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

May 2025

Commercial spillovers from academic research

- University-based scientific research has long been argued to be a central source of commercial innovation and economic growth

  [Nelson 1986; Jaffe 1989; Mansfield 1991]

- Yet at the same time, long-held concerns that many university-based discoveries never realize their potential social benefits
  - Policy debates over title rights surrounding the US Bayh-Dole Act
  - Cooperation of academic inventors often necessary for diffusion

    [Zucker-Darby-Brewer 1998; Jensen-Thursby 2001]

This paper

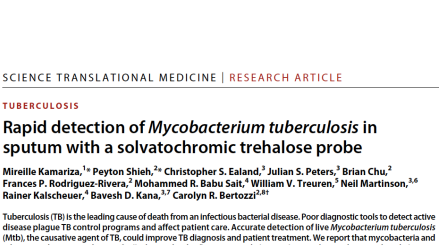
- Does *where* an academic discovery is made affect *whether* it diffuses into the economy?
  - Example: in 2016 Duke and Stanford had similar research expenditures ($910 million versus $990 million), but 9 versus 32 start-ups
  - But of course, field differences and sorting may be central

- As a first step towards unpacking this heterogeneity, we analyze academics who move across universities
  - How does the propensity of academic research to spill over to commercial innovation change when academics move?

Motivating example: DMN-Tre

DMN-Tre is a novel small molecule that allows tuberculosis to be detected rapidly, cheaply, and with high accuracy

- Developed in Professor Carolyn Bertozzi's lab at Stanford
- At the time, Mireille Kamariza was a PhD student of Bertozzi's
- Discovery led to both a scientific research paper and a patent filing



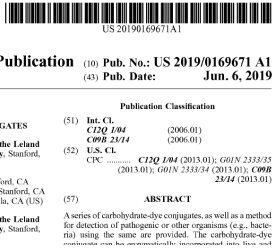
SCIENCE TRANSLATIONAL MEDICINE | RESEARCH ARTICLE

TUBERCULOSIS

Rapid detection of *Mycobacterium tuberculosis* in sputum with a solvatochromic trehalose probe

Mireille Kamariza,[1]* Peyton Shieh,[2]* Christopher S. Ealand,[3] Julian S. Peters,[3] Brian Chu,[2] Frances P. Rodriguez-Rivera,[2] Mohammed R. Babu Sait,[4] William V. Treuren,[5] Neil Martinson,[3,6] Rainer Kalscheuer,[4] Bavesh D. Kana,[3,7] Carolyn R. Bertozzi[2,8†]

Tuberculosis (TB) is the leading cause of death from an infectious bacterial disease. Poor diagnostic tools to detect active disease plague TB control programs and affect patient care. Accurate detection of live *Mycobacterium tuberculosis* (Mtb), the causative agent of TB, could improve TB diagnosis and patient treatment. We report that mycobacteria and



US 20190169671A1

(19) **United States**

(12) **Patent Application Publication**

Bertozzi et al.

(10) **Pub. No.: US 2019/0169671 A1**

(43) **Pub. Date: Jun. 6, 2019**

(54) **METHODS FOR DETECTING MYCOBACTERIA WITH SOLVATOCHROMIC DYE CONJUGATES**

(71) Applicant: **The Board of Trustees of the Leland Stanford Junior University**, Stanford, CA (US)

(72) Inventors: **Carolyn R. Bertozzi**, Stanford, CA (US); **Mireille Kamariza**, Stanford, CA (US); **Peyton Shieh**, La Jolla, CA (US)

(73) Assignee: **The Board of Trustees of the Leland Stanford Junior University**, Stanford, CA (US)

Publication Classification

(51) **Int. Cl.**
C12Q 1/04 (2006.01)
C09B 23/14 (2006.01)

(52) **U.S. Cl.**
CPC ***C12Q 1/04*** (2013.01); *G01N 2333/35* (2013.01); *G01N 2333/34* (2013.01); ***C09B 23/14*** (2013.01)

(57) **ABSTRACT**

A series of carbohydrate-dye conjugates, as well as a method for detection of pathogenic or other organisms (e.g., bacteria) using the same are provided. The carbohydrate-dye

Motivating example: OliLux Biosciences

OliLux Biosciences

Rapid and low-cost solution for tuberculosis detection

Portable diagnostics for anyone, anywhere.

Mireille Kamariza, PhD

Founder and CEO

Carolyn Bertozzi, PhD

Founder and Chair of Scientific Advisory Board

Motivating example: Bertozzi's lab

- Circa 2014: at UC-Berkeley
  - "Even as Bertozzi's lab came out with tool after tool...translating that information into actual diagnostics or therapies was further off."
  - "I kind of felt at Berkeley...It 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."
- Mid-2015: moved her lab to Stanford
  - Started four companies within five years of moving
  - "...attributes her entrepreneurial fervor to the natural collaborations that bubble up at Stanford in a way they didn't at Berkeley."

Sources:

- https://cen.acs.org/people/profiles/Carolyn-Bertozzis-glycorevolution/98/i5
- https://www.chemistryworld.com/features/carolyn-bertozzi/3008380.article

Commercial spillovers from academic research

A range of place-based factors have been argued to affect whether academic discoveries spill over into commercial innovation:

1. Universities, including but not limited to patenting and licensing practices
   [Henderson-Jaffe-Trajtenberg 1998; Hausman forthcoming]
2. Local market characteristics such as local availability of venture capital and technology clusters [Kenney 1986; Saxenian 1994; Moretti 2021]

We estimate that roughly 15-30% of geographic variation in spillovers from university-based research is attributable to place-specific factors

1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Measuring academic research: Web of Science data

We start by measuring academic research as publications catalogued in the Web of Science data from Clarivate Analytics

- Compiles records of peer-reviewed scholarly journals, conference proceedings, and editorially selected books
- Our version includes publications through the beginning of 2021, and contains more than 80 million publication records
- For each author in our sample, we measure annual publications based on the Web of Science-assigned (disambiguated) author ID ▸ example

1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

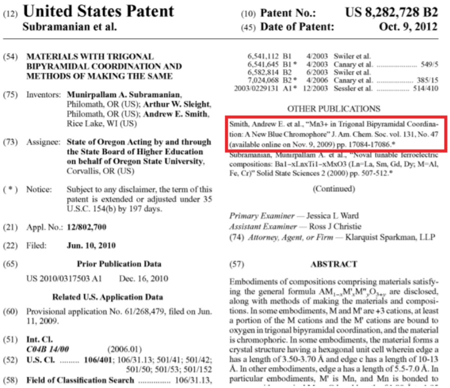
Measuring commercial linkages to academic research

- Commercial spillovers from university-based research can take many forms, not all of which are possible to measure

- Example: Langer Lab at MIT
  - He has authored over 1,500 scientific research papers
  - He holds over 1,400 granted or pending patents
  - These patents have been licensed or sublicensed to over 400 companies
  - Over 40 companies have spun out of the Langer lab, including Moderna (which Langer co-founded)
  - He sits on at least 45 Scientific Advisory Boards

Baseline outcome: Measuring citations from patents

Building on work by Bryan, Ozcan, and Sampat (2020) and Marx and Fuegi (2020), our baseline estimates measure the propensity of authors' published scientific research to be cited in granted US patents



(12) **United States Patent**
Subramanian et al.

(10) **Patent No.: US 8,282,728 B2**
(45) **Date of Patent: Oct. 9, 2012**

(54) **MATERIALS WITH TRIGONAL BIPYRAMIDAL COORDINATION AND METHODS OF MAKING THE SAME**

(75) Inventors: **Munirpallam A. Subramanian**, Philomath, OR (US); **Arthur W. Sleight**, Philomath, OR (US); **Andrew E. Smith**, Rice Lake, WI (US)

(73) Assignee: **State of Oregon Acting by and through the State Board of Higher Education on behalf of Oregon State University**, Corvallis, OR (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 197 days.

