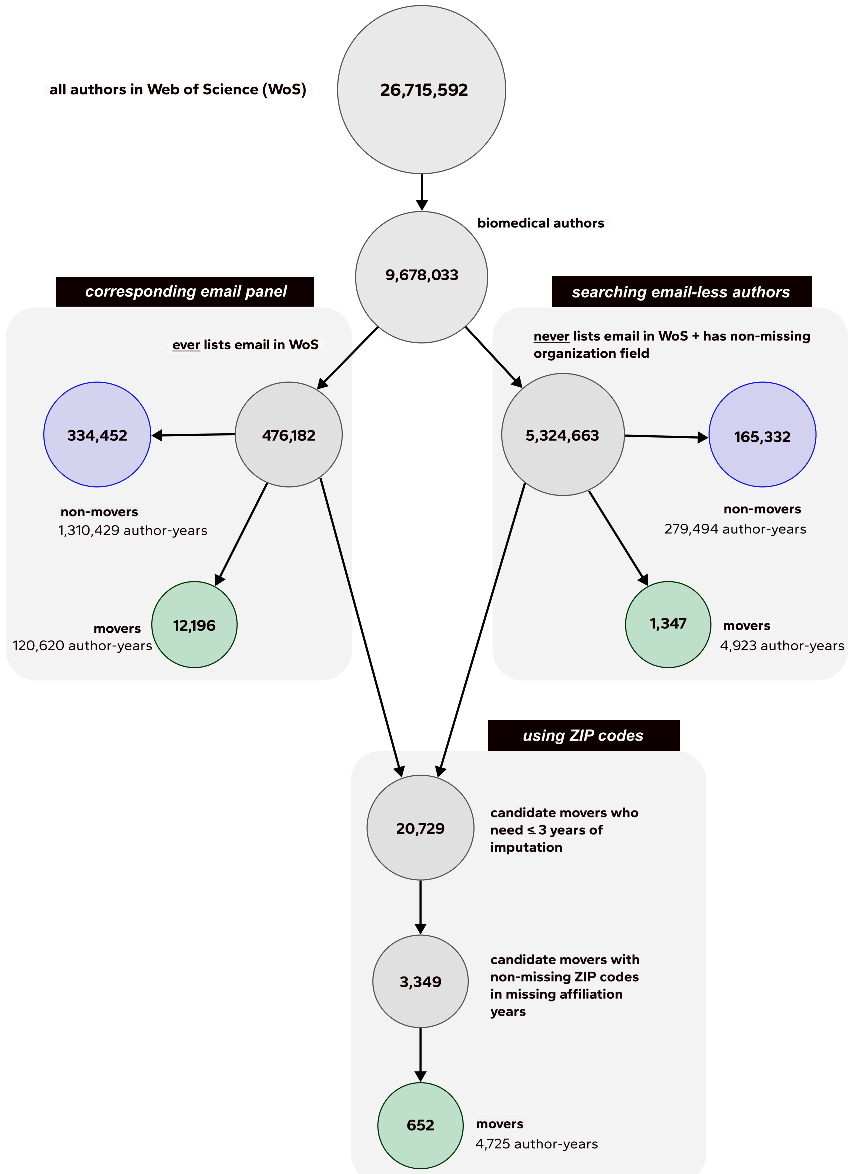


Figure A.1: Author-year panel construction

Notes: This diagram tracks the different ways we attempt to identify movers in our sample. This method begins with the author-year email skeleton shown in Table 1. The leftmost region shows the authors we capture from the basic email panel approach described in Section 2.4. 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. Summing across the values in the green nodes generates the total count of movers in our sample, 14,195. Summing across the values in the blue nodes equates the number of non-movers, 499,784. The same statements are true for the count of author-years, which are listed in the periphery of each colored node.

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 A.1 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. For each year we need to fill an affiliation for, we then 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 2.4, we code them to be at their destination. Only 3,349 of the 20,729 candidate movers ever have a ZIP code, of which 652 are coded as movers. Taken together, this procedure codes $12,196+1,347+652=14,195$ movers and $334,452+165,332=499,784$ non-movers, corresponding to a total of 1,720,191 author-years.

B Additive separability validation

We perform two exercises to validate our assumption that log commercialization is additively separable, as assumed in Equation 2.

B.1 Symmetry (Card, Heining, and Kline 2013)

Following the logic of Card, Heining, and Kline (2013), if the additive model

$$y_{it} = \alpha_i + \gamma_u + \varepsilon_{it}$$

is correct, then the effect on the outcome y of moving should be symmetric for two researchers moving in opposite directions. To see why, consider researcher i moving from u to u' and i' moving from u' to u:

$$y_{i2} - y_{i1} = (\gamma_{u'} - \gamma_u) + (\varepsilon_{i2} - \varepsilon_{i1})$$
$$y_{i'2} - y_{i'1} = (\gamma_u - \gamma_{u'}) + (\varepsilon_{i'2} - \varepsilon_{i'1})$$

If we average over all of the people moving from u to u' and u' to u, we get (assuming no serial correlation in the ε's):

$$E[\Delta y_{u \to u'}] = \gamma_{u'} - \gamma_u \text{ and } E[\Delta y_{u' \to u}] = \gamma_u - \gamma_{u'}$$

In other words, we get opposite and symmetric changes in the outcomes. Note that this is always true, even if very different groups of people (with different α's) are making the two different kinds of moves, *if additive separability holds*.

Now suppose additive separability does not hold. Instead, suppose the model has match effects between researchers and universities:

$$y_{it} = \alpha_i + \gamma_u + w(\alpha_i, \gamma_u) + \varepsilon_{it}$$

Again, consider worker i moving from u to u' and i' moving from u' to u:

$$y_{i2} - y_{i1} = (\gamma_{u'} - \gamma_u) + (w(\alpha_i, \gamma_{u'}) - w(\alpha_i, \gamma_u)) + (\varepsilon_{i2} - \varepsilon_{i1})$$
$$y_{i'2} - y_{i'1} = (\gamma_u - \gamma_{u'}) + (w(\alpha_{i'}, \gamma_u) - w(\alpha_{i'}, \gamma_{u'})) + (\varepsilon_{i'2} - \varepsilon_{i'1})$$

We can again take averages, but now the α terms no longer cancel. Let S be the set of workers who make $j \to j'$ moves and S' be the set of workers who make $j' \to j$ moves. Then taking averages over those sets, we have:

$$E[\Delta y_{u \to u'}] = (\gamma_{u'} - \gamma_u) + E_S[w(\alpha_i, \gamma_{u'}) - w(\alpha_i, \gamma_u)]$$
$$E[\Delta y_{u' \to u}] = (\gamma_u - \gamma_{u'}) + E_{S'}[w(\alpha_{i'}, \gamma_u) - w(\alpha_{i'}, \gamma_{u'})]$$

These are now no longer equal but opposite, because the $E_S[\cdot]$ and $E_{S'}[\cdot]$ values average over different sets of individuals with different values of α_i. Therefore, this suggests a natural test of the additive separability assumption: to examine whether equal and opposite moves lead to equal and opposite changes in our outcome measure.

