The Structure of a Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999

James Moody
The Ohio State University

Has sociology become more socially integrated over the last 30 years? Recent work in the sociology of knowledge demonstrates a direct linkage between social interaction patterns and the structure of ideas, suggesting that scientific collaboration networks affect scientific practice. I test three competing models for sociological collaboration networks and find that a structurally cohesive core that has been growing steadily since the early 1960s characterizes the discipline's coauthorship network. The results show that participation in the sociology collaboration network depends on research specialty and that quantitative work is more likely to be coauthored than non-quantitative work. However, structural embeddedness within the network core given collaboration is largely unrelated to specialty area. This pattern is consistent with a loosely overlapping specialty structure that has potentially integrative implications for theoretical development in sociology.

Science, carved up into a host of detailed studies that have no link with one another, no longer forms a solid whole.

Durkheim, 1933 [*1984*] *p. 294*

Direct correspondence to James Moody, The Ohio State University, Department of Sociology, 372 Bricker Hall, 190 N. Oval Mall, Columbus, OH, 43210 (moody.77@sociology.osu.edu). This work is supported in part by NSF ITR/SOC-0080860 and the Ohio State Initiative in Population Research. The author thanks jimi adams and Sara Bradley for help with data collection and manuscript preparation; and Mark Handcock, Mark Newman, and Paul von Hipple for help with data or analysis. The author thanks the following people for reading prior versions of the paper and providing helpful comments: Art Alderson, Susanne Bunn, Ben Cornwell, Bill Form, David Jacobs, Lisa Keister, Dan Lichter, Dan McFarland, Jason Owen-Smith, Woody Powell, and the participants of the Stanford University SCANCOR seminar and the University of Chicago Business School Organizations and Markets Workshop. The author thanks the ASR reviewers and editors (Camic and Wilson) for being extremely helpful; and reviewer John A. Stewart, in particular, for providing numerous suggestions that improved this paper.

Recent work in the sociology of knowledge suggests that the set of ideas one holds to be true is largely a function of the group of people one interacts with and references to authorities recognized by the group. This claim has been demonstrated in small groups (Martin 2002) and is consistent with literature on the social production of scientific knowledge (Babchuk et al. 1999; Crane 1972; Friedkin 1998; Kuhn 1970). Scientists embedded in collaboration networks share ideas, use similar techniques, and otherwise influence each other's work. Such effects have been studied in specific settings (Friedkin 1998) and implicated in lab ethnographies (Collins 1998; Owen-Smith 2001), but this social interaction structure has not been explored for entire disciplines. Although we might expect the link between networks and ideas to be strongest in small groups, a logical extension suggests that long-term trends in scientific work might depend on the broader pattern of disciplinary social networks.[1]

[1] Much literature on citation networks suggests similar subgroup effects and captures one facet of the

Commentaries on sociology often describe a lack of theoretical consensus without reference to social cohesion, though the two should be linked because structural cohesion is thought to generate coherent idea systems (Durkheim [1933] 1984; Hagstrom 1965; Hargens 1975; Martin 2002; Moody and White 2003; Whitley 2000). A network influence model suggests that if scientists exchange ideas, research questions, methods, and implicit rules for evaluating evidence with their collaborators, then structurally cohesive social networks should generate consensus, at least with respect to problems and methods if not on particular claims about the empirical world (Friedkin 1998).[2] This work suggests that understanding theoretical diversity within a discipline requires understanding its collaboration structure.

While a direct mapping from the idea space to network structure is often not transparent, claims about theoretical consensus in sociology suggest three distinct collaboration structures. First, many have noted that the discipline has no overarching theory but, instead, is theoretically fractured and composed of multiple disconnected research specialties. Authors argue that reactions against functionalism, rapid growth, institutional pressures for productivity, changing research techniques, and/or changes in the funding environment for social sciences have interacted to generate self-contained research specialties, with unique research techniques and standards for the evaluation of evidence (Collins 2001; Davis 2001; Lieberson and Lynn 2002; Stinchcombe 2001). This description suggests a highly clustered social network. Second, others have argued that scientific production depends crucially on a few scientific stars, whose work shapes the short-run course of a discipline (Allison et al. 1982; Cole and Cole 1973; Merton 1968; Zuckerman 1977; see also Crane 1972). Scientific stars attract a disproportionate level of research funding, high-profile appointments, and many students and collaborators. Star systems suggest an unequal distribution of involvement in collaboration networks. Finally, changes in research practice might interact with permeable theoretical boundaries to allow wide-ranging collaborations that are not constrained by research specialty (Abbott 2001; Hudson 1996). For example, an increase in sophisticated quantitative methods, which are (at least on the surface) substantively neutral, would allow collaboration between people with general technical skills and those working on particular empirical questions. This process suggests a wide-reaching structurally cohesive collaboration network.

In this paper, I describe the structure of a social science collaboration network over time, and I link this structure to claims about social scientific practice.[3] I first review literature on network structure and idea spaces, and link three descriptions of the current state of sociological practice to hypotheses about the structure of the network. I then describe collaboration trends and examine explanations for increasing collaboration over time. After describing the data source and measures, I first model *participation* in the collaboration network and then model *position* within the network given participation. I note two key findings. First, specialty areas differ in the likelihood of collaboration, and much (but not all) of this difference is due to use of quantitative methods. Second, the resulting collaboration network has a large structurally cohesive core that has been growing steadily since the late 1960s (both absolutely and relative to random expectation given growth in the discipline). While research specialty predicts having collaborated, specialty is only weakly related to position *within* the collaboration network, suggesting a relatively equal representation of specialties across the disciplinary network.

intellectual integration of scientific disciplines (Crane and Small 1992; Hargens 2000). Citation networks are not social networks, however, and thus do not capture the informal interaction structure described in work on social integration. Recent work has looked at the global structure of large-scale collaboration networks in the natural sciences (Newman 2001), but has not attempted to explain the features of such networks with respect to scientific practice.

[2] That is, scientific competition for distinction should produce a race toward new empirical explanations, but to gain status *within* a field those claims must conform to the general rules of evidence current within the scientific network.

[3] While the majority of all authors in the data I use are sociologists, the database does cover sociologically relevant work produced by other disciplines. As such, while it is technically more accurate to speak of "a social science" collaboration network, I will often use "sociology" for simplicity.