(21) Appl. No.: **12/802,700**

(22) Filed: **Jun. 10, 2010**

(65) **Prior Publication Data**

US 2010/0317503 A1 Dec. 16, 2010

Related U.S. Application Data

(60) Provisional application No. 61/268,479, filed on Jun. 11, 2009.

(51) **Int. Cl.**
C04B 14/00 (2006.01)

(52) **U.S. Cl.** **106/401**; 106/31.13; 501/41; 501/42; 501/50; 501/53; 501/152

(58) **Field of Classification Search** 106/31.13,

6,541,112 B1 4/2003 Swiler et al.
6,541,645 B1 * 4/2003 Canary et al. 549/5
6,582,814 B2 6/2003 Swiler et al.
7,024,068 B2 * 4/2006 Canary et al. 385/15
2003/0229131 A1 * 12/2003 Sessler et al. 514/410

OTHER PUBLICATIONS

Smith, Andrew E. et al., "Mn3+ in Trigonal Bipyramidal Coordination: A New Blue Chromophore" J. Am. Chem. Soc. vol. 131, No. 47 (available online on Nov. 9, 2009) pp. 17084-17086.*

Subramanian, Munirpallam A. et al., "Noval tunable ferroelectric compositions: Ba1-xLnxTi1-xMxO3 (Ln=La, Sm, Gd, Dy; M=Al, Fe, Cr)" Solid State Sciences 2 (2000) pp. 507-512.*

(Continued)

Primary Examiner — Jessica L Ward
Assistant Examiner — Ross J Christie
(74) *Attorney, Agent, or Firm* — Klarquist Sparkman, LLP

(57) **ABSTRACT**

Embodiments of compositions comprising materials satisfying the general formula $AM_{1-x}M'_xM''_yO_{3+y}$ are disclosed, along with methods of making the materials and compositions. In some embodiments, M and M' are +3 cations, at least a portion of the M cations and the M' cations are bound to oxygen in trigonal bipyramidal coordination, and the material is chromophoric. In some embodiments, the material forms a crystal structure having a hexagonal unit cell wherein edge a has a length of 3.50-3.70 Å and edge c has a length of 10-13 Å. In other embodiments, edge a has a length of 5.5-7.0 Å. In particular embodiments, M' is Mn, and Mn is bonded to

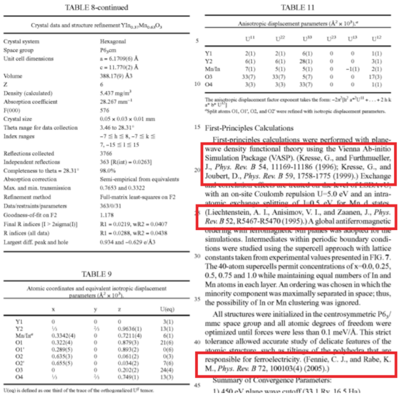


TABLE 8-continued

Crystal data and structure refinement $YIn_{0.37}Mn_{0.63}O_3$	
Crystal system	Hexagonal
Space group	$P6_3cm$
Unit cell dimensions	a = 6.1709(6) Å c = 11.770(2) Å
Volume	388.17(9) Å3
Z	6
Density (calculated)	5.437 mg/m^3
Absorption coefficient	28.267 mm^{-1}
F(000)	576
Crystal size	0.05 × 0.03 × 0.01 mm
Theta range for data collection	3.46 to 28.31°
Index ranges	−7 ≦ h ≦ 8, −7 ≦ k ≦ 7, −15 ≦ l ≦ 15
Reflections collected	3766
Independent reflections	363 [R(int) = 0.0263]
Completeness to theta = 28.31°	98.0%
Absorption correction	Semi-empirical from equivalents
Max. and min. transmission	0.7653 and 0.3322
Refinement method	Full-matrix least-squares on F2
Data/restraints/parameters	363/0/31
Goodness-of-fit on F2	1.178
Final R indices [I > 2sigma(I)]	R1 = 0.0219, wR2 = 0.0407
R indices (all data)	R1 = 0.0288, wR2 = 0.0438
Largest diff. peak and hole	0.934 and −0.629 e/Å3

TABLE 9

Atomic coordinates and equivalent isotropic displacement parameters (Å^2 × 10^3).

	x	y	z	U(eq)
Y1	0	0	0	3(1)
Y2	⅓	⅔	0.9636(1)	13(1)
Mn/In[a]	0.3342(4)	0	0.7211(4)	6(1)
O1	0.322(4)	0	0.879(3)	21(6)
O1'	0.289(5)	0	0.893(2)	0(6)
O2	0.635(3)	0	0.061(2)	0(3)
O2'	0.655(5)	0	0.034(2)	7(6)
O3	0	0	0.202(2)	24(4)
O4	⅓	⅔	0.749(1)	13(3)

U(eq) is defined as one third of the trace of the orthogonalized U^{ij} tensor.

TABLE 11

Anisotropic displacement parameters (Å^2 × 10^3).[a]

	U^{11}	U^{22}	U^{33}	U^{23}	U^{13}	U^{12}
Y1	2(1)	2(1)	6(1)	0	0	1(1)
Y2	6(1)	6(1)	28(1)	0	0	3(1)
Mn/In	7(1)	5(1)	5(1)	0	−1(1)	2(1)
O3	33(7)	33(7)	5(7)	0	0	17(3)
O4	3(3)	3(3)	33(7)	0	0	1(1)

The anisotropic displacement factor exponent takes the form: $-2\pi^2[h^2 a^{*2}U^{11} + \ldots + 2hk a^* b^* U^{12}]$
[a]Split atoms O1, O1', O2, and O2' were refined with isotropic displacement parameters.

First-Principles Calculations

First-principles calculations were performed with plane-wave density functional theory using the Vienna Ab-initio Simulation Package (VASP). (Kresse, G., and Furthmueller, J., *Phys. Rev. B* 54, 11169-11186 (1996); Kresse, G., and Joubert, D., *Phys. Rev. B* 59, 1758-1775 (1999).) Exchange and correlation effects are treated on the level of LSDA+U, with an on-site Coulomb repulsion U=5.0 eV and an intra-atomic exchange splitting of J=0.5 eV for Mn d states. (Liechtenstein, A. I., Anisimov, V. I., and Zaanen, J., *Phys. Rev. B* 52, R5467-R5470 (1995).) A global antiferromagnetic ordering with ferromagnetic Mn planes was adopted for the simulations. Intermediates within periodic boundary conditions were studied using the supercell approach with lattice constants taken from experimental values presented in FIG. 7. The 40-atom supercells permit concentrations of x=0.0, 0.25, 0.5, 0.75 and 1.0 while maintaining equal numbers of In and Mn atoms in each layer. An ordering was chosen in which the minority component was maximally separated in space; thus, the possibility of In or Mn clustering was ignored.

All structures were initialized in the centrosymmetric $P6_3$/mmc space group and all atomic degrees of freedom were optimized until forces were less than 0.1 meV/Å. This strict tolerance allowed accurate study of delicate features of the atomic structure, such as tiltings of the polyhedra that are responsible for ferroelectricity. (Fennie, C. J., and Rabe, K. M., *Phys. Rev. B* 72, 100103(4) (2005).)

