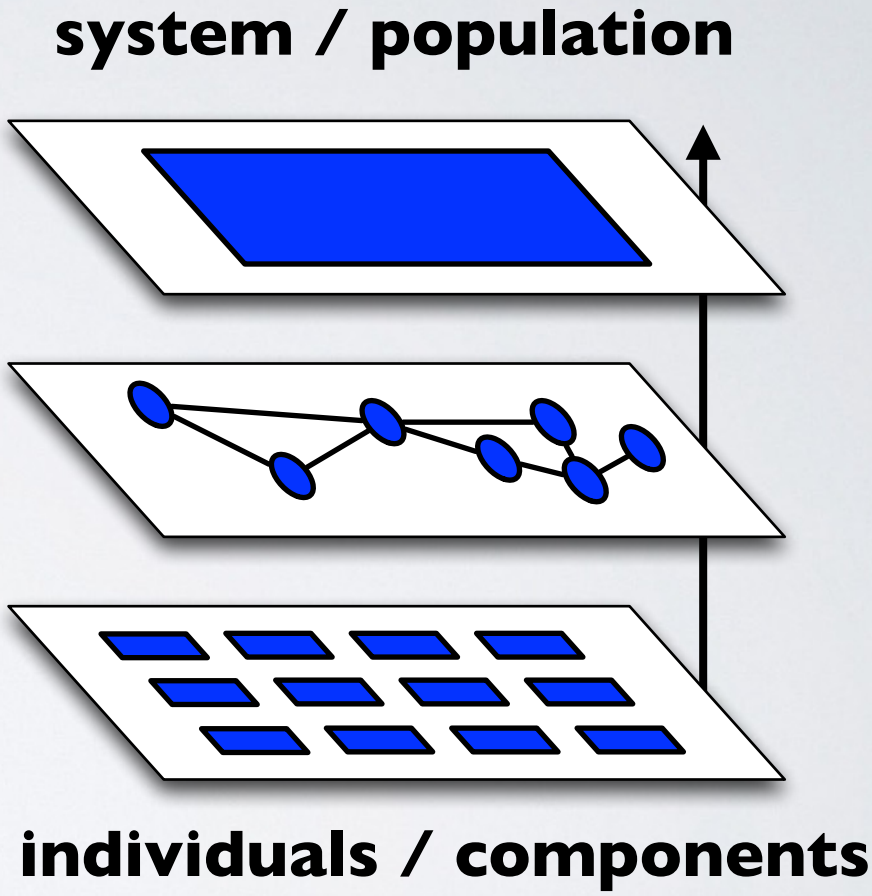
Networks, Prestige, and the Spread of Scientific Ideas

Aaron Clauset
@aaronclauset
Computer Science Dept. & BioFrontiers Institute
University of Colorado, Boulder
External Faculty, Santa Fe Institute

what are networks?

what are networks?

- an approach
- representation of structural complexity
- connect "micro" to "macro"
- *structure above*
  individuals / components
- *structure below*
  system / population



learning goals

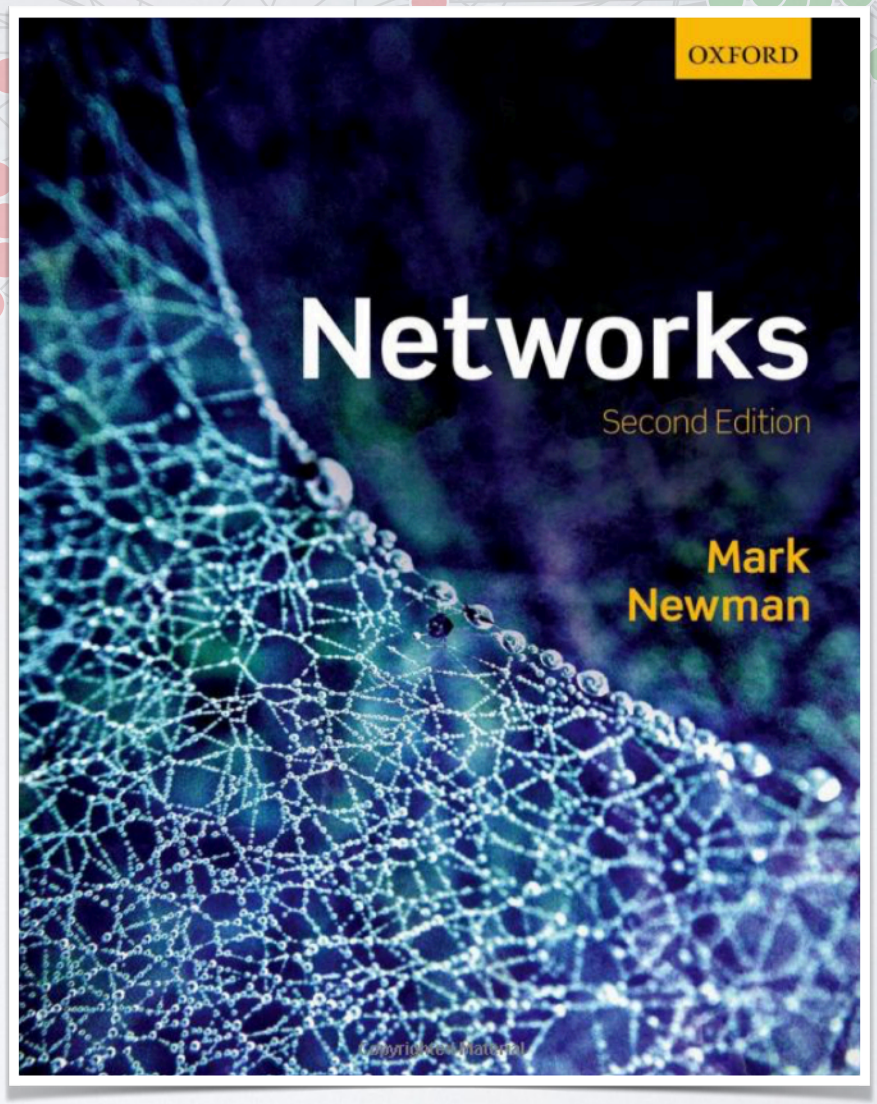
- build intuition
- expose key concepts
- highlight some big questions
- use "science" as a model complex system
- apply network ideas
- not a substitute for technical coursework

Mark Newman

Professor of Physics
University of Michigan

External Faculty
Santa Fe Institute

http://www-personal.umich.edu/~mejn/



University of Colorado **Boulder**

Network Analysis and Modeling

Instructor: Aaron Clauset *or* Daniel B. Larremore

This graduate-level course will examine modern techniques for analyzing and modeling the structure and dynamics of complex networks. The focus will be on statistical algorithms and methods, and both lectures and assignments will emphasize model interpretability and understanding the processes that generate real data. Applications will be drawn from computational biology and computational social science. No biological or social science training is required. (Note: this is not a scientific computing course, but there will be plenty of computing for science.)

Full lectures notes online (~150 pages in PDF)

https://aaronclauset.github.io/courses/5352/

University of Colorado **Boulder**

Biological Networks

Instructor: Aaron Clauset

This undergraduate-level course examines the computational representation and analysis of biological phenomena through the structure and dynamics of networks, from molecules to species. Attention focuses on algorithms for clustering network structures, predicting missing information, modeling flows, regulation, and spreading-process dynamics, examining the evolution of network structure, and developing intuition for how network structure and dynamics relate to biological phenomena.

Full lectures notes online (~150 pages in PDF)

https://aaronclauset.github.io/courses/3352/

Software

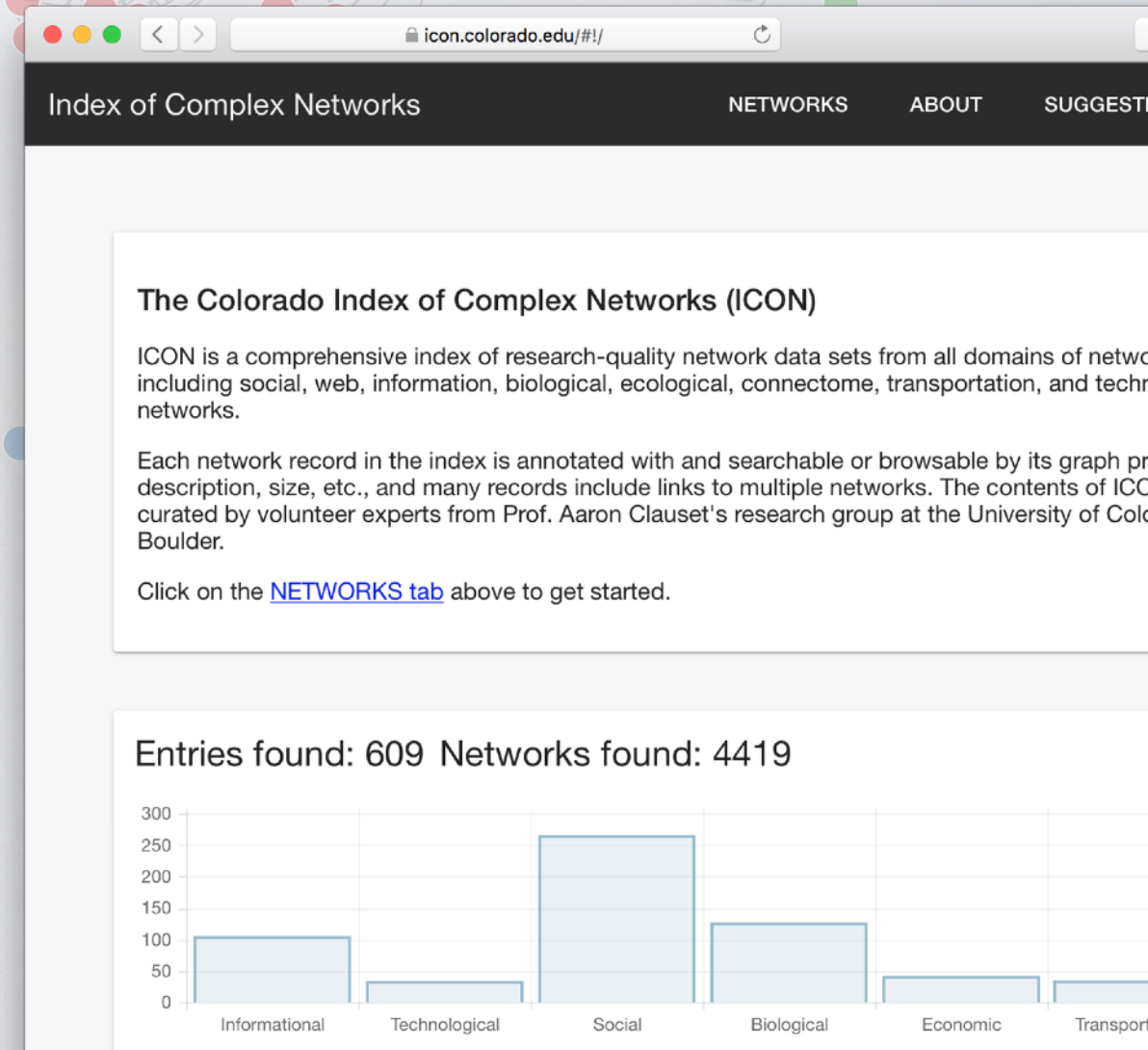
R
Python
Matlab
NetworkX [python]
graph-tool [python, c++]
GraphLab [python, c++]

Standalone editors

UCI-Net
NodeXL
Gephi
Pajek
Network Workbench
Cytoscape
yEd graph editor
Graphviz

Network data sets

Colorado Index of Complex Networks



little in academia makes sense except in the light of *prestige*

little in academia makes sense except in the light of *prestige*

prestige *pervades and structures* the scientific ecosystem

- shapes who gets jobs and where
- shapes who publishes where
- shapes who works on what ideas
- shapes who gets what resources (grants, students, attention, etc.)
- etc.

🤔 *how many good ideas are left unexplored because they lack the right markers of prestige?*

prestige in the sense of Burris (2004), as a form of social capital

what *is* prestige?

- prestige is related to *rankings*, but
- everyone agrees that USNews and NRC rankings are bad 
  [these measure *inputs* to science, not *outputs* or *value-added*]

NRC = National Research Council
vast literature: Bastedo & Bowman (2010), Cole (2011), Webster (2001), Leef & Lowrey (2004), Meredith (2004), O'Meara (2007), Tancredi et al. (2013), Grimson (2010)

what *is* prestige?

- prestige is related to *rankings*, but
- everyone agrees that USNews and NRC rankings are bad 🤢
- we can extract a *data-driven* ranking from the *network of who hires whose graduates as faculty* 👏

  vertices are PhD-granting universities

  consumers $\longleftrightarrow$ producers

  v hires from u, add an edge $u \longrightarrow v$

- *prestige* → *placement power* → *the ability to place your graduates as faculty at other universities*

- to measure prestige, we need faculty hiring data

prestige in the sense of Burris (2004), as a form of social capital; many papers have analyzed faculty hiring networks since, all showing similar results


COLLECT
ALL THE DATA
memegenerator.net

collect all the data (2015)

complete, hand-curated data

for 19,000 tenure-track faculty across 461 departments in

- Computer Science (205 depts)
- Business (112)
- History (144)

roughly 4000 hours of manual data collection

```
>>> record 1059
# facultyName : James H. Martin
# email       : [redacted]
# sex         : M
# department  : Computer Science
# place       : University of Colorado, Boulder
# current     : Full Professor
# [Education]
# degree      : BS
# place       : Columbia University
# field       : Computer Science
# years       : ????-1981
# [Education]
# degree      : PhD
# place       : UC Berkeley
# field       : Computer Science
# years       : ????-1988
# [Faculty]
# rank        : Assistant Professor
# place       : University of Colorado, Boulder
# years       : 1989-1995
# [Faculty]
# rank        : Associate Professor
# place       : University of Colorado, Boulder
# years       : 1995-2007
# [Faculty]
# rank        : Full Professor
# place       : University of Colorado, Boulder
# years       : 2007-2011
# recordDate  : 7/4/2011
```

CS data from 2011, Business schools from 2012, History from 2013
all data from public sources, mainly faculty CVs and homepages

collect all the data (2015)

complete, hand-curated data

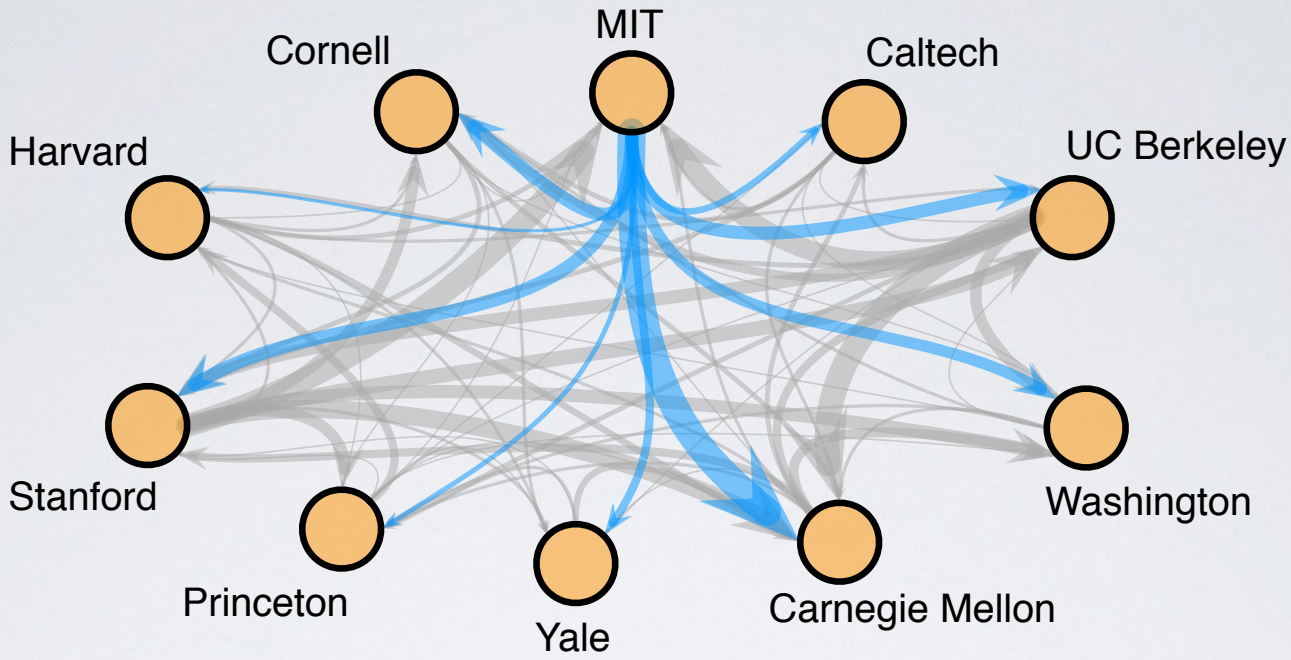
for 19,000 tenure-track faculty across 461 departments in

