Academic Employment Networks and Departmental Rank

Debra Hevenstone

University of Michigan, Public Policy, Sociology, & Complex Systems
ETH, visitor to Soziologie

European Conference on Complex Systems
October 3, 2007
## Persistent Top 10 Sociology Rankings

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Chicago (1)</td>
<td>Chicago (1)</td>
<td>Chicago (1)</td>
<td>Chicago (1)</td>
<td>Chicago (3)</td>
</tr>
<tr>
<td>Wisconsin (3)</td>
<td>Wisconsin (2)</td>
<td>Wisconsin (2)</td>
<td>Wisconsin (2)</td>
<td>Wisconsin (1)</td>
</tr>
<tr>
<td>Minnesota (4)</td>
<td>UCLA (9)</td>
<td>UCLA (6)</td>
<td>UCLA (5)</td>
<td>UCLA (8)</td>
</tr>
<tr>
<td>Missouri (7)</td>
<td>Chapel Hill (6)</td>
<td>Chapel Hill (4)</td>
<td>Chapel Hill (6)</td>
<td>Chapel Hill (4)</td>
</tr>
<tr>
<td>Berkeley (3)</td>
<td>Berkeley (2)</td>
<td>Berkeley (3)</td>
<td>Berkeley (3)</td>
<td>Berkeley (2)</td>
</tr>
<tr>
<td>Columbia (2)</td>
<td>Columbia (8)</td>
<td>Northwestern (9)</td>
<td>Northwestern (9)</td>
<td>UPenn (9)</td>
</tr>
</tbody>
</table>

* Hughes, Raymond M., *A Study of the Graduate Schools in America*  
+ National Research Council, *Research Doctorate Programs in the United States*  
** US News and World Report, *America’s Best Graduate Schools*
Central Questions

In the literature on academic networks & prestige

- Does trading faculty reinforce caste?
- Or are these faculty simply better trained?
- PhD exchange networks vs career networks?
- Determined by department/cohort size?
- Determined by methodological limitations?
Sociology Labor Market

+ low unemployment rates
+ retiring baby boomers
− movement towards adjunct / lecturers

Average Annual Salaries

- instructor/lecturer
- assistant professor
- associate professor
- full professor

ASA average
Michigan

Average Salary Comparison

Debra Hevenstone (University of Michigan)
Academic hiring literature

- **Determinants of career success** (Bair, Baldi, Burke, Hargens, etc.)
  - prestige -> first job
  - publications & years to graduation -> productivity
- **PhD exchange networks:** **centrality** and **prestige**
  - Computer Science & iSchool – PageRank (Wiggins)
  - Sociology – Eigenvector Centrality (Burris)
  - Political Science – Hubs and Authorities (Fowler)
- **Post PhD exchange networks**
  - Sociology (Grannis)
Role of department size?

- Experienced faculty
  - High turnover

- Citation rates
  - Faculty awards
  - Research budget
  - Student Placement

- Network degree

- Network Centrality

- Prestige
  - Rank
Introduction

Role of department size?

- Experienced faculty
  - High turnover
- Department size (faculty & students)
- Faculty trained by dept
- Faculty survey
- Citation rates
  - Faculty awards
  - Research budget
  - Student Placement
- Network degree
- Network Centrality
- Prestige Rank

Debra Hevenstone (University of Michigan)
Where were professors trained?

- Chicago trained > 10% of faculty at the top 10
- Conclusion: Top 10 hire from top 10
Where were professors trained?

- Chicago trained > 10% of faculty at the top 10
- Conclusion: Top 10 hire from top 10
- *So do others*, plus Austin, SUNY’s & local schools
Academic networks & methodological obstacles

- Network isolation (Grannis)
- Centrality measures (Barabasi et al, Goyal, Newman)
- Time (sample and coding) (Barabasi et al)
- Bipartite reduction (Borgatti & Everett, Robins & Alexander)
Academic networks & methodological obstacles

- Network isolation (Grannis)
- Centrality measures (Barabasi et al, Goyal, Newman)
- Time (sample and coding) (Barabasi et al)
- Bipartite reduction (Borgatti & Everett, Robins & Alexander)
Academic networks & methodological obstacles

- Network isolation (Grannis)
- Centrality measures (Barabasi et al, Goyal, Newman)
- Time (sample and coding) (Barabasi et al)
- Bipartite reduction (Borgatti & Everett, Robins & Alexander)
Data collection