Summary of Convergence Parameters:

1) 450 eV plane wave cutoff (33.1 Ry, 16.5 Ha)

Example paper: Levy et al. (*Science* 1985)

- 501 "front-page" patent citations
- 1,857 in-text patent citations

Inhibition of Calcification of Bioprosthetic Heart Valves by Local Controlled-Release Diphosphonate

Abstract. *Bioprostheses fabricated from porcine aortic valves are widely used to replace diseased heart valves. Calcification is the principal cause of the clinical failure of these devices. In the present study, inhibition of the calcification of bioprosthetic heart valve cusps implanted subcutaneously in rats was achieved through the adjacent implantation of controlled-release matrices containing the anticalcification agent ethanehydroxydiphosphonate dispersed in a copolymer of ethylene-vinyl acetate. Prevention of calcification was virtually complete, without the adverse effects of retarded bone and somatic growth that accompany systemic administration of ethanehydroxydiphosphonate.*

Bioprostheses fabricated from porcine aortic valves treated with glutaraldehyde are widely used to replace diseased cardiac valves (*1–3*). More than 300,000 of these bioprostheses have been used in clinical implants since 1971 (*4–7*). Cuspal calcification is the most frequent cause of the clinical failure of these devices (*4–6*) and necessitates removal of the prostheses after 5 years in more than 50 percent of the children (*5*) and 5 to 10 percent of the adults receiving the valve implants (*2, 6, 7*). The pathogenesis of cuspal calcification is not completely understood (*7*), and there is no effective therapy for its prevention (*7*). Its pathophysiology can be reproduced with orthotopic valve replacements in sheep (*8*) or calves (*9*), or with subcutaneous cuspal implants in rabbits (*10*), mice (*11*), or rats (*12*). Subcutaneous implants are used because the time course of the mineral-

Patent citation data: Reliance on Science

- In-text citations are only recently available via Reliance on Science (Marx and Fuegi 2020; 2022)
- Crosswalk Microsoft Academic Graph to Web of Science publications primarily using distinct object identifiers (DOIs)
- Arguably the most social welfare-relevant summary measure, although we also look at other outcomes which are potential mechanisms

July 19, 2022

Dataset Open Access

Reliance on Science in Patenting

Marx, Matt; Aaron Fuegi

This dataset contains both front-page and in-text citations from patents to scientific articles through 2020. *If you use the data, please cite* these two articles:

1. M. Marx & A. Fuegi, "Reliance on Science by Inventors: Hybrid Extraction of In-text Patent-to-Article Citations." *forthcoming in Journal of Economics and Management Strategy.* (http://doi.org/10.1111/jems.12455)

2. M. Marx, & A. Fuegi, "Reliance on Science: Worldwide Front-Page Patent Citations to Scientific Articles" (2020), *Strategic Management Journal 41(9):1572-1594.* **(**https://onlinelibrary.wiley.com/doi/full/10.1002/smj.3145**)**

Additional commercialization outcomes

1. Patent-paper pairs ▸ more detail
2. Scientific Advisory Board (SAB) membership ▸ more detail

1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Measuring university affiliations of authors

In order to identify movers – authors of research papers whose university affiliations change over time – we first need to determine each author's organizational affiliation in each year

- Web of Science links authors to one or more organization names
- However, "multiple affiliations" problem frequently arises detail
- We instead infer moves based on observed changes in the e-mail addresses authors choose to list on their publications

Measuring university affiliations of authors: An example

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

Measuring university affiliations of authors: An example

XML:

```
<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>
```

SQL:

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.

Defining universities

- Given the central role of technology transfer offices in facilitating spillovers of academic research to commercial innovation, we define "universities" as technology transfer offices (TTOs)
- Focus on $\sim$ 300 US-based TTOs participating in the Association of University Technology Managers (AUTM) surveys (1991-2019)
  - Primarily ($\sim$ 90%) university-based TTOs (e.g. Stanford University)
  - Some university system-wide TTOs (e.g. University of California)
  - Some affiliated research hospitals reporting separately (e.g. Brigham and Women's Hospital reports separately from Harvard University)
  - Some stand-alone research hospitals (e.g. St. Jude Children's)

Details: Measuring university affiliations of authors

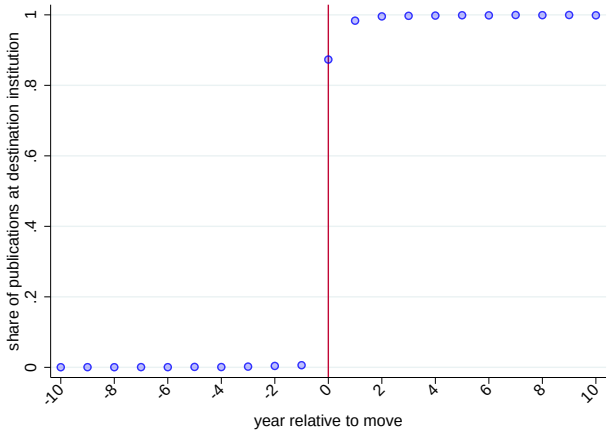
- Extracting data on e-mail domains:
  - Example: `dartmouth` from `heidi.lie.williams@dartmouth.edu`
  - Reported in Web of Science starting in 2000s ▸ reporting
  - Reported for $\sim 85\%$ of corresponding authors (vs. $\sim 15\%$)
  - Reported most often for first, solo, and last authors
- Link e-mail domains to universities (AUTM TTOs)
  - Example: how to know that `harvard.edu` and `hbs.edu` both map to the same technology transfer office at Harvard University
  - Crosswalk via Web of Science "preferred organizations" ▸ crosswalk
- For each author-article pair reporting an e-mail address, we can thus assign a unique AUTM organization
  - If within a year an author with multiple publications is linked to more than one AUTM organization, assign to most frequent
  - In case of ties, select at random
  - Generates a single affiliation for each author-year

Imputing moves and building our mover sample

- In any year in which an author is unobserved in the Web of Science (that is, in which they have no recorded publications), we assume authors maintain their most recent affiliation
- Move year is t if author's affiliation in t differs from t-1
- Sample restrictions:
  - Drop authors who change affiliation after a period of being unobserved (unclear what year to assign the move)
  - Drop authors who move more than once from our analysis
  - Only focus on authors in the biomedical sciences
  - Over 14k movers in total (1.7m mover-years)
  - More detail

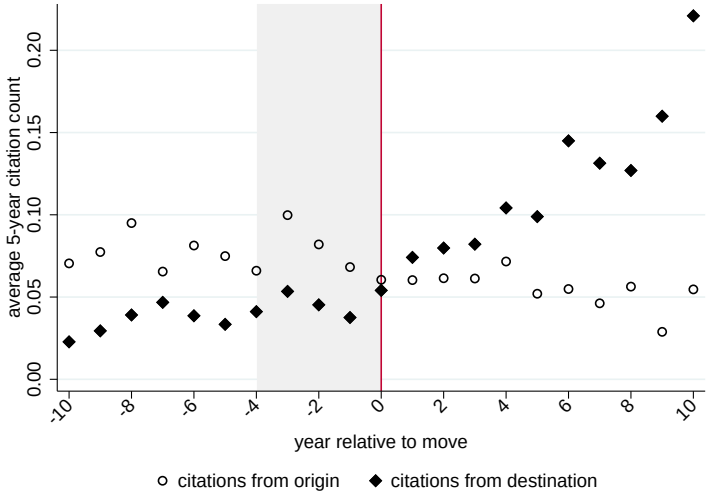
1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Share of publications associated with destination institution



Note: Sample restricted to origin and destination publications.