- Computer Science (205 depts)
- Business (112)
- History (144)

roughly 4000 hours of manual data collection

```
>>> record 1059
# facultyName : James H. Martin
# email       : [redacted]
# sex         : M
# department  : Computer Science
# place       : University of Colorado, Boulder
# current     : Full Professor
# [Education]
# degree      : BS
# place       : Columbia University
# field       : Computer Science
# years       : ????-1981
# [Education]
# degree      : PhD
# place       : UC Berkeley
# field       : Computer Science
# years       : ????-1988
# [Faculty]
# rank        : Assistant Professor
# place       : University of Colorado, Boulder
# years       : 1989-1995
# [Faculty]
# rank        : Associate Professor
# place       : University of Colorado, Boulder
# years       : 1995-2007
# [Faculty]
# rank        : Full Professor
# place       : University of Colorado, Boulder
# years       : 2007-2011
# recordDate  : 7/4/2011
```

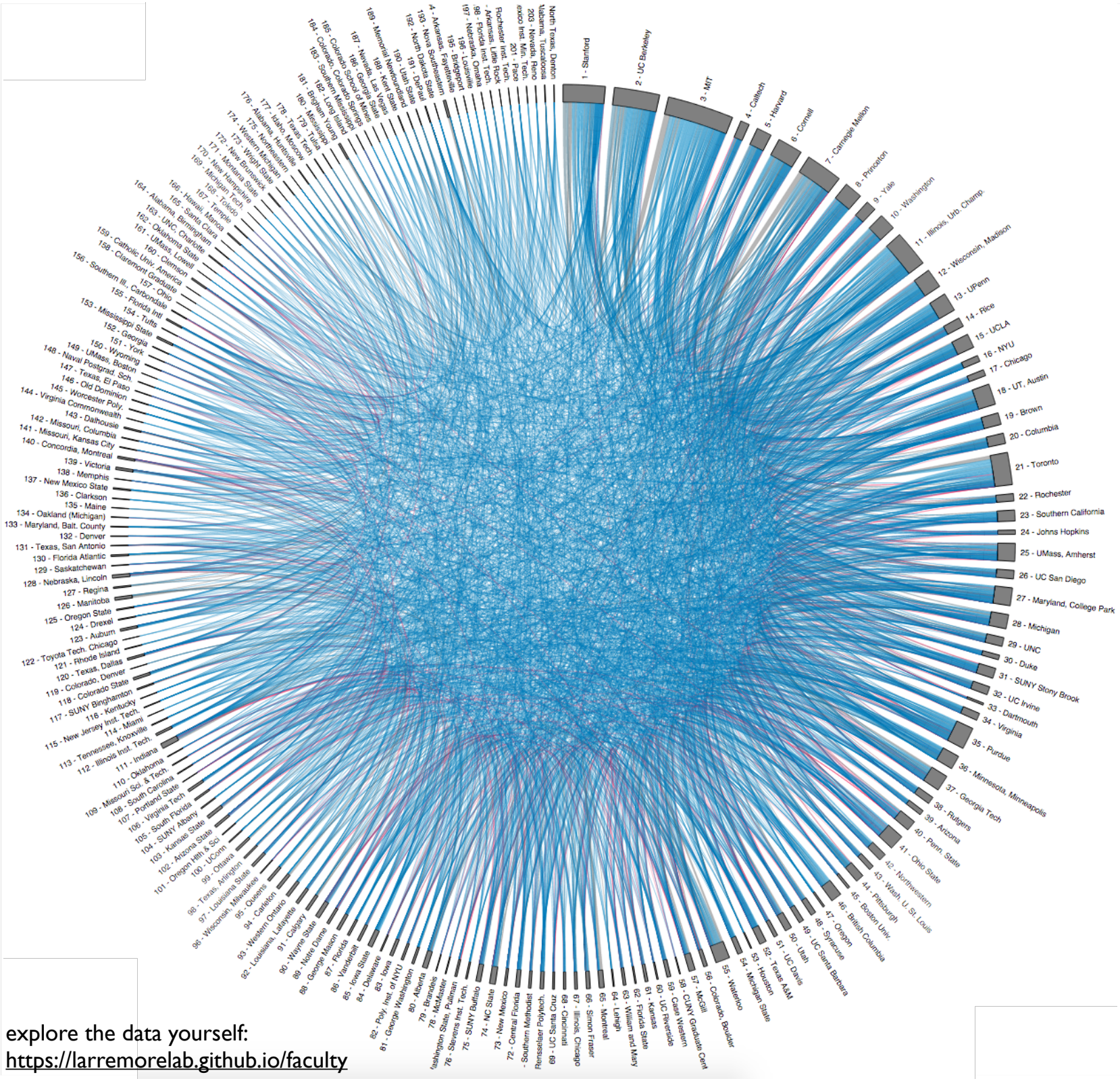
CS data from 2011, Business schools from 2012, History from 2013
all data from public sources, mainly faculty CVs and homepages



faculty hiring is a *network*

- vertices are PhD-granting universities
- consumers $\longleftrightarrow$ producers
- v hires from u, add an edge $u \longrightarrow v$

actual exchanges of 267 faculty among 10 elite CS departments, from our 2011 data, without self-hires



explore the data yourself:
https://larremorelab.github.io/faculty

huge inequalities in faculty production

huge inequalities in faculty production

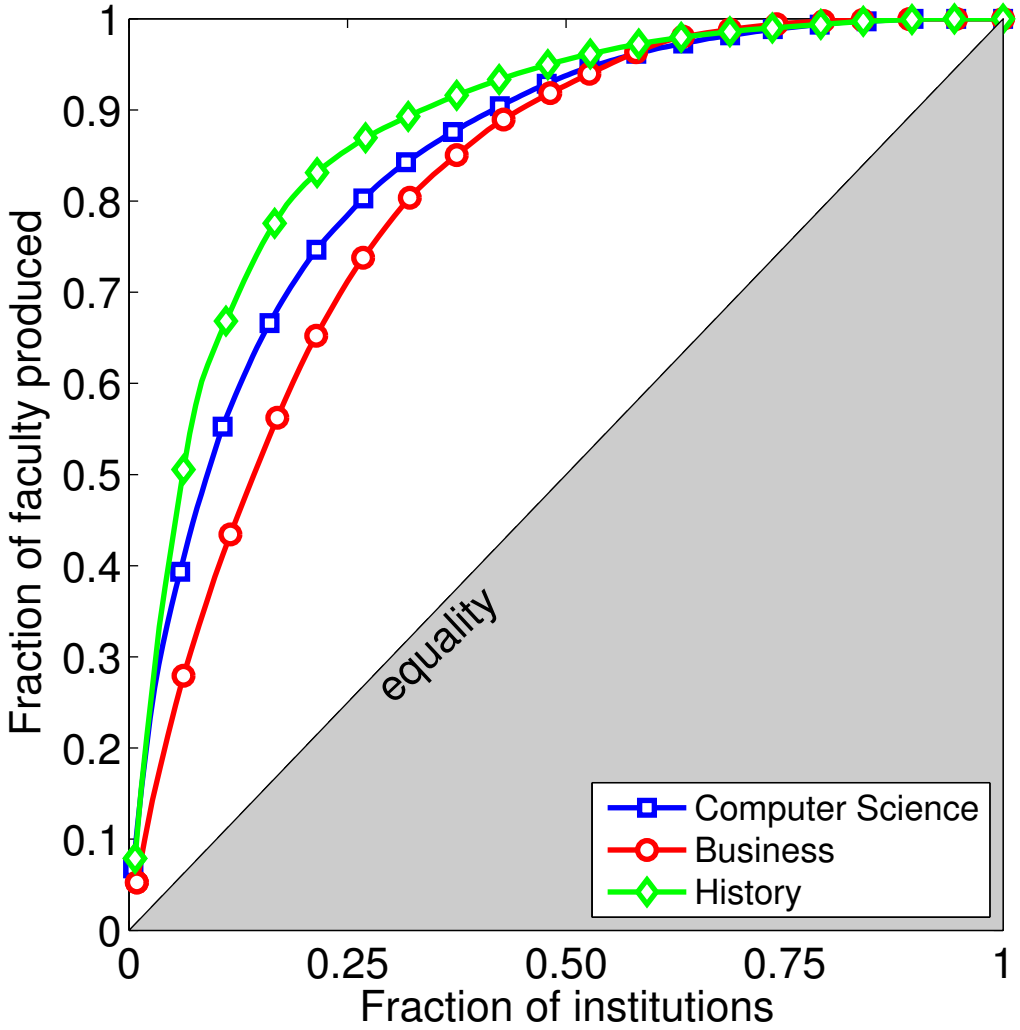
Computer Science (205 depts.)

faculty production, Gini = 0.69

18 depts. (9%) produced 50% of *all* faculty

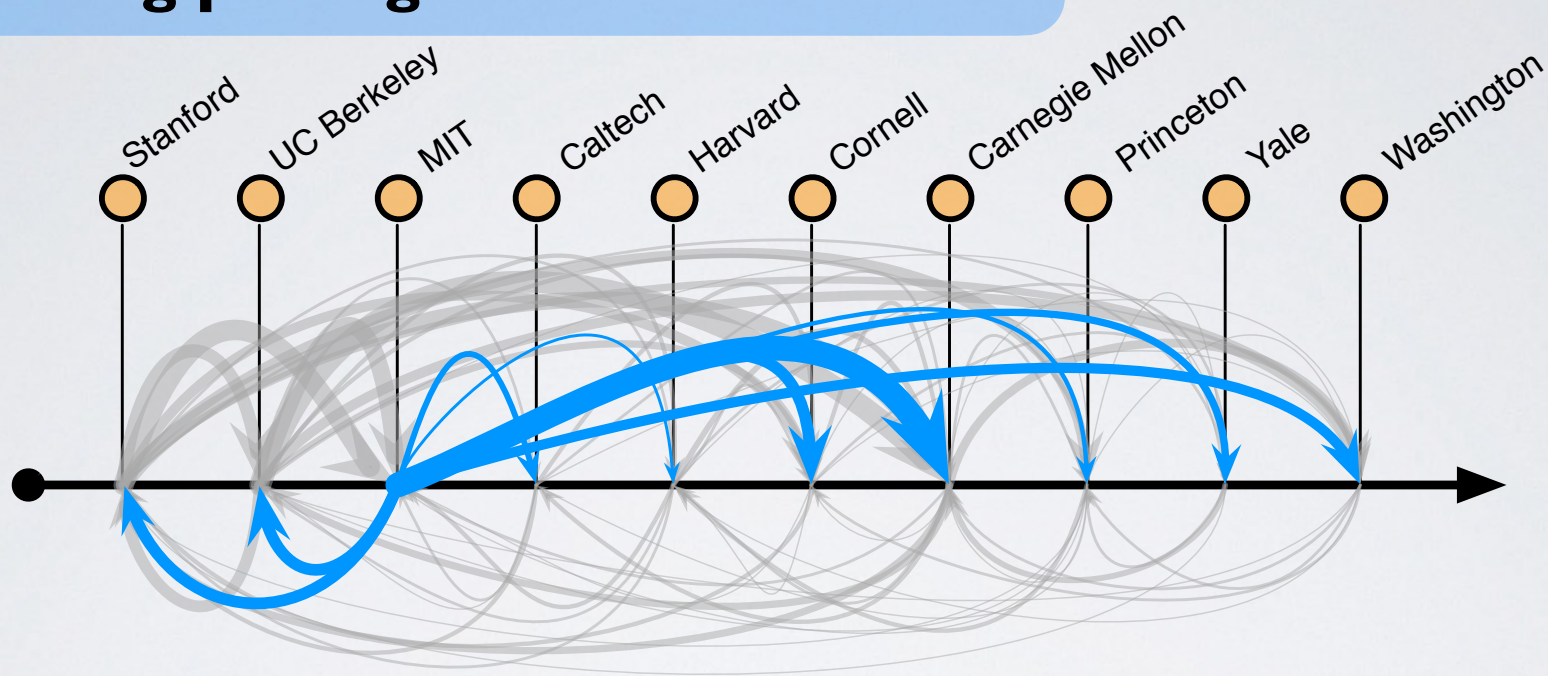
76% of depts. = net consumers
[more faculty hired than produced]

depts. ranked 1-10 produce :
1.6x *more* than ranks 11-20
3.1x *more* than ranks 21-30



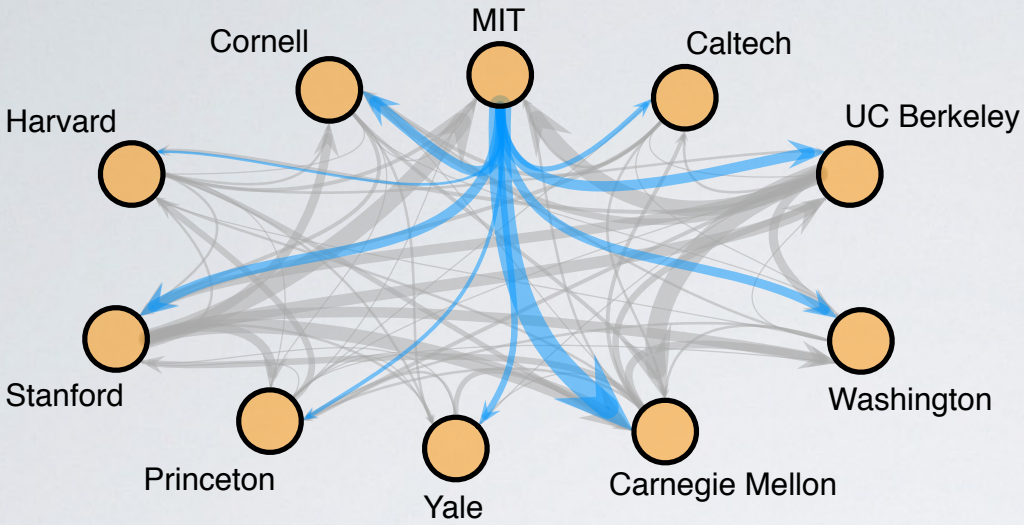
U.S. Income Gini coefficient = 0.45

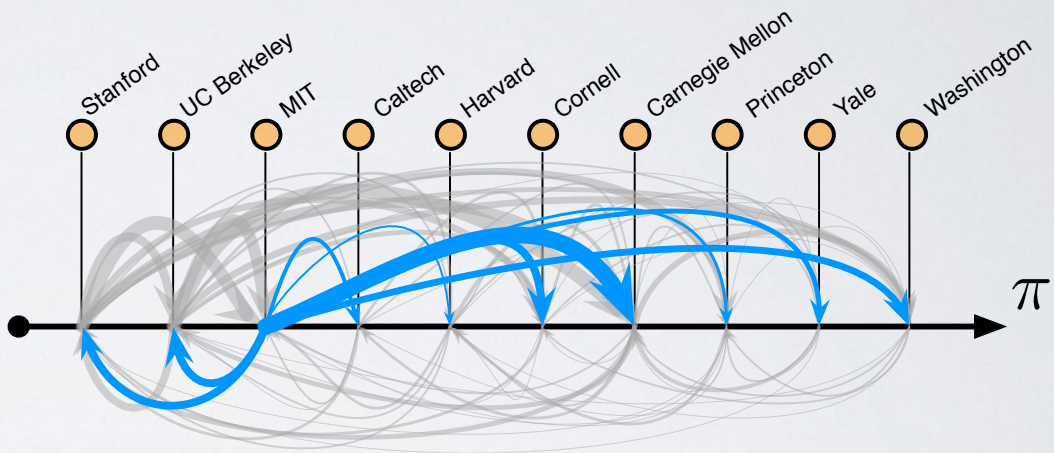
measuring prestige



- select permutation (a ranking) π that minimizes the number of "rank violations" : edges (u, v) where $\pi_v < \pi_u$
- higher-ranked nodes have greater *placement power*
- Minimum Violation Rankings (MVRs) are equivalent to solving *minimum feedback arc set* problem (NP-hard)

these "MVR"s have a deep history in social theory for extracting dominance or prestige hierarchies from data, especially in animal behavior (see de Vries (1999))
there are many equivalent MVRs for our network. we sample these using a zero-temp MCMC, and average across them to obtain $\langle \pi \rangle$





- the ordering π of the nodes that minimizes rank violations over arcs:

$$\mathrm{MVR}(A) = \inf_{\pi} \sum_{u,v} A_{u,v} \times \mathrm{sign}(\pi_u - \pi_v)$$