Of course, we do not have enough data to investigate individual, opposite moves (i.e., Berkeley to Stanford versus Stanford to Berkeley). So again following Card, Heining, and Kline (2013), we divide universities into quartiles based on their average commercialization measure. This leaves us with 16 possible types of moves. In Figure B.1, we compare the symmetry of researchers making opposite moves, focusing on the most "extreme" of these moves. The results are shown below. The sample sizes for each move type are relatively small (in the 100-200 person range for each move type). Still, the changes in the outcome upon moving are quite symmetric, consistent with additive separability being a reasonably good model.

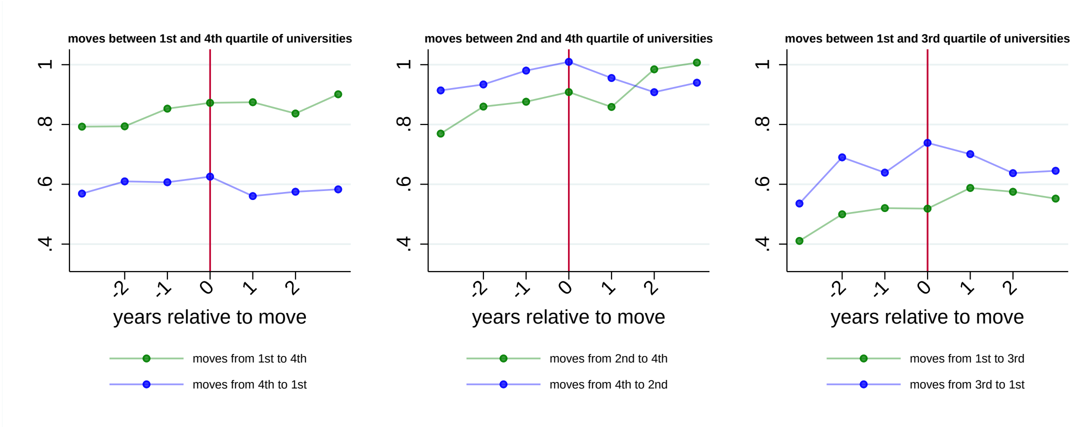


Figure B.1: Move symmetry

Notes: This figure shows the mean of IHS of the predicted patent citations by year for individuals who move from different types of universities, following Card, Heining, and Kline (2013). We restrict to a balanced panel.

B.2 Estimating complementarities (Bonhomme, Lamadon, and Manresa 2019)

Next, we try to estimate any complementarities directly, following Bonhomme, Lamadon, and Manresa (2019) (hereafter BLM).

We begin by constructing our sample. BLM use data from the entire working-age population in Sweden. They have 19,557 movers and 580,218 stayers the 2002-2004 time period. They estimate a two-period model where 2002 corresponds to $t = -1$ and 2004 corresponds to $t = 1$. Recall that we have 14,195 movers in total, but these moves occur across many different years. In an effort to preserve power, we drop the calendar time dimension in our data (in other words, moves in all years in our sample are combined). Still, we can only use movers who we observe for a year before and a year after their move. This leaves us with 8,610 movers. To construct stayers, we randomly assign placebo "move years" to our stayers, drawing from the distribution of real moves. Then we keep stayers that we observe on either side of the placebo move year. This leaves us with 66,669 stayers. Therefore, we have about 44% of the movers and 11% of the stayers that the original authors have in their sample.

We then cluster the universities into 10 groups. Following BLM exactly, we use a k-means algorithm, clustering the universities (using only stayers) according to their outcome distributions (where our outcome is our commercialization measure). We match on 20 points along the distribution, following BLM.

With these clusters in hand, we estimate two regression specifications. For author i moving in relative year t (note that calendar time and relative time are now synonymous), we start by estimating the linear regression:

$$y_{it} = \alpha_i + \gamma_{k(i,t)} + \varepsilon_{it}$$

where α_i is an individual fixed effect and γ_k is a university in group k fixed effect. This is a two-period model, so $t \in \{-1, 1\}$. This is similar to the model that we estimate in the main text, in the sense that it does not allow for researcher-university complementarities. However, it is worth noting that it is exactly not the same: it is estimated on a smaller sample, does not allow for calendar year effects, only uses two periods, and groups universities into ten different groups. Thus, we do not attempt to make direct comparisons between the results of this model and the model estimated in the main text.

Next, we estimate a model that allows for researcher-university interactions:

$$y_{it} = \alpha_i \cdot b_{k(i,t)} + \gamma_{k(i,t)} + \varepsilon_{it}.$$

The key difference is the interaction term $\alpha_i \cdot b_{k(i,t)}$, explicitly allowing for match effects between individuals and university groups. Following the original authors, the model is estimated using limited information maximum likelihood (LIML).

A variance decomposition from both models is shown below in Table B.1. The key takeaway is that the two models appear to offer very similar conclusions, suggesting that researcher-university complementarities are not substantively important. The fact that the R^2 values are so similar suggests that the complementarities add very little explanatory power.[3] This is consistent our Card, Heining, and Kline (2013) diagnostics above. It is also

[3] Careful readers might note that the R^2 of the linear model is actually *higher* than that of the interacted model, which

consistent with what Bonhomme, Lamadon, and Manresa (2019) find in their setting.

Model	$\frac{Var(\hat{\alpha})}{Var(y)}$	$\frac{Var(\hat{\gamma})}{Var(y)}$	$\frac{2Cov(\hat{\alpha},\hat{\gamma})}{Var(y)}$	$\frac{Var(\varepsilon)}{Var(y)}$	$Corr(\hat{\alpha},\hat{\gamma})$	R^2
Linear	65.49	0.16	1.25	33.09	0.19	0.67
Interacted	63.97	0.21	1.39	34.43	0.19	0.66

Table B.1: Linear vs. Interacted Model (BLM)

Notes: This table compares the variance decomposition from a linear model and an interacted model, following the methodology laid out above and in Bonhomme, Lamadon, and Manresa (2019). The first four columns show the percent of the total variance explained by each constituent part. The fifth column shows the correlation between the estimated researcher and university effects. The sixth column shows the R^2. N = 75,279 researchers, 8,610 of which are movers.