Citations from origin and destination

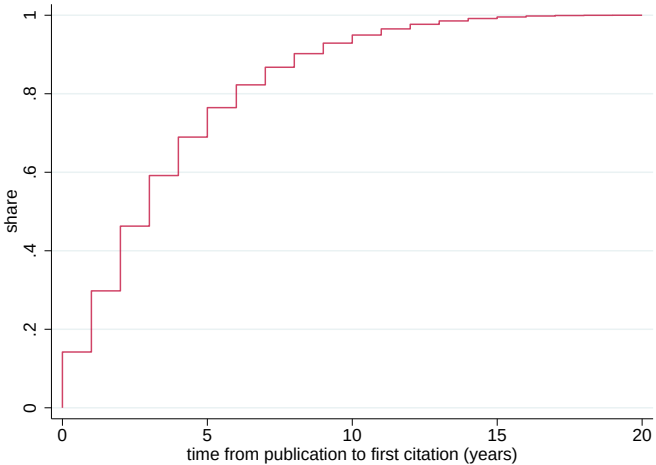


Author-year panel: Summary statistics

	(1) Non-movers mean	(2) Movers mean
Author metrics		
Total publications	12.52	60.18
Career length	17.05	22.92
Team size per publication	12.55	7.26
USPTO metrics		
Total USPTO patent citations	7.37	38.87
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 patent	0.06	0.05
USPTO citations per publication	0.38	0.47
Outcome measures		
Annual 5-year patent cites	0.63	1.43
Annual 5-year predicted patent cites	0.43	0.94
Annual 5-year paper cites	30.06	75.99
Annual 5-year predicted paper cites	18.20	41.64
Annual patent-paper pairs	0.74	1.98
Annual SAB memberships	0.00	0.00
Observations	500,296	14,242

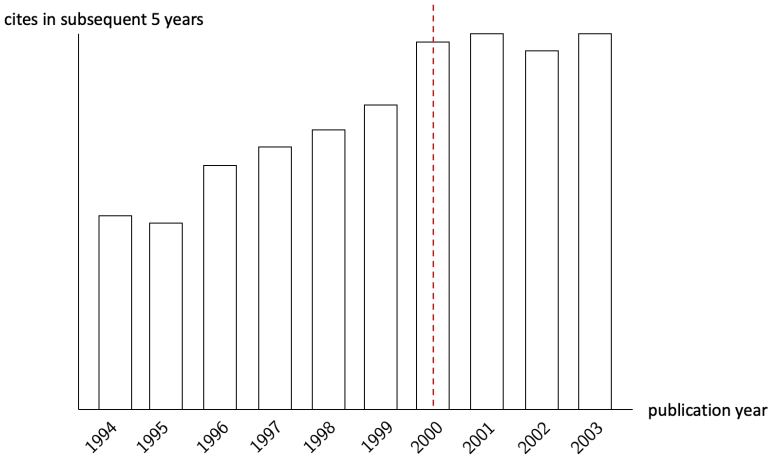
1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Challenge: Patent citations accrue slowly over time



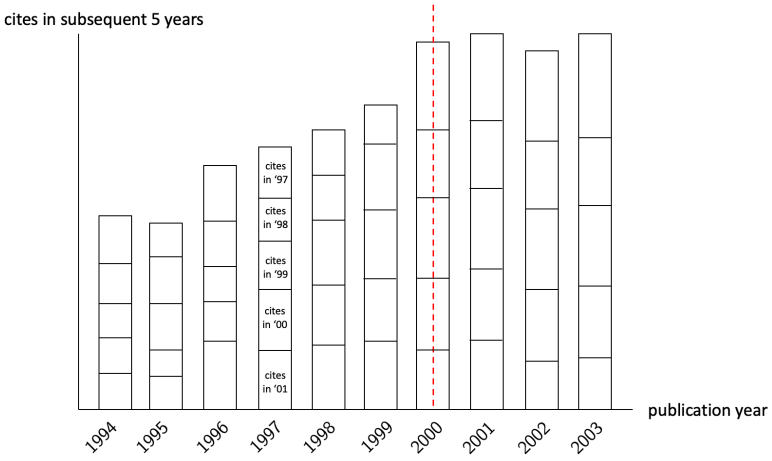
Pre-period contamination

Using cumulative cites poses a problem



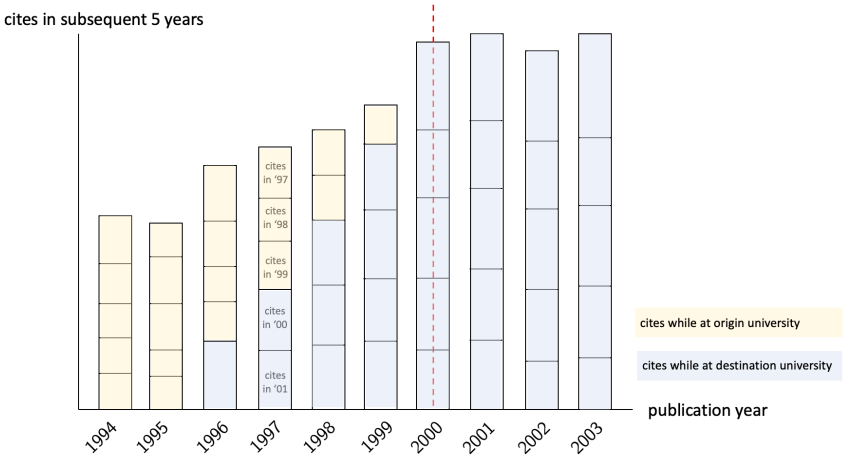
Pre-period contamination

Using cumulative cites poses a problem



Pre-period contamination

Using cumulative cites poses a problem



Solution: Predict citations at the time of publication

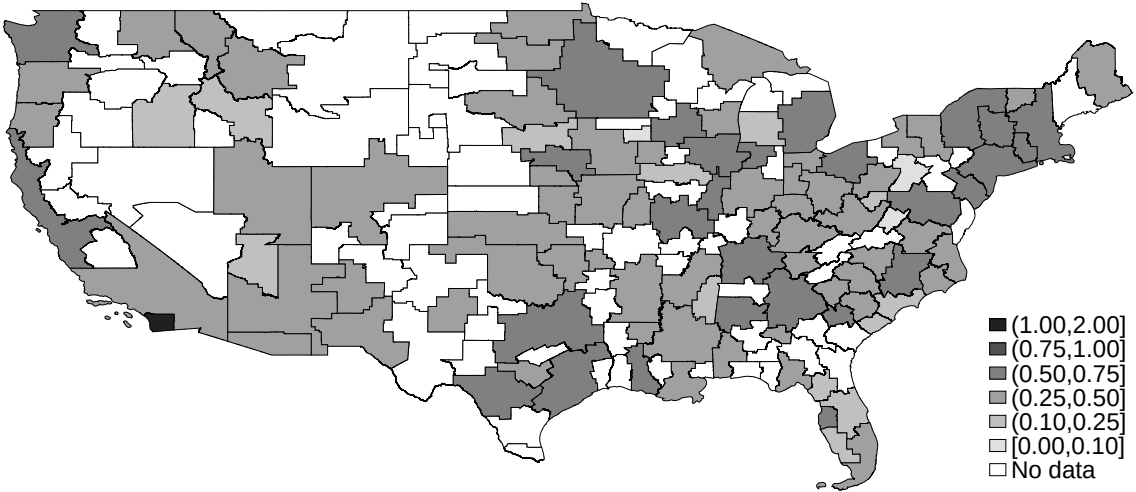
- We predict 5-year patent citations based on journal fixed effects
- Example: Papers in *Nature Biotechnology* are around 3 times more likely to be cited in a patent in the five years after publication than are scientific papers published in *PNAS*
- Inverse hyperbolic sine transformation

Important interpretation point: this outcome restricts us to observing an effect that operates via researchers changing the content of their research and/or the journals they target