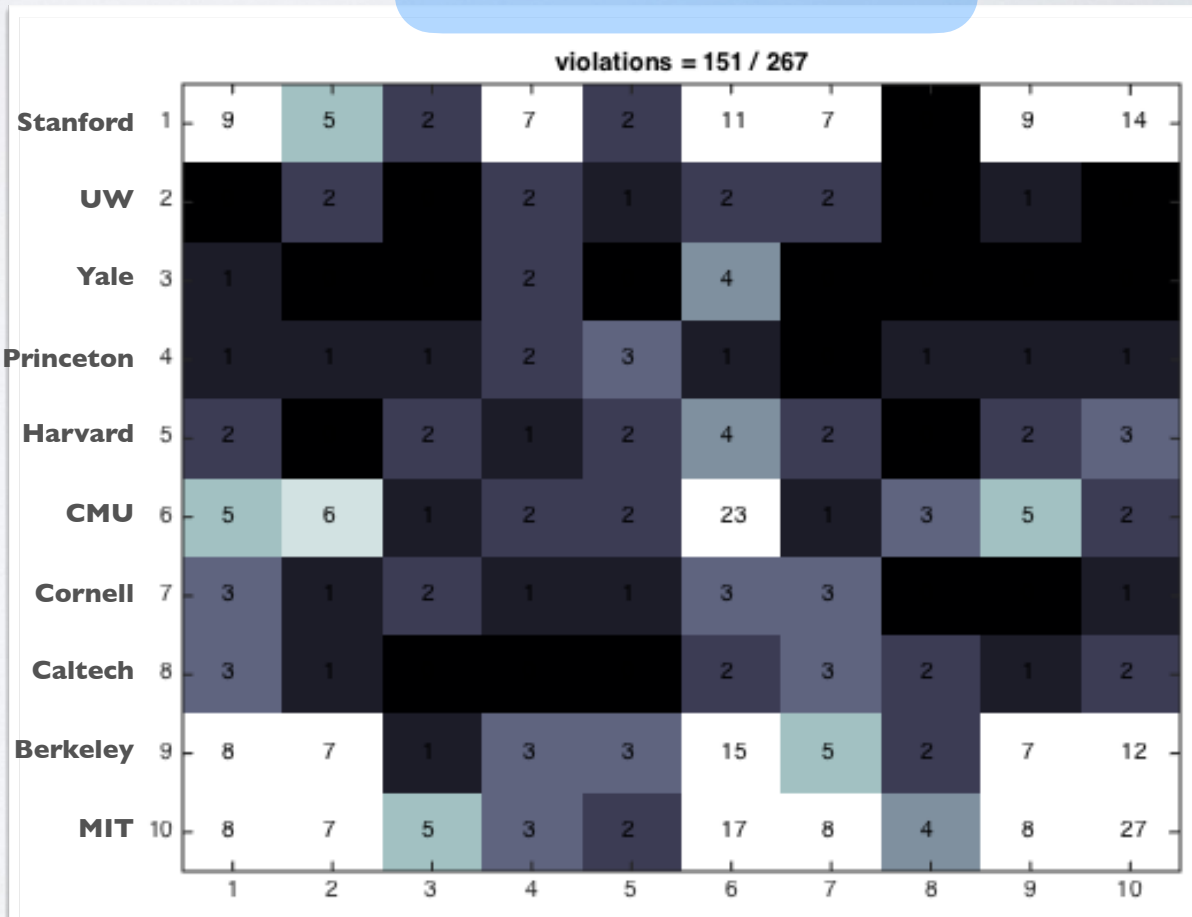
- actually, many such π's, so we sample them using MCMC and compute the mean ranking $\langle \pi \rangle$

these "MVR"s have a deep history in social theory for extracting dominance or prestige hierarchies from data, especially in animal behavior (see de Vries (1999))
there are many equivalent MVRs for our network. we sample these using a zero-temp MCMC, and average across them to obtain

- given an ordering π with $\psi(\pi, A)$ rank violations on network A

- for instance: 151 "violations" of 267 edges

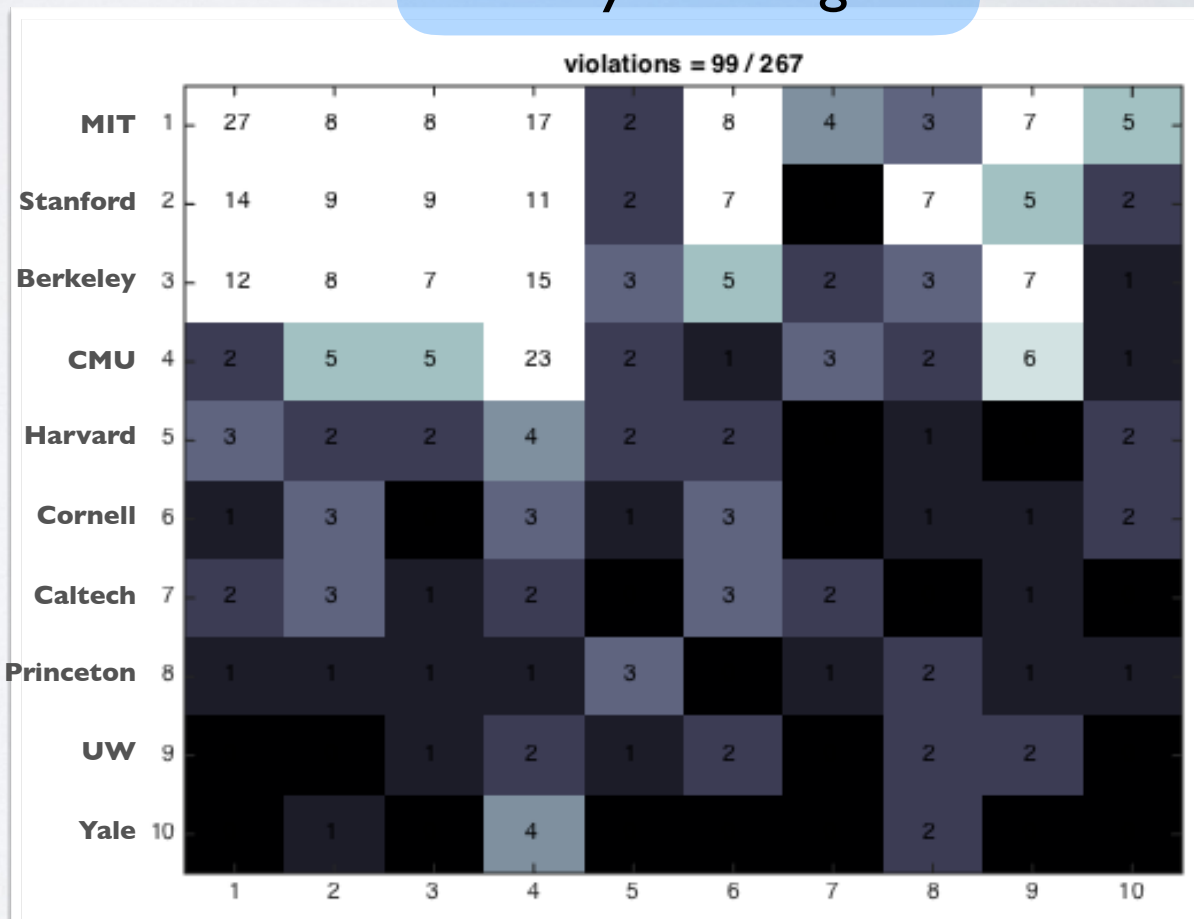
random



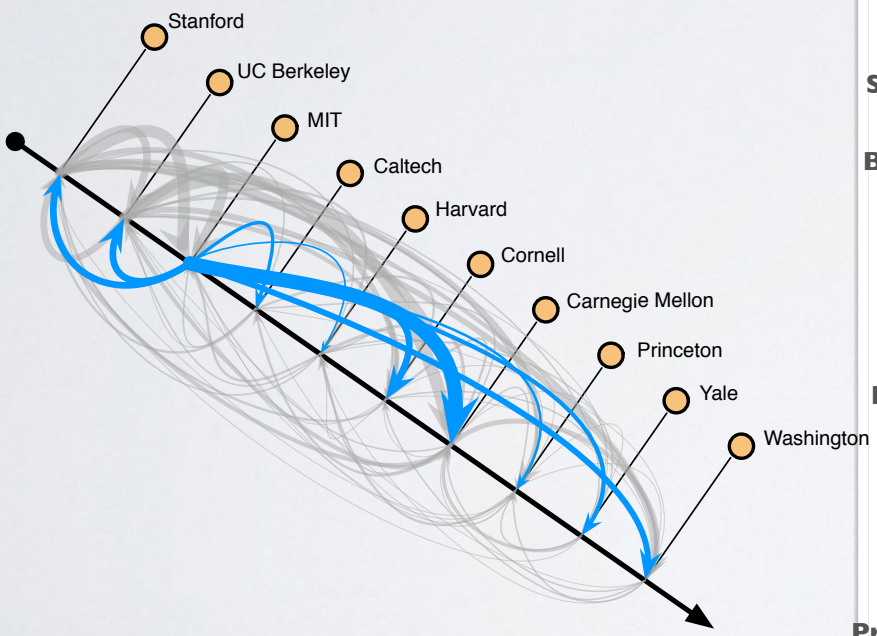
- given an ordering π with $\psi(\pi, A)$ rank violations on network A

- for instance: 99 "violations" of 267 edges

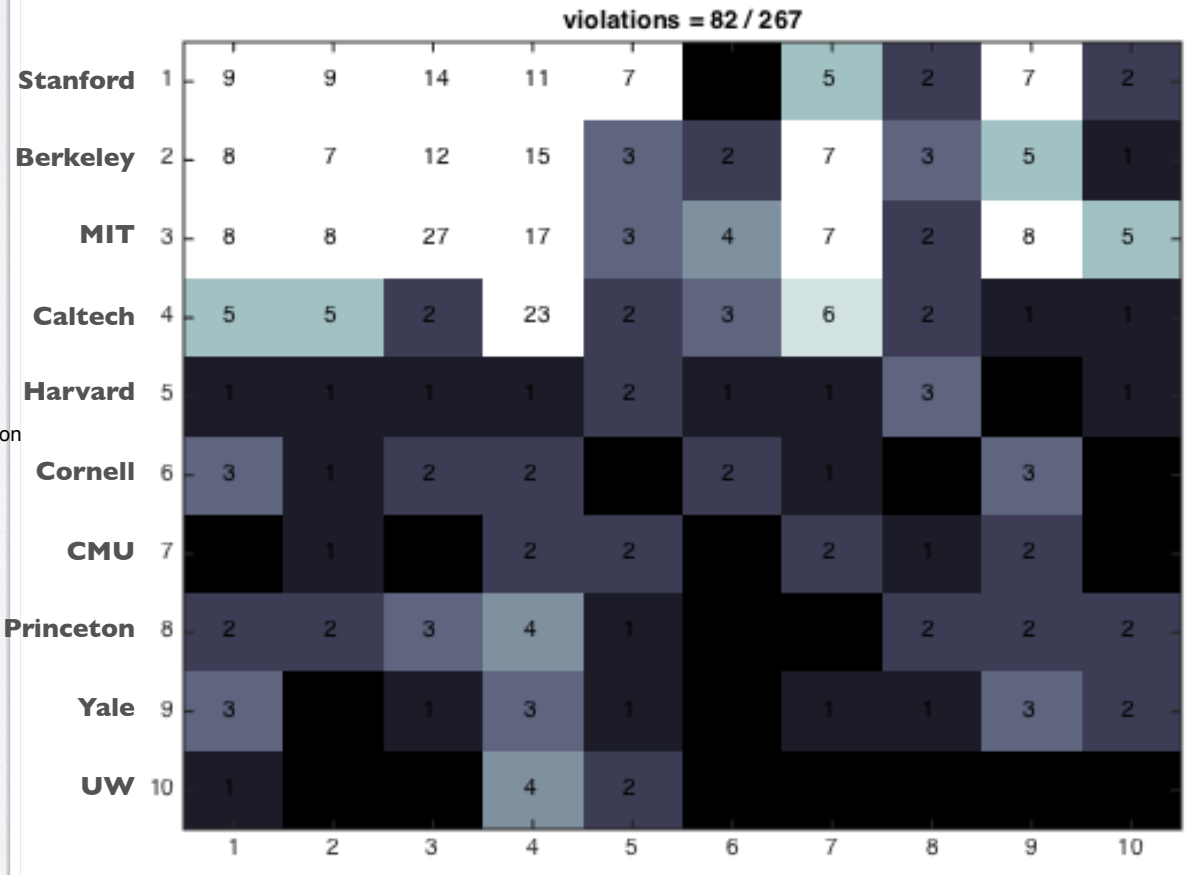
sort by out-degree



- given an ordering π with $\psi(\pi, A)$ rank violations on network A
- MCMC sampler : choose a pair (u, v), swap their ranks $\pi_u \leftrightarrow \pi_v$ to obtain π', compute $\psi(\pi', A)$, accept change if $\psi(\pi', A) \geq \psi(\pi, A)$
- for instance: 82 "violations" of 267 edges

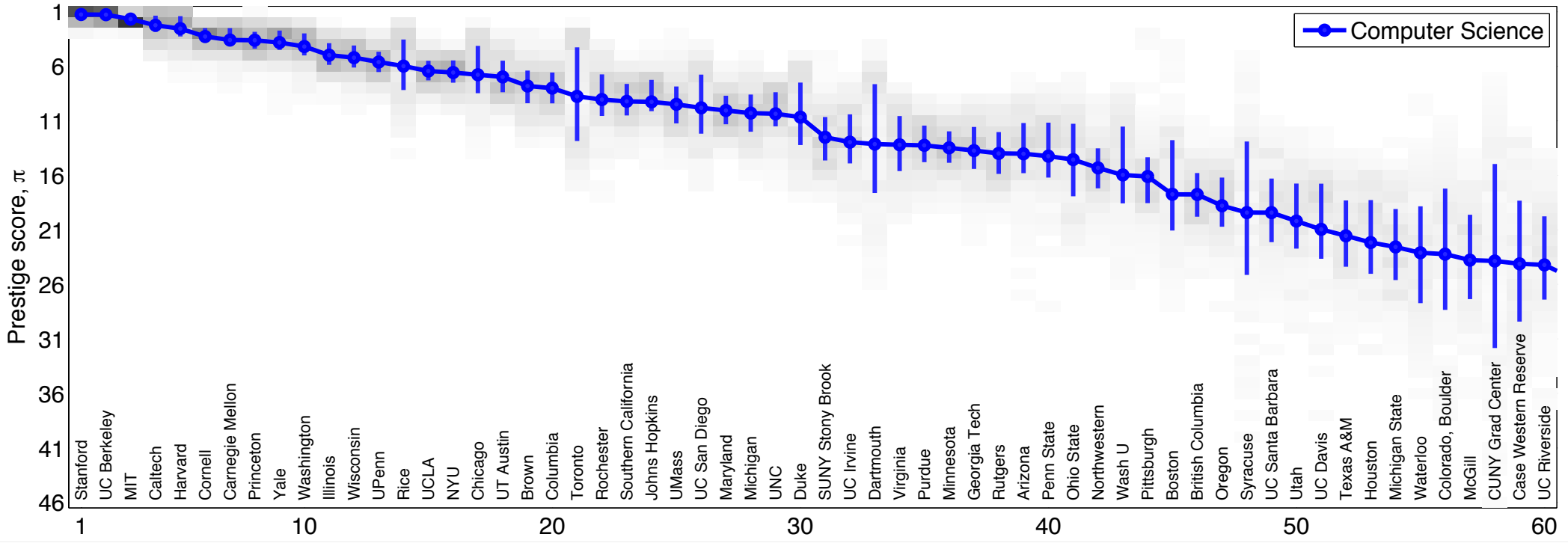


MVR



measuring prestige

- what do these "prestige" hierarchies look like?
- what do they tell us about the structure of the faculty market?



prestige correlates with USNews / NRC, but is *more predictive* of placement

Computer Science	π	US News	NRC	λ
π	—	0.79 (153)	0.80 (103)	0.80 (205)
US News	0.79 (153)	—	0.83 (103)	0.74 (153)
NRC	0.80 (103)	0.83 (103)	—	0.81 (103)
λ	0.80 (205)	0.74 (153)	0.81 (103)	—

