Measuring Scientific Communities with Network Analysis

Chung-Tang Cheng

National Taipei University

Email: ctcheng@gm.ntpu.edu.tw

@National Taiwan University, 2 Mar 2023

Outline

- Network analysis in economic history

- Network analysis in the history of economics

- Historical background

- Research question and aims

- Theoretical framework

- Data, empirical strategies, and some results

- Picturing knowledge transfers

Network analysis in economic history

- Esteves and G. Mesevage (2019, *EEH*)

- **Case 1**: Network of the international monetary system from 1890 to 1910 (Flandreau and Jobst, 2005, *JEH*, Fig. 4)

1. Node: global financial centers

2. Relation: FX quotations

3. England, France, and Germany as "core" nations of FX markets
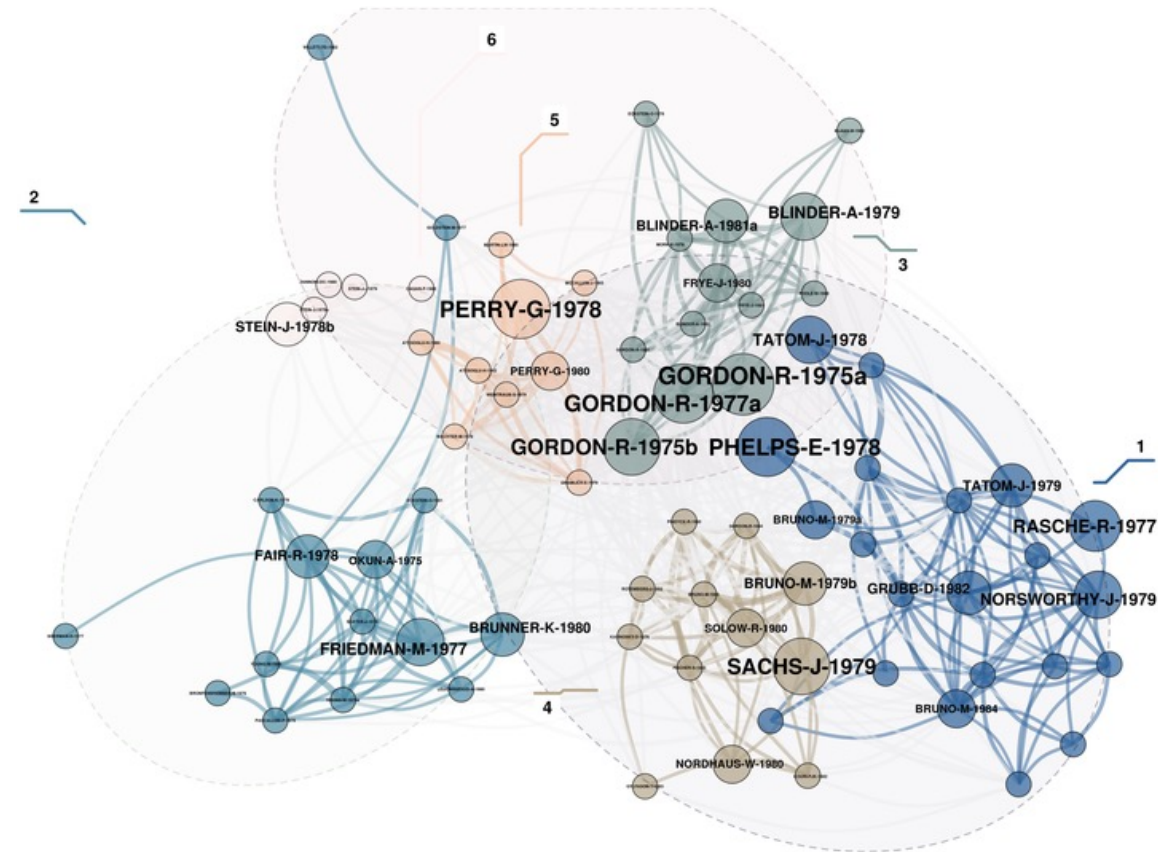
Cheng

Network analysis in economic history

- **Case 2**: The commenda network in medieval Genoa from 1296 to 1315 (van Doosselaere, 2009, Fig. 3.10)