C Variance decomposition

Like in Finkelstein, Gentzkow, and Williams (2016), we provide results from a variance decomposition of our main outcome: the IHS of predicted five-year patent citations. To begin with, we use the estimates generated by Equation (2)— $\hat{y}_{it}$, $\hat{\gamma}_u$, and $\hat{\alpha}_i$— and average each to the university-year level, and then again to the university level. This produces a dataset with $\hat{y}_u$, $\hat{\gamma}_u$, and $\hat{\alpha}_u$ for each university with a non-zero count of movers ($N = 250$). The share of cross-university variance that would be reduced by equalizing researchers across universities is written as

$$S^{var}_{researcher} = 1 - \frac{Var(\hat{\gamma}_u)}{Var(\hat{y}_u)} \tag{12}$$

where the *reduction* in variance is modeled to depend on only differences in university effects, scaled by the overall variance across universities. 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

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

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 (2) 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

seems implausible. However, this happens due to the different estimation methodologies for the two models (OLS vs. LIML). Our overall takeway is simply that the R^2 values are very similar.

cross-university variance in $\hat{y}_u$ is driven by noise resulting from few movers, or "limited mobility bias." In particular, we know from Equation (12) 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 university and researcher effects as additive, if $\hat{\gamma}_u > \gamma_u$ we know $\hat{\alpha}_u < \alpha_u$.

Empirically, Figure E.22 shows that the number of movers varies across schools. To minimize noise in $\hat{\gamma}_u$, and in accordance with Andrews et al. (2012), we restrict the variance decomposition to include only universities with 25 movers.

The results from this exercise are shown in Table 4 and are generally consistent with estimates presented by the event study (Figure 6) and additive (Table 3) analogs. Specifically, it is estimated that about 20% of the cross-university variance in predicted citations is attributable to variance in university effects.

D Empirical Bayes adjustment

One additional step we take to minimize noise in our estimate of $\hat{\gamma}_u$ is to employ a non-parametric Empirical Bayes correction (Kline, Saggio, and Sølvsten, 2020; Morris, 1983). This approach pools information across universities to improve our estimated university fixed effects, while allowing for potential correlation in noise across universities, the latter of which can be important when using mover designs (Walters, 2024).

To begin, we assume that $\hat{\Gamma} = (\hat{\gamma}_1, ..., \hat{\gamma}_U)'$ is an unbiased estimate of the true university effects, Γ_u. A $U \times U$ matrix V describes the sampling variance of Γ_u, with squared standard error terms s_u^2 of the individual $\hat{\gamma}_u$'s along the diagonal. $\hat{V}$ is an unbiased estimator of this matrix.

Next, we assume that the latent parameters γ_u are randomly drawn from the distribution G_γ, which is defined in the population of universities and is normally distributed. Together, these assumptions yield the hierarchical model:

$$\begin{aligned} \hat{\gamma}_u | \gamma_u, s_u^2 &\sim N(\gamma_u, s_u^2) \\ \gamma_u | s_u^2 &\sim N(\mu_\gamma, \sigma_\gamma^2). \end{aligned} \tag{14}$$

This model has two hyperparameters, μ_γ and σ_γ^2, which we estimate, following Walters (2024), as:

$$\hat{\mu}_\gamma = \frac{1}{U} \sum_{u=1}^{U} \hat{\gamma}_u, \quad \hat{\sigma}_\gamma{}^2 = \hat{\Gamma}' A_0 \hat{\Gamma} - tr(A_0 \hat{V}). \tag{15}$$

where

$$A_0 = \frac{1}{U-1}\left(I_U - \frac{1}{U}1_U 1_U'\right) \tag{16}$$

Once these two hyperparameters are estimated, the empirical Bayes method produces posteriors as:

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

Figure D.1 shows the result of this adjustment on the distribution of $\hat{\gamma}_u$.

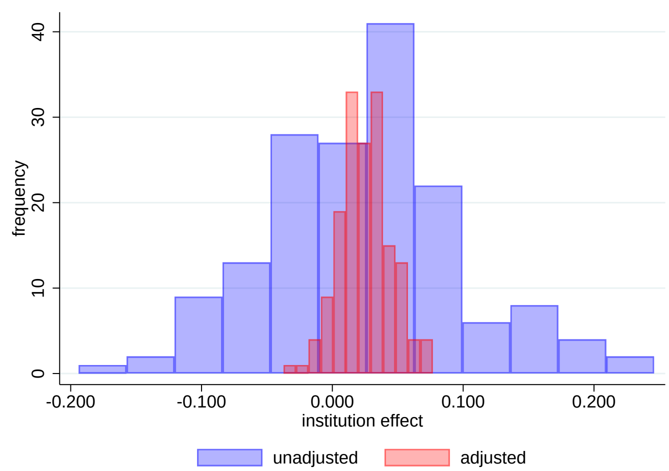


Figure D.1: Empirical Bayes: Distribution of university effects

Notes: This figure is based on estimation of Equation (2), where the dependent variable y_{iut} is the IHS of the predicted count of five-year patent citations to author i's papers in year t whose main academic affiliation is at university u. In particular, this figure displays the results of an Empirical Bayesian shrinkage of the causal university effects ($\hat{\gamma}_u$). The x-axis is the estimated value of the university effect for a given school, and the y-axis is the frequency of schools within an university effect bin. The blue bars show the distribution of the university effects estimated from Equation (2). The red bars show the distribution of the adjusted university effects, following the Bayesian shrinkage ($\hat{\gamma}_u^{EB}$). To center the mean of this distribution near zero, we pre-omit Syracuse University when estimating Equation (2). While our full sample includes 264 universities, we restrict our attention to the 164 with at least 25 movers. This leaves us with 1,663,147 author-years.

E Supplemental figures and tables

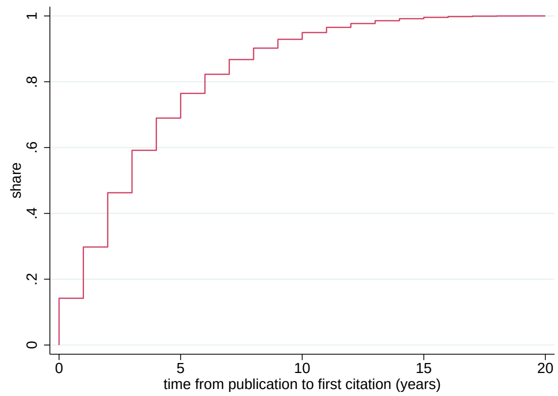
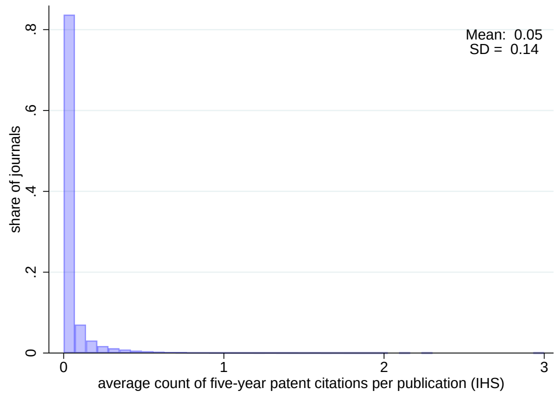


Figure E.1: Lag between paper publication and first patent citation

Notes: This plot 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 sample is all publications with at least one patent cite, written by authors in our analysis sample (N = 1,706,178).