- prestige π directly measures *placement power*
- uncertainty increases as prestige decreases
- similar results, but different orderings for *Business* and *History*

uncertainties derived from the set of MVRs
for prediction task results, see Clauset et al. Science Advances (2015)
λ is eigenvector centrality

most placements far down the hierarchy

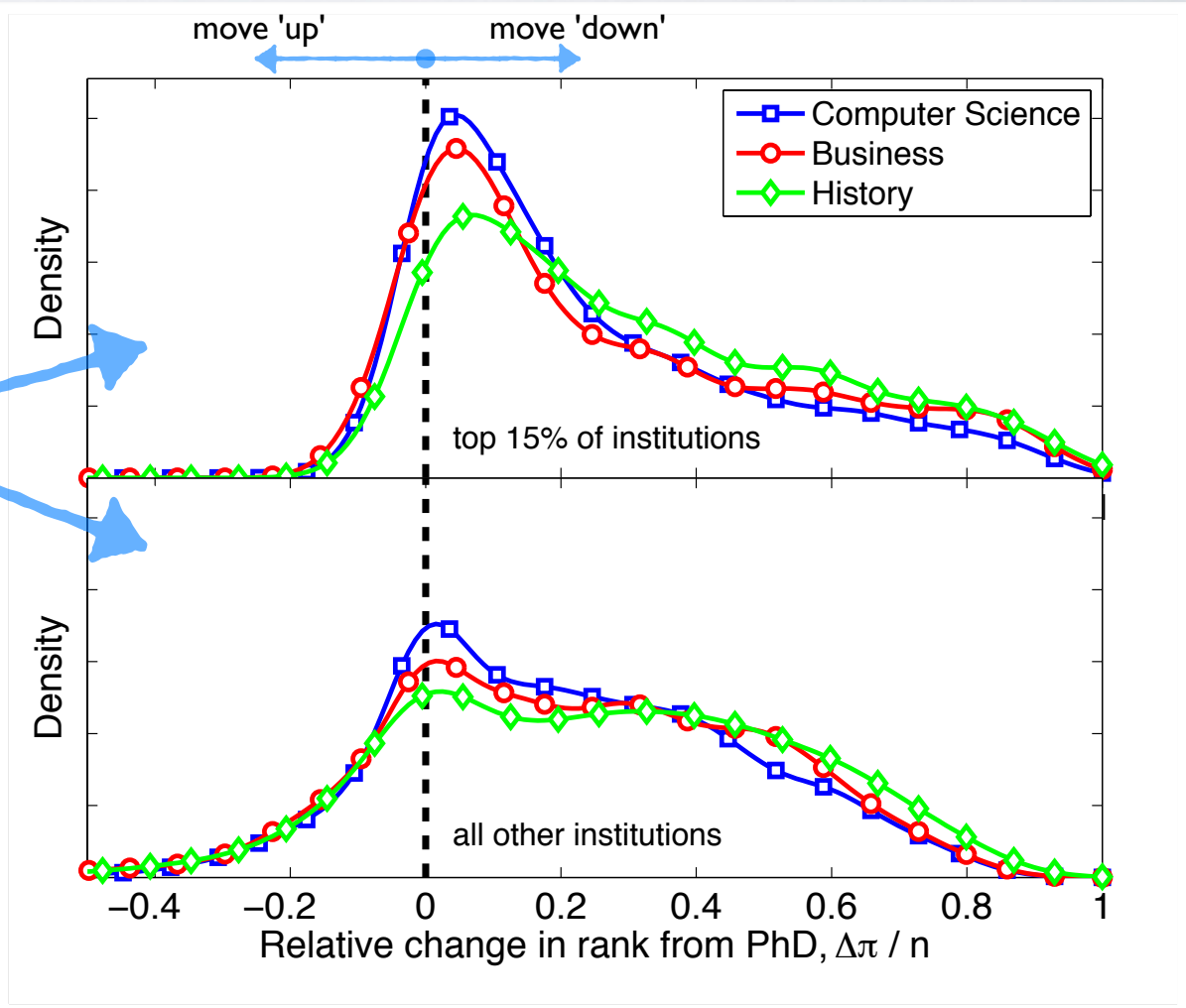
Computer Science

80% move down
8% self-hires
12% move up

average = 47 steps down

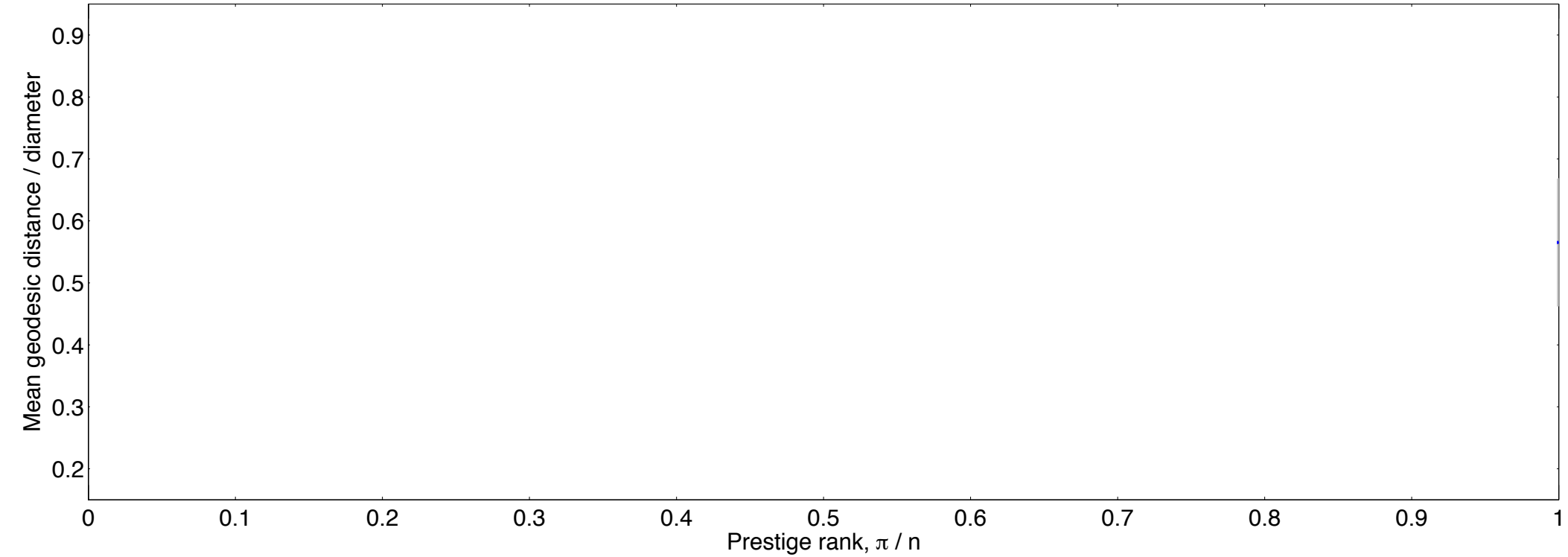
a rich club:

- top 15% of depts. produce 68% of their own faculty
- and hire only 7% from outside top 25% of depts.



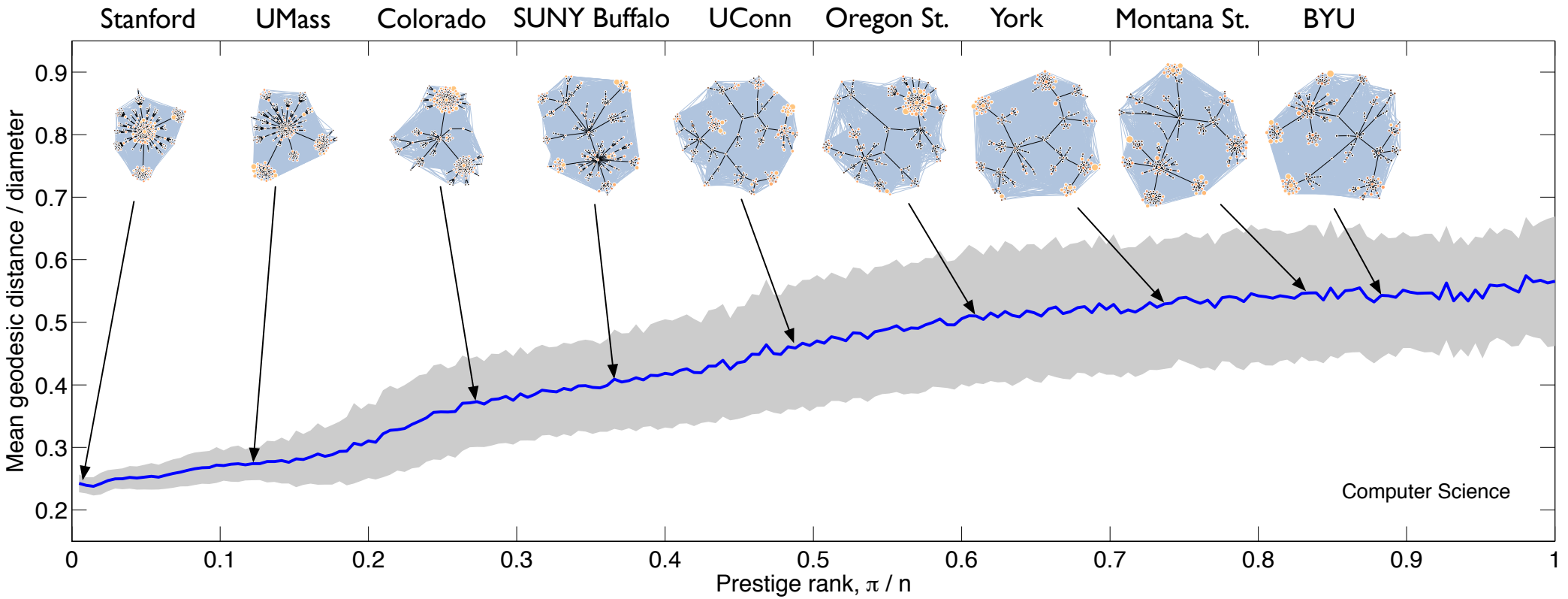
architecture of faculty hiring networks

- how *far* is some v from a node u, on average?
- how does this distance vary with prestige?



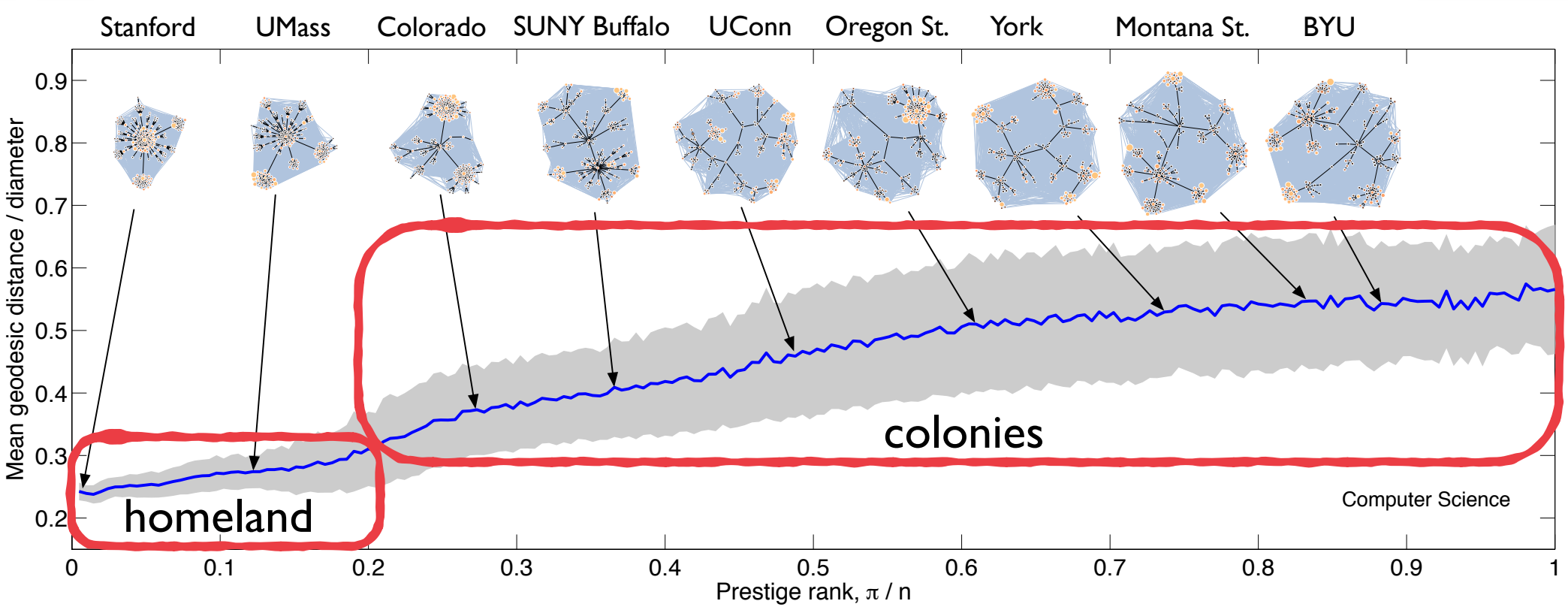
architecture of faculty hiring networks

- core and periphery



architecture of faculty hiring networks

- ~~core and periphery~~ *homeland* and *colonies*
- **prestige is *influence***, via doctoral placement, over research agendas, research communities, and departmental norms across the discipline



prestige and faculty hiring networks

- *prestige pervades and structures the academic ecosystem*
- placement power quantifies reputation via *outcomes* (not inputs)
- reveals *core-periphery structure* of academia

  faculty flow from core → periphery ("the colonies")

  modest fraction stays inside core ("homeland")

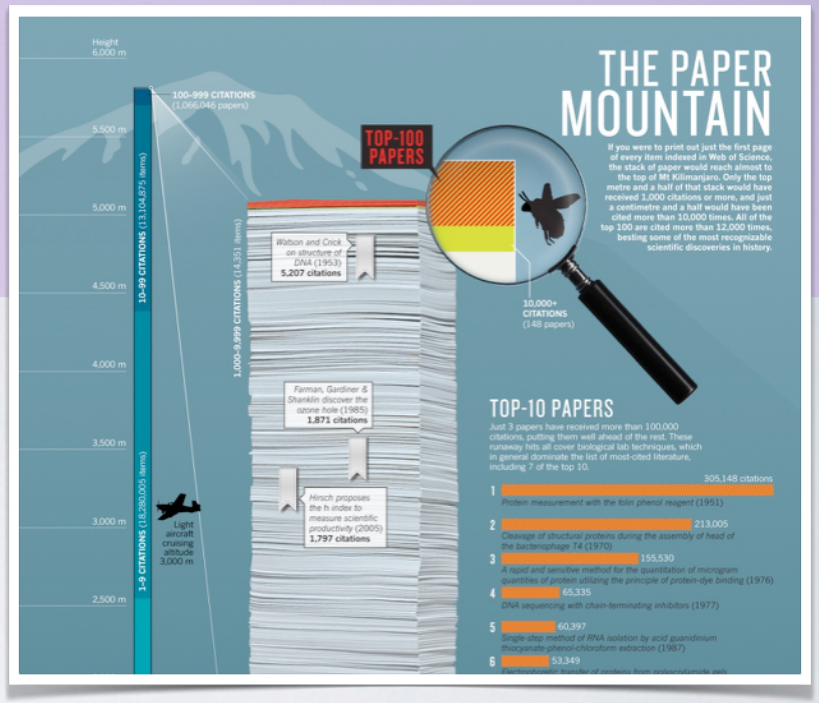
  small fraction flows "upstream"
- prestige is *influence*, via placement

🤔 *how many discoveries are left unmade because hiring hierarchies so strongly determine who gets faculty positions?*

one might argue that these strong prestige hierarchies reflect true meritocratic differences, but I think that's a hypothesis that needs to be proved, rather than disproved

prestige drives epistemic inequality

some ideas spread further than others — why?