SOCIAL AND THEORETICAL INTEGRATION

Network Structures and Idea spaces

There is increasing interest in linking the distribution of cultural ideas and practices to the interaction structure of social communities (Bearman 1993; Burt 1987; Crane 1972; Martin 2002; Swidler and Arditi 1994). Theorists have long argued that one's ideas are a function of position in a social setting, which is deeply structured by interaction patterns (Durkheim [1933] 1984; Mannheim 1936; Simmel 1950). Kuhn (1970), for example, argued that belief in the empirical validity of theory could be sustained long past the available empirical evidence if scientists were embedded in research communities who systematically interpreted data in similar ways. This implicit perspective was made clear in Crane's (1972) work linking the rapid development of new ideas to the social structure of small "invisible colleges." Crane found that research specialties were characterized by a core group of scientists who collaborated with each other and generated a disproportionate volume of new ideas.

Recent work has built on these ideas to directly link network structure to the distribution of ideas. Martin (2002) argues that while predicting the specific content of ideas is often not possible, we can link the shape of an idea space to the structure of a network. In the small groups that Martin studied, belief consensus depended on the authority structure within the group's social network. Similar work on ideational diffusion suggests that people influence the beliefs of their social contacts. As part of a broader project devoted to understanding the distribution of ideas, Friedkin (1998) shows that agreement across different groups of scientists depends on loosely overlapping cohesive groups in the underlying social network. Thus, while Friedkin's mechanism differs from Martin's (interpersonal influence as opposed to hierarchical authority), the general point is quite clear: Belief consensus depends critically on the shape of the underlying social network.[4] If this work is correct, then we can draw hypotheses about the structure of interaction networks in the social sciences from descriptions of scientific practice and speculate about the potential for scientific consensus in a field based on the observed collaboration pattern.

[4] There is much literature within social networks on the various mechanisms that generate consensus from networks (see Burt 1987 for a review).

Theory and Practice in Social Science

THEORETICAL FRAGMENTATION. There is much literature on the lack of theoretical consensus in sociology (Abbott 2000; Collins 1986; Connell 2000; Davis 2001).[5] For example, Stinchcombe (2001) argues that sociology likely has a dim future, because "[first] it is unlikely to develop much consensus on who best represents the sociologists' sociologist to be hired in elite departments. Second, …, it is unlikely to be able to argue with one voice about what is 'elementary'" (p. 86). This problem is compounded by multiple empirical specialties in the discipline. "The wide variety of substantive subject matter in disintegrated disciplines, and the strong boundaries around substantive specialties, means that people cannot get interested in each other's work." (p.89), and many authors comment on this basic state of intellectual anomie.

Sociology's rapid growth also contributes to perceptions of fractionalization. Simspon and Simpson (2001) report a nearly 5-fold increase in ASA membership since the 1950s, and a rise in the number of ASA sections (5 in 1961, 25 in 1987, 44 in 2003). However, fears about disciplinary fractionalization based on number of sections and multitudes of topics risk mistaking growth on the *margins* for *separation*. When viewed as a potential mixing space defined by the intersection of research areas (Crane and Small 1992; Daipha 2001; Ennis 1992), simple increases in the number of sections tell us little about how people mix across these areas.

Visions of a theoretically fractured social science suggest a highly clustered social network. If substantive boundaries mean that people are not interested in each other's work, then people should turn to fellow specialists as potential

[5] Not all of this literature is negative, and several people comment on what is good about the discipline, including the diversity of ideas and interest of topics, and many attempts are made at theoretical integration or reformulations (cf Lieberson and Lynn 2002; Skvoretz 1998).

collaborators. Graduate students will be trained within particular specialties and a shop-production model should build distinct communities surrounding particular topics. The resulting network will admit to clear clusters with little collaboration crossing specialty boundaries.

The social network model that best fits this description is the *small-world model* (Milgram 1969; Watts 1999; Watts and Strogatz 1998). Intuitively, a *small-world network* is any network where the level of local clustering (one's collaborators are also collaborators with each other) is high, but the average number of steps between actors is small. An archetypical small-world network will have many distinct clusters, connected to each other by a small number of links. Distinct research clusters will likely inhibit broad theoretical integration, since theory will progress largely within distinct research groups.

STAR PRODUCTION. Past research on stratification in the sciences has identified large inequality in the returns to scientific labor (Allison, Long and Krauze 1982; Cole and Cole 1973; Merton 1968). Although most scientists labor in obscurity, a small number of scientists receive disproportionate recognition. This has been clearly demonstrated for indicators such as citations, number of publications, or grants. However, research suggests that collaboration is also unequally divided. Crane (1972) found that a small number of very prominent scientists form the core of each specialty's collaboration network and that most others were connected to the rest of the community through these highly active individuals. This central position helps explain why core scientists were able to so rapidly diffuse their ideas through the community, and we would expect that those with many collaborators are likely to be influential (at least locally). Newman (2001) turns collaboration itself into a status marker and asks, "Who is the Best Connected Scientist?"[6]

The large inequality in numbers of collaborators can be explained through a process of *preferential attachment*. High-status scientists make attractive collaborators since one's own status is a function of the status of those to whom one is connected (Bonacich 1987; Gould 2002; Leifer 1988). This implies that people will seek to work with high-status scientists, and this process will be self-reinforcing. The typical preferential attachment process suggests that as new people enter the network, they collaborate with those already in the network with probability proportional to their current number of partners. The critical structural feature for the preferential attachment model is that star actors are responsible for connecting the network.

The network model that best fits a star production process is a *scale-free model* (Barabasi and Albert 1999; Newman 2000). Barabasi (1999) proved that when networks are constructed through a preferential attachment process, the resulting distribution of the number of unique collaborators (called the *degree* distribution) will have a scale-free power-law distribution, such that the probability of having k partners is distributed as $k^{-\gamma}$, and we can thus use the degree distribution to test for a preferential attachment process.[7] Theoretical integration in such networks will likely depend crucially on ideas generated by star producers, as collaborators follow the lead of those responsible for connecting the entire network.