1. Node: participants in the network

2. Relation: business partnerships

3. Increasing collaborations and decreasing centralizations between the Genoese elites

Cheng
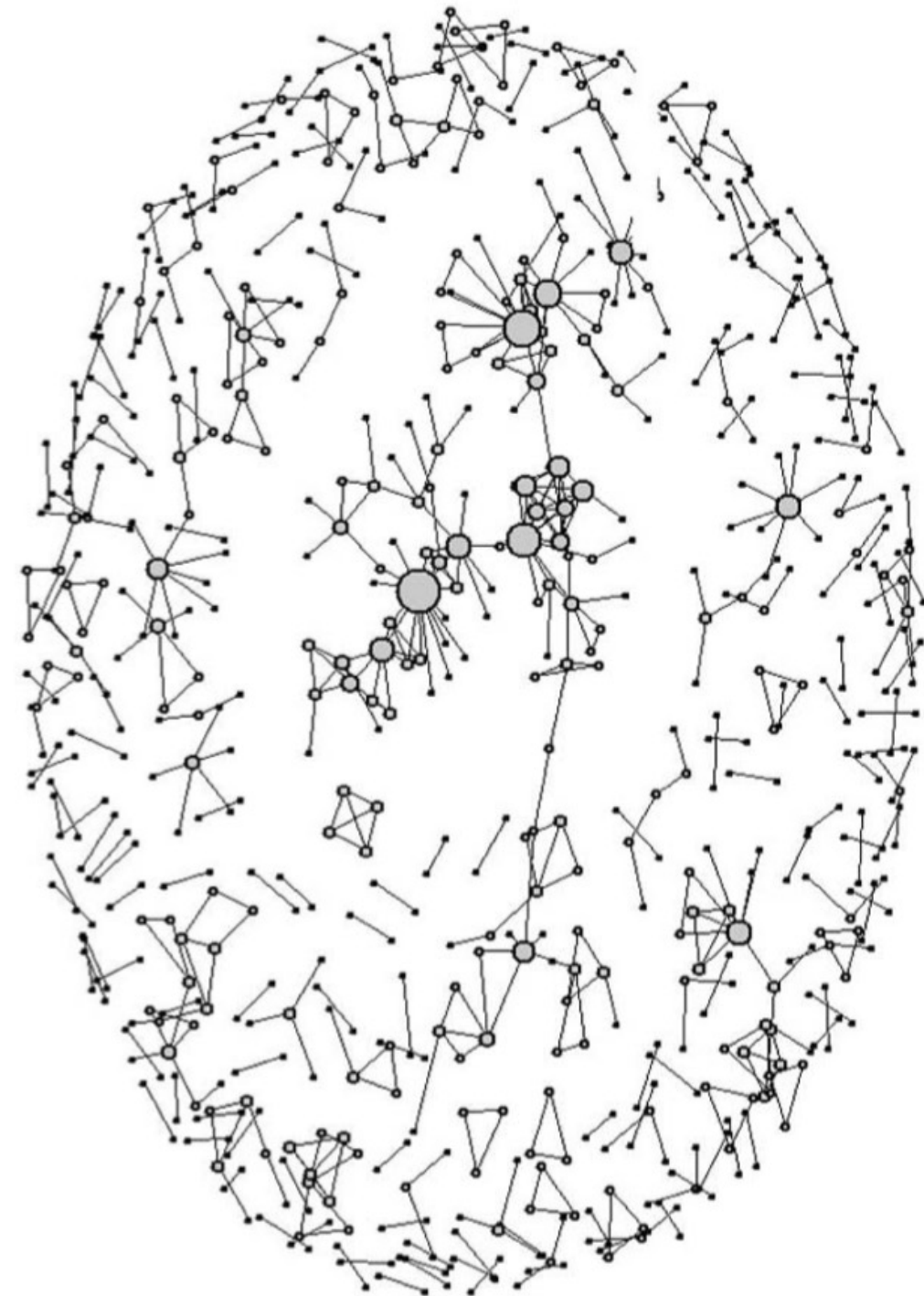
Network analysis in the history of economics

- Claveau and Herfeld (2018, *HOPE*)

- **Case**: The bibliometric coupling network on the US stagflation, 1975–86 (Goutsmedt, 2021, *Rev. d'Écon. Politi.*, Fig. 3)

1. Node: documents explaining the 1970s stagflation

2. Relation: no. of shared references

3. Changing dominant explanations

Source: https://aurelien-goutsmedt.com/publication/stagflation-great-inflation/

Historical background

Richard Stone
(1913–1991)

- The formation and consolidation of microeconometrics from the 1920s to the 1960s (Cheng, 2021)

- Main themes:

1. Interwar analysis of consumer surveys

2. **Richard Stone's demand analysis**

3. **Guy Orcutt's microsimulation**

Source: National Portrait Gallery and Watts (1991,173)

Guy Orcutt
(1917-2006)

Historical background

- Sir Richard Stone (1913–1991), Cambridge economist, founding **Director of the DAE (1946–55)**, Professor (1956–80), Nobel laureate (1984)

- **The Department of Applied Economics (DAE)**: a "statistical realistic" research institute launched in 1946

- Main tasks: national accounting, time-series econometrics, regional sampling, and microeconometrics

- New microdata and computer: two interwar household budget surveys and the Electronic Delay Storage Automatic Calculator

Historical background

- Guy Orcutt (1917-2006), Cambridge DAE (1946-48), Harvard (1949-56), **Wisconsin (1956-66)** and founding **Director of the SSRI (1958-66)**, Yale (1968–90)

- The creator of the first demographic microsimulation in economics (Orcutt et al., 1961) and microsimulation of the US economy (Orcutt et al., 1976; Cheng, 2020)

- **The Social Systems Research Institute (SSRI)** in 1959

- Main tasks: Bayesian statistics, microsimulation, and household behaviours on durable goods

The Stone-Orcutt managerial contrast

- "…his [Stone] personal influence has been **extraordinarily strong**, partly because of the compelling lucidity of his writings, but also by the **example** he set to the stream of economists and statisticians who spent time in the Department of Applied Economics with him…all of this **work owe much to Stone's presence** and to the existence of the Department of Applied Economics…" (Angus Deaton, 2008).

