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

1925 (Hughes)*	1982 (NRC)+	1995 (US News)**	1995 (NRC)+	2005 (US News)**
Chicago (1)	Chicago (1)	Chicago (1)	Chicago (1)	Chicago (3)
Wisconsin (3)	Wisconsin (2)	Wisconsin (2)	Wisconsin (2)	Wisconsin (1)
Michigan (5)	Michigan (4)	Michigan (4)	Michigan (4)	Michigan (3)
Harvard (6)	Harvard (5)	Harvard (6)	Harvard (7)	Harvard (8)
Minnesota (4)	UCLA (9)	UCLA (6)	UCLA (5)	UCLA (8)
Missouri (7)	Chapel Hill (6)	Chapel Hill (4)	Chapel Hill (6)	Chapel Hill (4)
	Stanford (7)	Stanford (8)	Stanford (8)	Stanford (6)
	Berkeley (3)	Berkeley (2)	Berkeley (3)	Berkeley (2)
Columbia (2)	Columbia (8)	Northwestern (9)	Northwestern (9)	UPenn (9)
	Arizona (10)	Princeton (9)	Washington (10)	Princeton (6)

∗ Hughes, Raymond M., *A Study of the Graduate Schools in America*

+ National Research Council, *Research Doctorate Programs in the United States*

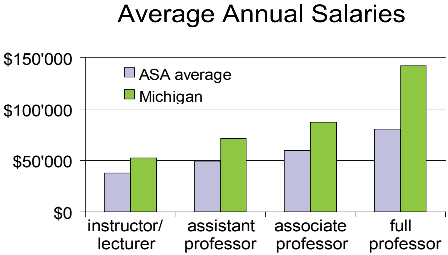
∗∗ US News and World Report, *America's Best Graduate Schools*

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?