PERMEABLE THEORETICAL BOUNDARIES AND GENERIC METHODS. Abbott (2001) also argues that the social sciences have little theoretical consensus, but he does not suggest that this generates a clustered discipline. Instead, he suggests that the nature of sociology creates permeable theoretical boundaries that make it impossible for sociology to exclude ideas from the discipline once they are introduced (p. 6; see also Daipha 2001). Moreover, the process of theoretical development is not linear, but instead follows a "fractal walk" through the available idea space. Pushed by competition for status, proponents of one set of ideas attempt to van-

[6] Competition among mathematicians over having the smallest Erdös number speaks similarly to the status attached to one's position in a collaboration network.

[7] Not finding a power-law distribution will falsify a preferential attachment model, but the opposite direction does not hold: finding a power-law distribution is consistent with preferential attachment, but other processes can also generate power-law distributions. This asymmetry has lead to some misunderstandings in the current literature over large-scale networks.

quish another, only to find that they need to reinvent those same ideas later. This results in a constant revisiting of ideas and interests in the discipline (though usually from a different direction) as actors continuously loop through wide sections of the available idea space.

Permeability allows for cross-topic collaboration, since the same theoretical frame (ecological, competition, diffusion through networks, etc.) can be applied to multiple empirical questions. This implies that while people might specialize in techniques or approaches, these techniques and approaches are transferable across research questions. If this is the case, then the boundless character of sociology would promote wide-ranging collaboration. Authors with particular technical, empirical or theoretical skills will mix freely with those who have worked in different research areas, in an attempt to establish a new position by combining previous work. If many engage in this kind of cross-fertilization, mixing across multiple areas, the result will be a social structure with few clear divisions.

The social network model that best fits this description is *structural cohesion* (White and Harary 2001; Moody and White 2003; White et al. 2002). A network is *structurally cohesive* when ties are distributed evenly across the network, implying no clear fissures in the underlying structure (Markovsky 1998). Moody and White (2003) show that this network feature can be exactly characterized as *the extent to which a network will remain connected when nodes are removed from the network*. Such networks are, topologically, the structural opposite of those implied by a preferential attachment process. In preferential attachment networks, most relational paths pass through highly active nodes, which if removed would disconnect the network. This key structural location gives star actors unique control over the spread of ideas in a network. In structurally cohesive networks, in contrast, ties are distributed such that stars in the network are not crucial for connecting the network, and ideas are more likely to spread over the entire network. A structurally cohesive network suggests increasing theoretical integration, at least within the multiply-connected core collaboration network.

The three models for large-scale networks correspond theoretically to expectations about social scientific production. If authors in well-defined research specialties collaborate with each other, then we would expect to find distinct clusters in the knowledge production network, which corresponds to a small-world network structure. If the network was generated by preferential attachment, where young authors write with well-established stars, then we would expect to find a scale-free network structure. If the multiple theoretical perspectives found in the social sciences admit to permeable boundaries allowing specialists to mix freely, then we would expect no strong fissures in the network, but instead find a structurally cohesive network. Before looking at the structure of the coauthorship network, we need to first examine the trends and determinants of coauthorship.

SCIENTIFIC COLLABORATION TRENDS

The probability of coauthoring differs across disciplines and over time. Coauthorship is more common in the natural sciences than in the social sciences, but has been increasing steadily across all fields (Endersby 1996; Fisher et al. 1998; Hargens 1975; Laband and Tollison 2000). The changing likelihood of coauthorship is evident in Figure 1, which shows the proportion of all articles coauthored in *ASR* from inception and in *Sociological Abstracts* from 1963 to 1999.

Several explanations have been given for the increase in coauthorship over time (Laband and Tollison 2000; McDowell and Michael 1983). Funding requirements, particularly in large lab settings, might induce collaboration (Laband and Tollison 2000; Zuckerman and Merton 1973). While social scientists are rarely as dependent on labs, the rise of large-scale data collection efforts suggests a similar team-production model. Training differences between disciplines might also account for coauthorship differences. Advanced work by PhD students in the natural sciences is usually closely related to an advisor's work, and commonly results in collaboration. Social science students, in contrast, tend to work on projects that are more independent.

Other explanations focus on the division of labor among scientists. In high-growth, fast-changing specialties, we would expect to see more coauthorship because it is easier to bring in a new author than it is to learn new material oneself. Hudson (1996) argues that the increase

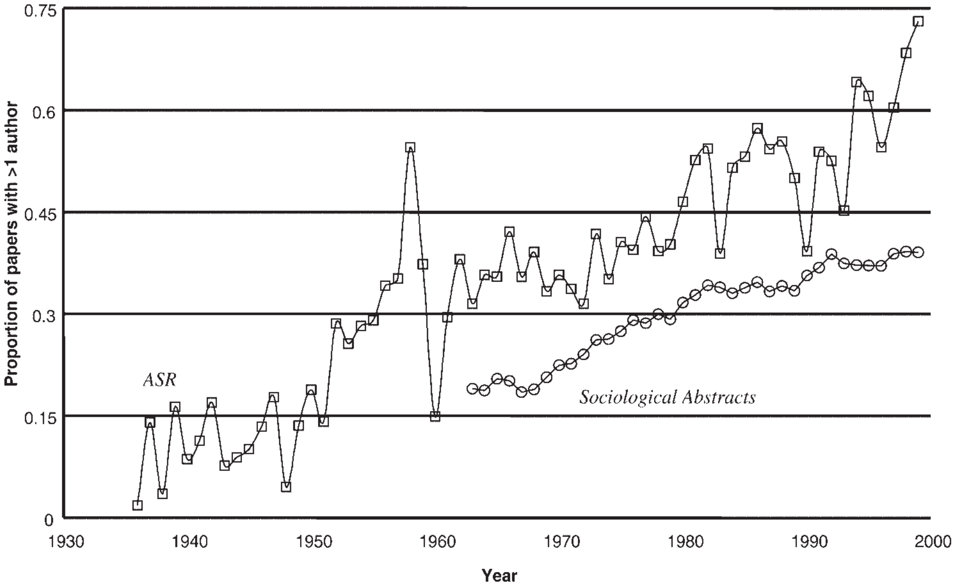


Figure 1. Coauthorship Trends in Sociology

in coauthorship in economics is due to the rise in quantitative methods. As quantitative techniques become more complicated, specialists are often added to research teams to do the analyses. These points are well supported by the lower rates of coauthorship among theoretical or historical specialties compared to coauthorship in quantitative work (Endersby 1996; Fisher et al. 1998). For social scientists, this suggests that work that is difficult to divide, such as ethnography, will be coauthored less often (Babchuk et al. 1999).