- Chose 2 samples of 9 and 6 departments
- Downloaded CVs from department web sites
- Included faculty, excluded adjuncts, visitors, emeritus
- Augmented missing data with google searches
- Coded 3 types of edges:
  - PhD training institution
  - Non-tenure track job (visiting, post-doc, non-academic)
  - Tenure-track job
- 7% missing CV’s (current job and training coded)
Network data

- Sample One
  - Wisconsin, U Michigan, Harvard, Berkeley, UCLA, U Chicago, Brown, Stanford, U Arizona
  - 193 institutions
    - 99 ranked
  - 256 professors
    - 886 links

- Sample Two
  - Yale, U Penn, Northwestern, Princeton, Johns Hopkins, NYU
  - 241 institutions
    - 89 ranked
  - 182 professors
    - 882 links
## Graph specification

<table>
<thead>
<tr>
<th>Method</th>
<th>Edge Inclusion</th>
<th>Graph Reduction</th>
<th>Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all edges, no PhD, PhD &amp; tenure</td>
<td>bipartite or reduced</td>
<td>one or two</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

**12 graphs**
Centrality measures

- **Standardized Degree**
  \[ S_b(i) = d_i \]
  \[ S_r(i) = d_i - 1 \]

- **Closeness**
  \[ C_b(i) = \frac{2}{n_p + n_d - 1} \left( \sum_k D_{ik} + \sum_j D_{ij} \right) \]
  \[ C_r(i) = \frac{n_d - 1}{\sum_k D_{ik}} \]

- **Eigenvector Centrality**
  \[ E_i = \alpha \left( \frac{1}{n_d} + \frac{1}{n_p} \sum_j A_{ij} E_j \right) \]

- **bipartite**
- **r reduced**

- **d_i** degree
- **n_p** number professors
- **n_d** number departments
- **D_{ik}, D_{ij}** distance between i & (k or j)
- **k** set of institutions
- **j** set of professors
- **A_{ij}** = 0, if not connected
  = 1, if connected (bipartite)
  = n, if connected (reduced)
Centrality measures

**Standardized Degree**

\[
S_{bi}^b = \frac{d_i}{n_p}
\]

\[
S_{ri}^r = \frac{d_i}{n_d-1}
\]

**Closeness**

\[
C_{bi}^b = \frac{2(n_p+n_d-1)}{\left(\sum_k D_{ik} + \sum_j D_{ij}\right)}
\]

\[
C_{ri}^r = \frac{n_d-1}{\left(\sum_k D_{ik}\right)}
\]

**Eigenvector Centrality**

\[
E_i = \alpha \sum_{j=1}^{n_d+n_p} A_{ij} E_j
\]

- \(b\) bipartite
- \(r\) reduced
- \(d_i\) degree
- \(n_p\) number professors
- \(n_d\) number departments
- \(D_{ik}\) or \(D_{ij}\) distance between i & (k or j)
- \(k\) set of institutions
- \(j\) set of professors

- \(A_{ij} = 0\), if not connected
- \(A_{ij} = 1\), if connected (bipartite)
- \(A_{ij} = n\), if connected (reduced)
Exogenous variables

- **Domestic Sociology Rank: National Research Council**
  - size, tenure track, faculty & student funding, student demographics, peer assessment

- **International University Rank: US News and World Report**
  - student retention, students' qualifications, proportion accepted, faculty resources, student to faculty ratio, alumni giving, peer assessment

- **Faculty Size:**
  - domestic: NRC report
  - foreign: website listings

Correlation = 0.625
Networks, pictures
Sample 2, reduced, PhD & tenure

Plots based on Kamada Kawai spring algorithm
- Minimizes length of connections
- Treats thicker connections as stronger
Networks, pictures
Sample 2, reduced, PhD & tenure

Sample 1, bipartite, all edges

- Plots based on Kamada Kawai spring algorithm
  - Minimizes length of connections
  - Treats thicker connections as stronger
## Correlation between centrality measures