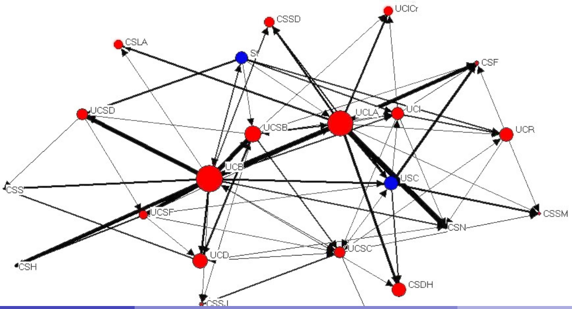
Sociology Labor Market

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

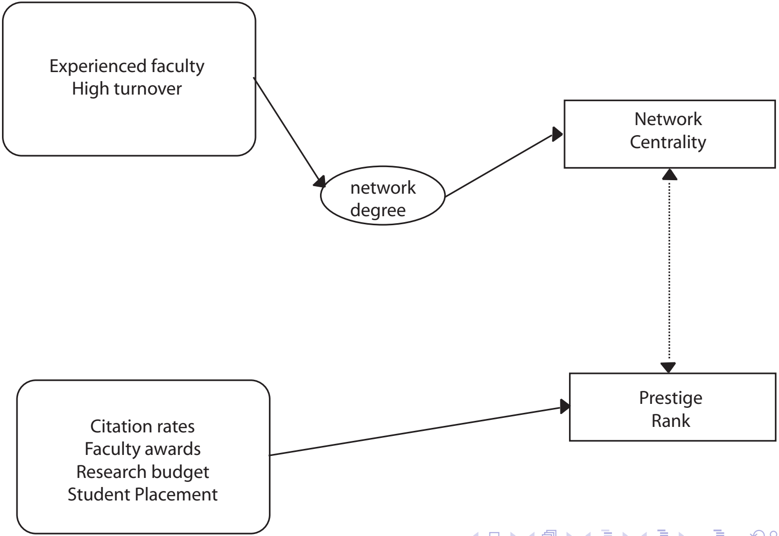


Academic hiring literature

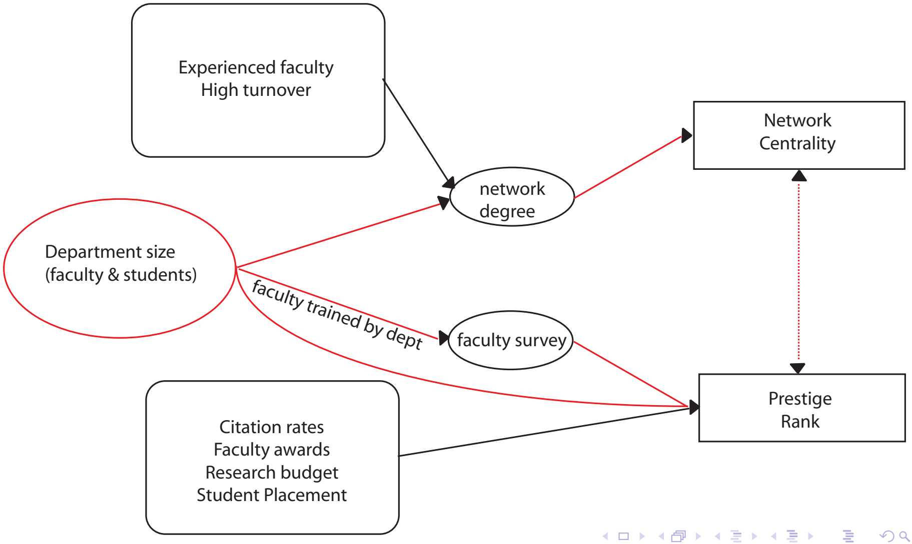
- Determinants of career success (Bair, Baldi, Burke, Hargens, etc.)
  - prestige -> first job
  - publications & years to graduation -> productivity
- PhD exchange networks: **centrality** and **prestige**
  - Computer Science & iSchool – PageRank (Wiggins)
  - Sociology – Eigenvector Centrality (Burris)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



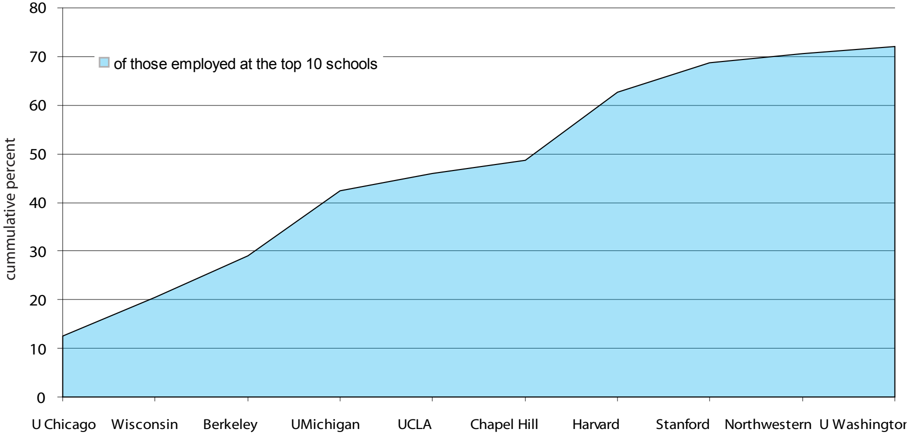
Role of department size?



Role of department size?

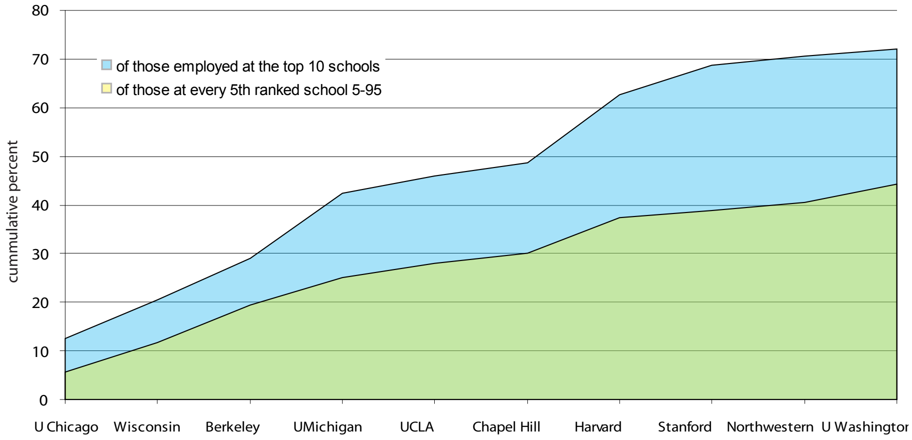


Where were professors trained?



- Chicago trained $>$ 10% of faculty at the top 10
- Conclusion: Top 10 hire from top 10

Where were professors trained?