DATA AND METHODS

My primary interest is to identify the observed structure of the social science collaboration network to distinguish between the three models suggested by commentaries on sociological practice. Because participation is a necessary minimum requirement for influence in the network, I first model participation in the network, and then examine the structure of the network among those who have coauthored.

Sample and Source

To examine network participation and embeddedness, I use all English journal articles listed in *Sociological Abstracts* that were published between 1963 and 1999. Nineteen sixty three is the earliest date listed in the database, and 1999 was chosen to ensure complete coverage within years. The *Sociological Abstracts* database covers all journals in sociology proper (all ASA journals, for example), and many journals publishing sociologically relevant work in other fields (such as anthropology, political science and economics), and coverage has followed the growth in social science over the last 36 years.[8] *Sociological Abstracts* limits coverage to journal articles, neglecting conference presentations, book reviews, essays, or books. While these types of collaboration represent social contact, each has only spotty coverage in the database. The exclusion of books is perhaps most troubling, to the degree that books are more common in particular specialties such as social movements or theory. While unfortunate, the generally lower rate of coauthorship in books

[8] There is no published universe of journals to compare with, so it is impossible to know definitively if change in the number of journals reflects growth in the discipline or changes in database inclusion. *SA* provides a coverage indicator, however, that tells how often articles from a particular journal are indexed. I use this indicator in the models below to help account for possible inclusion differences.

may offset some of the error introduced by their exclusion.

Authors are identified by name, which can lead to problems when names are inconsistent over time. Errors usually occur due to inconsistent use of middle initials or when two people have the same name. Based on the observed distribution of names, first and last names were coded as either common or uncommon.[9] If two records differed only in their middle initials and had either the same uncommon last name or the same uncommon first name (Howard (a common name) Aldrich (an uncommon name) and Howard E. Aldrich, for example), they were coded as being the same person. Second, the coauthorship pattern was used to identify papers where the same author might use slightly different names. For example, I assumed that two papers written by "David Jacobs" and "David R. Jacobs", with identical sets of coauthors, were written by the same person and, *for that paper*, "David Jacobs" is recoded to "David R. Jacobs."[10]

Measures

Actor information available from *Sociological Abstracts* is somewhat limited, but we can get the total number of unique publications, the number of unique coauthors, cohort, and time in the discipline (date of last publication minus date of first publication). Publication volume accounts for productivity and increased opportunities to coauthor. If the network structure were random, embeddedness within the core of the network would be determined entirely by number of collaborators. Due to changing trends over time, those who enter the network later should be more likely to coauthor than those entering earlier. Star production models suggest that those who have been in the discipline longer should be more deeply embedded than those who just entered. Finally, by linking first names to the Census' gender distribution of first names, we can estimate the effects of gender on collaboration and position in the network.[11]

Every article in *Sociological Abstracts* is assigned to one of 149 detailed subject codes, which are nested within 36 broad specialty areas. These 36 areas are used to capture research specialties.[12] *Sociological Abstracts* also lists the number of tables in every article, which provides a simple proxy for whether the paper uses quantitative methods.[13] To control for changes in journals covered by *Sociological Abstracts*, I include an indicator for how completely *Sociological Abstracts* indexes the journals where people publish. Coverage is indicated at three levels: complete (100% of the articles in that journal are indexed), priority (more than 50% of the articles are indexed), and selective (less than 50% are indexed).

I construct the collaboration network by assigning an edge between any two people who wrote a paper together, regardless of the how often they have coauthored. Figure 2 demon-

[9] A cutoff of 15 appearances of a name was used to distinguish common from uncommon.

[10] Prior work on large collaboration networks has not attempted to identify these types of errors. An alternative source for name cleaning would be to use author affiliation. Unfortunately, *Sociological Abstracts*, only lists affiliation for the first author, and authors may have multiple affiliations (such as a research center and an academic department). While such corrections are important to help ensure accurate measures, the general graph features examined here do not differ significantly if I use the corrected versus the non-corrected data.

[11] This results in a probability score based on the proportion of people with a given first name who are male. Since not all first names can be matched, using this measure results in missing data for 32% of cases, likely predominantly among rare and foreign names. The substantive model results for the other variables do not differ with the inclusion of the gender measure. Tables not using gender are available from the author upon request.

[12] The sociological abstract area categories are likely not substantively ideal, being subject to both errors of misclassification and internal heterogeneity. However, they remain the only *tractable* information on substantive area. While each record contains keywords describing content, the sheer number of such words (over 8000 unique keywords) would require some sort of categorization routine, the development of which is not transparent.

[13] This is also an imperfect measure, since while all quantitative papers include tables some non-quantitative papers also include textual tables, so this measure over-estimates the number of quantitative papers. At the aggregate level used here, the small number of nonquantitative papers with tables washes out relative to differences across specialties.