- "In the University [of Wisconsin] there is **remarkably little discipline possible in forming teams**…Guy [Orcutt] had in mind something like the Manhattan Project, which brought together a group of people of different disciplines and they all did fit into the plan sent down from the top…That kind of discipline was perhaps possible there, but **not in an ordinary university setting**" (Interview with Robert Lampman, 1981).

Research question and aims

- Are Deaton and Lampman right **empirically**?

- **The aims of this paper**:

1. how scientists in their community were interpersonally close and intellectually connected to each other?

2. how typical knowledge is transferred between scientific works?

- **Method**: network visualizations and measurements, text and citation analysis

- **Source**: academic publications by the community members

Theoretical framework

- Mark Granovetter's strength-of-weak-tie theory (1973)

  "the strength of a tie is a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie" (*ibid*, 1361).

- **Prediction**: people in a social network benefit from their weaker connections. Those weak ties play important roles in facilitating communications and information transmissions between groups connected with strong ties.

Theoretical framework

- Thomas Kuhn's exemplars as **shared examples** of the scientific community (1970)

  "… these shared examples should, however, be added at least some of the technical problem-solutions found in the periodical literature that scientists encounter during their post-educational research careers and that also show them by example how their job is to be done" (*ibid*, 187).

- How classic articles and books transmit as exemplars in the community?

Data

- The DAE reprint series under Stone's directorship (1946-48, 1948-51, 1951-53, and 1954-57) and the SSRI reprint series under Orcutt's directorship (1961-68)

- Published journal articles in economics and statistics excluding notes, replies, book chapter, and survey papers

- Citation records from Web of Science and Scopus

- References in all samples are digitized under three criteria: (1) working paper and conference paper, (2) to be published, and (3) multiple edition.

