Faculty hiring and the spread of scientific ideas

Allison Morgan, Dimitrios Economou, Samuel Way, Aaron Clauset

Building Better Epistemic Networks Workshop, Sept 5th 2019
Goals for this talk:

1. Quantify prestige in academia
2. Identify a structural mechanism for how prestige affects research progress
3. Discuss implications
Goals for this talk:

1. Quantify prestige in academia
2. Identify a structural mechanism for how prestige affects research progress
3. Discuss implications
Academic workforce

Trajectories of individuals form a network
Trajectories of individuals form a network

- BS, Physics, Haverford College
- BA, Physics, Reed College
- PhD, Computer Science, University of New Mexico
- Fellow, Santa Fe Institute
- MS, Computer Science, University of Colorado
- PhD, Computer Science, University of New Mexico
- Assistant Professor, University of Colorado
- Associate Professor, University of Colorado
Faculty hiring networks

Each directed edge $u \rightarrow v$
PhD from $u \rightarrow$ faculty at $v$

[US academia: big, mobile, self-contained, competitive]

Dramatic inequality in PhD production [80/20 rule holds]

Common large-scale structure:
influential, well-connected core

Small percentage of edges are self-loops [8% in CS]

Assumption: reveals collective preferences. Hiring committees want to hire the best candidates

Computer science faculty hiring network; http://tuvalu.santafe.edu/~aaronc/facultyhiring/
Quantifying prestige

Many rankings exist:


But they often describe the inputs to the system, or don't reflect what universities truly think of each other.

Let's construct a ranking based on revealed university preferences through hiring.

https://en.wikipedia.org/wiki/College_and_university_rankings
Quantifying prestige

Compute an average **Minimum Violation Rank** (MVR) to find an order of nodes (universities) that minimizes “upsets.”

Select permutation (a ranking) $\pi$ that minimizes the number of “rank violations”: edges $(u, v)$ where $\pi_v < \pi_u$.

Higher-ranked universities have greater placement power.

**Systematic inequality and hierarchy in faculty hiring networks**

Aaron Clauset,1,2,3* Samuel Arbesman,4 Daniel B. Larremore5,6

Aside: Application of prestige

We can use these rankings to generate predictions for individuals and the system.

Consider the annual matching process of candidates to openings.

Each year $t$, has $\{u_i\}$ candidate “stubs” and $\{v\}_t$ opening “stubs.”

Given a pair $(u_i, v)$, the probability of matching depends on that pair’s features $x[u_i, v]$.
Aside: Application of prestige

We can use these rankings to generate predictions for individuals and the system.

Consider the annual matching process of candidates to openings.

Each year $t$, has $\{u_i\}$ candidate “stubs” and $\{v\}_t$ opening “stubs.”

Given a pair $(u_i, v)$, the probability of matching depends on that pair’s features $x[u_i, v]$. 
Aside: Application of prestige

We can use these rankings to generate predictions for individuals and the system.

Consider the annual matching process of candidates to openings.

Each year $t$, has $\{u_t\}$ candidate “stubs” and $\{v_t\}$ opening “stubs.”

Given a pair $(u_i, v)$, the probability of matching depends on that pair’s features $x[u_i, v]$. 
Aside: Application of prestige

We can use these rankings to generate predictions for individuals and the system.

Consider the annual matching process of candidates to openings.

Each year $t$, has $\{u_t\}$ candidate “stubs” and $\{v_t\}$ opening “stubs.”

Given a pair $(u_i, v)$, the probability of matching depends on that pair’s features $x[u_i, v]$.

The most important modeling features are **differences in prestige** and productivity.
Features of hierarchy

**systematic**
90% of hiring movement is “down” the hierarchy

**steep**
< 7% of faculty have PhD from lower 75% of universities

**biased**
median change for women ~3 ranks worse than men

http://danlarremore.com/faculty/
Explore the data for History, CS, Business
Core-periphery position changes with rank

What are the implications?

University Prestige (π / n)

Mean geodesic distance / diameter
Shape of the faculty hiring network

Universities in the core are (obviously) very close to all other core universities.

**Next:** core position enables substantial influence over research **agendas**, research **communities**, and departmental **norms** throughout a discipline.
Large inequalities in placement power

Faculty flow out of core, into periphery

Modest fraction stays inside core

Small fraction flows “upstream”

Prestige describes influence via individuals placement

Next: How does prestige affect science as a system?
Goals for this talk:

1. Quantifying prestige in academia
2. Identify a structural mechanism for how prestige affects research progress
3. Discuss implications
Science is a meritocracy... right?
Science is a meritocracy… right?
Some scientists and institutions are far more influential than others

**The Matthew Effect in Science**

The reward and communication systems of science are considered.