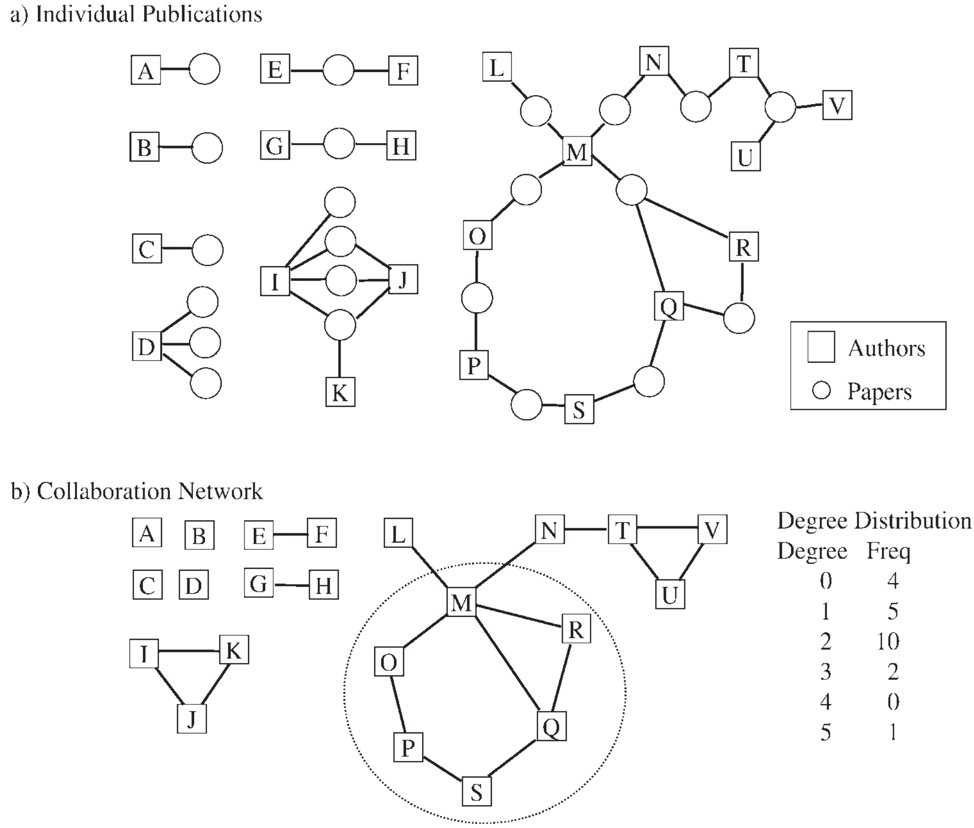


Figure 2. Constructing Collaboration Networks

strates how the networks are constructed from the authorship data.[14]

The top panel of figure 2 is a schematic representation of data as given in *Sociological Abstracts*, with authors (squares) connected to the papers (circles) they write. The data include single authored papers (persons A,B,C and D) as well as those with more authors. The structure on the top right of figure 2 represents a large connected set of authors, each of whom has coauthored with someone who has coauthored with someone else. The bottom panel of figure 2 provides the resulting collaboration network. Those who have written only single authored papers do not participate in the collaboration network, but can be represented as structural isolates. Pairs of people who have only coauthored with each other are represented as isolated dyads {EF, GH}.

[14] Thanks to a reviewer for suggesting this figure.

The *largest connected component* is the maximal set of people who are connected by a chain of any length to each other. The large structure at the bottom right of figure 2 is the largest connected component. Nested within this component is a *bicomponent* (circled). While a component requires only a single traceable path between each actor, a bicomponent requires that there be at least two node-independent paths connecting every pair of actors in the network. Simmel (1950) argued that the necessary condition for a group is that a supra-individual body remains even if a person leaves. Bicomponents meet this criterion, since the group remains connected even if a single person is deleted (Moody and White 2003). This conception scales, as tricomponents (3-node independent paths), 4-components, and higher order k-components identify increasingly cohesive subgroups in a network.

The degree distribution of the network is used to test the preferential attachment model.

An actor's degree is the number of unique people they are directly connected to, in this context the number of unique collaborators. The degree distribution for the example network is given in the lower right of figure 2. Geodesic paths define network *distance*, as the number of intermediaries on the shortest path connected two nodes in a network. So, for example, nodes L and S are 3 steps apart.

PUBLICATION TRENDS

The primary constraints on the shape of a collaboration network are the distributions of the number of papers people publish and the number of authors on a paper. Table 1 below gives these distributions for all papers in the dataset (including those with only a single author).

Of all authors that appear in *Sociological Abstracts*, 66% appear only once, and an additional 15% appear only twice, with the number of publications dropping quickly after that. Publication volume has increased slightly over time. The percent of authors with only one publication has dropped, from 71% in the 75–85 period to 67% in the 89–99 period and the tail of the distribution is a little fatter.[15] About 67% of papers have 1 author, and 22% (66% of all coauthored papers) have only 2 authors. Coauthorship increases over time, both in instance (31% in the early period compared to 38% in the later period) and extent (the average number of authors per coauthored paper was 2.40 in the early period, compared to 2.70 in the late period). Even with the increase over time, these levels are low compared to the physical sciences, which range from an average of 2.2 authors per paper in computer science to 8.9 authors per paper in high-energy physics (Newman 2001). A low number of authors per paper decreases the size of complete clusters formed through common authorship on a single paper.

SPECIALTY AREA AND NETWORK PARTICIPATION

Having ever coauthored a paper is a necessary condition for being embedded in the larger collaboration network. If the collaboration network shapes commitment to particular ways of doing science, then identifying systematic differences in who collaborates will identify key differences in those exposed to the information

[15] Note that the period-specific distributions only count publications within that period. Because the periods do not cover all dates and individuals can publish in both periods, the three columns do not necessarily sum.

Table 1. Sociology Publication Patterns: Distributions of Publications, Coauthorship and Number of Collaborators

	Publications per Author			Authors per Paper			Unique Collaborators[a]		
Count	Total	1975–1985	1989–1999	Total	1975–1985	1989–1999	Total	1975–1985	1989–1999
0	—	—	—	—	—	—	35.27%	41.06%	32.35%
1	65.80%	70.57%	67.71%	66.83%	68.85%	62.57%	23.88	27.55	22.99
2	15.05	14.80	15.59	21.88	22.78	22.79	14.48	14.53	15.01
3	6.46	5.88	6.40	7.06	6.20	8.47	8.73	7.40	9.73
4	3.63	3.05	3.44	2.49	1.73	3.41	5.34	3.74	6.19
5	2.18	1.74	2.05	.93	.45	1.45	3.53	2.18	4.04
6	1.50	1.21	1.34	.42	.17	.66	2.22	1.14	2.67
7	1.08	.77	.87	.19	.07	.31	1.54	.67	1.77
8	.82	.54	.63	.09	.03	.14	1.11	.45	1.30
9	.60	.38	.46	.05	.01	.08	.81	.36	.92
10	.46	.24	.33	.02	.01	.04	.59	.24	.63
11	.37	.18	.26	.01	.01	.02	.48	.22	.47
>11	2.04	.65	.80	.03	.01	.04	2.02	.46	1.93
N	197,976	59,567	123,766	281,090	68,934	141,497	197,976	59,567	123,766

[a] Only people with coauthorships.

and ideas flowing through the collaboration network.

Table 2 lists coauthorship level by broad specialty area, the number of papers within each category, growth in both number of articles and coauthorship over time, and specialties sorted by the coauthorship rate.