Journals with the highest mean 5-year patent count

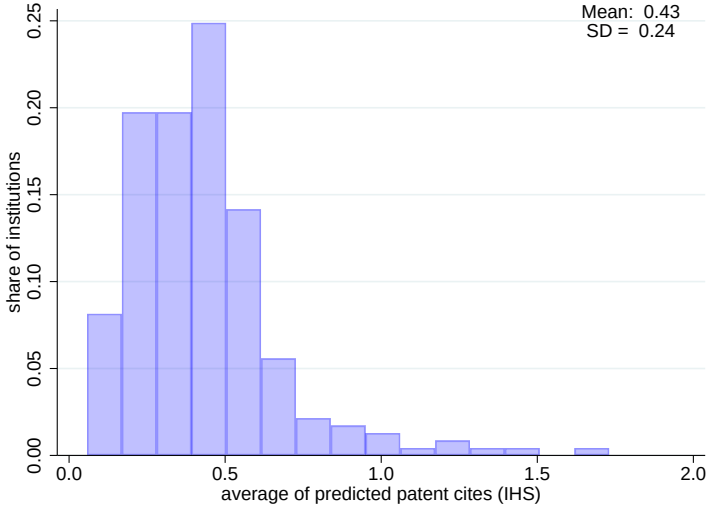
rank	*journal*	*publication count*	*mean patent citations*
1	2008 SID International Symposium, Technical Digest	545	3.00
2	2007 SID International Symposium, Technical Digest	478	2.27
3	Annual Review Of Immunology	270	2.26
4	IEE International Electron Devices Meeting 2000, Technical Digest	201	2.25
5	ACM Transactions On Computer Systems	226	2.09
6	mAbs	1220	2.00
7	IEEE International Electron Devices Meeting 2004, Technical Digest	251	1.97
8	Nature Biotechnology	8493	1.93
9	2007 IEEE International Electron Devices Meeting	240	1.91
10	IBM Systems Journal	502	1.78

Map of commercialization propensities



▸ Stability over time

Commercialization propensities across institutions



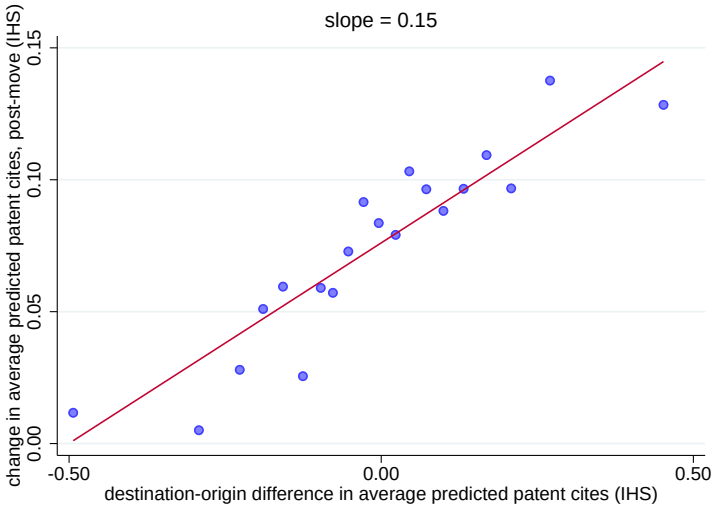
Variation in research commercialization

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

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

1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Change in predicted patent citations by size of move



1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Empirical framework

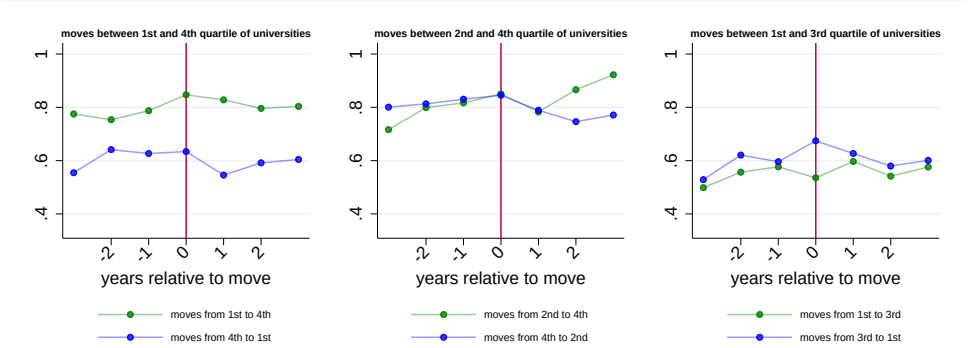
Analogous to Finkelstein et al. (2016), author i at university u in year t:

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

- α is an individual fixed effect
- γ is a university fixed effect
- τ is a calendar year fixed effect

Additive separability

In an effort to test the additively separable model, we look at movers who go between schools near the top / bottom of the commercialization distribution [Card-Heining-Kline 2013]



Symmetric responses to equal but opposite moves suggest additively separable model is a reasonable approximation

Additive decomposition

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

Share of the difference between u and u' attributable to the place is:

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

while the share attributable to differing researchers is:

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

Decomposition: Results from estimation

	Above/Below Median	*Top & Bottom 25%*	*Top & Bottom 10%*	*Top & Bottom 5%*
Difference in IHS commercialization				
Overall	0.36	0.44	0.64	0.82
Researchers	0.31	0.34	0.49	0.75
Universities	0.05	0.10	0.15	0.07
Share of difference attributable to				
Researchers	0.86	0.77	0.76	0.92
Universities	0.14	0.23	0.24	0.08
	($\pm$0.07)	($\pm$0.08)	($\pm$0.10)	($\pm$0.11)

1. Data and measurement
   - Measuring academic research
   - Measuring commercial linkages to academic research
   - Measuring university affiliations and cross-university moves
2. Descriptive statistics
   - Descriptive statistics: Movers and their moves
   - Descriptive statistics: Commercialization
   - Descriptive statistics: On-move changes in commercialization
3. Empirical strategy and results
   - Additive decomposition
   - Event study estimation

Size and direction of move

Authors are making different types of moves:

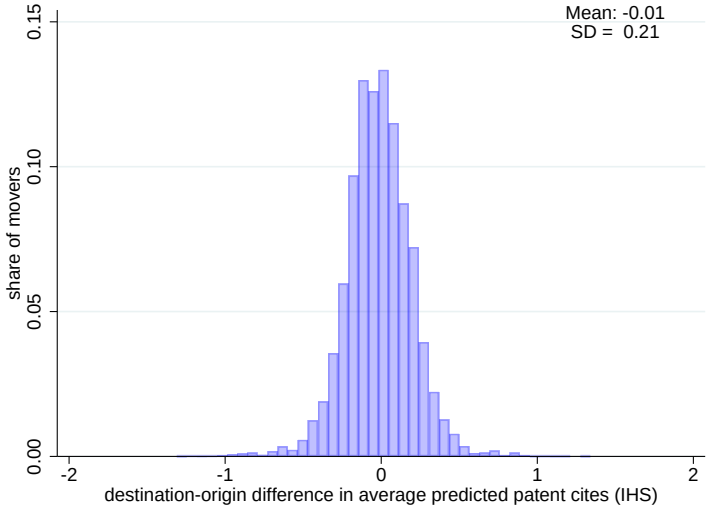
- In our simple example, if half of the researchers went from u to u' and the other half from u' to u, any effects would cancel each other out if we simply plotted an event study graph
- We want to scale by size and direction of move

For an author i who moves from university $o(i)$ to $d(i)$:

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

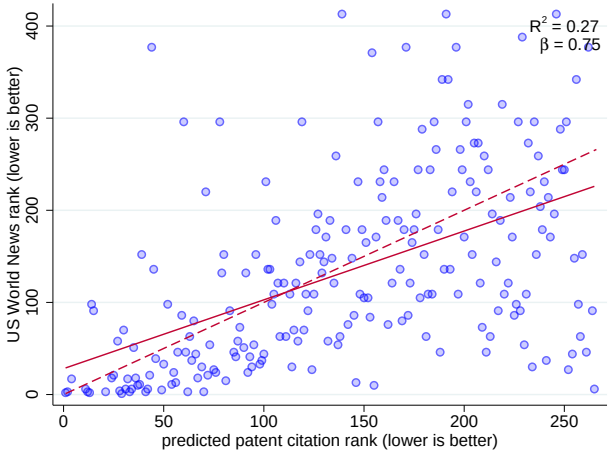
University-level measures are created by averaging across faculty-years

Distribution of destination-origin differences



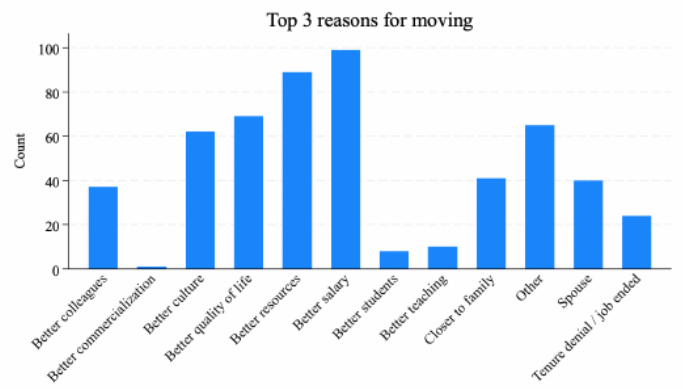
Additional data on moves

- Surveys suggest most academics do not know the royalty-sharing policies at their own institutions [Love 2014]
- Commercialization propensity not highly correlated with school rank



Survey evidence on movers

In addition, we surveyed 3,000 of our movers, asking *why* they moved (received 311 responses)



Commercialization opportunities were the least likely factor to be cited

Event study estimating equation

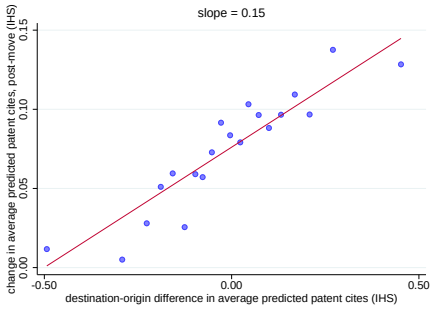
Let $r(i,t)$ define years relative to the move. Our estimating equation:

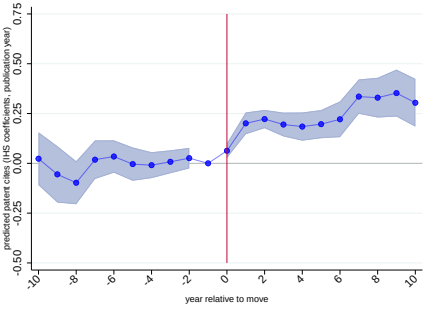
$$y_{it} = \hat{\alpha}_i + \theta_{r(i,t)}\hat{\delta}_i + \tau_t + \varepsilon_{it} \tag{6}$$

$\hat{\theta}_{r(i,t)}$'s are the coefficients of interest

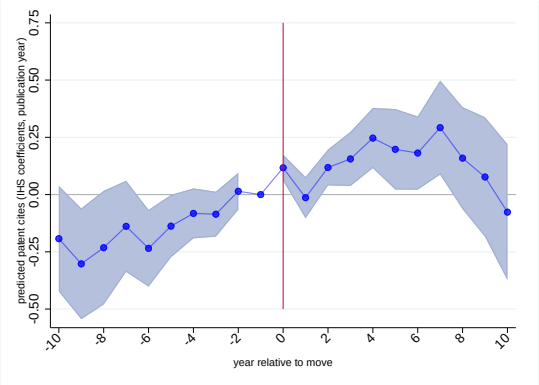
- Should be zero if $r < 0$
- Tell us the degree of convergence when $r \geq 0$

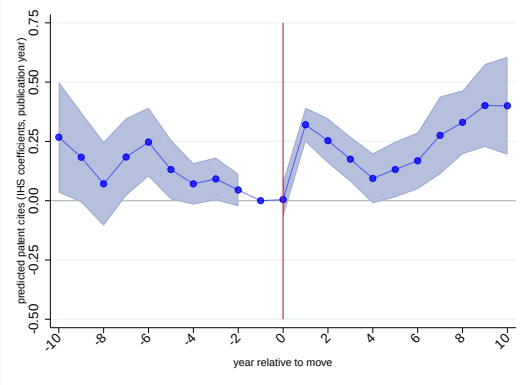
Estimates: Predicted patent citations



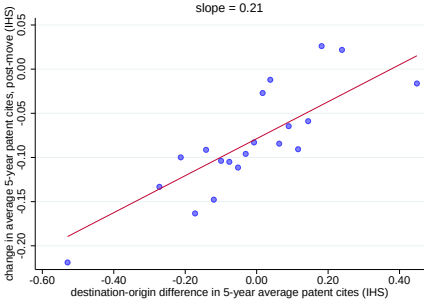


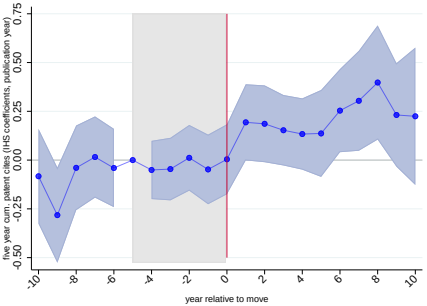
Estimates: Up moves vs. down moves





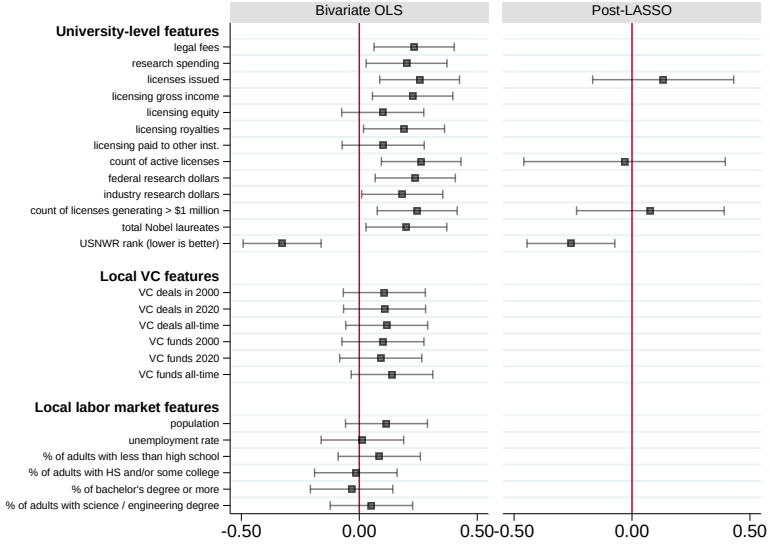
Estimates: Observed patent citations





- Alternative citation outcomes: ▸ front page citations ▸ in-text citations
- Additional outcomes: ▸ SAB ▸ patent-paper pairs
- Robustness: ▸ last author