Data

|	The DAE Reprints	The SSRI Reprints
No. of docs.	100	131
No. of references	1,040	1,324
No. of referenced*	15,413 from 87 documents	9,448 from 127 documents
Top-5 sources (no.)	J. Royal Stat. Soc. (21); Econ. J. (15); Biometrika (8); Econometrica (8); Rev. of Econ. Stud. (8)	J. Am. Stat. Assoc. (15); Rev. Econ. Stat (13); Econometrica (11); Am. Econ. Rev. (11); Q. J. Econ. (9)
No. of micro. docs.	28	35

*Until Oct 2020

Empirical strategies

• Information needed:

1. Collaboration (from names)

2. Bibliometric coupling (from cited references)

3. Acknowledgement (from footnotes)

4. ~~Supervision (from biography or prosopography)~~

5. Co-citation (from cited references)

Empirical strategies

- Bibliometric coupling (of scholars) vs. co-citation (of papers)
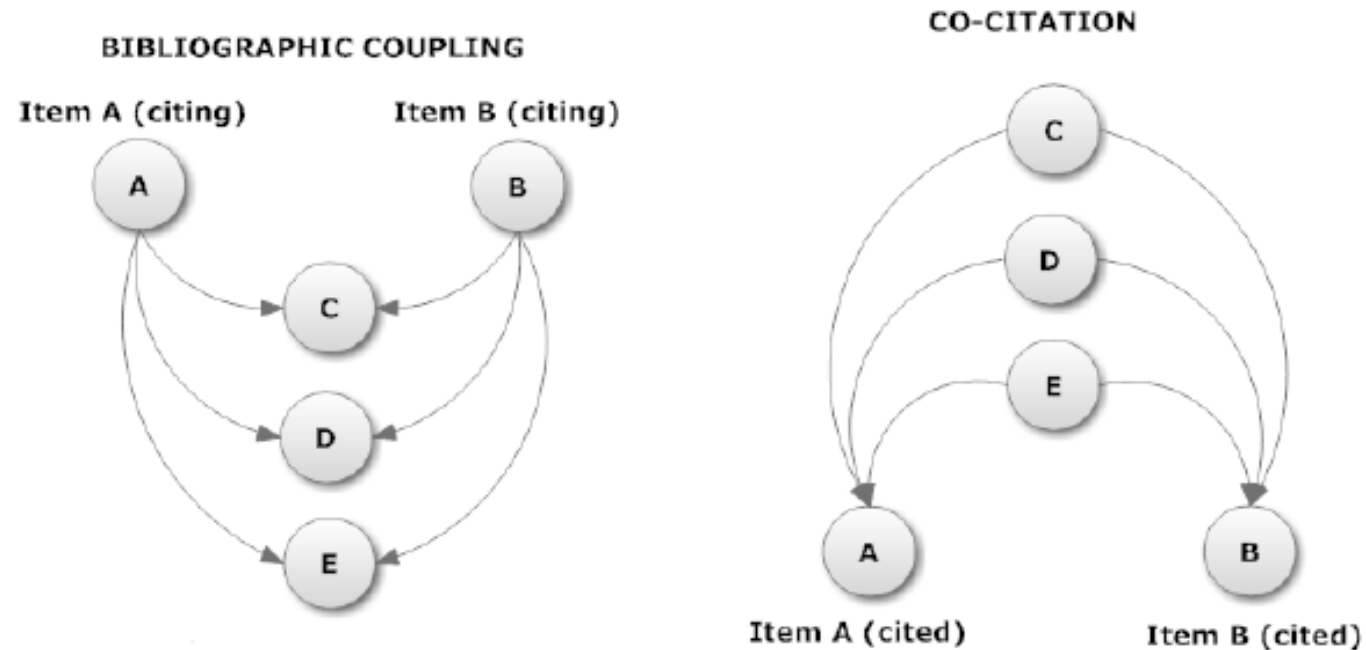
Source: Garfield, E. 2001. From bibliographic coupling to co-citation analysis via algorithmic historio-bibliography. Technical report. Drexel University, Philadelphia.

Empirical strategies

- Network measurements used:

1. Degree centrality
   (D, building connections)

2. Closeness centrality
   (B, clustering abilities)

3. Betweenness centrality
   (A, bridging other nodes)

- Other unused: Eigenvector (C), Harmonic (E), Katz (F),...
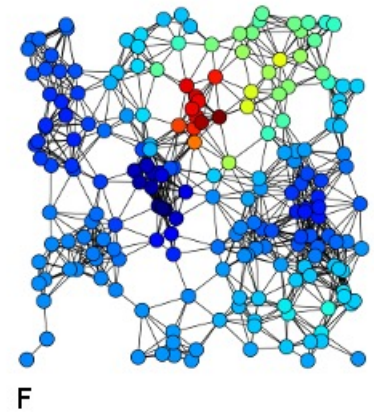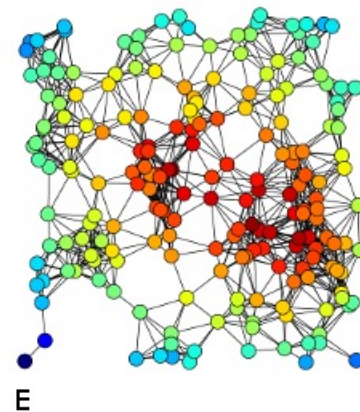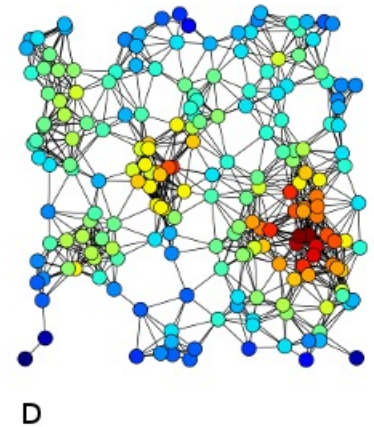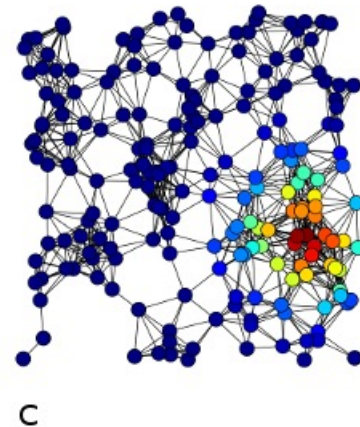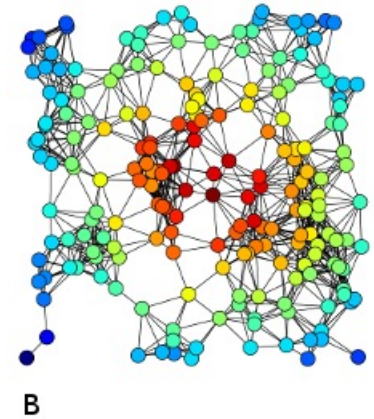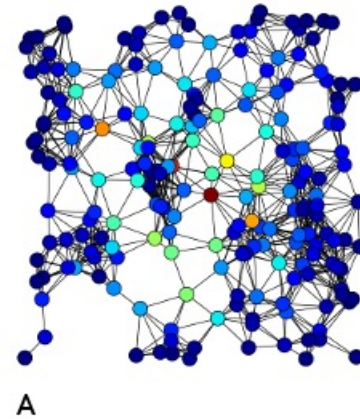
Source: Wikipedia