While about 33% of all papers have been coauthored, the range is high, from a low of 8% for Marxist Sociology to 53% for Social

Table 2. Growth in Number of Articles and Coauthorship Rates, by Specialty, 1963–1999

Areas of Sociology	N	Papers (%)	Paper Growth	Coauthored (%)	Coauthorship
Growth					
All Areas	281,090	100	346.68	33.2	.013
Individual Areas					
Marxist	1,044	.37	1.66	8.0	–.009[a]
Radical	908	.32	2.82	8.0	–.001[c]
Knowledge	3,406	1.21	1.91	8.2	.000[c]
History & Theory	17,231	6.13	17.03	12.9	.001[c]
Culture and Society	7,040	2.50	3.67	14.3	.003
Visual	141	.05	.25[c]	14.8	–.009[b,c]
Language and Arts	5,673	2.02	7.44	17.8	.008
Political	14,412	5.13	15.93	18.1	.002
Science	4,518	1.61	5.28	20.0	.007
Change & Economic Development	6,632	2.36	7.12	20.5	.007[a]
Religion	5,569	1.98	4.84	23.6	.009[a]
Group Interactions	8,611	3.06	13.48	24.4	.006
Urban	4,444	1.58	.06[c]	25.6	.004
Community Development	1,694	.60	–.39[c,d]	26.0	.009
Feminist Gender Studies	7,225	2.57	11.9	27.2	.003[a,c]
Social Development	9,805	3.49	15.93	27.9	.016
Social Control	7,804	2.78	6.16	28.4	.009
Policy & Planning	3,243	1.15	6.75	28.6	.008
Clinical	280	.10	–.15[c]	29.8	.000[b,c]
Mass Phenomena	12,069	4.29	14.98	30.0	.006
Rural	3,746	1.33	.65[c,d]	30.5	.008
Education	10,628	3.78	8.71	31.9	.010
Environmental Interactions	3,102	1.10	7.93	32.1	.012
Methodology	8,897	3.16	6.68	32.1	.009
Studies in Violence	1,521	.54	3.11	33.4	.010
Demography	6,542	2.33	4.38	33.6	.010
Social Differentiation	9,769	3.48	–.58[c]	34.4	.011
Studies in Poverty	1,393	.50	2.06	34.9	.014
Social Planning/Policy	12,232	4.35	20.21	35.8	.014
Complex Organizations	13,986	4.98	20.22	37.4	.012
Business	195	.07	1.79	40.0	.054[b,c]
Social Psychology	13,527	4.81	4.67	44.2	.013
Problems & Welfare	10,674	3.80	16.35	45.4	.019
The Family	19,806	7.05	20.13	46.1	.019
Health/Medicine	14,634	5.20	23.16	49.5	.025
Social Welfare	28,689	10.21	90.58	53.2	.045

[a] Trend levels off in recent years.

[b] Recent category, less than 10 years covered.

[c] Trend is not statistically significant.

[d] Although the overall trend is not significant, growth is nonlinear, dropping sharply in the 1970s, and rising steadily since about 1980.

Welfare.[16] The number of articles in the database has increased by about 350 papers per year over the past 30 years (column 4, row 1), and overall growth by specialty can be seen in the rest of the table (the sum of the specialty cells equals the total). Culture and theory papers are among the least likely to be coauthored, as are papers in the sociology of knowledge, sociology of science, language and arts, radical sociology, Marxist sociology, political sociology, and visual sociology. Papers on social welfare are the most likely to be coauthored, followed closely by those in social psychology, the family, sociology of health and medicine, and social problems and welfare. Work in these areas often involves large datasets and cumbersome analyses that lend themselves to a substantive division of labor. There is a moderate correlation between growth in a specialty and the proportion of papers that are coauthored, though fields such as social psychology and social theory / historical sociology clearly buck the trend. The proportion of papers that are coauthored has increased by about 1% per year over the time span covered in the database, and this trend has been largely linear. Growth in coauthorship has been relatively steady within specialties as well, with most having growth levels close to the average.

While the *Sociological Abstracts* data are too limited to test many of the proposed explanations for increases in coauthorship, we can identify the contribution of research method and specialty area.[17] Figure 3 plots the proportion of papers within a specialty area that are coauthored against the proportion that have tables, for two time periods.

The figure shows a strong correlation between the proportion of papers with tables in

[16] Social Welfare is a large category including topics such as AIDS, health care, addiction, adolescence, illness and health care, marital and family problems, crime & public safety, etc. This category has the highest coauthorship rate and is among the fastest growing areas in the dataset. To ensure that the observed results were not simply an artifact of this category, I have replicated the main findings for the paper on a dataset that excludes these papers from the set. There are no substantively meaningful differences when these papers are excluded.

[17] Previous versions of this paper included data on a limited set of journals and examined funding, mentorship and location explanations for coauthorship. These tables are available on request.

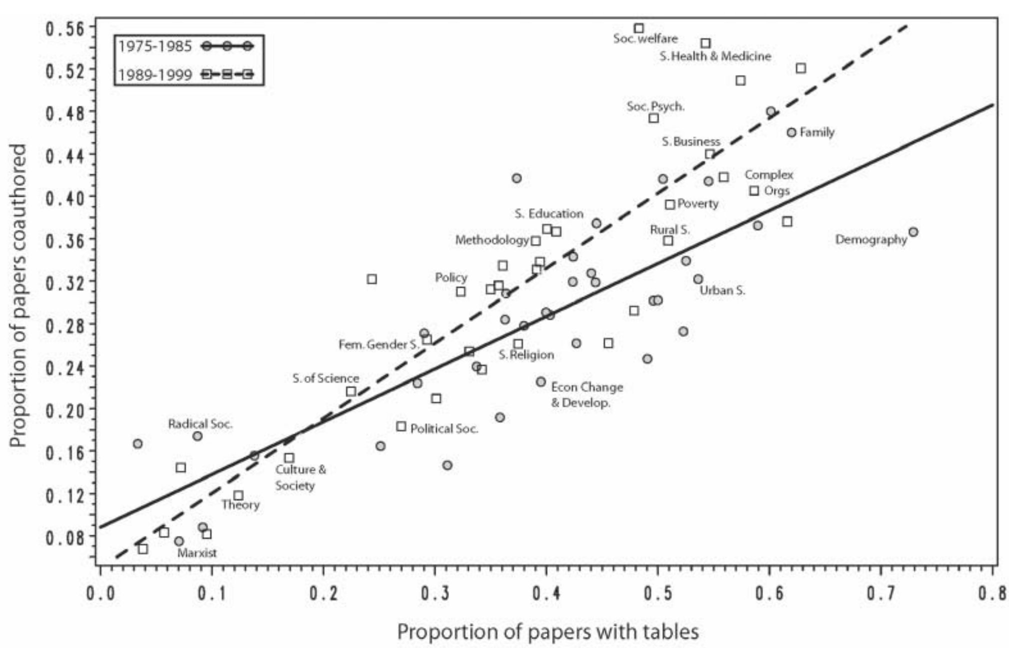


Figure 3. Coauthorship and Quantitative Work 1975–1999 by Specialty.

each specialty and the proportion of papers that are coauthored.[18] The returns to quantitative work increased between the two periods, suggesting a stronger effect of quantitative work on coauthorship in the most recent decade. A simple explanation for the change over time is that as quantitative work becomes more sophisticated, methodological specialists are brought into more projects. The increase in specialized knowledge needed for advanced techniques increases the value to dividing labor in papers (Hudson 1996).