(a) Distribution of commercialization across journals

rank	*journal*	*publication count*	*mean patent citations*
1	2008 SID International Symposium, Technical Digest	545	3.00
2	2007 SID International Symposium, Technical Digest	478	2.27
3	Annual Review Of Immunology	270	2.26
4	IEE International Electron Devices Meeting 2000, Technical Digest	201	2.23
5	ACM Transactions On Computer Systems	226	2.15
6	mAbs	1220	2.13
7	IEE International Electron Devices Meeting 2004, Technical Digest	251	1.97
8	Nature Biotechnology	8493	1.96
9	Acm Computing Surveys	1004	1.92
10	2007 IEEE International Electron Devices Meeting	240	1.91

(b) Ranked journals by commercialization propensity

Figure E.2: 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. Panel (b) lists the ten journals whose publications have the highest average count of five-year patent citations.

(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.

Figure E.3: Measuring university affiliations of authors: An example

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

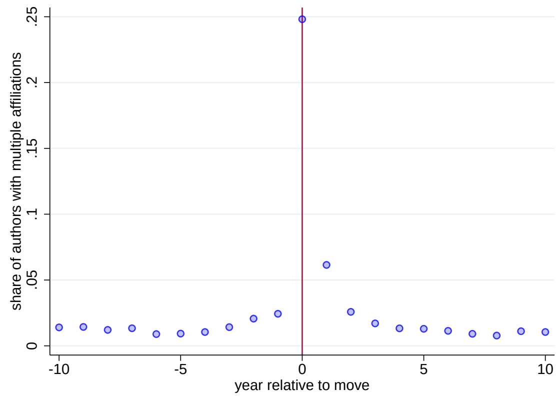


Figure E.4: Multiple affiliations in the WoS by relative year

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 relative year. A red vertical line denotes the move year. The sample of all mover-years is used to calculate this share (N =119,214).

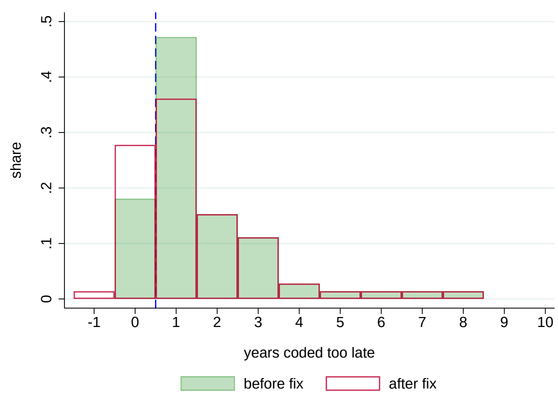
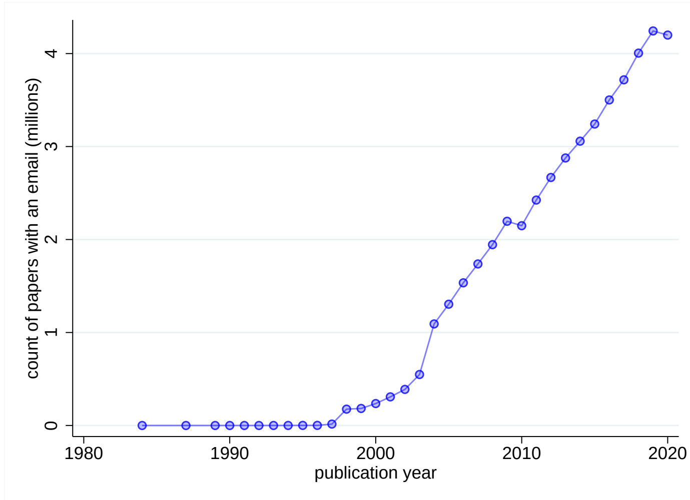
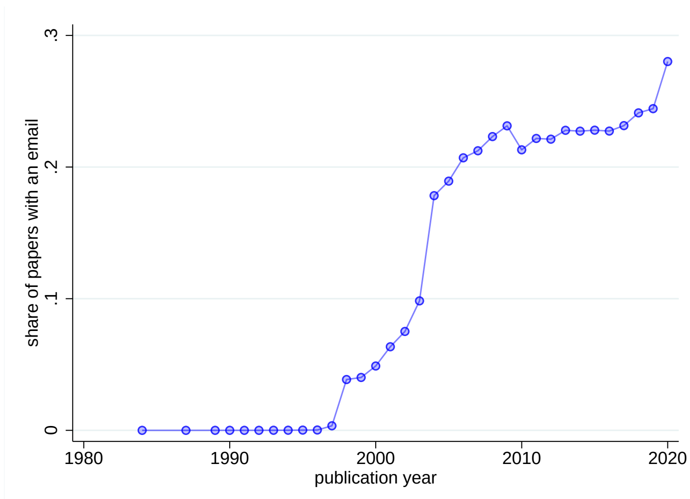


Figure E.5: Move timing correction

Notes: This figure plots the distribution of move-timing errors found in a hand-coded, randomly selected sample of 100 movers from our full sample of 14,195. By collecting public information from CVs, 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 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. The y-axis in each plot is the count of the hand-coded movers. Per the green bars, the modal move timing error is one year.



(a) Count



(b) Share

Figure E.6: 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 to 2020.

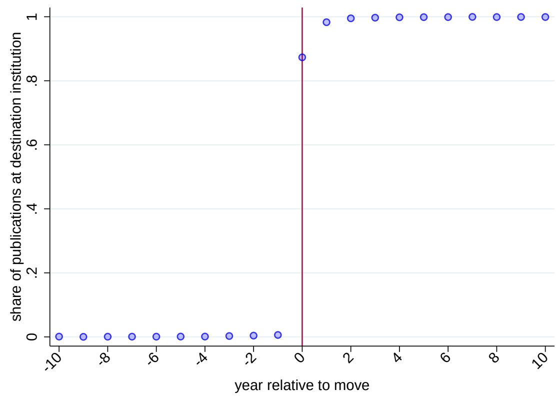
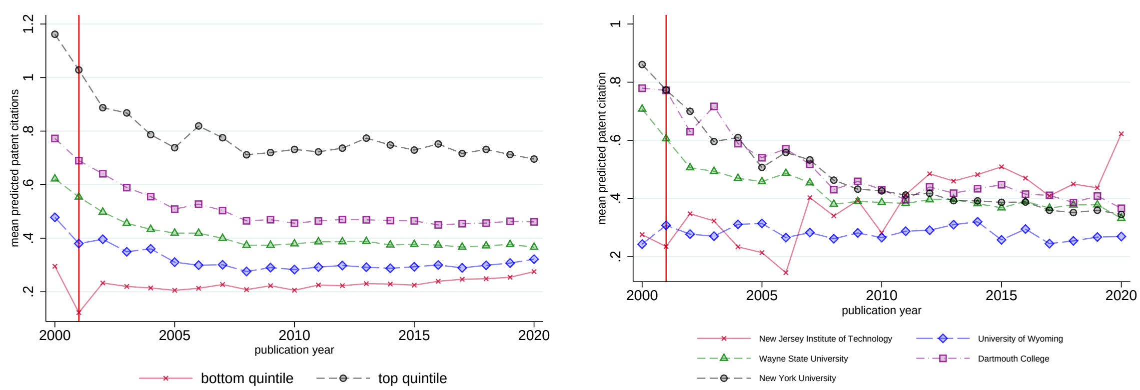


Figure E.7: Share of publications associated with destination university by relative year

Notes: This figure plots the average share of a mover's publications per relative year that list their destination university as their academic affiliation, restricting to affiliations that are either the destination or origin. 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. The sample is the full sample of movers (N = 14,195).