Cheng

Collaboration networks of reprint series of the DAE and the SSRI
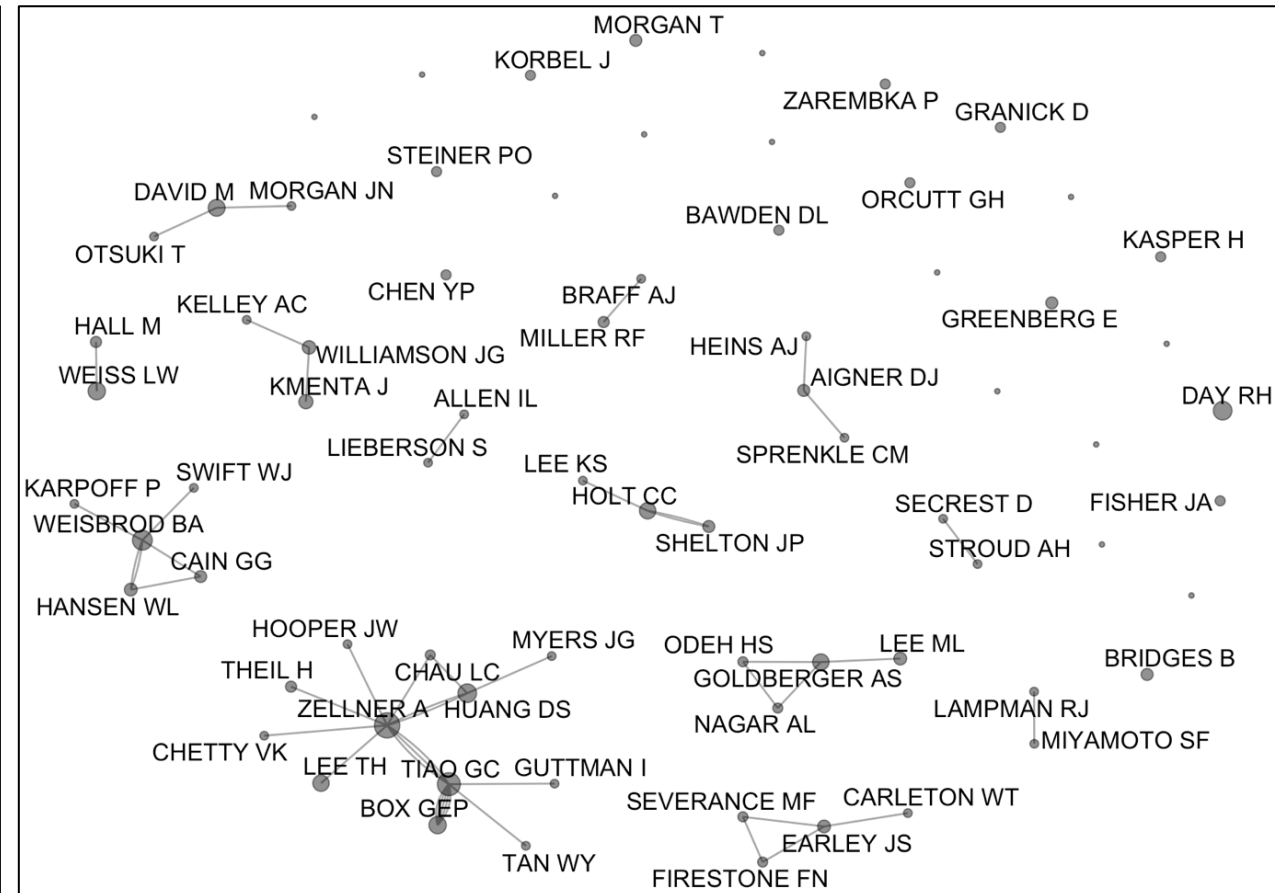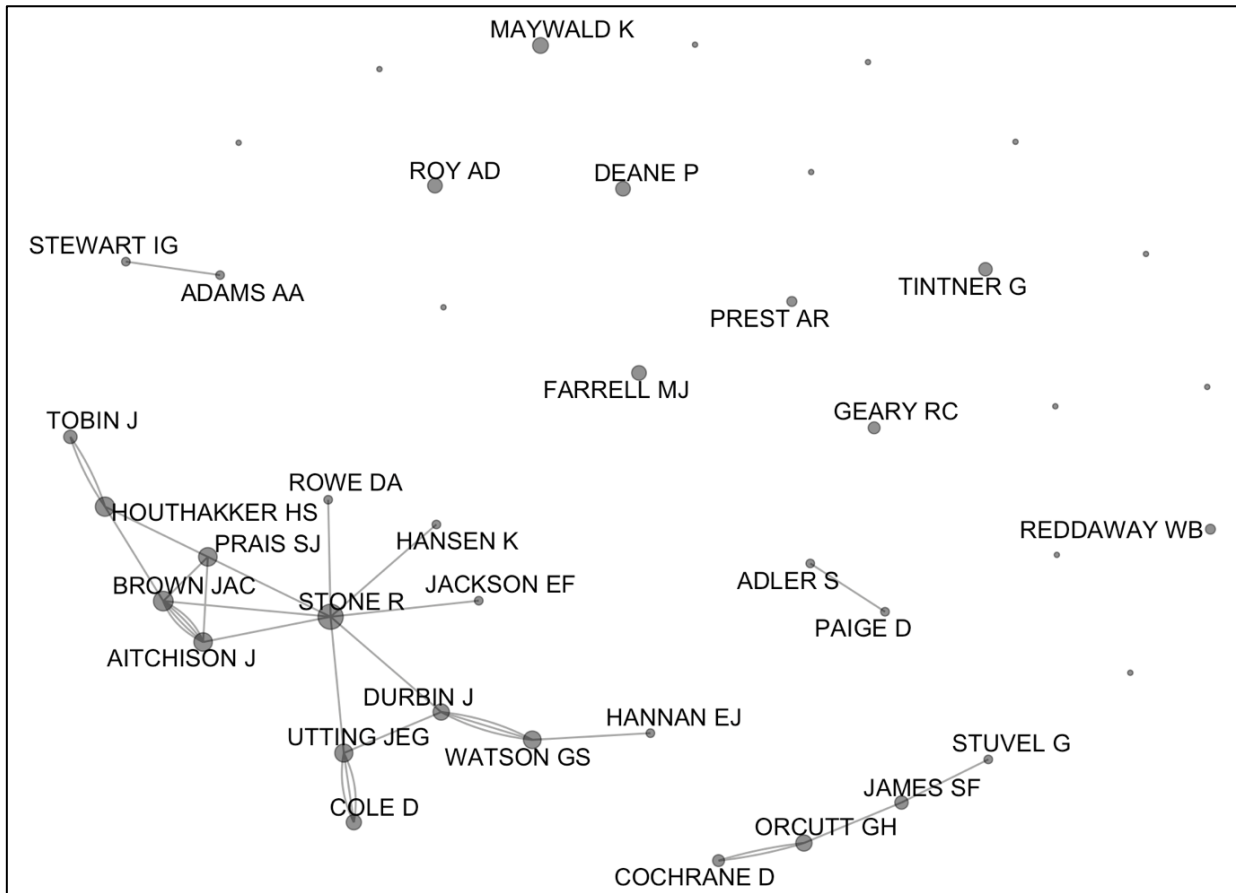
Node degrees: number of articles co-authored; Layout: Fruchterman-Reingold