Table 3 provides an individual level model of having ever coauthored. Substantive area is coded from the *Sociological Abstracts* as a count of the number of publications for each author in each specialty area. Within each of the three time windows, Model 1 presents a baseline model containing only the actor history, demographic, publication and journal coverage variables (the first 6 rows of the table) and a single specialty area. The specialty area coefficients in model 1 are thus the coefficients for *separate* models with the single specialty area and the first 6 variables listed in the table. We can thus interpret the odds ratio as the change in the odds of coauthorship for each publication in that specialty, relative to *all other* specialty areas. Since people can write in multiple areas, and we would expect some areas to be much more closely related than others (Crane and Small 1992; Daipha 2001), model 2 presents a multivariate version of model 1, entering all of the specialty areas simultaneously.[19] Model 3 adds an indicator of quantitative work, the proportion of an author's papers that have tables, to model 2, allowing us to identify specialty difference net of research method.

The models show that those with greater time in the discipline (exposure) are slightly more likely to coauthor, though the magnitude of this effect is small.[20] The cohort effects are consistent, with those publishing later having a higher likelihood of having coauthored, though again the effect is relatively small. The strongest publication effect is the simple number of publications. Each additional publication increases the odds of coauthoring by 1.28. Consistent with prior research in economics, the odds of men coauthoring are about .64 times the odds of women coauthoring, though this effect decreases (to .73) once you control for specialty area.[21]

The actor-attribute effects remain largely constant when controlling for specialty or examined over time. The clearest exception is that the effect of the number of publications increases when controlling for specialty area, and is stronger in the latter period than in the early period. Similarly, cohort effects are less pronounced in the later period than in the early period, likely reflecting the greater ubiquity of coauthorship over time.

There is a clear effect of specialty on the likelihood of having coauthored. Authors who write in historical, qualitative, radical and interpretive specialties are less likely to coauthor than those writing in more positivist and quantitative specialties. For example, the odds of coauthorship controlling for specialty overlap in Marxist sociology are about half (.57) those in sociology of education, and those writing in social history and theory are about .58 times as likely. In contrast, those writing on the family

[18] The correlation in the early period is .82, in the late period it is .89. The change in slope between the two periods is statistically significant at $p = .017$. To avoid clutter in the figure, each specialty is only labeled for one time period. An alternative to using the specialty areas would be to aggregate within journals, because with the exception of a few general journals, journals specialize in particular topics. This figure is available on request, and substantive results are the same.

[19] Since the categories are exhaustive, at least one must be omitted. There is no substantive reason to pick one area over another as the reference category, but a consistent reference category aids in evaluating change over time. In all models I used the *sociology of education* as the reference category, because the odds of coauthorship were very close to average in both time periods, the rate of change in coauthorship mirrors the rate of change overall, and the area is large enough to provide a stable reference category.

[20] To avoid redundancy in the text, I will focus comments mainly on the pooled 1963–1999 models, and mention the other models only to the extent that the patterns differ from this model.

[21] In addition, the control variables for *Sociological Abstracts* coverage are also significant, showing that journals included only incidentally (the omitted category) are slightly more likely to be coauthored, though this effect either becomes insignificant or changes direction when substantive area is included.

Table 3. Logistic Regression of Having Ever Coauthored on Publication Characteristics