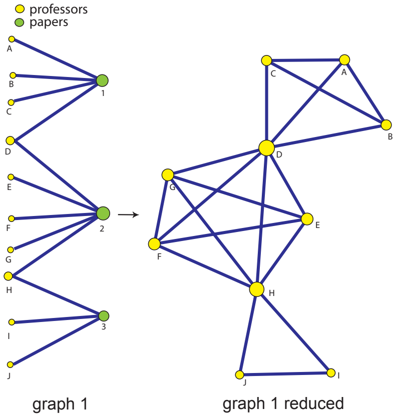
- Chicago trained $>$ 10% of faculty at the top 10
- Conclusion: Top 10 hire from top 10
- *So do others*, plus Austin, SUNY's & local schools

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)

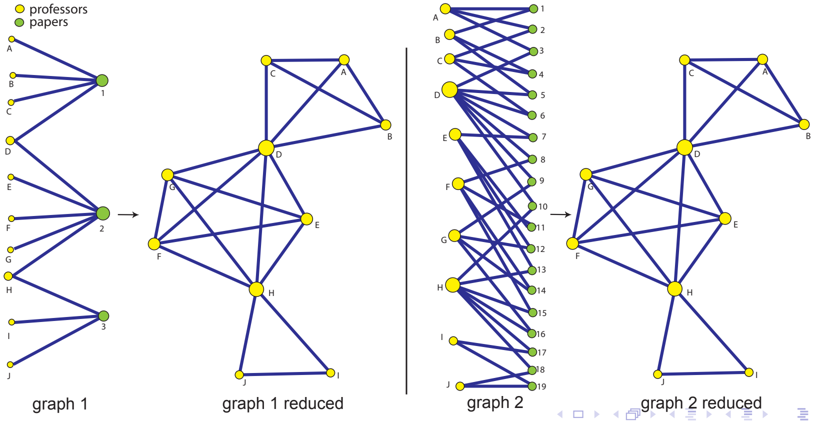
Academic networks & methodological obstacles

- 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)



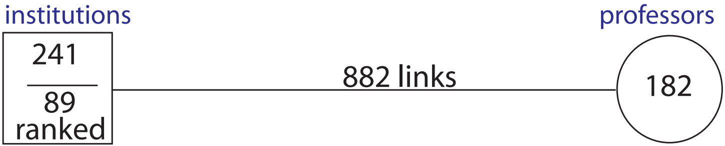
Data collection

- 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

- Sample Two
  - Yale, U Penn, Northwestern, Princeton, Johns Hopkins, NYU

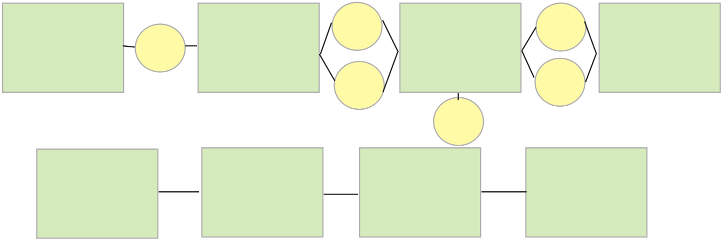


Graph specification

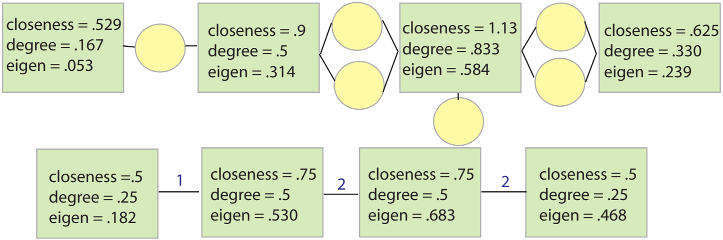
Edge Inclusion		Graph Reduction		Sample
all edges, no PhD, PhD & tenure		bipartite or reduced		one or two
3	$*$	2	$*$	2