(a) Quintiles over time

(b) Sampled university within quintile over time

Figure E.8: Stability in commercialization quintiles over time

Notes: Panel (a) plots quintiles of university-level IHS of predicted patent citations over publication years from 2000 to 2020. These quintiles are calculated by estimating the distribution of the 2001 university-level average count of predicted patent citations per publication written by authors at a given TTO. Mean counts of predicted patent citations are then calculated over the author-years assigned to each quintile, by publication year. These quintile means are calculated using both movers and non-movers, totaling 513,979 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, the New Jersey Institute of Technology is in the lowest quintile and New York University is in the highest. In each panel, a red vertical line denotes 2001 as the publication year used to construct these quintiles.

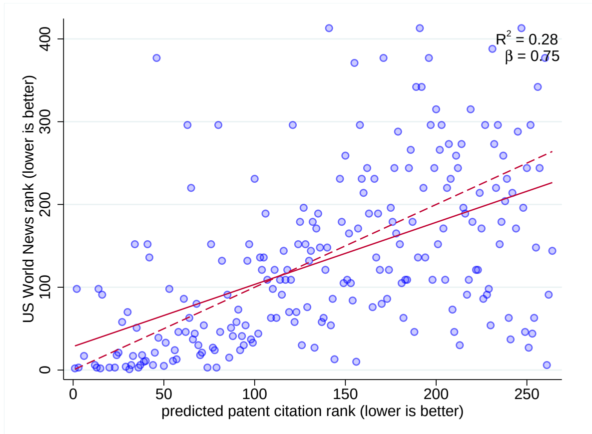


Figure E.9: U.S. News and World Report ranking vs. commercialization rank

Notes: This figure plots university-level commercialization rank against *U.S. News and World Report*'s (USNWR) 2025 university rank. To match more TTOs to USNWR ranks, hospitals and institutes are assigned to their parent universities where appropriate. University systems are assigned to their highest-ranking campus. Commercialization rank is calculated as the average count of predicted citations that researchers affiliated with a given university 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. 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 264 universities, only 223 can be linked back to USNWR rankings.

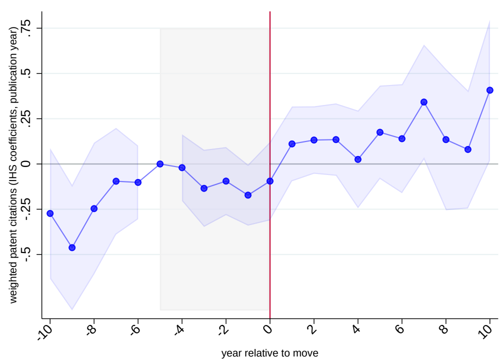
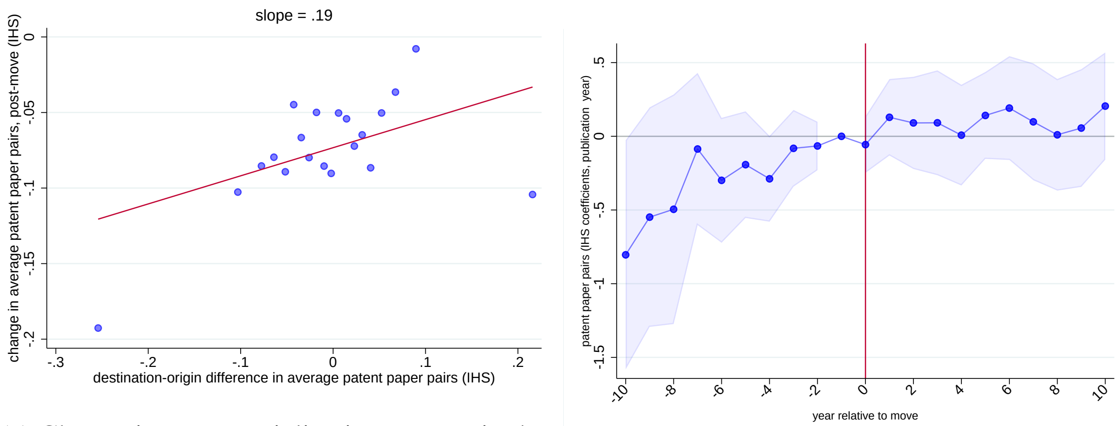


Figure E.10: Event study: Weighted observed five-year patent citations

Notes: This figure is identical to Figure 8b, except that the outcome replaces our patent cite measure with a weighted patent cite measure. Patent citations are weighted by the cumulative count of patent citations they themselves 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. There are 81,725 author years in our sample, coming from the sample of 11,441 movers for whom we can calculate actual five-year patent citation measures.

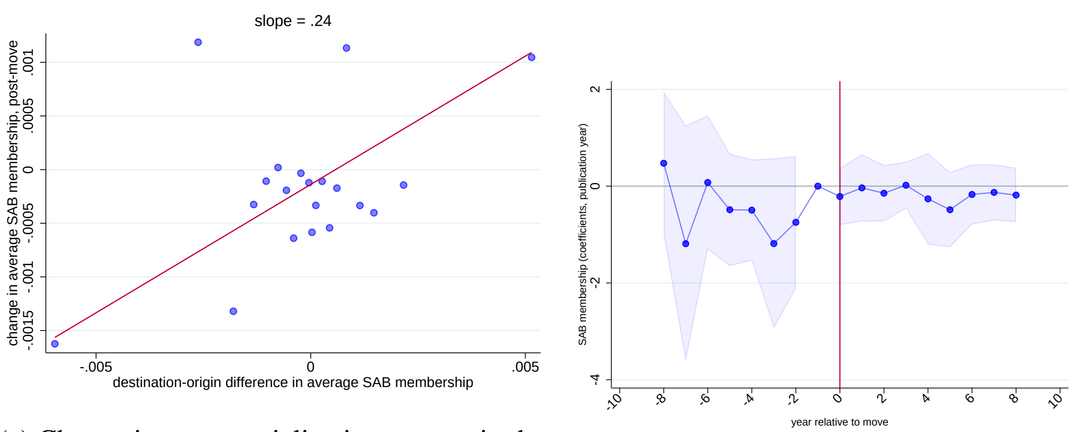


(a) Change in commercialization propensity by size of move

(b) Event study

Figure E.11: Patent-paper pairs

Notes: This figure is identical to Figure 8, except that our outcome is patent-paper pairs. There are N = 119,214 author years in our sample, coming from the sample of 14,195 movers.