Nature "The top 100 papers." (2014)

The term "epistemic inequality" due to Wellmon and Piper (2017)
See work by Crane (1965), Cole & Cole (1967), Zuckerman (1967), Merton (1968), Pelz & Andrews (1976), Allison et al. (1982), Allison & Long (1990), Moed (2006), Way et al. (2017)

prestige drives epistemic inequality

1. genuine differences in merit } sorting is unbiased
2. non-meritocratic social processes } biases in who gets credit & opportunities
   → fame, prestige, seniority, discrimination, history, etc.
3.

prestige drives epistemic inequality

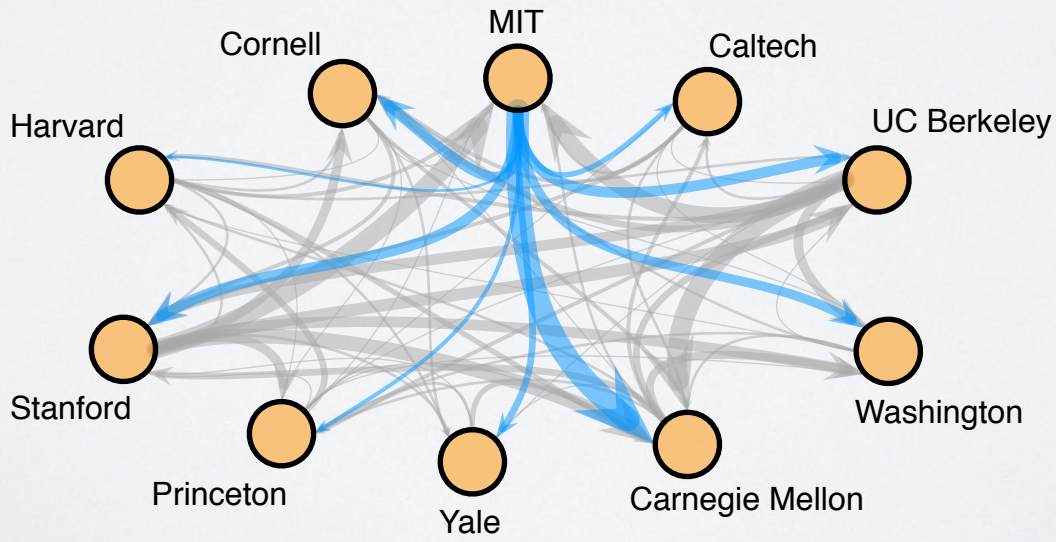
1. genuine differences in merit
2. non-meritocratic social processes
   - fame, prestige, seniority, discrimination, history, etc.
3. non-meritocratic structural factors
   - *scientists carry ideas from PhD to faculty institution*
   - *difference in placement power drives epistemic inequality*

} sorting is unbiased

} biases in who gets credit & opportunities

} a *mechanism*: biases in who does what work where

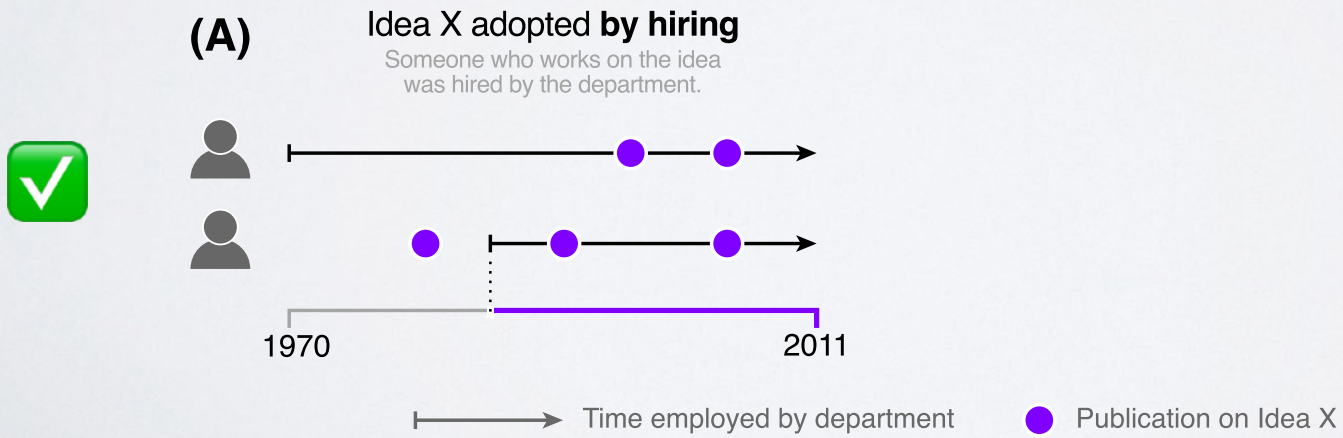
agenda setting theory



test 1: does faculty hiring drive spread of ideas?

3. non-meritocratic structural factors
   - *scientists carry ideas from PhD to faculty institution*
   - *difference in placement power drives epistemic inequality*

- using data on **hiring events** & **publication topics**, test whether **where an idea is being worked on** can be **explained by faculty placement** (and hence the prestige hierarchy)
- two possibilities:



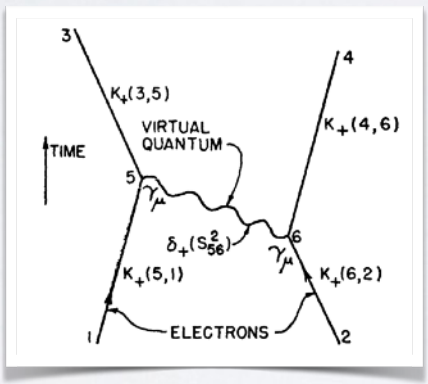
We use the same CS hiring data as before, but now augmented with timing and topic of publications

test 1: does faculty hiring drive spread of ideas?

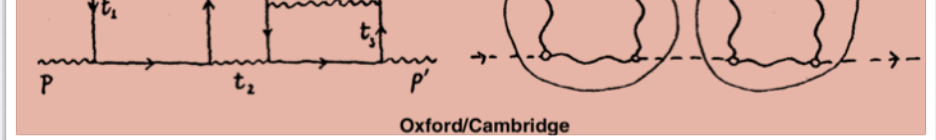
3. non-meritocratic structural factors
   - *scientists carry ideas*
   - *for example*

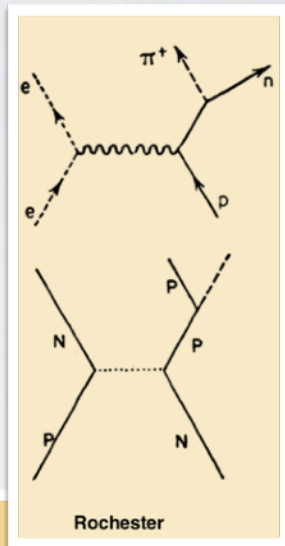
Feynman diagrams, born 1948

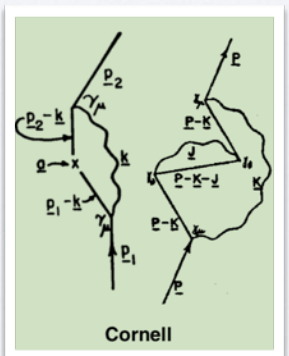
Lamb, Wheeler, Pais, Feynman, Feshbach & Schwinger

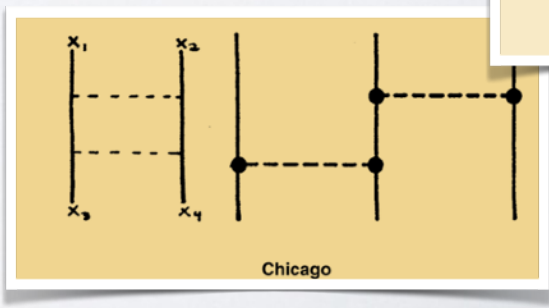


then spread by postdocs who learned the technique at the Institute for Advanced Study







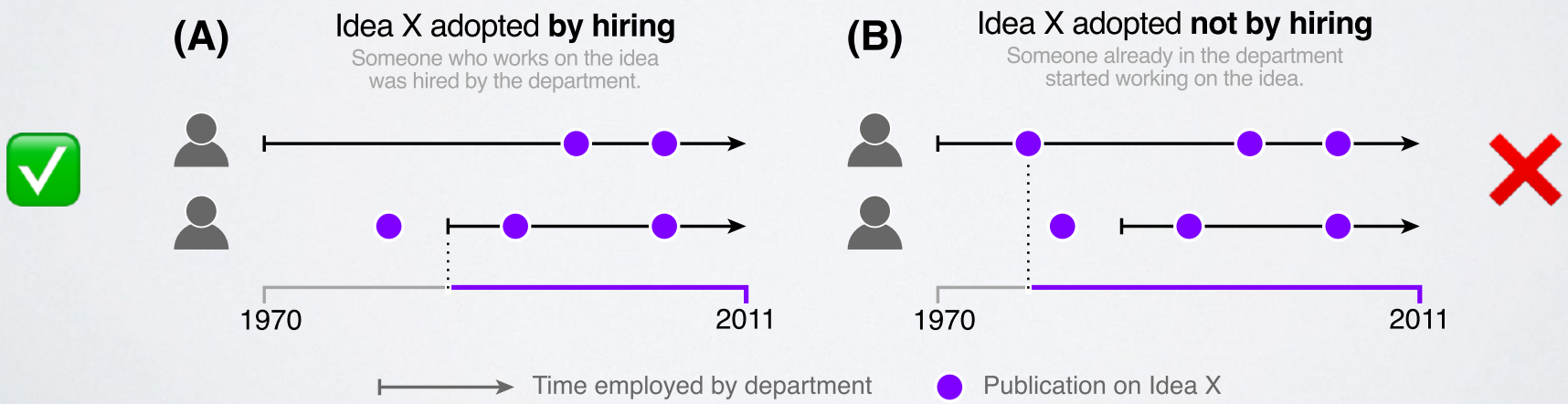


see Kaiser, American Scientist 93 (2005) for more on this fascinating case study
see Rotabi et al. ICWSM (2017) for lighter analysis of the spread of LaTeX macros across collaboration networks

test 1: does faculty hiring drive spread of ideas?

3. non-meritocratic structural factors
   - *scientists carry ideas from PhD to faculty institution*
   - *difference in placement power drives epistemic inequality*

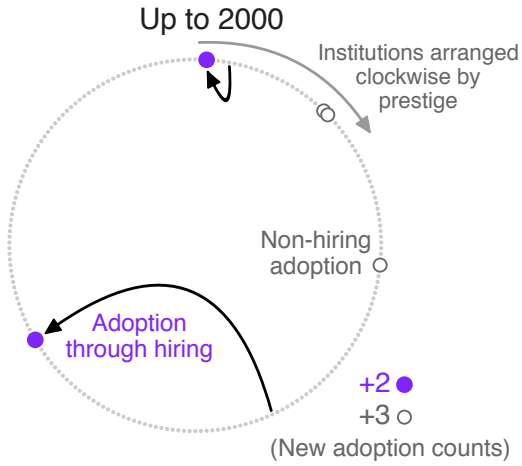
- using data on **hiring events** & **publication topics**, test whether **where an idea is being worked on** can be **explained by faculty placement** (and hence the prestige hierarchy)
- two possibilities:



We use the same CS hiring data as before, but now augmented with timing and topic of publications