12 graphs

Centrality measures



Centrality measures



Standardized Degree

$$S_i^b = \frac{d_i}{n_p}$$

$$S_i^r = \frac{d_i}{n_d - 1}$$

Closeness

$$C_i^b = \frac{2*(n_p+n_d-1)}{\left(\sum_k D_{ik} + \sum_j D_{ij}\right)}$$

$$C_i^r = \frac{n_d-1}{\left(\sum_k D_{ik}\right)}$$

Eigenvector Centrality

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

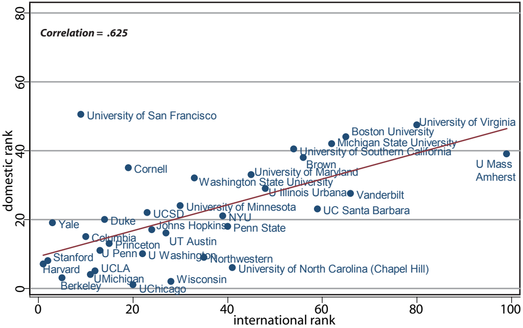
b	bipartite
r	reduced
d_i	degree
n_p	number professors
n_d	number departments

D_{ik} or D_{ij}	distance between i & (k or j)
k	set of institutions
j	set of professors

$$A_{ij} = \begin{cases} 0, & \text{if not connected} \\ 1, & \text{if connected (bipartite)} \\ n, & \text{if connected (reduced)} \end{cases}$$

Exogenous variables

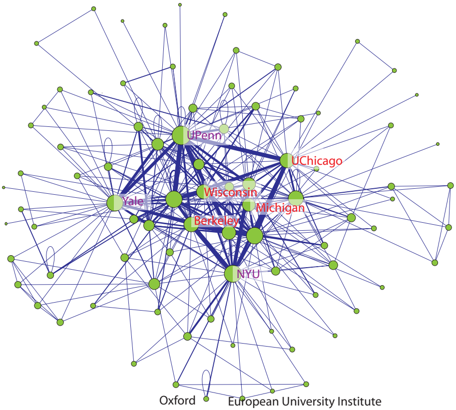
- Domestic Sociology Rank: National Research Council
  - size, tenure track, faculty & student funding, student demographics, peer assessment
- International University Rank: US News and World Report
  - student retention, students' qualifications, proportion accepted, faculty resources, student to faculty ratio, alumni giving, peer assessment



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

Networks, pictures

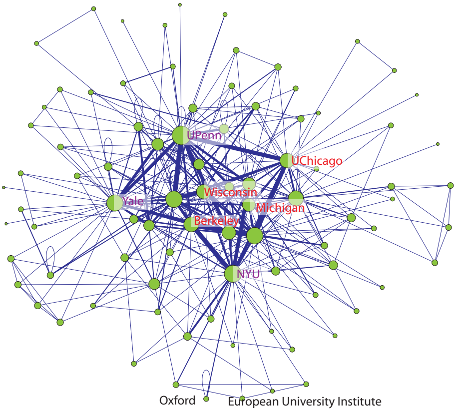
Sample 2, reduced, PhD & tenure



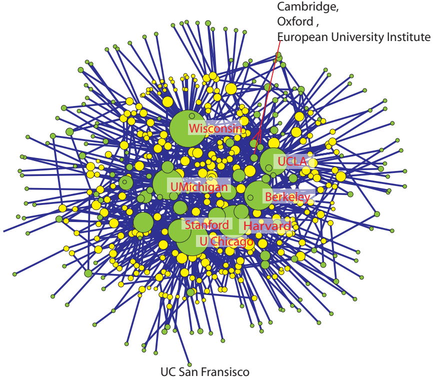
- Plots based on Kamada Kawai spring algorithm
  - Minimizes length of connections
  - Treats thicker connections as stronger

Networks, pictures

Sample 2, reduced, PhD & tenure



Sample 1, bipartite, all edges



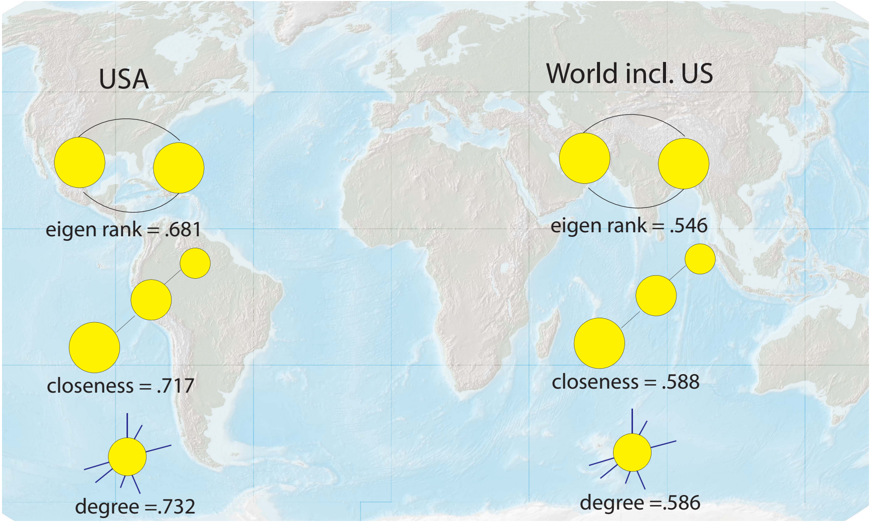
- Plots based on Kamada Kawai spring algorithm
  - Minimizes length of connections
  - Treats thicker connections as stronger