(a) Change in commercialization propensity by size of move

(b) Event study

Figure E.12: Scientific advisory board membership

Notes: This figure is identical to Figure 8, except that our outcome is a 0/1 indicator for whether a researcher joined an SAB in a given year. 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 extreme years due to few authors joining SABs. As a result, there are 110,920 author years in our sample, coming from the sample of 14,195 movers.

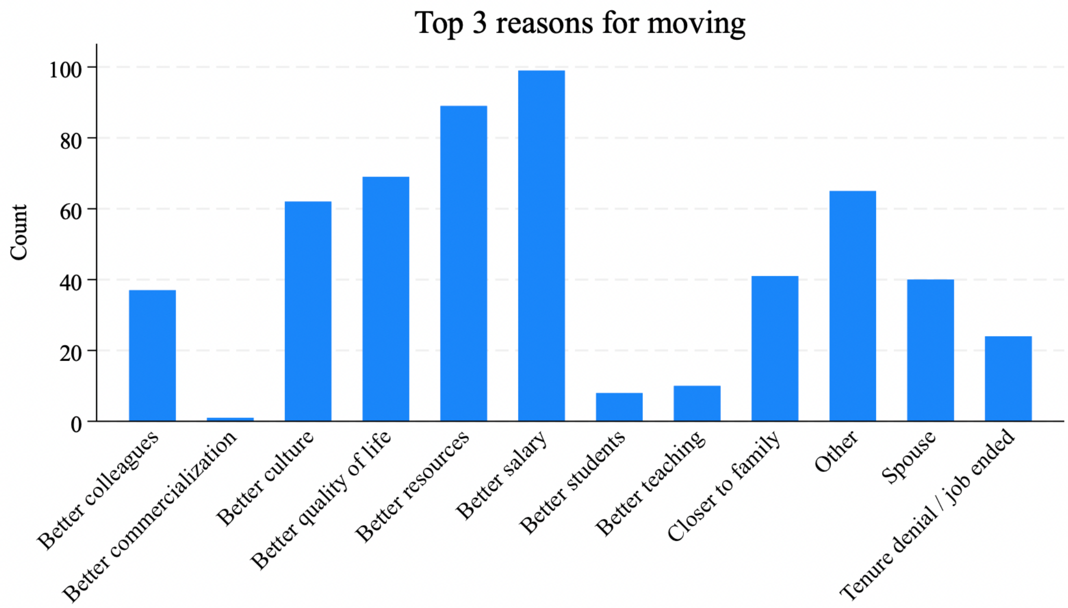


Figure E.13: Survey results: reasons for moving

Notes: This figure shows the count of top three reasons that respondents listed for their most recent academic move. $N = 211$. Respondents could provide up to 3 reasons, so the total counts sum to more than 211.

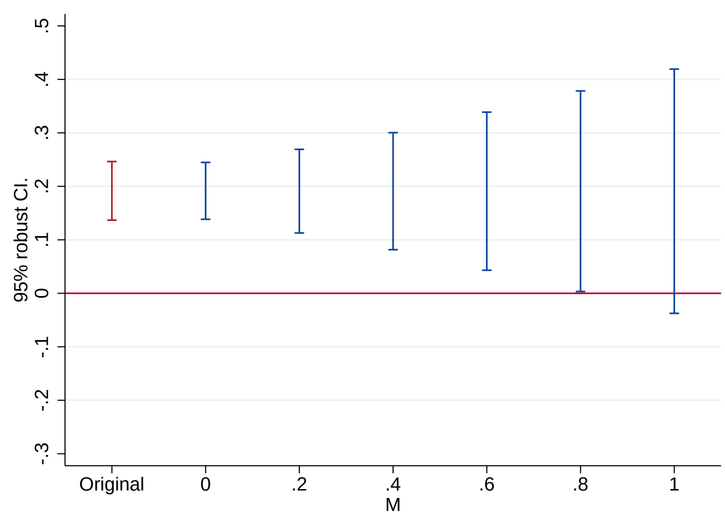
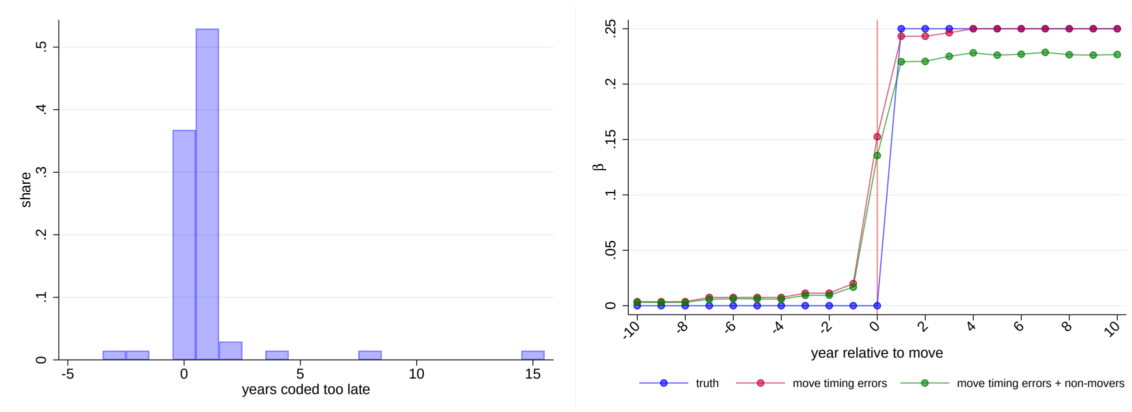


Figure E.14: Rambachan-Roth pre-trend sensitivity

Notes: This figure shows the point estimate and standard error for θ_1 from Figure 6 in the main text as M increases, following the methodology from Rambachan and Roth (2023). $M = 0$ corresponds to no pre-trends. $M = 1$ corresponds to the worst-case pre-trends that are possible given the observed data.



(a) Move timing errors

(b) Simulated event study

Figure E.15: 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 14,195 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 the treatment, but remain coded as movers to match the mistake we make in our sample construction.