Robert K. Merton

*Science* 159.3810, 56-63 (1968)

**Reputation and impact in academic careers**


**DEPARTMENTAL EFFECTS ON SCIENTIFIC PRODUCTIVITY**

Paul D. Allison

J. Scott Long


**Publication, Power, and Patronage: On Inequality and Academic Publishing**

Chad Wellmon and Andrew Piper

*Critical Inquiry* (2017)

**Inputs, Outputs, and the Prestige of University Science Departments**

Warren O. Hagstrom


**Professional Standing and the Reception of Scientific Discoveries**

Stephen Cole

Three explanations

(1) genuine differences in merit
(2) non-meritocratic social processes
(3) non-meritocratic structural factors
Three explanations

1. genuine differences in merit
2. non-meritocratic social processes
3. non-meritocratic structural factors

Faculty hiring as a mechanism

**R1:** Are research ideas carried by faculty hiring?
Faculty hiring as a mechanism

**R1:** Are research ideas carried by faculty hiring?
Faculty hiring as a mechanism

**R1:** Are research ideas carried by faculty hiring?

*W. Lamb, J. Wheeler, A. Pais, R. Feynman, H. Feshbach, J. Schwinger*
Faculty hiring as a mechanism

**R1:** Are research ideas carried by faculty hiring?

**R2:** Does the structure of the faculty hiring network affect the spread of ideas?
Data

Education & employment for 4,388 faculty from 205 U.S. and Canadian CS departments employed between 2011-2012

- Institution (node) $u$ with unique prestige $\pi$
- Edge $(u, v)$ represents a single PhD candidate from $u$ who got an assistant faculty position at $v$

Over 2M publication records for tenure-track faculty.

- Title, author list, venue, and date
- Matched with employment start dates
Data

Education & employment for 4,388 faculty from 205 U.S. and Canadian CS departments employed between 2011-2012
- Institution (node) $u$ with unique prestige $\pi$
- Edge $(u, v)$ represents a single PhD candidate from $u$ who got an assistant faculty position at $v$

Over 2M publication records for tenure-track faculty.
- Title, author list, venue, and date
- Matched with employment start dates

Publication records

R1: Are research ideas carried by faculty hiring?

For each department that has adopted a research idea, either:
R1: Are research ideas carried by faculty hiring?

For each department that has adopted a research idea, either:

(A) Idea X adopted **by hiring**
Someone who works on the idea was hired by the department.

(B) Idea X adopted **not by hiring**
Someone already in the department started working on the idea.
**R1**: Are research ideas carried by faculty hiring?

**Test**: choose five research topics and evaluate the fraction of times those topics spread via (A) in real life, compared to the expected fraction under a permutation of publication titles.

### Diagram

- **(A)**: Idea X adopted by hiring
  - Someone who works on the idea was hired by the department.
  - Timeline:
    - 1970
    - 2011
    - Time employed by department
    - Publication on Idea X

- **(B)**: Idea X adopted not by hiring
  - Someone already in the department started working on the idea.
  - Timeline:
    - 1970
    - 2011
    - Time employed by department
    - Publication on Idea X
R1: Are research ideas carried by faculty hiring?

Test: choose five research topics and evaluate the fraction of times those topics spread via (A) in real life, compared to the expected fraction under a permutation of publication titles

Recover 241 spreading events for the five topics, each affecting between 11-58% of departments

- 88 (37%) of these happen by way of hiring
  - 71 (81%) of those, move via from high to low prestige universities
R1: Are research ideas carried by faculty hiring?

**Test:** choose five research topics and evaluate the fraction of times those topics spread via (A) in real life, compared to the expected fraction under a permutation of publication titles.

<table>
<thead>
<tr>
<th>Research Areas</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>topic modeling</td>
<td>✓</td>
</tr>
<tr>
<td>incremental computing</td>
<td>✓</td>
</tr>
<tr>
<td>quantum computation</td>
<td>✓</td>
</tr>
<tr>
<td>mechanism design</td>
<td>✓</td>
</tr>
<tr>
<td>deep learning</td>
<td>✗</td>
</tr>
</tbody>
</table>
R1: Are research ideas carried by faculty hiring?

Test: choose five research topics and evaluate the fraction of times those topics spread via (A) in real life, compared to the expected fraction under a permutation of publication titles.