Bibliometric coupling networks of reprint series of the DAE and the SSRI
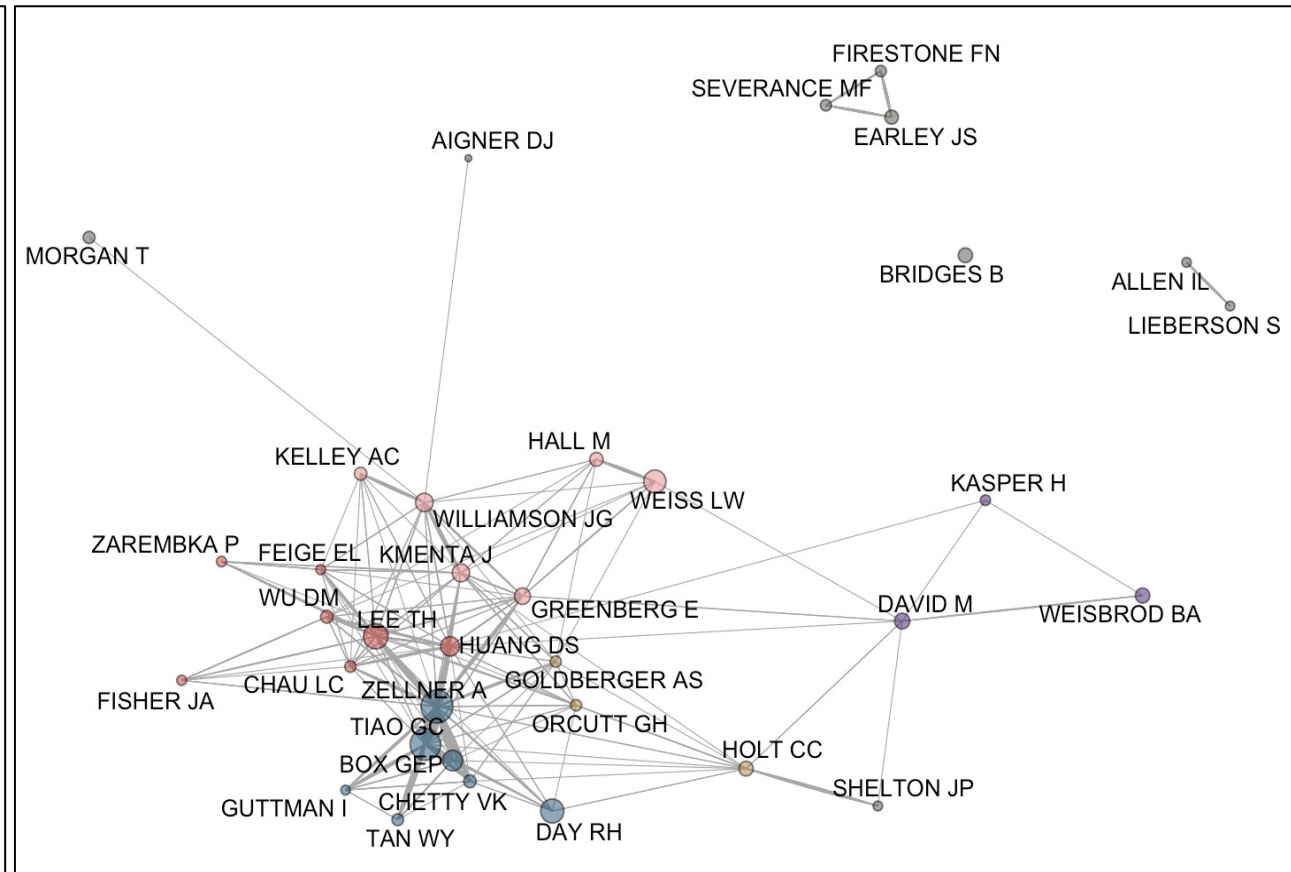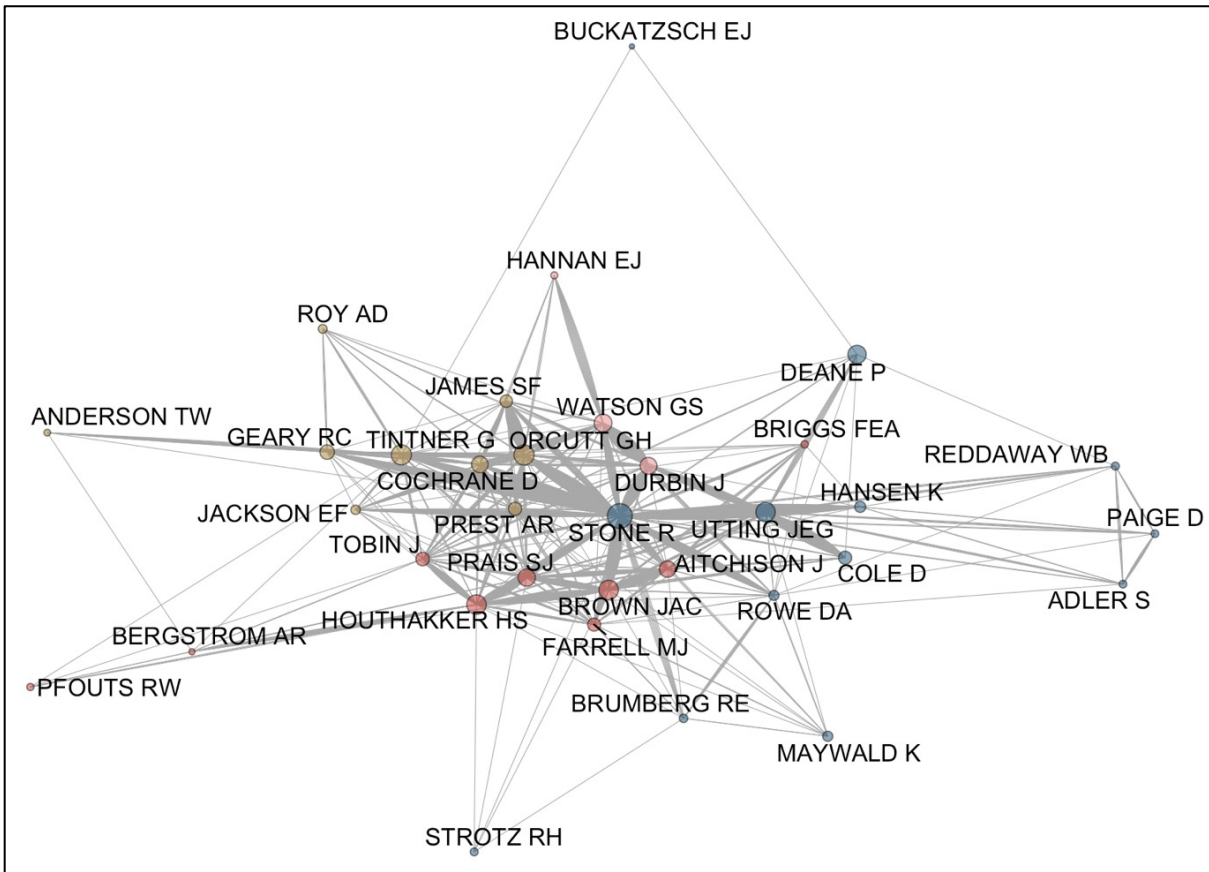
Node degrees: number of references overlapped; Layout: Fruchterman-Reingold;
Community-detection algorithm: Louvain