test

Topic Modeling

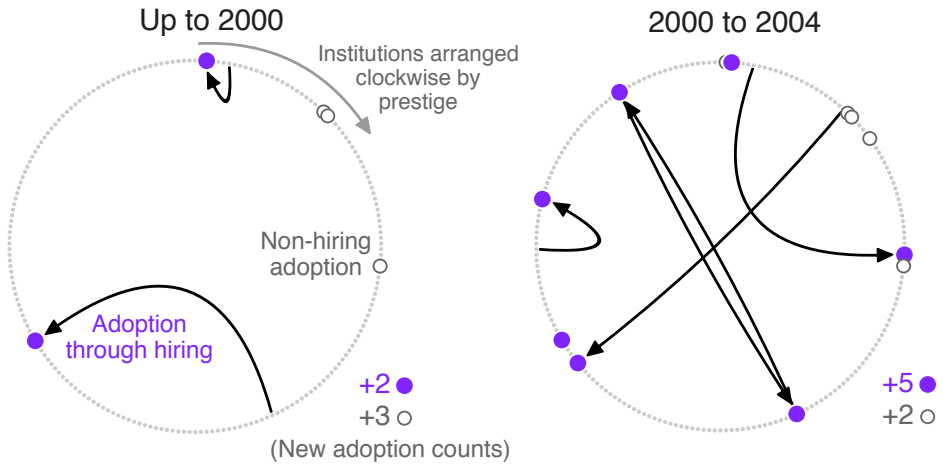


Incremental Computing

Deep Learning

test

Topic Modeling

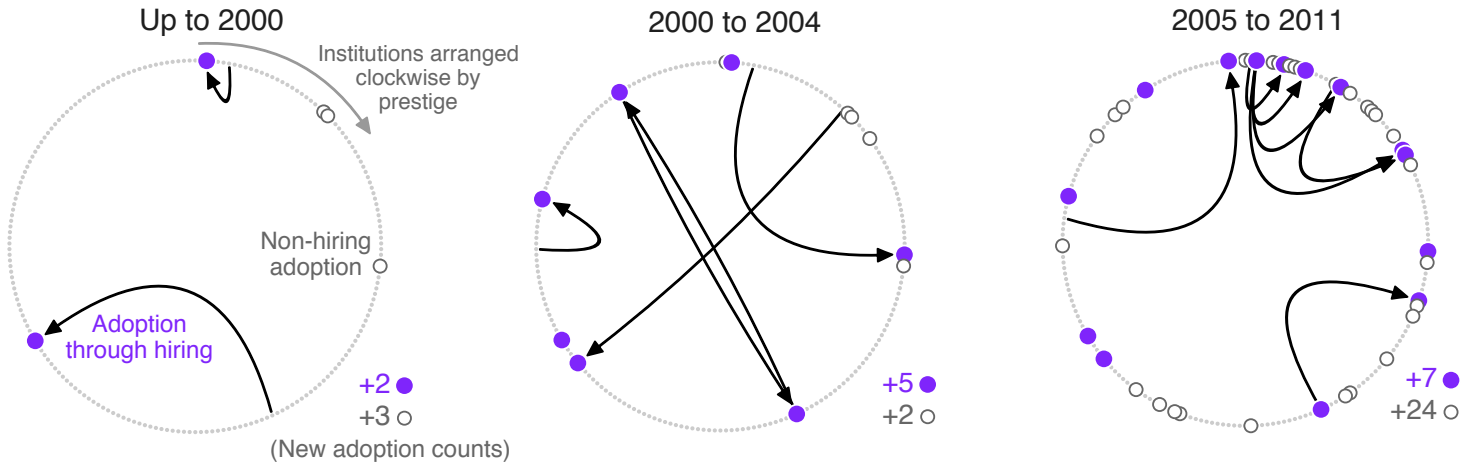


Incremental Computing

Deep Learning

test

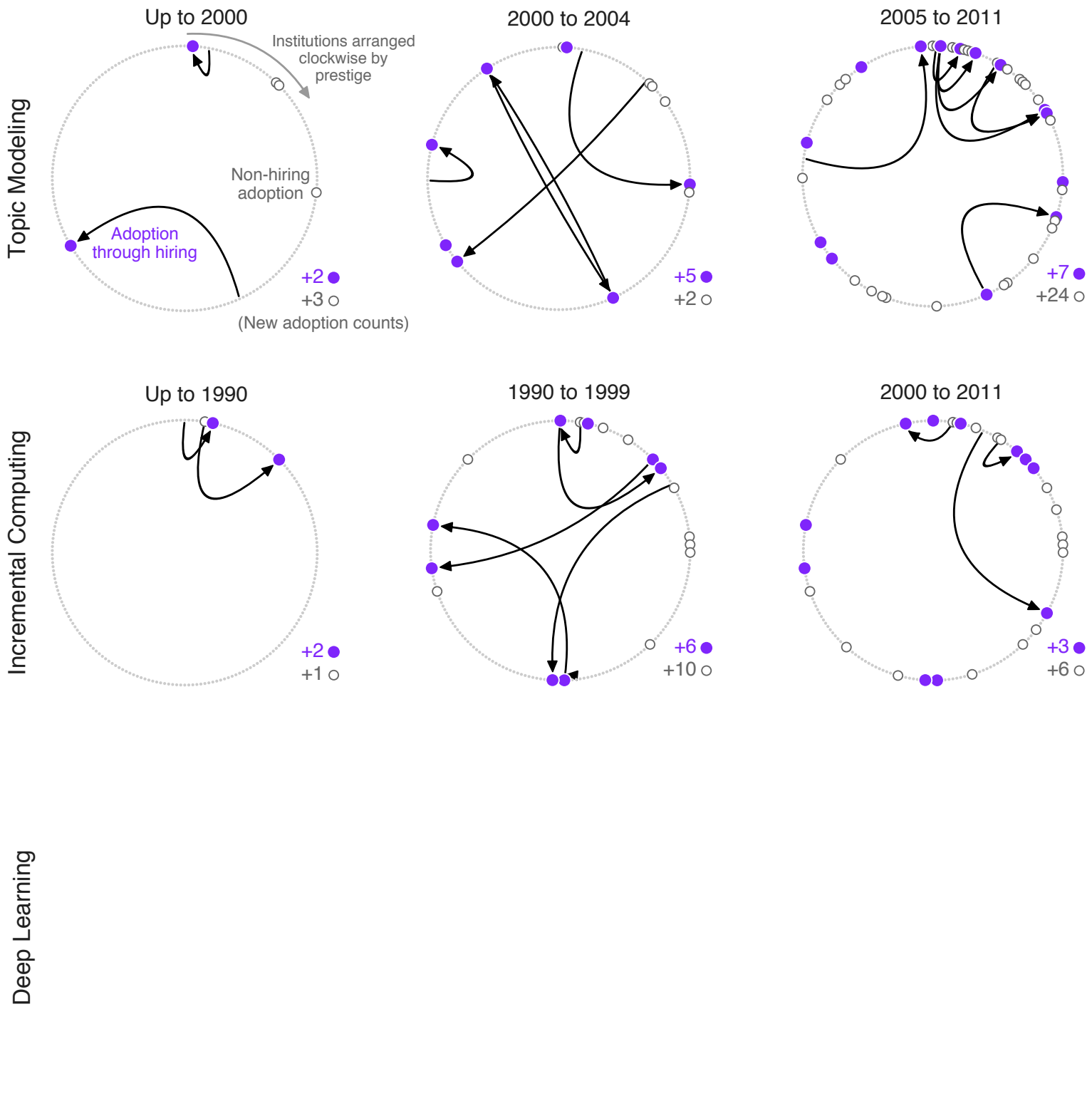
Topic Modeling



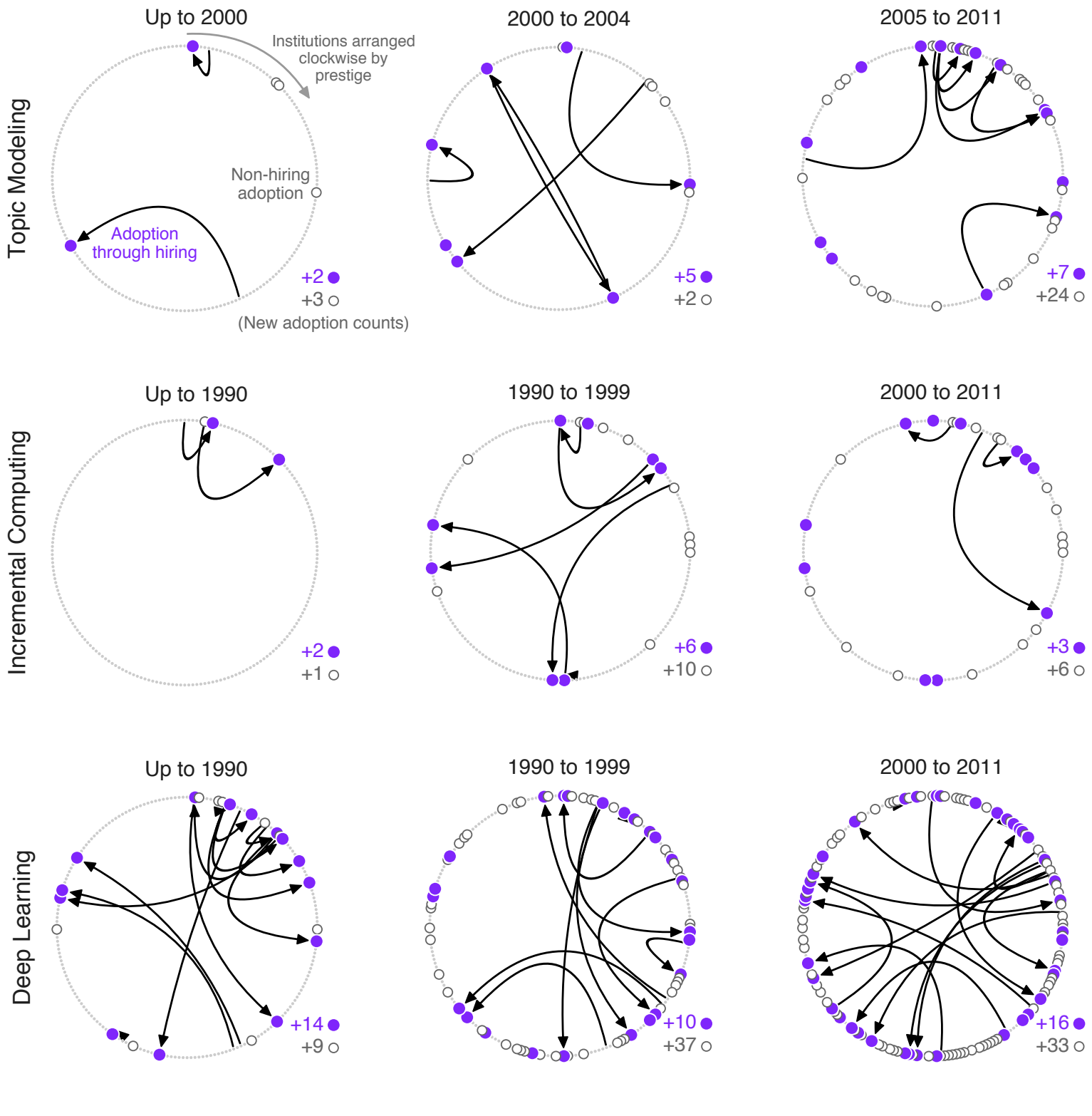
Incremental Computing

Deep Learning

test


Topic Modeling
Up to 2000
Institutions arranged clockwise by prestige
Non-hiring adoption
Adoption through hiring
+2
+3
(New adoption counts)
2000 to 2004
+5
+2
2005 to 2011
+7
+24
Incremental Computing
Up to 1990
+2
+1
1990 to 1999
+6
+10
2000 to 2011
+3
+6
Deep Learning

test


Topic Modeling
Up to 2000
Institutions arranged clockwise by prestige
Non-hiring adoption
Adoption through hiring
+2
+3
(New adoption counts)
2000 to 2004
+5
+2
2005 to 2011
+7
+24
Incremental Computing
Up to 1990
+2
+1
1990 to 1999
+6
+10
2000 to 2011
+3
+6
Deep Learning
Up to 1990
+14
+9
1990 to 1999
+10
+37
2000 to 2011
+16
+33

test 1: does faculty hiring drive spread of ideas?

topic X	f_{obs}	f_{exp}	p
topic modeling	0.35		
incremental computing	0.39		
deep learning	0.35		
quantum computing	0.32		
mechanism design	0.48		

fraction of real hires that spread topic X

fraction under permutation test

test 1: does faculty hiring drive spread of ideas?

topic X	f_{obs}	f_{exp}	p
topic modeling	0.35	0.23	0.01 ± 0.01
incremental computing	0.39	0.20	0.01 ± 0.01
deep learning	0.35	0.34	0.34 ± 0.01
quantum computing	0.32	0.22	0.01 ± 0.01
mechanism design	0.48	0.21	0.01 ± 0.01

fraction of real hires that spread topic X

fraction under permutation test

→ ***faculty hiring is a significant driver of the spread of ideas for some topics***

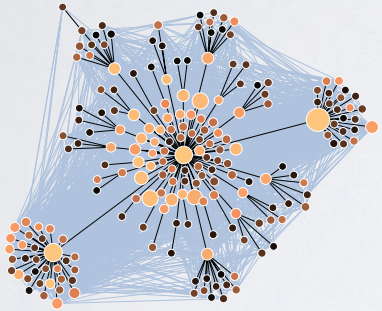
[1] expected fraction under null model where everyone's pub titles chosen randomly from across all papers (permutation test)

test 2: spread ideas across faculty hiring network

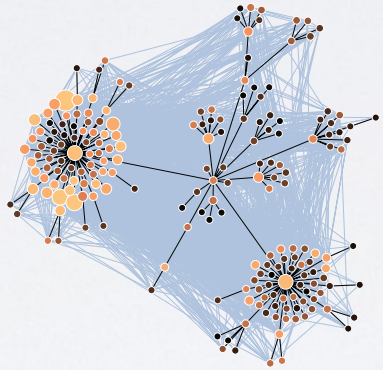
how does point of origination shape spread of an idea?

- prestige drives → core-periphery structure → drives spread

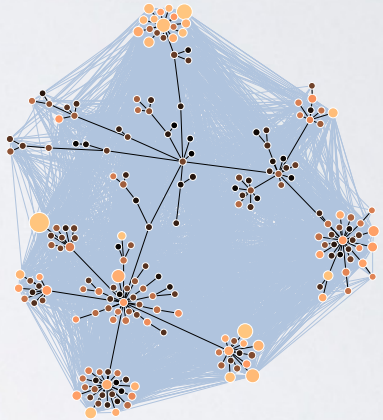
Stanford



Colorado



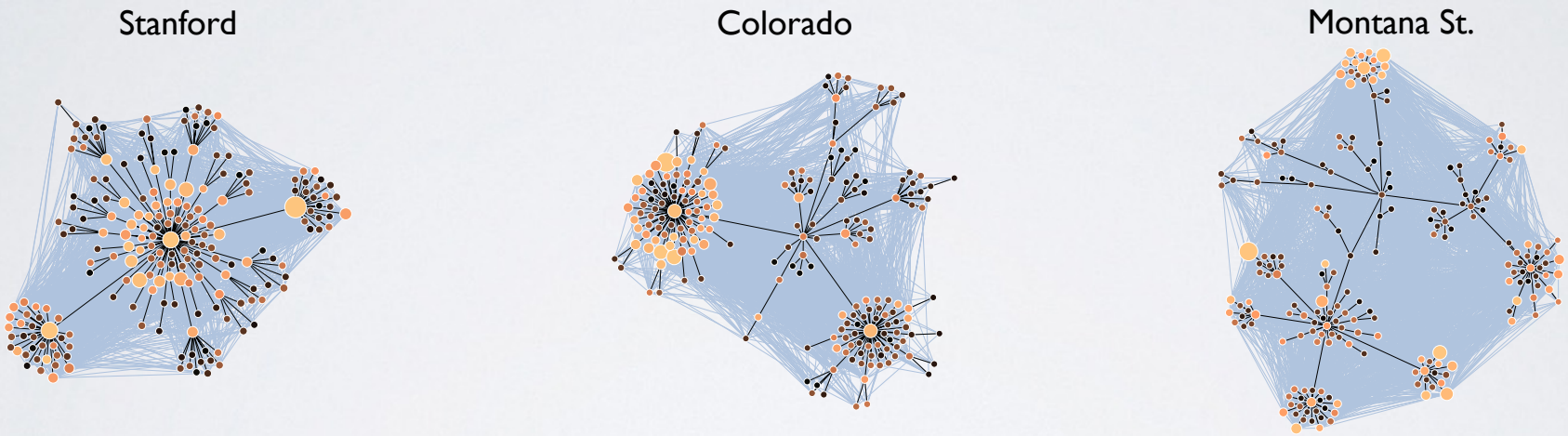
Montana St.



test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

- prestige drives → core-periphery structure → drives spread



model spread as Susceptible-Infected (SI) model on faculty hiring network

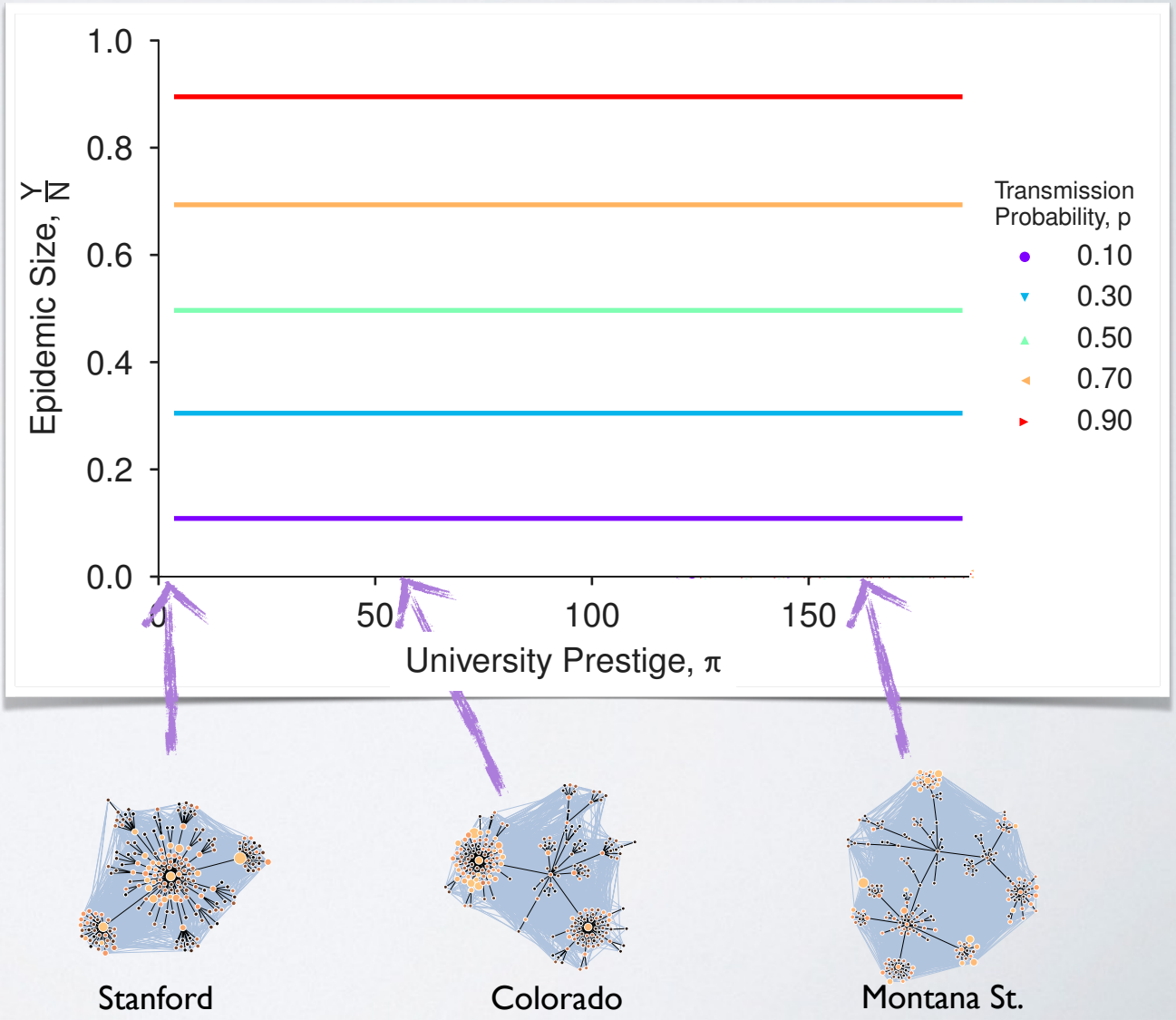
- seed simulation at each university, by prestige π
- vary idea transmissibility $p \in \{0.1, 0.3, 0.5, 0.7, 0.9\}$
- measure normalized size Y/N

test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

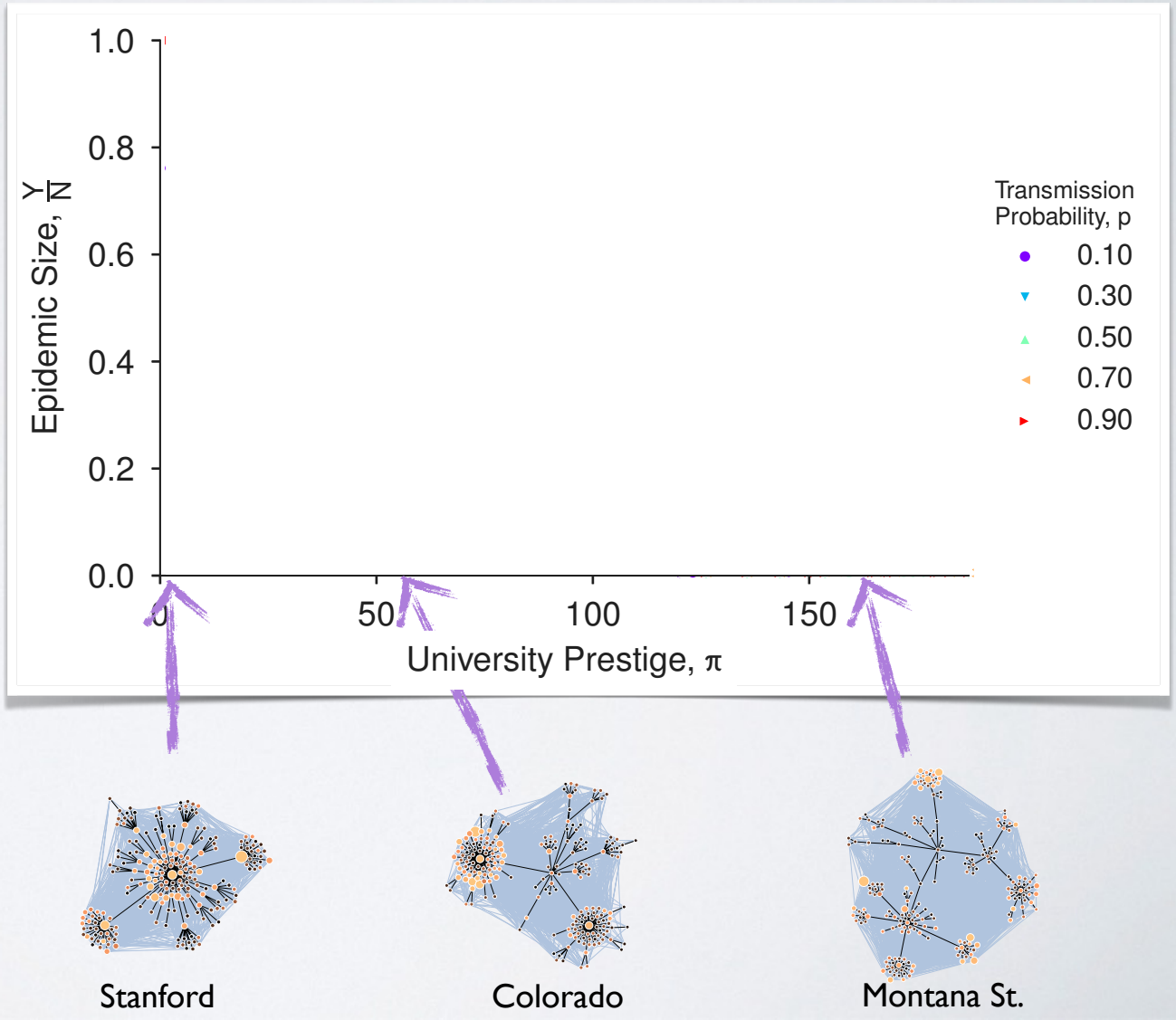
meritocracy?