Correlation between centrality measures

	graph type		eigen to	eigen to	degree to
sample	reduce	edges	closeness	degree	closeness
1	**yes**	**all**	**.614 (.733)**	**.589 (.888)**	**.889 (.840)**
1	yes	PhD & tenure	.915 (.866)	.948 (.989)	.922 (.878)
1	yes	no PhD	.929 (.891)	.935 (.977)	.927 (.870)
1	no	all	.979 (.759)	.865 (.990)	.836 (.792)
1	no	PhD & tenure	.983 (.681)	.935 (.981)	.930 (.732)
1	no	no PhD	.949 (.648)	.787 (.873)	.827 (.810)
2	yes	all	.959 (.828)	.931 (.987)	.958 (.863)
2	**yes**	**PhD & tenure**	**.977 (.897)**	**.955 (.944)**	**.969 (.951)**
2	yes	no PhD	.975 (.854)	.940 (.924)	.955 (.928)
2	no	all	.985 (.811)	.903 (.988)	.905 (.817)
2	no	PhD & tenure	.961 (.721)	.936 (.974)	.952 (.764)
2	no	no PhD	.908 (.767)	.841 (.944)	.911 (.795)
overall			.924 (.696)	.877 (.650)	.913 (.523)

- *entries are rank correlations*
- *(...) are continuous correlations*

Centrality & prestige, rank correlations



The “best” predictions by graph type

- Compare centrality measures’ rank versus NRC rank

$$\epsilon = \sum_{i=1}^{i=10} (S_i^r - p^r)^2 + (C_i^r - p^r)^2 + (E_i^r - p^r)^2$$

S_i^r	degree centrality rank
C_i^r	closeness centrality rank
E_i^r	eigenvector centrality rank
p^r	exogenous rank
i	school

$\Longrightarrow$ The graph with the least ϵ is the best graph

Centrality ranks for the top 10 NRC departments

	best graph			worst graph			average rank	
Institution	eigen rank	close rank	degree rank	eigen rank	close rank	degree rank	sample one	sample two
1. UChicago	2	2	2	2	3	3	2	12
2. Wisconsin	1	1	1	2	1	1	1	8
3. Berkeley	5	4	5	4	6	5	5	5
4. UMichigan	4	5	4	1	4	2	3	13
5. UCLA	6	6	6	5	2	6	4	14
6. Chapel Hill	15	15	13	55	21	16	15	15
7. Harvard	3	3	3	57	5	4	6	7
8. Stanford	7	7	7	6	7	7	7	9
9. Northwestern	11	10	12	10	12	12	11	4
10. UW	37	28	29	81	35	45	41	23

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

	domestic ranking			international ranking		
eigen-vector	.542***			.608***		
close		.584***			.669***	
degree			.577***			.686***
faculty						
R^2	.426	.493	.469	.239	.293	.314
coeff tests						