Acknowledgement networks of reprint series of the DAE and the SSRI
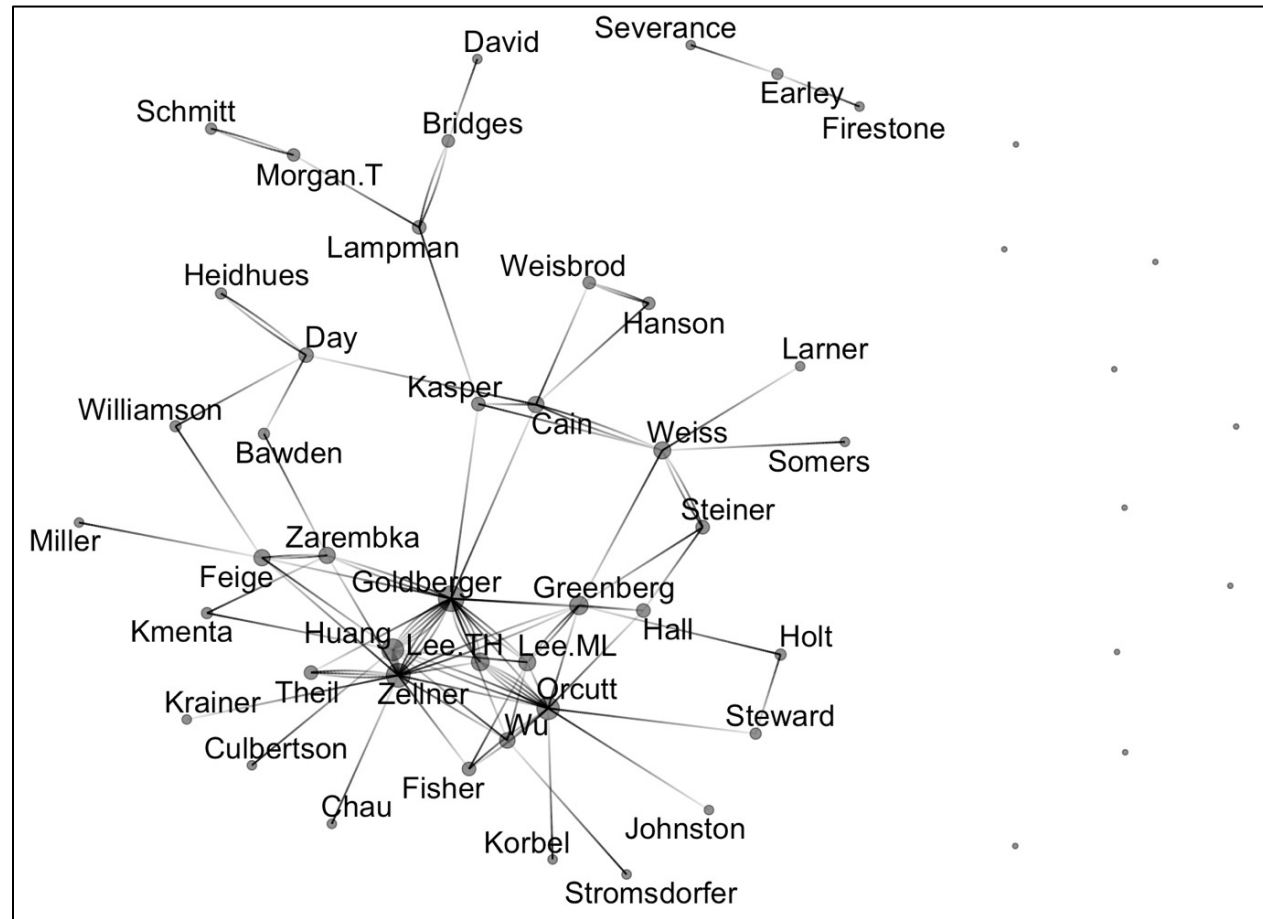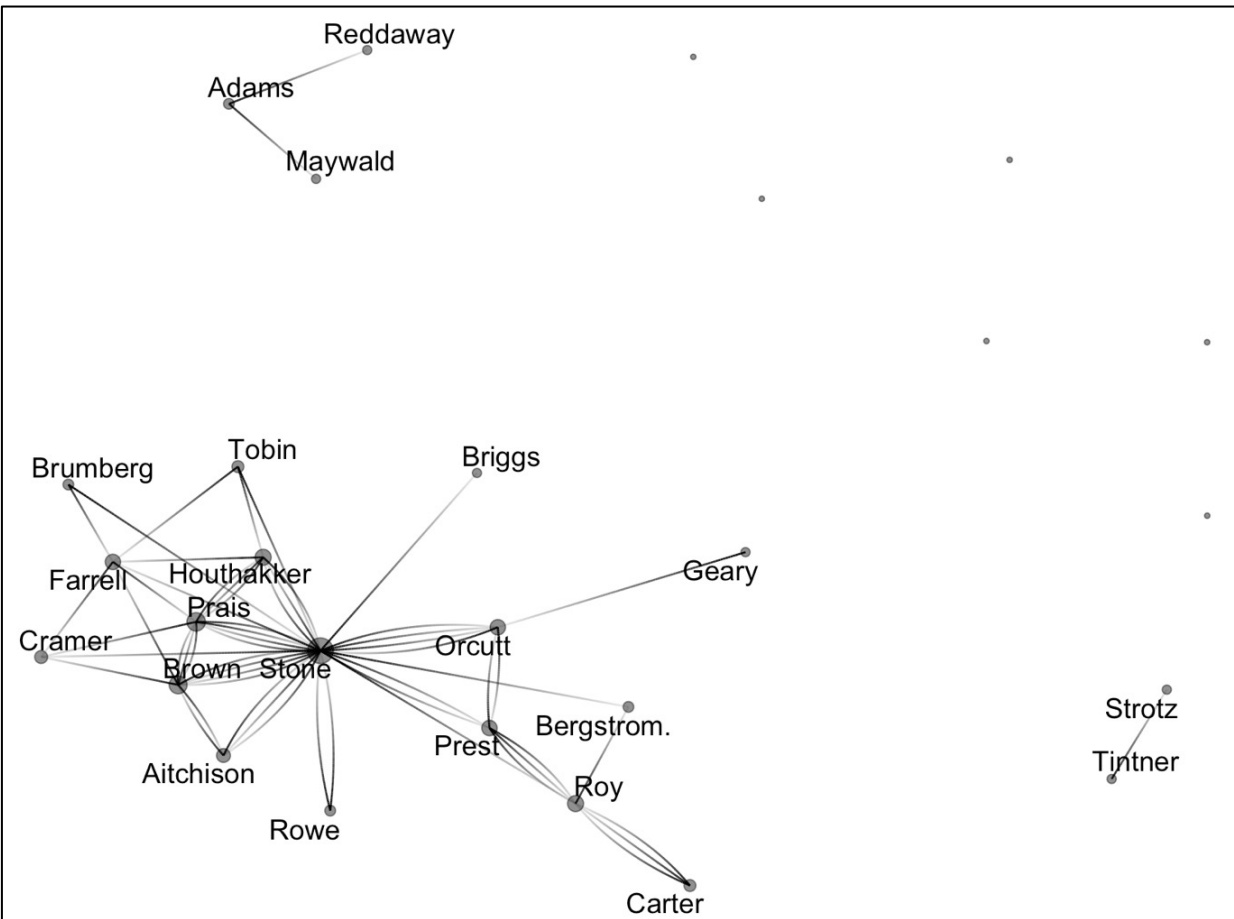
Node degrees: number of acknowledgement received; Layout: Fruchterman-Reingold