- idea spread is proportional to idea "quality"
- spread is *independent* of birthplace



test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

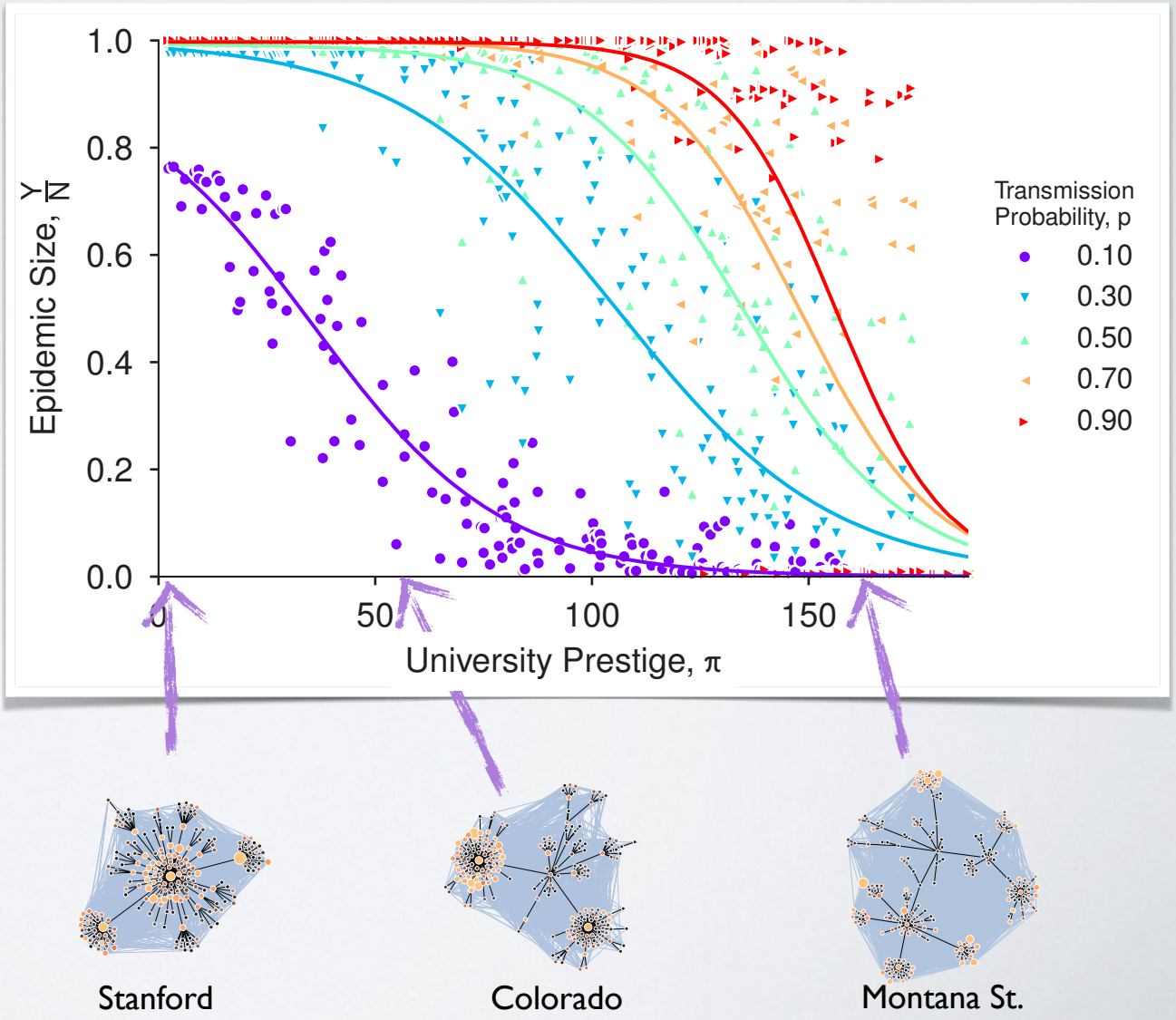


test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

not-a-meritocracy.

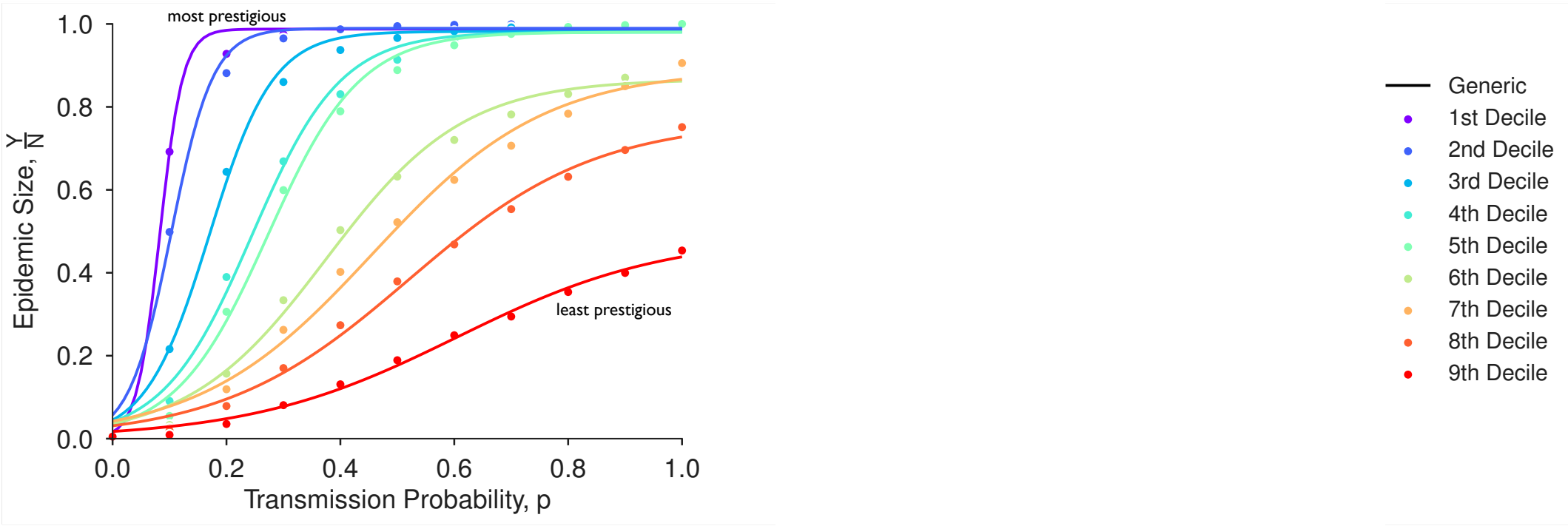
- 👍 • best ideas spread independent of origin
- 👎 • bad ideas spread easily from elite departments
- 👎 • good ideas from mid-prestige spread less well than bad ideas from high-prestige



test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

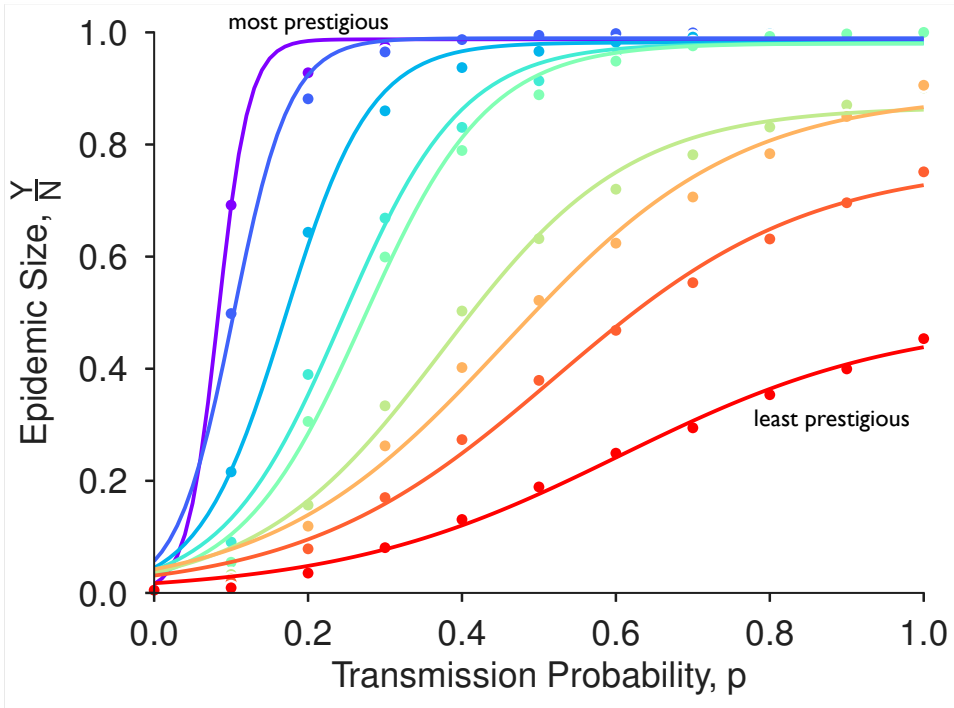
- classic logistic growth → higher p, bigger spread Y/N
- rescale $p^* = -p\,/\log(1-d)$ to collapse curves for prestige deciles d

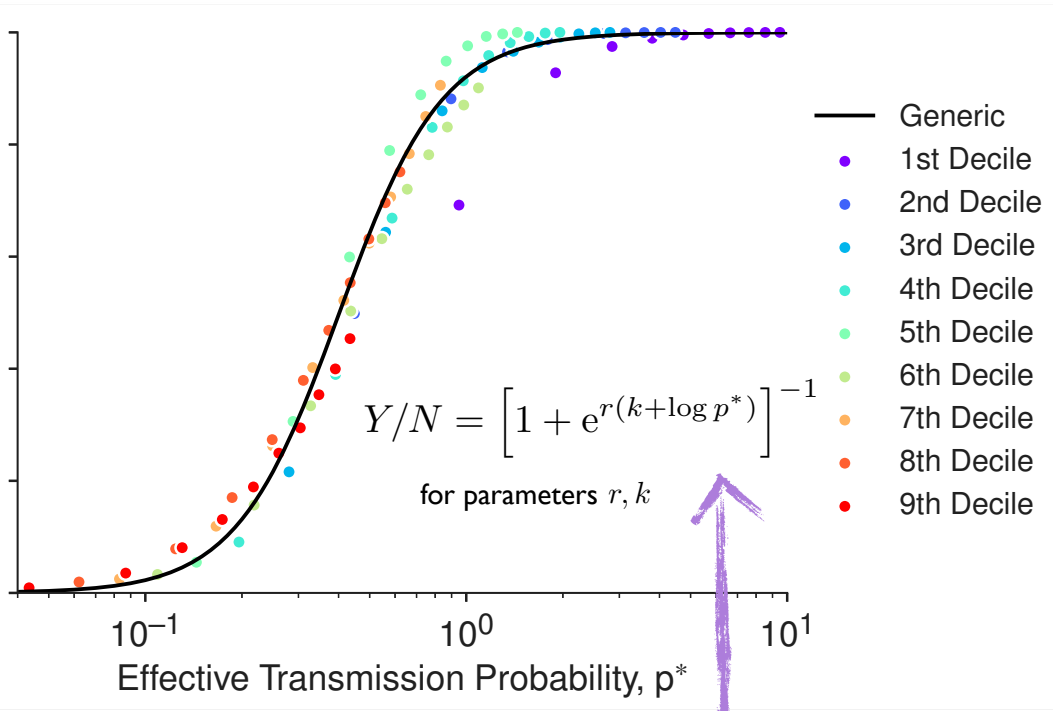


test 2: spread ideas across faculty hiring network

how does point of origination shape spread of an idea?

- classic logistic growth → higher p, bigger spread Y/N
- rescale $p^* = -p\,/\log(1-d)$ to collapse curves for prestige deciles d





prestige is an attention amplifier

epistemic inequalities

3. inequalities reflect non-meritocratic structural factors
   - ✅ *scientists carry ideas from PhD to faculty institution*
   - ✅ *difference in placement power drives epistemic inequality*

---

prestige → placement → spread of ideas

- → simulations suggests an *exponential dependence* of "impact" on increased prestige
- → faculty hiring network is a *structural* mechanism: it creates & maintains epistemic inequality
- → *your location in network shapes how far your ideas spread*

Morgan, Economou, Way, Clauset, *EPJ Data Science* **7**, 40 (2018)

networks, prestige, and the spread of scientific ideas

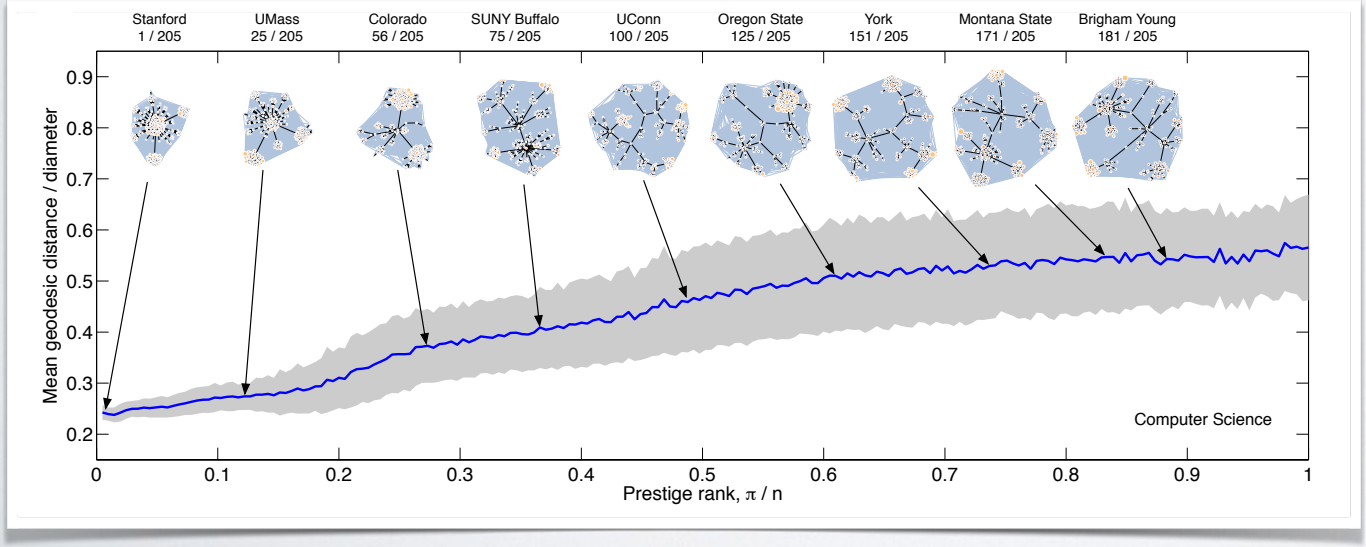
prestige *pervades and structures* the scientific ecosystem

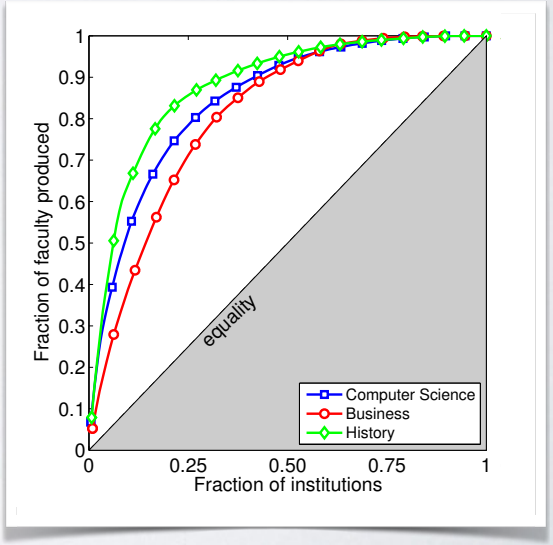
- jobs, publications, spread of ideas, resources…

networks, prestige, and the spread of scientific ideas

prestige *pervades and structures* the scientific ecosystem

- jobs, publications, spread of ideas, resources…
- hiring networks → core-periphery structure (homeland/colonies)
  - → highly-skewed faculty production
  - → latent one-D embedding (a hierarchy)