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

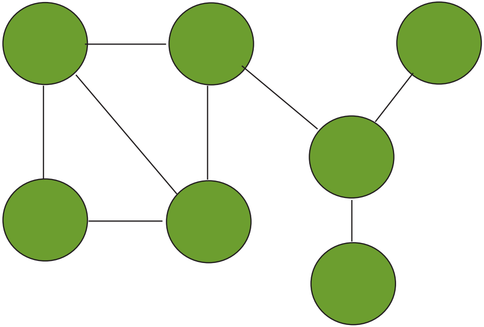
Predicting rank with centrality, OLS

	domestic ranking				international ranking			
eigen- vector	.477*** .542***			-.0399	.584*** .608***			-.199*
close		.521*** .584***		.336***		.659*** .669***		.438**
degree			.516*** .577***	.244***			.665*** .686***	.432***
faculty	-.494*	-.462***	-.499***	-.467***	-.119	-.079**	-.065	-.061
R^2	.567 .426	.567 .493	.557 .469	.580	.250 .239	.311 .293	.318 .314	.328
coeff tests	$\beta_{eig} = \beta_{degree}$		(.0013)		$\beta_{eig} = \beta_{degree}$		(.0001)	
	$\beta_{eig} = \beta_{closeness}$		(.0004)		$\beta_{eig} = \beta_{closeness}$		(.0036)	
	$\beta_{closeness} = \beta_{degree}$		(.0004)		$\beta_{closeness} = \beta_{degree}$		(.9789)	

- 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

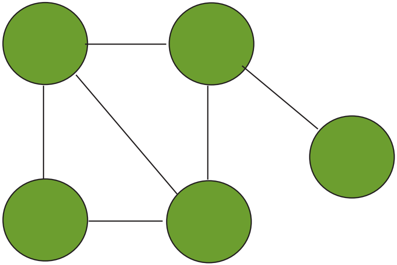
K-Core, definition

- 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



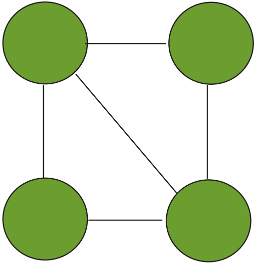
K-Core, definition

- 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



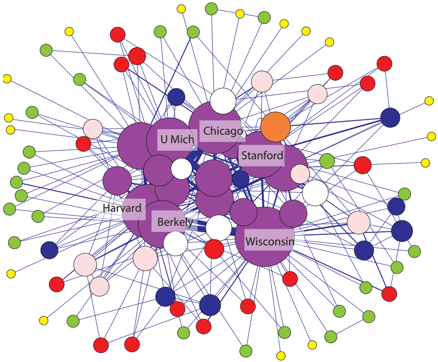
K-Core, definition

- 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

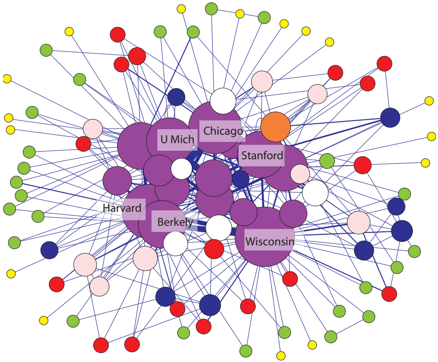
Typical result



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

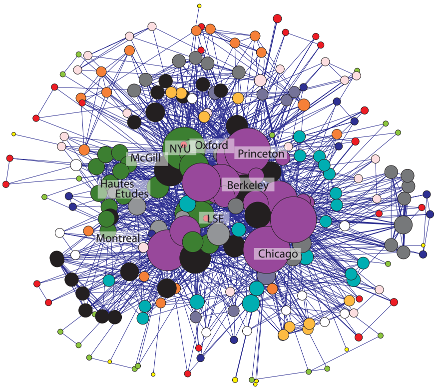
K-Core, two example graphs

Typical result



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

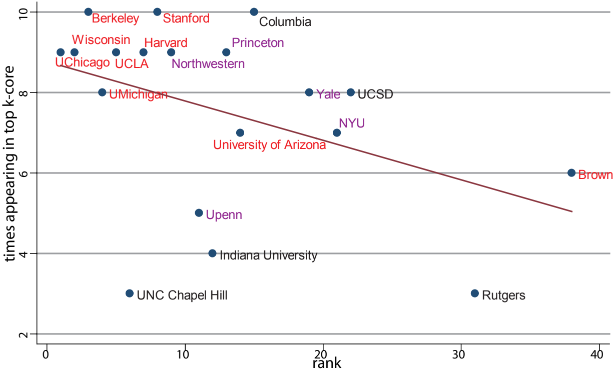
Atypical result



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

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

- 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

	eigen vector centrality	degree centrality	closeness centrality
bipartite	.152*	-.468***	.621***
sample 1	.322***	.425***	.228**
all edges	-.198*	.137	.0427
no student edges	-.278**	-.163*	-.214**
R square	.188	.467	.495

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)