Aggregate centrality measurements of the DAE and the SSRI networks

	DAE	SSRI
(a) Collaboration		
degree	0.168	0.083
closeness	0.014	0.004
betweenness	0.068	0.018
(b) Coupling		
degree	0.521	0.281
closeness	0.088	0.020
betweenness	0.161	0.086
(c) Acknowledgement		
in-degree	0.351	0.202
closeness	0.042	0.030
betweenness	0.254	0.204
The degree-of-academic-tie index*	0.313	0.168
The closeness-of-academic-tie index*	0.037	0.013
The betweenness-of-academic-tie index*	0.141	0.068

* Calculates from the geometric mean of each measurement from (a), (b), and (c)

Individual centrality measurements in the DAE networks (standardized)

Author	Collaboration			Coupling			Acknowledgement			Sum
	degree	closeness	betweenness	degree	closeness	betweenness	in-degree	closeness	betweenness	
Stone R	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**900.00**
Prais SJ	50.00	99.50	17.24	69.70	94.41	26.86	36.36	96.74	0.74	491.55
Brown JAC	50.00	99.50	17.24	63.64	92.86	16.18	45.45	96.74	1.66	483.27
Houthakker HS	37.50	98.76	20.69	66.67	93.89	22.08	27.27	96.45	0.37	463.68
Aitchison J	37.50	99.34	0.00	63.64	92.86	13.01	18.18	95.88	0.00	420.40
Orcutt GH	25.00	74.75	3.45	57.58	92.35	19.24	18.18	96.45	16.57	403.57
Tobin J	12.50	97.79	0.00	57.58	92.35	10.15	27.27	96.17	0.00	393.81
Utting JEG	37.50	99.42	20.69	66.67	92.86	28.21	0.00	43.12	0.00	388.47
Farrell MJ	0.00	69.45	0.00	66.67	93.89	22.83	18.18	97.31	4.97	373.31
Durbin J	37.50	99.50	37.93	51.52	91.35	6.05	0.00	43.12	0.00	366.97
Rowe DA	12.50	99.01	0.00	42.42	89.42	7.60	9.09	95.60	0.00	355.65
Prest AR	0.00	69.45	0.00	60.61	92.86	11.41	18.18	96.74	2.21	351.45
Tintner G	0.00	69.45	0.00	63.64	93.37	45.92	9.09	44.72	0.00	326.19
Roy AD	0.00	69.45	0.00	21.21	86.67	0.58	9.09	96.74	17.13	300.87
Geary RC	0.00	69.45	0.00	36.36	88.95	2.88	9.09	92.35	0.00	299.09
Brumberg RE	0.00	69.45	0.00	24.24	85.79	0.64	18.18	95.88	0.00	294.18
James SF	25.00	74.75	3.45	42.42	89.89	9.87	0.00	43.12	0.00	288.50
Cochrane D	12.50	74.66	0.00	54.55	91.85	10.35	0.00	43.12	0.00	287.02
...	...	...	...	...	...	...	...	...	...	...

Individual centrality measurements in the SSRI networks (standardized)

Author	Collaboration			Coupling			Acknowledgement			Sum
	degree	closeness	betweenness	degree	closeness	betweenness	in-degree	closeness	betweenness	
Zellner A	**100.00**	**100.00**	**100.00**	**100.00**	99.74	56.48	91.67	98.71	44.73	791.32
Goldberger AS	42.86	88.65	4.35	55.56	98.97	15.10	**100.00**	**100.00**	**100.00**	605.49
Huang DS	42.86	99.84	21.74	96.30	**100.00**	96.77	0.00	97.45	22.73	577.68
Williamson JG	28.57	87.40	2.17	70.37	99.10	91.06	16.67	93.85	2.83	492.03
Orcutt GH	0.00	85.01	0.00	59.26	98.78	16.52	75.00	98.32	44.62	477.52
David M	28.57	87.40	2.17	51.85	98.78	**100.00**	8.33	86.41	0.00	463.53
Kmenta J	14.29	87.39	0.00	85.19	99.48	43.56	16.67	93.62	0.34	440.53
Holt CC	28.57	87.40	2.17	51.85	98.91	58.18	16.67	93.39	0.79	437.93
Weiss LW	14.29	86.19	0.00	37.04	98.78	38.31	16.67	96.58	32.90	420.75
Cain GG	28.57	89.90	0.00	14.81	95.90	7.17	33.33	98.07	48.89	416.64
Lee TH	14.29	99.78	0.00	74.07	99.16	18.37	8.33	96.22	0.00	410.23
Greenberg E	0.00	85.01	0.00	51.85	99.16	19.49	16.67	97.70	33.07	402.94
Tiao GC	57.14	99.91	58.70	66.67	98.85	14.79	0.00	0.00	0.00	396.06
Day RH	0.00	85.01	0.00	48.15	98.53	31.35	16.67	94.67	19.40	393.77
Weisbrod BA	57.14	89.94	10.87	25.93	96.19	15.65	0.00	93.62	0.00	389.34
Kasper H	0.00	85.01	0.00	22.22	97.84	14.81	8.33	97.95	59.79	385.95
Chau LC	28.57	99.82	0.00	48.15	98.59	2.95	8.33	94.08	0.00	380.50
Theil H	14.29	99.78	0.00	51.85	98.85	10.72	8.33	95.73	0.00	379.55
...	...	...	...	...	...	...	...	...	...	...

Cheng

Strength-of-weak-tie interpretations

- Definition of a "strong" tie: collaboration and supervision

- Stone built 14 out of 28 acknowledgements while Orcutt built 10 out of 54.

- Orcutt received 7.5 out of 13 acknowledgements from his strong ties while Stone received 8 out of 22.

- Orcutt received 11 out of 13 acknowledgements from micro-econometricians while Stone received 12.5 out of 22.

- As a weak tie, Stone benefit his staff members more than Orcutt. Orcutt tended to benefit his strong ties other than weak ties.

Picturing knowledge transfers

- How are the most co-cited materials (exemplars) transferred?

- Data: microeconometric studies in the DAE (28 in total) and SSRI series (35 in total)

- The DAE's most co-cited materials: R. G. D. Allen and Arthur Bowley (1935), *Family Expenditure*; Hendrik Houthakker (1952), "The Econometrics of Family Budgets"; James Tobin (1950), "A Statistical Demand Function for Food in the U.S.A."

- The SSRI's most co-cited materials: Guy Orcutt et al. (1961), *Microanalysis of Socioeconomic Systems: A Simulation Study*

Picturing knowledge transfers

- **Listening** citations vs. **talking** citations

    "…the assumption itself was made by Allen and Bowley (1935), although, since they only applied it to linear individual demand functions, they **missed its many interesting implications**" (Farrell 1954, 176n).

    "…it is possible to tabulate the values of consumption corresponding to different income levels for the purposes of comparison with nutritional standards. Data and discussion on this relationship **are to be found**, for this country, in the works of Allen & Bowley (1935)…" (Brown 1955, 65).

- Knowledge maintenance vs. knowledge progression

Compositions of citations

	Listening citations			Talking citations			Total
	total	in-community*	in-sample	total	in-community*	in-sample	
DAE	48	33	19	25	5	4	73
SSRI	49	23	5	11	2	0	60

* Indicates that the authors and cited authors belong to the same community.

- Paradigm consolidation vs. paradigm refutation

- "Progressive" DAE vs. "degressive" SSRI?

- Another proof of the Stone-Orcutt contrast?

- Diversity in research topics?

- Empirical economics vs. theoretical economics

Some unsolved issues

- An integrated software package

- A better way to extract bibliometric information

- Dealing with large-size data

- Difficulties in locating communal products and in quantifying supervision ties

- Only valid for similar communities, i.e., research culture, subfield, collaboration pattern, citation preference, acknowledgement habit, time span,…

Thank you