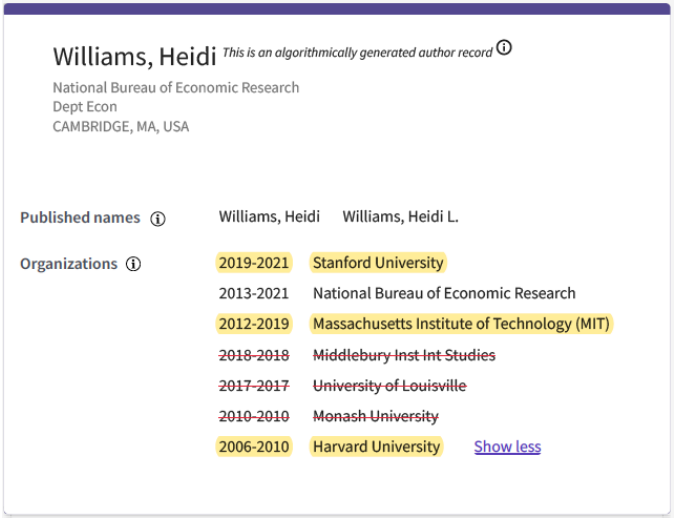
Correlates of high-commercialization place effects



Conclusion

- Do academic discoveries made at different universities have a similar propensity to spill over to commercial innovation?
- Our estimates: 15-30% of geographic variation in spillovers from university-based research is attributable to place-specific factors

Author disambiguation



Return

Measuring venture capital-backed patents

- Data from Refinitiv on all U.S. firms that raised venture capital funds from 1968-2021 [Akcigit et al. 2020; Bernstein et al. 2016]
  - Firm name-patent assignee linkage, also using geocode information
- Define venture-backed patents as those that were applied for within five years of the firm's first round of financing

Return

Measuring patent-paper pairs

Building on work by Ducor (2000), Murray (2002), Murray-Stern (2007), we measure the propensity of authors' published scientific research to be part of a "patent-paper pair"

- A given idea can be disclosed in both a scientific paper and a patent
- While rare in some fields (e.g. computer science), patent-paper pairs are increasingly common in the life sciences
- Key advantage: timing
  - Patenting is generally concomitant with publication
  - One-year grace period during which the novelty of an invention is not destroyed by prior publication (35 U.S.C. §102(b))

Return

Measuring patent-paper pairs: An example

"A method has been developed for control of molecular weight and molecular weight dispersity during production of polyhydroxyalkanoates in genetically engineered organisms by control of the level and time of expression of one or more PHA synthases in the organisms. The method was demonstrated by constructing a synthetic operon for PHA production in E. coli …Modulation of the total level of PHA synthase activity in the host cell by varying the concentration of the inducer…was found to effect the molecular weight of the polymer produced in the cell." (Snell; Kristi D. (Belmont, MA); Hogan; Scott A. (Troy, MI); Sim; Sang Jun (Seoul, KR); Sinskey; Anthony J. (Boston, MA); Rha; Chokyun (Boston, MA) 1998, Patent No. 5,811,272)

"A synthetic operon for polyhydroxyalkanoate (PHA) biosynthesis designed to yield high levels of PHA synthase activity in vivo was constructed …by positioning a genetic fragment … behind a modified synthase gene containing an Escherichia coli promoter and ribosome binding site. Plasmids containing the synthetic operon …were transformed into E. coli DH5 alpha and analyzed for polyhydroxybutyrate production… Comparison of the enzyme activity levels of PHA biosynthetic enzymes in a strain encoding the native operon with a strain possessing the synthetic operon indicates that the amount of polyhydroxyalkanoate synthase in a host organism plays a key role in controlling the molecular weight and the polydispersity of polymer." (Sim SJ, Snell KD, Hogan SA, Stubbe J, Rha CK, Sinskey AJ, Nature Biotechnology 1997)

Note: Example drawn from Murray-Stern (2007).

▸ Return

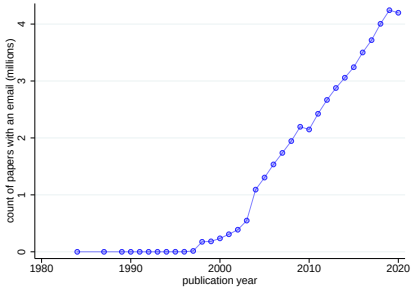
Measuring scientific advisory board memberships

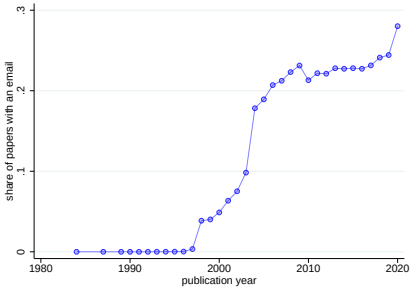
Building on work by Ding-Murray-Stuart (2013), we measure the propensity of authors to join corporate scientific advisory boards (SABs)

- Typically formed very early in a company's development
- 5-10 members who join by invitation
- Very low exit/turnover
- Paid as a fee and granted stock options
- Provide expertise and signal scientific quality to external inventors
- Data from Capital IQ
  - Name/affiliation-based match on author-year panel
- Examples: Langer has 45 SAB memberships; Bertozzi has 5

Return

E-mail addresses in the Web of Science data





Note: The sample is all author-year observations in the Web of Science data (N=300,691,264).

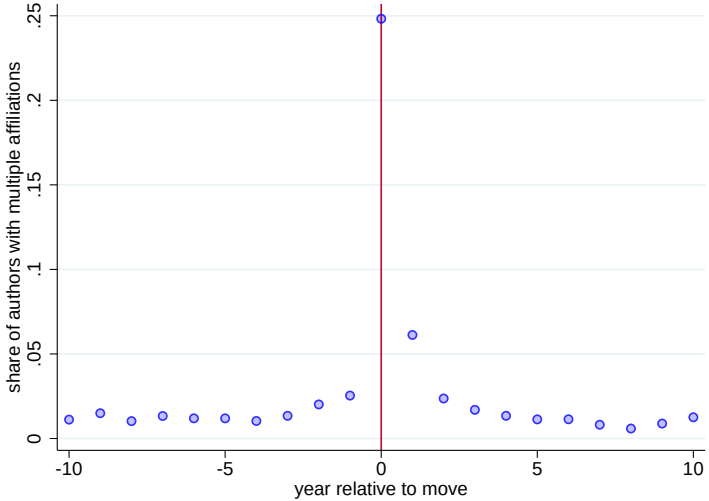
Return

Linking e-mail domains to organizations

- Web of Science lists author-publication-level organization names
- Clarivate Analytics attempts to standardize these organization names into "preferred organizations"
  - Example: preferred organization for Harvard University aggregates institutions such as Harvard Stem Cell Institute, Harvard School of Engineering and Applied Sciences, and Harvard Society of Fellows
- For each e-mail domain, we select the Web of Science preferred organization most frequently associated with that domain
  - Exception: manually correct e.g. `bws.harvard.edu`, which our method would default to pool with Harvard University but which should be kept separate because the hospital has its own independent TTO
- Link Web of Science preferred organizations to AUTM TTOs by hand

Return

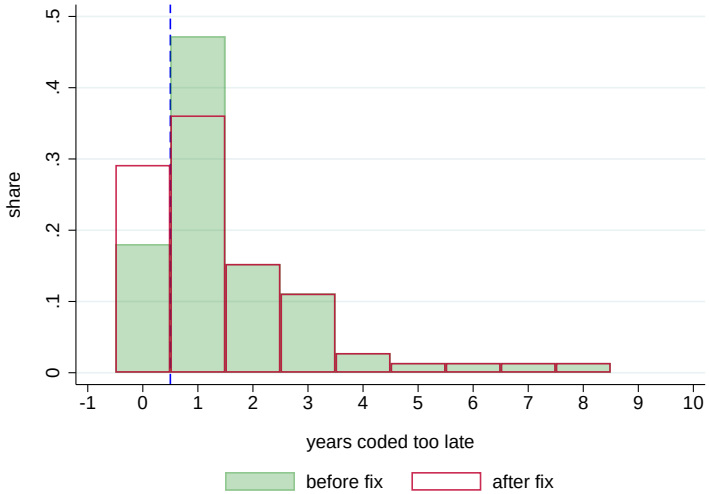
Share of authors with multiple affiliations