<table>
<thead>
<tr>
<th>sample</th>
<th>graph type</th>
<th>edges</th>
<th>eigen to closeness</th>
<th>eigen to degree</th>
<th>degree to closeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>yes</td>
<td>all</td>
<td>.614 (.733)</td>
<td>.589 (.888)</td>
<td>.889 (.840)</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>PhD &amp; tenure</td>
<td>.915 (.866)</td>
<td>.948 (.989)</td>
<td>.922 (.878)</td>
</tr>
<tr>
<td>1</td>
<td>yes</td>
<td>no PhD</td>
<td>.929 (.891)</td>
<td>.935 (.977)</td>
<td>.927 (.870)</td>
</tr>
<tr>
<td>1</td>
<td>no</td>
<td>all</td>
<td>.979 (.759)</td>
<td>.865 (.990)</td>
<td>.836 (.792)</td>
</tr>
<tr>
<td>1</td>
<td>no</td>
<td>PhD &amp; tenure</td>
<td>.983 (.681)</td>
<td>.935 (.981)</td>
<td>.930 (.732)</td>
</tr>
<tr>
<td>1</td>
<td>no</td>
<td>no PhD</td>
<td>.949 (.648)</td>
<td>.787 (.873)</td>
<td>.827 (.810)</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>all</td>
<td>.959 (.828)</td>
<td>.931 (.987)</td>
<td>.958 (.863)</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>PhD &amp; tenure</td>
<td>.977 (.897)</td>
<td>.955 (.944)</td>
<td>.969 (.951)</td>
</tr>
<tr>
<td>2</td>
<td>yes</td>
<td>no PhD</td>
<td>.975 (.854)</td>
<td>.940 (.924)</td>
<td>.955 (.928)</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>all</td>
<td>.985 (.811)</td>
<td>.903 (.988)</td>
<td>.905 (.817)</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>PhD &amp; tenure</td>
<td>.961 (.721)</td>
<td>.936 (.974)</td>
<td>.952 (.764)</td>
</tr>
<tr>
<td>2</td>
<td>no</td>
<td>no PhD</td>
<td>.908 (.767)</td>
<td>.841 (.944)</td>
<td>.911 (.795)</td>
</tr>
</tbody>
</table>

- entries are rank correlations
- (...) are continuous correlations
Centrality & prestige, rank correlations

eigen rank = .681
closeness = .717
degree = .732

USA

eigen rank = .546
closeness = .588
degree = .586

World incl. US
The “best” predictions by graph type

- Compare centrality measures’ rank versus NRC rank

\[
\epsilon = \sum_{i=1}^{i=10} (S^r_i - p^r)^2 + (C^r_i - p^r)^2 + (E^r_i - p^r)^2
\]

- $S^r_i$ degree centrality rank
- $C^r_i$ closeness centrality rank
- $E^r_i$ eigenvector centrality rank
- $p^r$ exogenous rank
- $i$ school

$\implies$ The graph with the least $\epsilon$ is the best graph
## Centrality ranks for the top 10 NRC departments

<table>
<thead>
<tr>
<th>Institution</th>
<th>best graph</th>
<th>worst graph</th>
<th>average rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>eigen rank</td>
<td>close rank</td>
<td>degree rank</td>
</tr>
<tr>
<td>UChicago</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Berkeley</td>
<td>5</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>UMichigan</td>
<td>4</td>
<td>5</td>
<td>4</td>
</tr>
<tr>
<td>UCLA</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Chapel Hill</td>
<td>15</td>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>Harvard</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Stanford</td>
<td>7</td>
<td>7</td>
<td>7</td>
</tr>
<tr>
<td>Northwestern</td>
<td>11</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>UW</td>
<td>37</td>
<td>28</td>
<td>29</td>
</tr>
</tbody>
</table>

*best = sample 1, bipartite, no non-tenure edges*

*worst = sample 1, reduced, all edges*

- **sample 1**
- **sample 2**

Sampled schools always make it into the top 10
### Predicting rank with centrality, OLS

<table>
<thead>
<tr>
<th></th>
<th>domestic ranking</th>
<th>international ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>eigen-vector close</td>
<td>.542***</td>
<td>.608***</td>
</tr>
<tr>
<td>degree</td>
<td>.584***</td>
<td>.669***</td>
</tr>
<tr>
<td>faculty</td>
<td>.577***</td>
<td>.686***</td>
</tr>
<tr>
<td>$R^2$</td>
<td>.426 .493 .469</td>
<td>.239 .293 .314</td>
</tr>
</tbody>
</table>

- **bivariate coefficient**

- **Sample size**
  - 58 institutions with domestic ranks, 58 controlling for faculty size
  - 82 institutions with foreign ranks, 80 controlling for faculty size
## Predicting rank with centrality, OLS