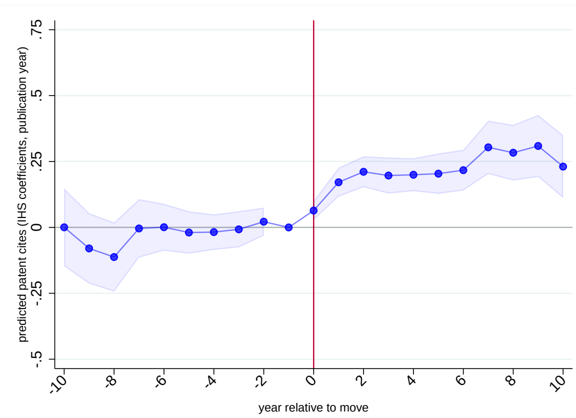
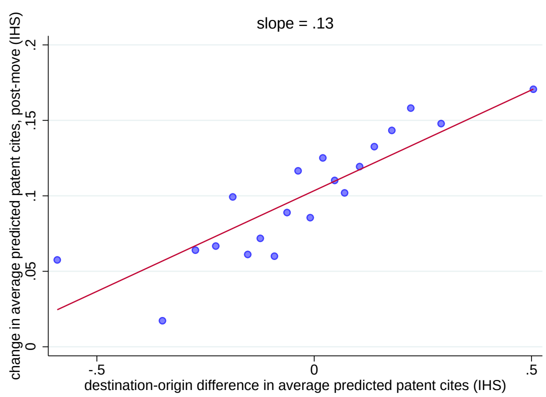
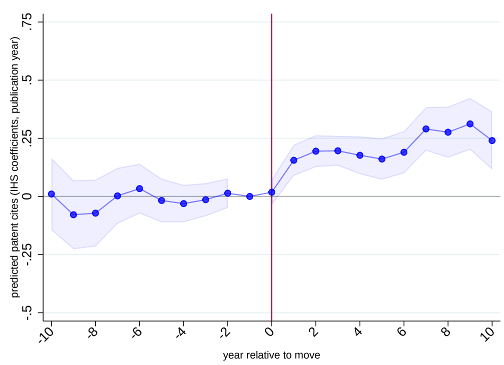


Figure E.16: Event study: Predicted five-year patent citations with relative year fixed effects

Notes: This figure is identical to Figure 6, except that it adds uninteracted relative year fixed effects, estimating Equation (11). There are 119,214 author years in our sample, coming from the sample of 14,195 movers.



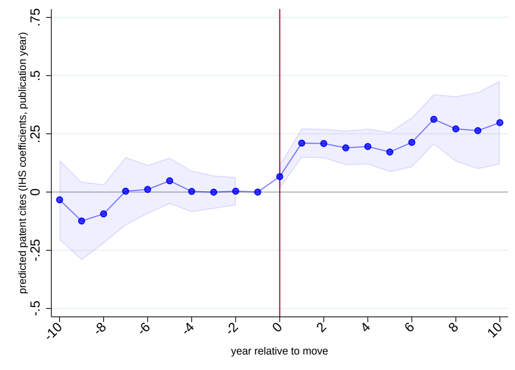
(a) Change in commercialization propensity by size of move



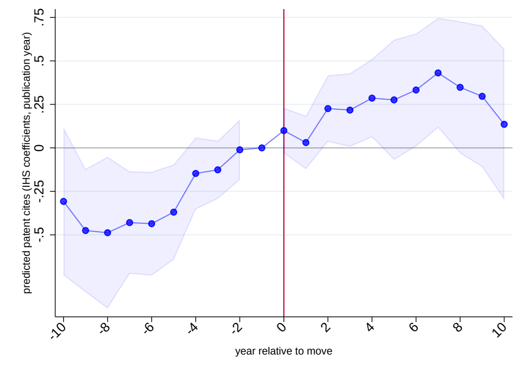
(b) Event study

Figure E.17: Last authors

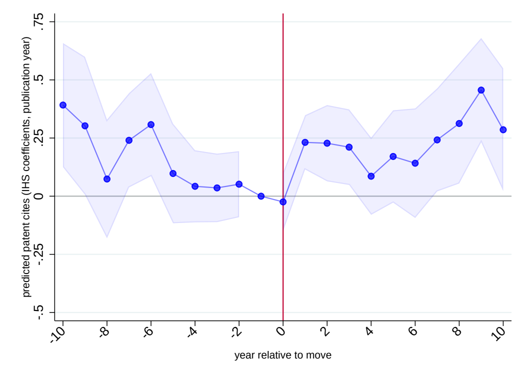
Notes: This figure is identical to Figures 5 and 6, except that we restrict to authors who are last authors at both their origin and destination. There are 7,552 last author movers in our sample, which corresponds to a total of 78,880 author-years.



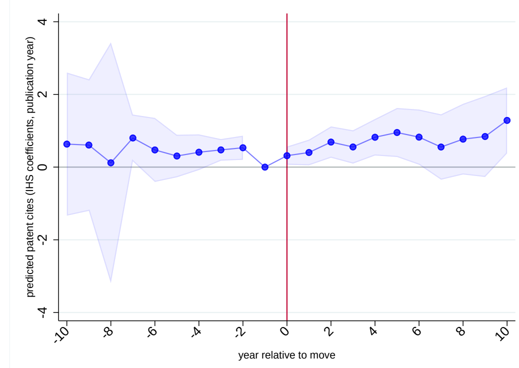
(a) University → university
(12,873 movers)



(b) University → non-university
(739 movers)



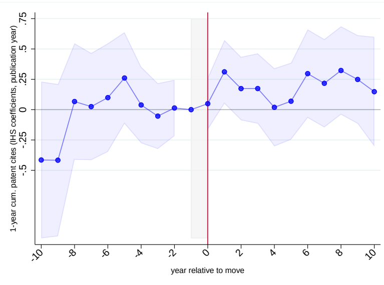
(c) Non-university → university
(537 movers)



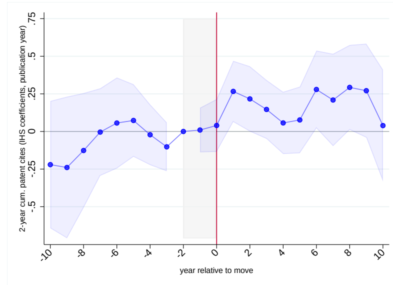
(d) Non-university → non-university
(46 movers)

Figure E.18: Event study for different move types

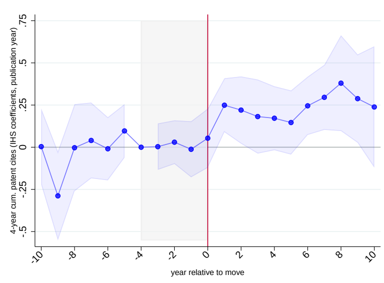
Notes: This figure is identical to Figure 6 from the main text, splitting the movers into four mutually-exclusive groups based on move type. Number of movers is provided in each subpanel.