Return

Move timing correction

For authors whose move we have coded one year too late



Return

Sample construction

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]	4,169,556
Author-years after randomly breaking ties [d]	4,138,812
Author-years after including only biomedical [e] authors	1,722,456
Author-years after including only non-movers and movers ± 10 years around move	1,711,406

Return

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

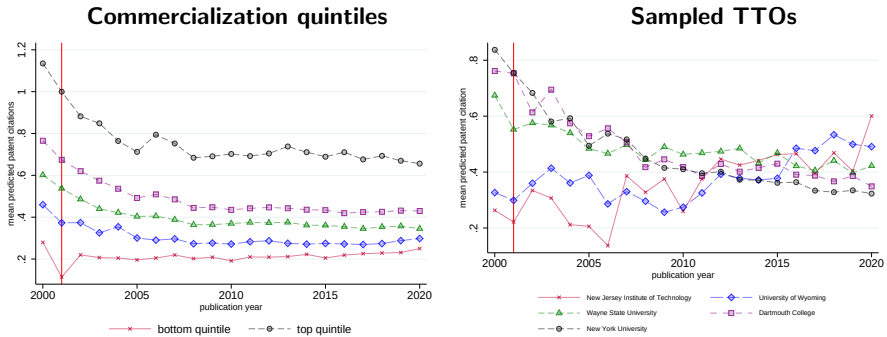
[b] Note the vast majority ($> 99\%$) of observations discarded here are non-.edu emails

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

[d] 61,063 author-years appear with multiple affiliations. 59,810 are two-way ties, 1,209 are three-way ties, and 44 are four-way ties. This yields 30,744 author-years with at least two conflicting affiliations

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

Stability of predicted citations over time



Return

Scaling effect sizes

Given these different "sized" moves, it is helpful to scale the effect sizes:

$$y_{it}^{scaled} = \frac{y_{it} - \bar{y}_{o(i)}}{\delta_i} = S_{uni} \tag{7}$$

- If $y_{it}^{scaled} = 0$, then i does not respond to the move
- If $y_{it}^{scaled} = 1$, then i commercializes at avg rate of his destination
- If $y_{it}^{scaled} \in (0, 1)$, then i adopts some (but not all) of the university's commercialization propensity

Estimating equation

Start with the data-generating process:

$$y_{iut} = \alpha_i + \gamma_u + \tau_t + \varepsilon_{iut} \tag{8}$$

Let $r(i,t)$ define years relative to the move. Substitute for γ_u:

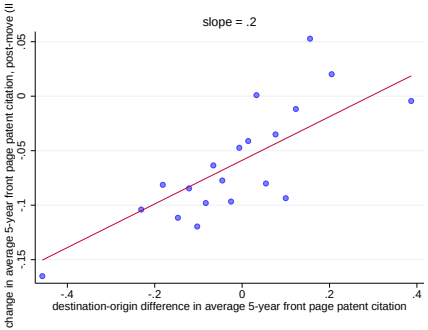
$$y_{it} = \alpha_i + \gamma_{o(i)} + I_{r(i,t)>0} S^i_{uni} \delta_i + \tau_t + \varepsilon_{it} \tag{9}$$

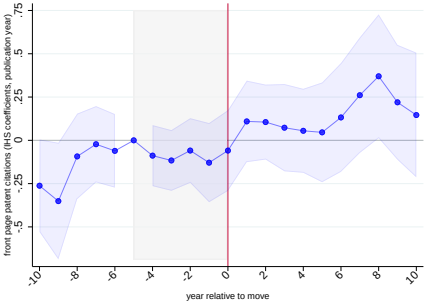
Combine $\alpha_i + \gamma_{o(i)}$ into single FE; plug in sample analogues:

$$y_{it} = \hat{\alpha}_i + \theta_{r(i,t)} \hat{\delta}_i + \tau_t + \varepsilon_{it} \tag{10}$$

Return

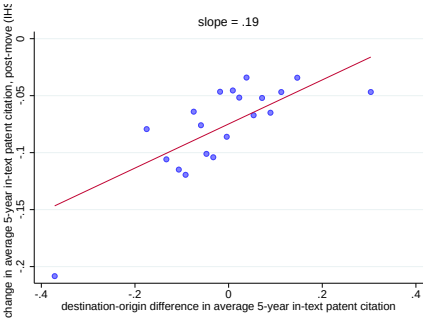
Front page patent citations

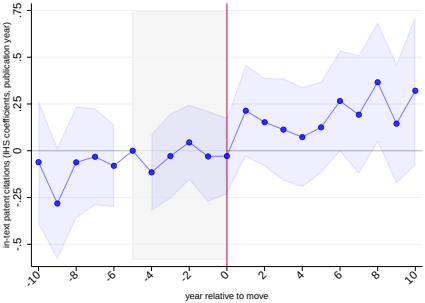




Return

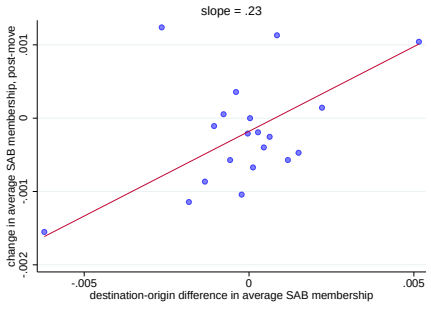
In-text patent citations

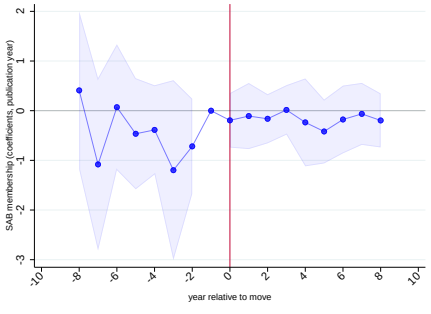




Return

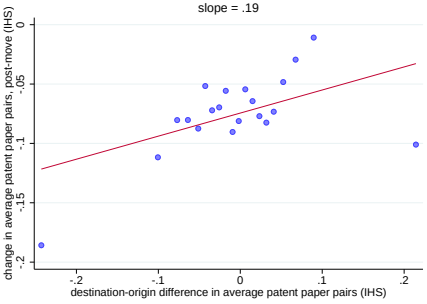
SAB membership

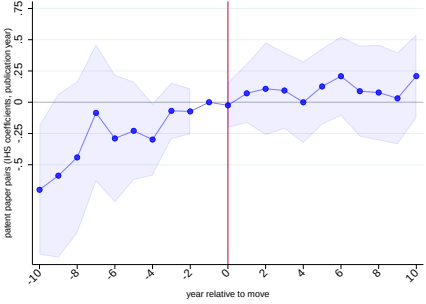




Return

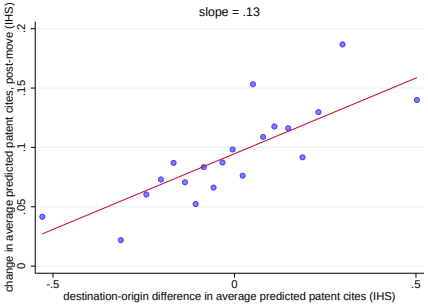
Patent-paper pairs

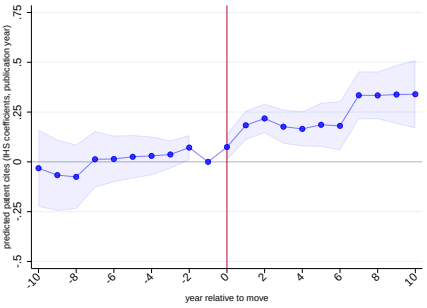




Return

Last authors: A proxy for principal investigators





Return