<table>
<thead>
<tr>
<th></th>
<th>domestic ranking</th>
<th>international ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>eigenvector</strong></td>
<td>.477***</td>
<td>.584***</td>
</tr>
<tr>
<td>close</td>
<td>.542***</td>
<td>.608***</td>
</tr>
<tr>
<td></td>
<td>.521***</td>
<td>.659***</td>
</tr>
<tr>
<td></td>
<td>.584***</td>
<td>.669***</td>
</tr>
<tr>
<td><strong>degree</strong></td>
<td>.516***</td>
<td>.665***</td>
</tr>
<tr>
<td></td>
<td>.577***</td>
<td>.432***</td>
</tr>
<tr>
<td></td>
<td>.557***</td>
<td>.438**</td>
</tr>
<tr>
<td></td>
<td>.587***</td>
<td>.467***</td>
</tr>
<tr>
<td><strong>faculty</strong></td>
<td>-.494*</td>
<td>-.119</td>
</tr>
<tr>
<td></td>
<td>-.462***</td>
<td>-.079**</td>
</tr>
<tr>
<td></td>
<td>-.499***</td>
<td>-.065</td>
</tr>
<tr>
<td></td>
<td>-.467***</td>
<td>-.061</td>
</tr>
<tr>
<td><strong>$R^2$</strong></td>
<td>.567</td>
<td>.250</td>
</tr>
<tr>
<td></td>
<td>.567</td>
<td>.311</td>
</tr>
<tr>
<td></td>
<td>.557</td>
<td>.318</td>
</tr>
<tr>
<td></td>
<td>.580</td>
<td>.328</td>
</tr>
<tr>
<td><strong>coeff tests</strong></td>
<td>$\beta_{eig} = \beta_{degree}$</td>
<td>$\beta_{eig} = \beta_{degree}$</td>
</tr>
<tr>
<td></td>
<td>(.0013)</td>
<td>(.0001)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{eig} = \beta_{closeness}$</td>
<td>$\beta_{eig} = \beta_{closeness}$</td>
</tr>
<tr>
<td></td>
<td>(.0004)</td>
<td>(.0036)</td>
</tr>
<tr>
<td></td>
<td>$\beta_{closeness} = \beta_{degree}$</td>
<td>$\beta_{closeness} = \beta_{degree}$</td>
</tr>
<tr>
<td></td>
<td>(.0004)</td>
<td>(.9789)</td>
</tr>
</tbody>
</table>

- **bivariate coefficient**
- **controlling for faculty size**
- **Sample size**
  - 58 institutions with domestic ranks, 58 controlling for faculty size
  - 82 institutions with foreign ranks, 80 controlling for faculty size
K-Core, definition

- Largest subgraph \( S \) in which all vertices have degree \( \geq k \) within \( S \)
  - Calculated by recursively pruning vertices with degree \( < k \)
  - Members of cores \( k=n \) are also defacto members of \( k = n-1 \)
  - Members of core \( k =n \) are those who are not in \( k = n+1 \) core
  - Members need not be connected to all other members
K-Core, definition

- Largest subgraph $S$ in which all vertices have degree $\geq k$ within $S$
  - Calculated by recursively pruning vertices with degree $< k$
  - Members of cores $k=n$ are also defacto members of $k = n-1$
  - Members of core $k = n$ are those who are *not* in $k = n+1$ core
  - Members need not be connected to all other members
K-Core, definition

- Largest subgraph $S$ in which all vertices have degree $\geq k$ within $S$
  - Calculated by recursively pruning vertices with degree $< k$
  - Members of cores $k=n$ are also defacto members of $k = n-1$
  - Members of core $k =n$ are those who are *not* in $k = n+1$ core
  - Members need not be connected to all other members
K-Core, two example graphs

Typical result

Sample 1, reduced, excluding non-tenure
Biggest k = 8
K-Core Analysis

K-Core, two example graphs

Typical result

Atypical result

Sample 1, reduced, excluding non-tenure
Biggest $k = 8$

Sample 2, reduced, excluding PhD training
Biggest $k = 19$, then $14$
K-Core, across all 12 graphs

- How often is each university in the top k-core?

- High ranked and sampled departments are often in the top k-core
- No visible differences among top schools
- Columbia is usually in the top core, despite not being sampled
Conclusions

- Findings
  - Robust prestige – centrality correlation
  - Top schools enhance relative prestige through PhD exchange
  - Independent of department size
  - Prestigious international vs. average domestic institutions

- But, is it a caste system?
  - Selected attrition would show this
  - Career ladder would show this
  - Two independent tiers would not
Further Research

- **Academic Labor Market**
  - Attrition & transition hazards
  - Visiting appointments’ role
  - European sample
  - Better sampling
  - Time code edges/ dynamic network

- **General Labor Market**
  - Are prestigious employers more central?
  - Core vs non-core (support) occupations?
## Networks, descriptive statistics