Faculty hiring acts as a mechanism for the spread of ideas, with differential effects by topic, across the computer science community.
R2: Does the structure of the faculty hiring network affect the spread of ideas?

To simulate the diffusion of ideas, use a **Susceptible-Infected** (SI) model.

Seed an epidemic at a university with unique prestige $\pi$, varying the transmissibility $p$ (quality of an idea)

Quality of idea relates to how many nodes will adopt an idea (on average)

Measure the fraction of universities which adopted the idea
**R2:** Does the structure of the faculty hiring network affect the spread of ideas?

To simulate the diffusion of ideas, use a **Susceptible-Infected** (SI) model.

Seed an epidemic at a university with unique prestige $\pi$, varying the transmissibility $p$ (quality of an idea).

Quality of idea relates to how many nodes will adopt an idea (on average).

Measure the fraction of universities which adopted the idea.
**R2:** Does the structure of the faculty hiring network affect the spread of ideas?

To simulate the diffusion of ideas, use a **Susceptible-Infected** (SI) model.

Seed an epidemic at a university with unique prestige $\pi$, varying the transmissibility $\rho$ (quality of an idea).

Quality of idea relates to how many nodes will adopt an idea (on average).

Measure the fraction of universities which adopted the idea.
R2: Does the structure of the faculty hiring network affect the spread of ideas?

Remember: Core-periphery position changes with prestige
R2: Does the structure of the faculty hiring network affect the spread of ideas?

High Prestige ($\pi = 2.23$)  
Medium Prestige ($\pi = 68.17$)  
Low Prestige ($\pi = 130.66$)

Explore more simulations  
https://pikawolfy.github.io/epistemicInequality/
**R2:** Does the structure of the faculty hiring network affect the spread of ideas?

High Prestige ($\pi = 2.23$)  
Medium Prestige ($\pi = 68.17$)  
Low Prestige ($\pi = 130.66$)

Explore more simulations  
https://pikawolfy.github.io/epistemicInequality/
R2: Does the structure of the faculty hiring network affect the spread of ideas?
R2: Does the structure of the faculty hiring network affect the spread of ideas?

Average of 10K simulations originating from $u$ with particular quality $\rho$

Infection Rate $\rho$ ("Idea Quality")
- $0.1$
- $0.3$
- $0.5$
- $0.7$
- $0.9$
**R2:** Does the structure of the faculty hiring network affect the spread of ideas?

*Assuming independence between quality of idea and origin; https://en.wikipedia.org/wiki/Meritocracy#Etymology*
R2: Does the structure of the faculty hiring network affect the spread of ideas?
R2: Does the structure of the faculty hiring network affect the spread of ideas?

Good ideas spread more easily from high-prestige universities
**R2:** Does the structure of the faculty hiring network affect the spread of ideas?

Great ideas can spread regardless of starting place
R2: Does the structure of the faculty hiring network affect the spread of ideas?

We may lose medium quality research ideas because the system structurally disallows their spread.

![Graph showing the fraction of the network infected against university prestige, indicating a decrease in infected fraction as prestige increases.](graph.png)
Aside: What about other fields?

Gini coefficient for history is 0.72, business is 0.62, and computer science is 0.69.
Aside: What about other mechanisms?

Allowing a single jump to a disconnected node. Transmission probability is held constant at 0.1
R2: Does the structure of the faculty hiring network affect the spread of ideas?

Good ideas can spread further and faster from prestigious universities, but great ideas can spread from any university.
R2: Does the structure of the faculty hiring network affect the spread of ideas?

Good ideas can spread further and faster from prestigious universities, but great ideas can spread from any university.

So what?
Goals for this talk:

1. Quantifying prestige in academia
2. Identify a structural mechanism for how prestige affects research progress
3. Discuss implications
Ideas spread in academia via faculty hiring. The structure of this network can privilege elite institutions.

Caveats: Model assumes quality is independent of institution and hiring decisions.

To mitigate this, we could try to remove signals of prestige from our evaluations of quality (e.g., double-blind review).

Systems which incentivize a large quantity of incremental ideas will tend promote the visibility of prestigious researchers.

Thought experiments: What if hiring was random? What if the lowest ranked universities chose first? What other non-meritocratic mechanisms might be at play? Could we validate these findings empirically?
Thanks!

Collaborators: Dimitrios Economou, Samuel Way, Aaron Clauset, Daniel Larremore, McKenzie Mae Weller

Email: allison.morgan@colorado.edu
Twitter: @alliecmorgan

University of Colorado Boulder

NSF