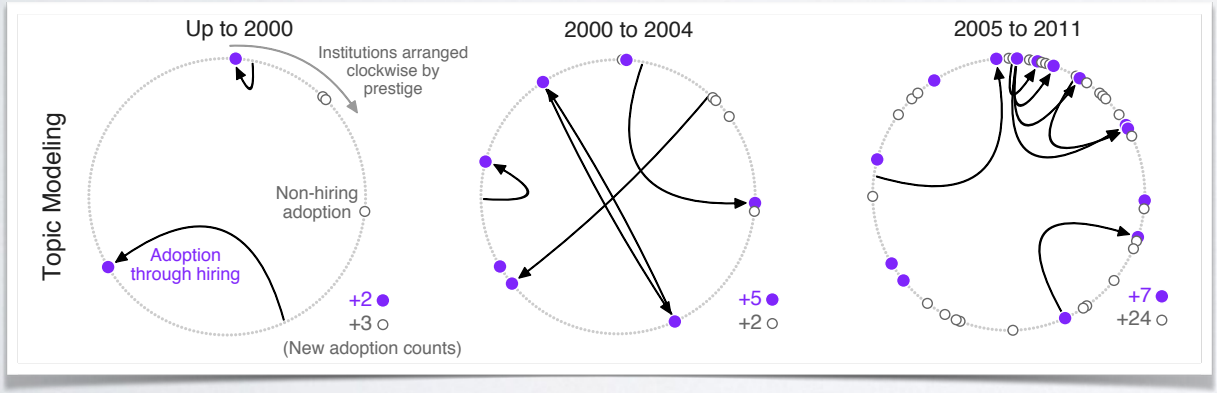


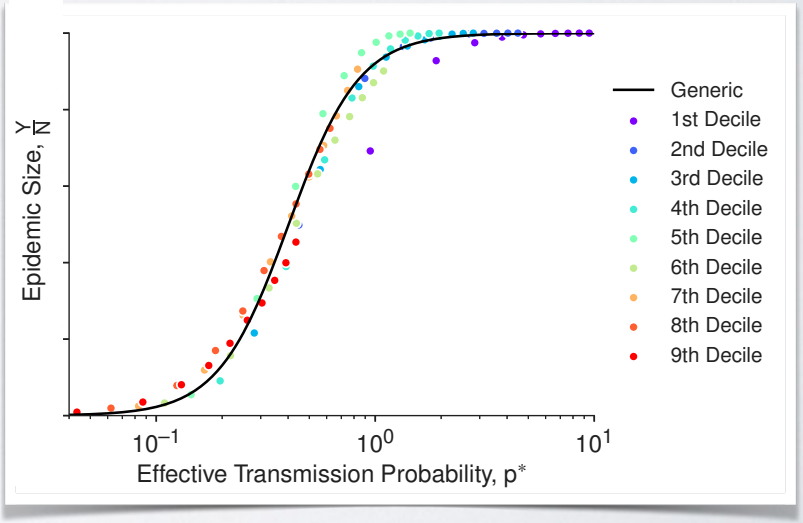


networks, prestige, and the spread of scientific ideas

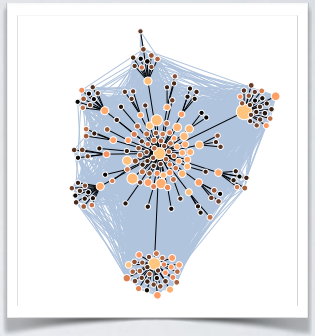
prestige *pervades and structures* the scientific ecosystem

- jobs, publications, spread of ideas, resources…
- hiring networks → drives spread of ideas
  - → prestige is an idea amplifier
  - → non-meritocratic bias





networks, prestige, and the spread of scientific ideas



prestige *pervades and structures* the scientific ecosystem

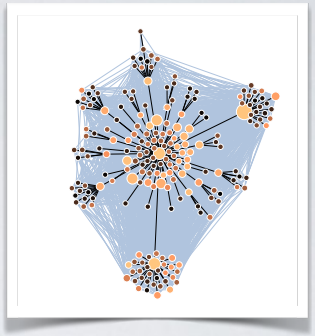
- jobs, publications, spread of ideas, resources…
- hiring networks → prestige.

science is a *model* complex system

- competition, cooperation, feedback, incentives, networks, multi-scales
- hypothesis: *scientific discovery is an emergent property of the system*

🤔 big questions:

networks, prestige, and the spread of scientific ideas



prestige *pervades and structures* the scientific ecosystem

- jobs, publications, spread of ideas, resources…
- hiring networks → prestige

science is a *model* complex system

- competition, cooperation, feedback, incentives, networks, multi-scales
- hypothesis: *scientific discovery is an emergent property of the system*

🤔 big questions:

- → *does prestige help or hinder scientific progress?*
- → *how much competition is too much or too little?*
- → *what changes would accelerate discoveries?*
- → *what discoveries are not being made because of system bias?*

NETWORK SCIENCES

Systematic inequality and hierarchy in faculty hiring networks

Aaron Clauset,[1,2,3]* Samuel Arbesman,[4] Daniel B. Larremore[5,6]

Science Advances 1, e1400005 (2015)

Prof. Aaron Clauset
(Colorado)

Prof. Daniel Larremore
(Colorado)

Dr. Sam Arbesman
(now Lux Capital)

Prestige drives epistemic inequality in the diffusion of scientific ideas

Allison C. Morgan[1*], Dimitrios J. Economou[1], Samuel F. Way[1] and Aaron Clauset[1,2,3]

EPJ Data Science 7, 40 (2018)

Dr. Allison Morgan
(now Twitter)

Dimitrios Economou
(Queen's U.)

Dr. Sam Way
(now Spotify)

Productivity, prominence, and the effects of academic environment

Samuel F. Way[a,1], **Allison C. Morgan**[a], **Daniel B. Larremore**[a,b,2], **and Aaron Clauset**[a,b,c,1,2]

[a]Department of Computer Science, University of Colorado, Boulder, CO, USA; [b]BioFrontiers Institute, University of Colorado, Boulder, CO, USA; [c]Santa Fe Institute, Santa Fe, NM, USA

Proc. Natl. Acad. Sci. USA 116, 10729 (2019)

ESSAY

Data-driven predictions in the science of science

Aaron Clauset,[1,2]* Daniel B. Larremore,[2] Roberta Sinatra[3,4]

Science 355, 477 (2017)

Prof. Roberta Sinatra
(ITU Copenhagen)

Ewing Marion KAUFFMAN Foundation

NSF

fin

see also:

Morgan et al., "Socioeconomic roots of academic faculty." Preprint (2021)

Morgan et al. "The unequal impact of parenthood in academia." *Science Advances* 7 (2021)

Way et al., "Productivity, prominence, and the effects of academic environment." *PNAS* 116 (2019)

Morgan et al., "Prestige drives epistemic inequality in the diffusion of scientific ideas." *EPJ Data Science* 7 (2018)

Way et al., "The misleading narrative of the canonical faculty productivity trajectory." *PNAS* 114 (2017)

Clauset et al., "Data-driven predictions in the science of science." *Science* 355 (2017)

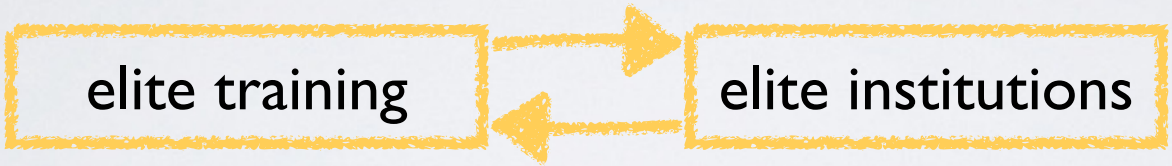
Way et al., "Gender, productivity, and prestige in computer science faculty hiring networks." *Proc. WWW* (2016)

Clauset et al., "Systematic inequality and hierarchy in faculty hiring networks." *Science Advances* 1 (2015)

prestige shapes productivity — why?

- **scientific discourse** is dominated by **elite scientists** and **elite institutions**

but:



prestige shapes productivity — why?

- **scientific discourse** is dominated by **elite scientists** and **elite institutions**

but: **elite training** ⇄ **elite institutions**

- this dynamic reflect **endogenous cumulative advantage** : past achievements correlate with future achievements

🤔 **what causes differences in individual or institutional scholarly achievement?**

prestige shapes productivity — why?

- *what drives future productivity and prominence?*

  idea 1: where a scientist trained.

  - skill, talent, training, temperament, etc.
  - faculty hiring market *sorts* individuals by their natural or previously acquired characteristics that correlate with outcomes
  - e.g., most productive scientists have elite pedigree, Harvard, Yale, Penn, Stanford, etc.

  *where you **trained** causes future productivity and prominence*

prestige shapes productivity — why?

- *what drives future productivity and prominence?*

  idea 1: where a scientist trained.

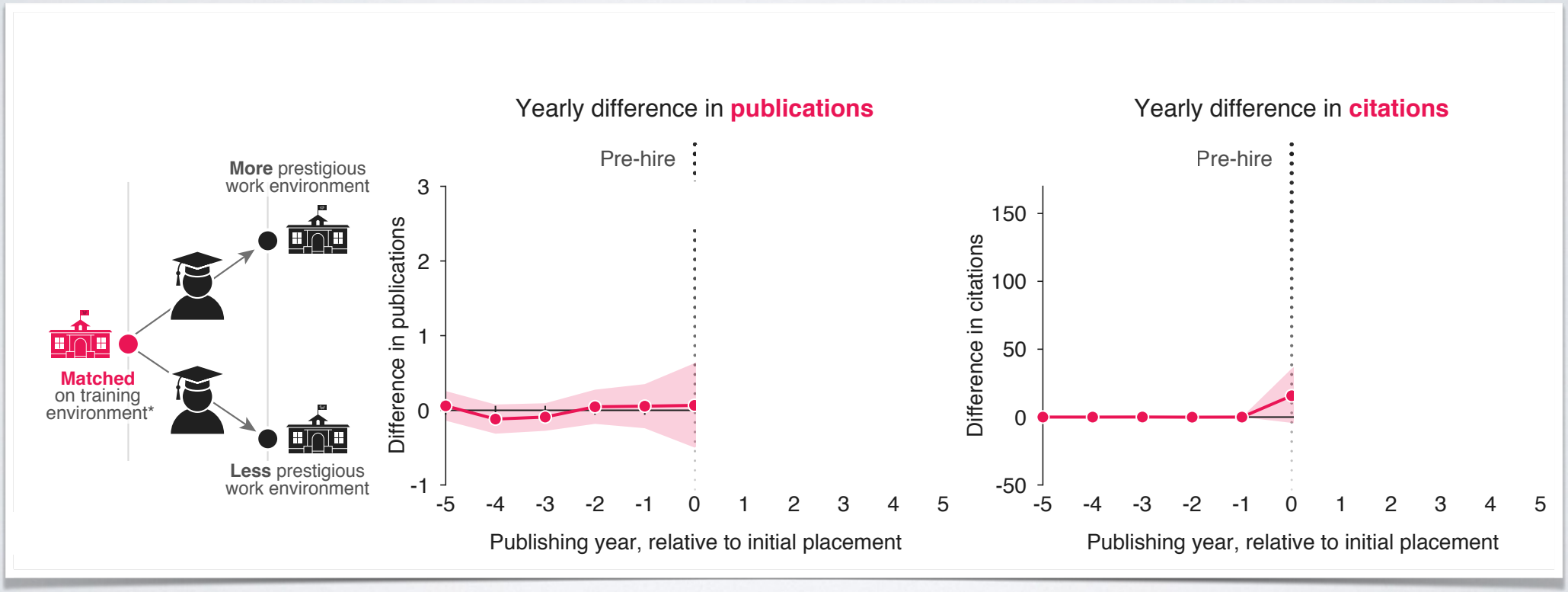
  idea 2: where a scientist works.

  - environmental factors, resources, people, support
  - beyond a basic training, a scientist's output is driven by local environment
  - e.g., moving from poor to rich environments improves output, and vice versa (Allison & Long 1990)

*where you **work** causes future productivity and prominence*

quasi-natural experiments : faculty hiring

similar training, different placement

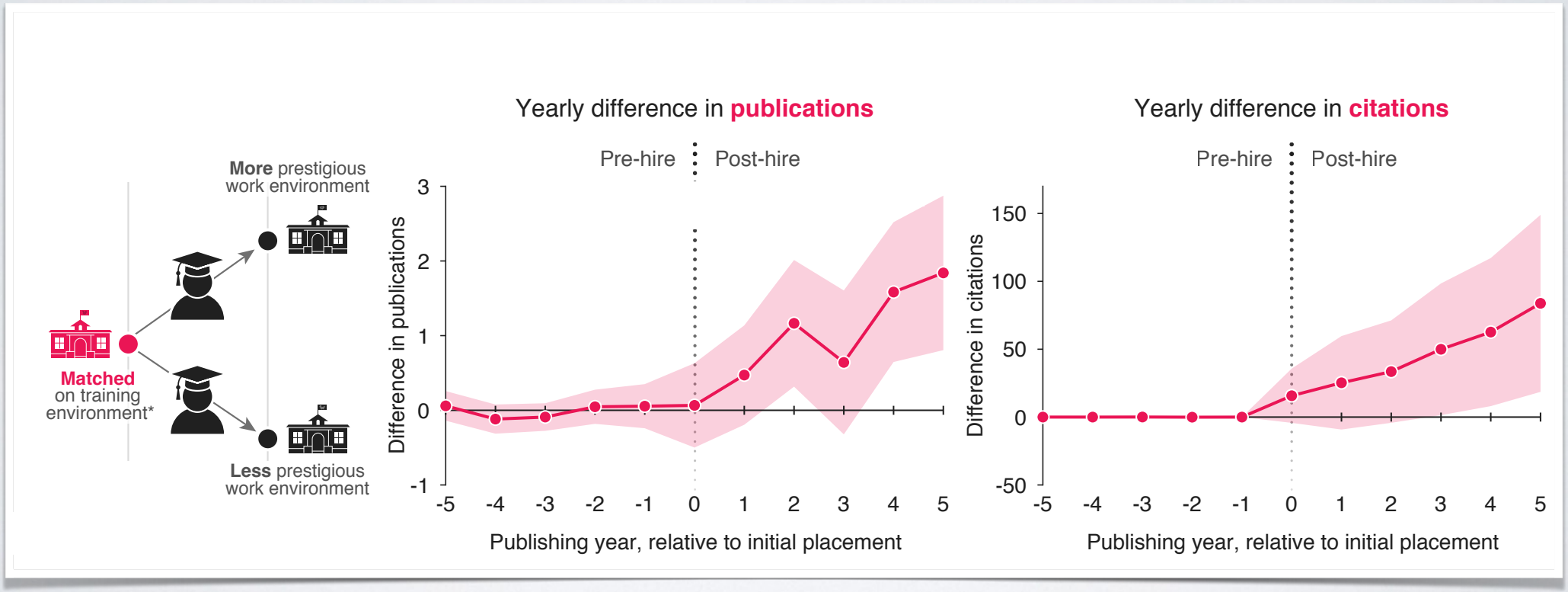


caliper matched faculty on {gender, subfield, hiring prestige OR phd prestige, year of placement, postdoctoral training}. results robust to caliper variations

publications: N = 196 pairs citations: N = 96 pairs

quasi-natural experiments : faculty hiring

similar training, different placement → different productivity & prominence



caliper matched faculty on {gender, subfield, hiring prestige OR phd prestige, year of placement, postdoctoral training}. results robust to caliper variations

publications: N = 196 pairs citations: N = 96 pairs

prestige shapes productivity — why?

conditioned on holding a faculty position:

more elite training *does not drive* greater scholarly impact *relative to peers in similar faculty positions*

work environment, not training appears to drive impact

we couldn't control for recommendation letters, job talk quality, charisma, future funding prospects, etc., which may correlate with greater future productivity / prominence. however, these should correlate with pre-hire productivity, which we control for. further, research on mid-career moves supports our finding (see Allison & Long 1990)

prestige shapes productivity — why?

conditioned on holding a faculty position:

more elite training *does not drive* greater scholarly impact *relative to peers in similar faculty positions*

work environment, not training appears to drive impact

why do elite institutions dominate science?

doctoral prestige → faculty location

Clauset et al. (2015), Way et al. (2016)

we couldn't control for recommendation letters, job talk quality, charisma, future funding prospects, etc., which may correlate with greater future productivity / prominence. however, these should correlate with pre-hire productivity, which we control for. further, research on mid-career moves supports our finding (see Allison & Long 1990)

prestige shapes productivity — why?

conditioned on holding a faculty position:

more elite training *does not drive* greater scholarly impact *relative to peers in similar faculty positions*

work environment, not training appears to drive impact

why do elite institutions dominate science?

doctoral prestige → faculty location → scholarly impact

these results, Way et al. (2019)

why? because of **working environment**
[not selection, not retention, not expectations]

we couldn't control for recommendation letters, job talk quality, charisma, future funding prospects, etc., which may correlate with greater future productivity / prominence. however, these should correlate with pre-hire productivity, which we control for. further, research on mid-career moves supports our finding (see Allison & Long 1990)

prestige shapes productivity

- where you work is an ***environmental mechanism*** to explain cumulative advantage
- individual productivity and prominence ***cannot be separated*** from their place in the academic system

future work:

how does this effect vary across different *fields*? why?

what department-level *facilitation* mechanisms drive productivity?

what role for differences in *students*?

Way, Morgan, Larremore, Clauset, *Proc. Natl. Acad. Sci. USA* **116**, 10729 (2019)