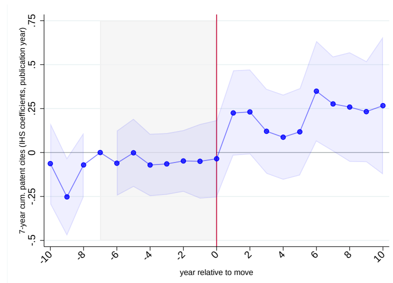
(a) 1-year cites



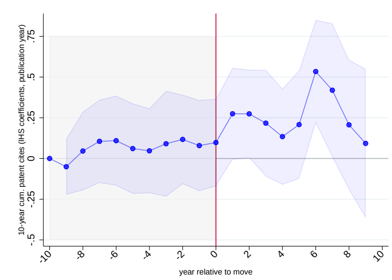
(b) 2-year cites



(c) 4-year cites



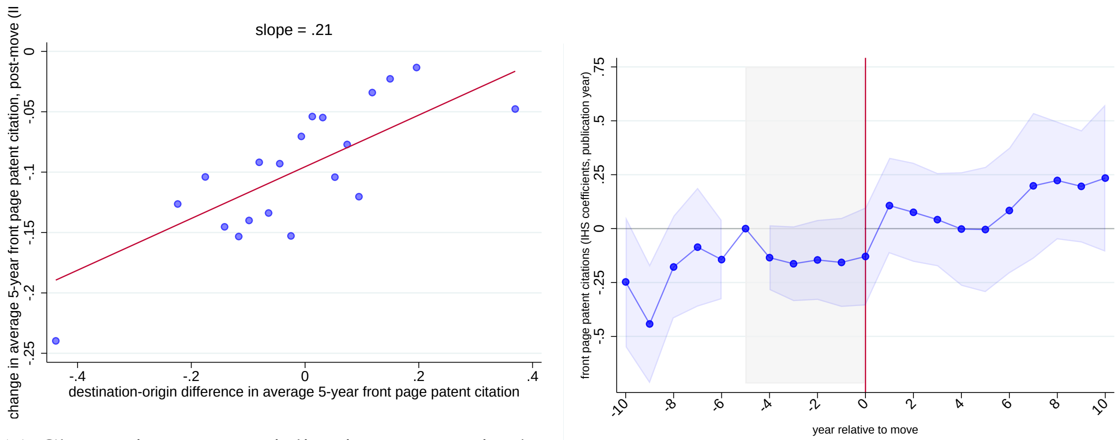
(d) 7-year cites



(e) 10-year cites

Figure E.19: Citation windows

Notes: These event study plots are identical to Figure 8b, but the outcomes look at patent citations over different time horizons (one year, two years, etc.) 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. Thus we do not report it.

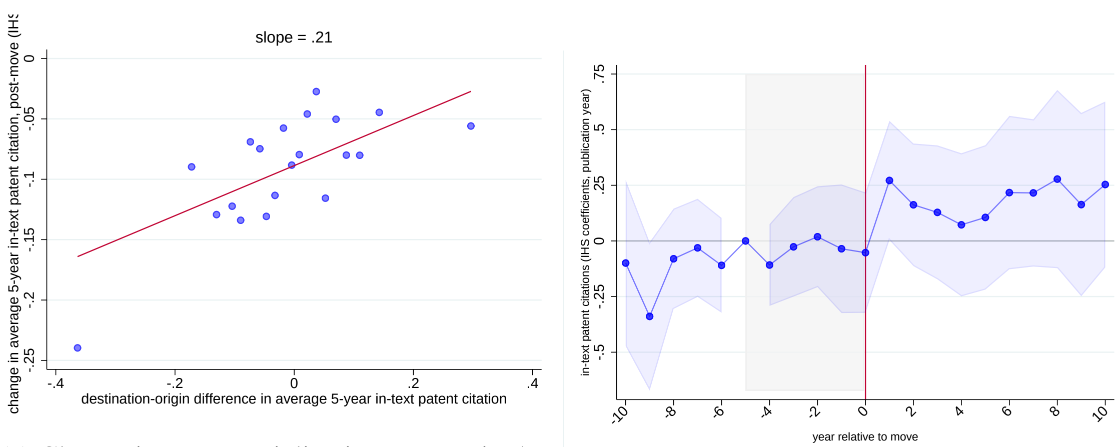


(a) Change in commercialization propensity by size of move

(b) Event study

Figure E.20: Five-year front page patent citations

This figure is identical to Figure 8, replacing all five-year patent citations with five-year front page patent citations. Each panel is estimated using 11,441 movers and 81,725 mover-years: this sample is defined by dropping the years between 2016 and 2020 from our main estimation sample with 14,195 movers and 119,214 mover-years.



(a) Change in commercialization propensity by size of move

(b) Event study

Figure E.21: Five-year in-text patent citations

This figure is identical to Figure 8, replacing all five-year patent citations with five-year in-text patent citations. Each panel is estimated using 11,441 movers and 81,725 mover-years: this sample is defined by dropping the years between 2016 and 2020 from our main estimation sample with 14,195 movers and 119,214 mover-years.

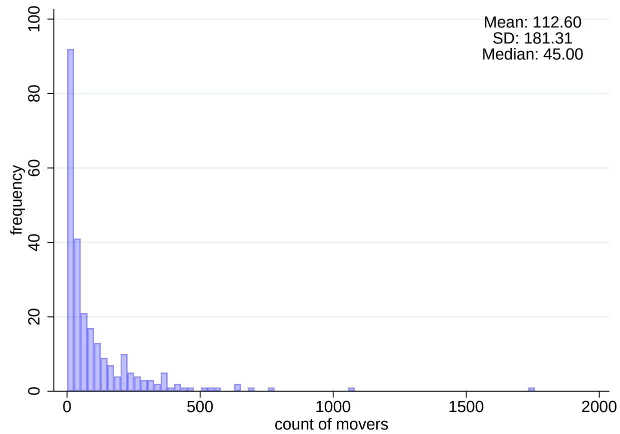


Figure E.22: Movers by university

Notes: This figure shows the distribution of the number of movers per university, conditional on observing at least one move. Here, a mover is counted once for their origin and once for their destination. 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 $14,195$ movers is used; they span 250 universities.

References

Andrews, Martyn, Len Gill, Thorsten Schank, and Richard Upward. 2012. "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias." *Economics Letters* 117 (3): 824–827.

Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa. 2019. "A Distributional Framework for Matched Employer Employee Data." *Econometrica* 87 (3): 699–739.

Card, David, Jörg Heining, and Patrick Kline. 2013. "Workplace Heterogeneity and the Rise of West German Wage Inequality." *Quarterly Journal of Economics* 128 (3): 967–1015.

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.

Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten. 2020. "Leave-Out Estimation of Variance Components." *Econometrica* 88 (5): 1859–1898.

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.

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

Rambachan, Anesh and Jonathan Roth. 2023. “A More Credible Approach to Parallel Trends.” *Review of Economic Studies* 90 (5): 2555–2591.

Walters, Christopher. 2024. “Empirical Bayes Methods in Labor Economics.” In: *Handbook of Labor Economics*. Ed. by Thomas Lemieux and Christian Dustman. Vol. 5. Elsevier. Chap. 3: 183–260.