<table>
<thead>
<tr>
<th></th>
<th>nodes</th>
<th>orgs</th>
<th>edges</th>
<th>avg degree</th>
<th>avg distance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>sample 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bipartite all edges</td>
<td>479</td>
<td>193</td>
<td>886</td>
<td>4.59</td>
<td>1.92</td>
</tr>
<tr>
<td>reduced all edges</td>
<td>193</td>
<td>193</td>
<td>952</td>
<td>9.87</td>
<td>2.30</td>
</tr>
<tr>
<td>bipartite no non-tenure</td>
<td>386</td>
<td>99</td>
<td>642</td>
<td>6.57</td>
<td>1.73</td>
</tr>
<tr>
<td>reduced no non-tenure</td>
<td>99</td>
<td>99</td>
<td>321</td>
<td>6.48</td>
<td>2.35</td>
</tr>
<tr>
<td>bipartite no student</td>
<td>457</td>
<td>178</td>
<td>631</td>
<td>3.56</td>
<td>2.08</td>
</tr>
<tr>
<td>reduced no student</td>
<td>178</td>
<td>178</td>
<td>631</td>
<td>7.97</td>
<td>2.45</td>
</tr>
<tr>
<td><strong>sample 2</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>bipartite all edges</td>
<td>425</td>
<td>241</td>
<td>882</td>
<td>3.66</td>
<td>1.98</td>
</tr>
<tr>
<td>reduced all edges</td>
<td>241</td>
<td>241</td>
<td>1712</td>
<td>21.83</td>
<td>2.28</td>
</tr>
<tr>
<td>bipartite no non-tenure</td>
<td>273</td>
<td>89</td>
<td>509</td>
<td>5.79</td>
<td>3.83</td>
</tr>
<tr>
<td>reduced no non-tenure</td>
<td>89</td>
<td>89</td>
<td>331</td>
<td>7.44</td>
<td>2.37</td>
</tr>
<tr>
<td>bipartite no student</td>
<td>421</td>
<td>237</td>
<td>700</td>
<td>2.95</td>
<td>2.07</td>
</tr>
<tr>
<td>reduced no student</td>
<td>237</td>
<td>237</td>
<td>1533</td>
<td>12.90</td>
<td>2.35</td>
</tr>
</tbody>
</table>

- diameter = 4 for all graphs
Centrality and graph specification: hypotheses

- Eigenvector centrality rank will inflate sampled departments less
- Graph reduction will increase density and consequently uniformity
- Excluding non-tenure track edges will increase stratification
Centrality and graph specification: hypotheses

- Eigenvector centrality rank will inflate sampled departments less
  - Closeness centrality ranks the sampled departments 1.2 positions higher than eigenvector centrality.

- Graph reduction will increase density and consequently uniformity

- Excluding non-tenure track edges will increase stratification
Centrality and graph specification: hypotheses

- Eigenvector centrality rank will inflate sampled departments less
  - Closeness centrality ranks the sampled departments 1.2 positions higher than eigenvector centrality.

- Graph reduction will increase density and consequently uniformity
  - Eigenvector and closeness centrality have significantly wider standard deviations for bipartite graphs.

- Excluding non-tenure track edges will increase stratification
Centrality and graph specification: hypotheses

- Eigenvector centrality rank will inflate sampled departments less
  - Closeness centrality ranks the sampled departments 1.2 positions higher than eigenvector centrality.

- Graph reduction will increase density and consequently uniformity
  - Eigenvector and closeness centrality have significantly wider standard deviations for bipartite graphs.

- Excluding non-tenure track edges will increase stratification
  - Top ten institutions’ eigenvector centralities significantly increase excluding non-tenure edges and excluding PhD training edges.
Predicting centrality, multivariate analysis

<table>
<thead>
<tr>
<th></th>
<th>Eigen vector centrality</th>
<th>Degree centrality</th>
<th>Closeness centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bipartite</td>
<td>.152*</td>
<td>-.468***</td>
<td>.621***</td>
</tr>
<tr>
<td>Sample 1</td>
<td>.322***</td>
<td>.425***</td>
<td>.228**</td>
</tr>
<tr>
<td>All edges</td>
<td>-.198*</td>
<td>.137</td>
<td>.0427</td>
</tr>
<tr>
<td>No student edges</td>
<td>-.278**</td>
<td>-.163*</td>
<td>-.214**</td>
</tr>
<tr>
<td>R square</td>
<td>.188</td>
<td>.467</td>
<td>.495</td>
</tr>
</tbody>
</table>

**Standardized betas for top 10 schools**

- reducing the graph
  - + degree centrality but – others (overconnected effect)

- using sample 1
  - + all types of centrality (sample bias)

- excluding student edges
  - – all types of centrality (overtraining effect)