	1963–1999			1975–1985			1989–1999		
Variable	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3
Exposure	1.02	1.01	1.01	1.03	1.01[a]	1.02[a]	1.02	.99[a]	1.00[a]
Number of Publications	1.28	1.45	1.44	1.33	1.31	1.29	1.34	1.61	1.56
Year of 1st Publication	1.03	1.02	1.01	1.04	1.03	1.01[a]	1.01	1.01	1.01[a]
Male Author (probability)	.64	.73	.73	.70	.74	.76	.64	.76	.75
Complete Coverage	.92	1.22	1.19	1.16	1.32	1.22	.66	1.02[a]	1.03[a]
Priority Coverage	.87	1.00[a]	.94[a]	.98[a]	1.01[a]	.94[a]	.80	.98[a]	.91[a]
Quantitative Work	—	—	5.45	—	—	4.17	—	—	6.38
Specialty Area of Sociology (code)									
Radical (25)	.26	.48	.61	.35	.58	.70[a]	.20	.34	.47
Marxist (30)	.32	.57	.62	.34	.50	.66	.20	.35	.49
Knowledge (22)	.24	.38	.42	.26	.37	.48	.16	.26	.40
History & Theory (2)	.49	.58	.64	.54	.68	.81	.35	.42	.54
Culture & Society (5)	.52	.56	.62	.42	.49	.52	.37	.44	.57
Visual (33)	.47	.57	.69[a]	.44[a]	.59[a]	.86[a]	.32	.35	.48[a]
Language & Arts (13)	.56	.59	.60	.51	.61	.69	.57	.59	.63
Political (9)	.58	.68	.69	.60	.75	.74	.48	.59	.63
Science (17)	.64	.69	.73	.67	.81	.92[a]	.58	.64	.75
Social Change (7)	.63	.77	.76	.65	.87[a]	.93[a]	.61	.75	.75
Religion (15)	.74	.72	.72	.77	.85	.92[a]	.76	.73	.72
Group Interaction (4)	.64	.68	.71	.79	.90[a]	.89[a]	.58	.62	.66
Urban (12)	.89	.91[a]	.89	1.06[a]	1.20	1.14[a]	.74	.78	.73
Community Development (23)	.85	.90[a]	.94[a]	.85[a]	1.02[a]	1.16[a]	.91[a]	1.04[a]	1.14[a]
Female Gender (29)	.68	.69	.71[a]	.93[a]	1.01[a]	1.04[a]	.55	.58	.66
Social Development (83=36)	.82	.86	.76	.76	.87	.80	.78	.86	.79
Social Control (16)	.88	.82	.84	.96[a]	1.02[a]	1.07[a]	.87	.81	.86
Policy & Plan (24)	.81	.87	.95[a]	.79	.91[a]	.96[a]	.74	.79	.91[a]
Clinical (31)	1.34	1.24[a]	1.21[a]	.93[a]	1.01[a]	1.09[a]	.76[a]	.83[a]	.98[a]
Mass Phenomena (8)	.90	.97[a]	.95[a]	.98[a]	1.13[a]	1.08[a]	.82	.87	.88[a]
Rural (11)	.99[a]	1.07[a]	1.01[a]	.86[a]	1.02[a]	.96[a]	1.07[a]	1.08[a]	.98[a]
Education (14)	.95	—	—	.89	—	—	.93	—	—
Methodology (1)	1.07	1.15	1.10	1.00[a]	1.13[a]	1.12[a]	1.17	1.25	1.22
Environmental (26)	.95[a]	1.03[a]	1.02[a]	.86[a]	.97[a]	.95[a]	.96[a]	1.02[a]	1.02[a]
Violence (28)	.88	.89[a]	.84	1.17[a]	1.24[a]	1.14[a]	.73	.75	.75
Demography (18)	1.09	1.07[a]	.93[a]	1.21	1.29	1.04[a]	1.03[a]	1.01[a]	.83
Social Difference (10)	1.12	1.14	1.02[a]	1.17	1.28	1.13[a]	1.18	1.22	.97[a]
Poverty (27)	1.07[a]	1.01[a]	.94[a]	1.08[a]	1.30[a]	1.05[a]	1.12[a]	1.06[a]	.96[a]
Social Plan/Policy (72)/35	1.19	1.21	1.13	1.23	1.35	1.34	1.05[a]	1.10	1.13
Complex Organizations (6)	1.17	1.19	1.06[a]	1.25	1.38	1.23	1.17	1.20	1.03[a]
Business (32)	1.39[a]	1.45[a]	1.26[a]	.69[a]	.81[a]	.59[a]	1.76	1.99	1.80[a]
Social Psychology (3)	1.88	1.91	1.63	2.21	2.34	1.95	1.60	1.77	1.50
Social Problems (21)	1.76	1.65	1.46	1.53	1.37	1.46	1.90	1.77	1.49
Family (19)	1.71	1.61	1.35	1.68	1.72	1.46	1.83	1.75	1.37
Health (20)	2.02	1.92	1.64	1.53	1.62	1.48	2.24	2.13	1.72
Social Welfare (61)	2.35	2.19	2.04	1.56	1.66	1.80	2.67	2.40	2.24
R-Square	.096 [b]	.212	.322	.067 [b]	.154	.254	.067 [b]	.215	.352
N		130,141			41,386			82,475	

Note: Data shown as odds ratios. Unless otherwise noted, all cell values are significant at $p \leq .01$ (tables with detailed significance levels are available from the author). Mod = Model.

[a] Value is *not* significant at the $p < .01$

[b] Pertains only to the publication and demographic characteristics.

are 1.61 times more likely to coauthor, those in social psychology are 1.91 times more likely to coauthor and social welfare writers are 2.19 times as likely to coauthor, net of individual attributes, index coverage, publication levels and overlaps among area specialties. The importance of specialty area for collaboration is clearly indicated in the increase in model fit. The pseudo-R^2 increases significantly with the addition of specialty area.

The likelihood of coauthorship by specialty area does evidence some interesting changes over time. Many of the fields that are unlikely to coauthor are comparatively more unlikely to coauthor in the later period than in the early period, suggesting that the importance of specialty for coauthorship has increased over time. For example, while those writing in feminist gender studies had about average coauthorship levels in the early period, the odds of coauthorship in this specialty are about half the average in the later period. Similarly, those writing in historical sociology and theory had an odds ratio of .68 in the early period, compared to .42 in the later period. The trend is replicated in reverse for fields where coauthorship is more common, though perhaps not as uniformly. For example, while methodologists had average levels of coauthorship in the early period (not statistically different from feminist gender studies), odds in the later period are about 1.25 times that of average groups. Similar patterns are evident for work on the family. Social Psychology, on the other hand, is relatively more likely to coauthor in the early period than the later period, likely reflecting its 'first-mover' status in coauthorship.

Model 3 adds the indicator of quantitative work, and shows that quantitative work has a strong association with coauthorship. Overall, those writing quantitative papers are over 5 times more likely to have coauthored than someone who has not written quantitative work. This effect has increased over time, from 4.17 in the early period to 6.38 in the later period. Adding coauthorship to the model significantly improves the model fit, attenuating the magnitude of the specialty effects, though they are generally substantively similar to those in model 2. This suggests that, while research method is clearly important, research specialty still makes a unique contribution to the odds of coauthorship.

These models suggest that coauthorship might be bifurcating across specialties, which is likely due to an increasing rate of coauthorship growth in the less interpretive fields. Once a division of labor process takes hold, it likely propagates as mentors teach students that coauthoring and collaboration is normative. As such, we should expect continual growth in the incidence of coauthorship within these specialty areas. This bifurcation suggests that, to the extent that theoretical integration follows social integration, a theoretical gulf will mirror the network participation pattern. Those specialties that are least likely to collaborate are thus less likely to be theoretically integrated with those that are more likely to collaborate.

COLLABORATION NETWORK STRUCTURE

While the models presented above tell us who coauthors, they tell us nothing about the structure of the network given coauthorship. The discussion presented above suggest three competing models: A preferential attachment model, suggesting a structure reliant on star producers; a small-world model, where specialty areas cluster into distinct social groups; and a structural cohesion model that suggests a broad overarching connectivity among a large portion of the network.

DOES THE NETWORK DEPEND ON STAR COLLABORATORS?

If the observed network were generated through a preferential attachment process, the distribution of number of coauthors would follow a power-law distribution, which will be seen as a straight line when plotted on a log-log scale. The distribution of the total number of collaborators is given in the last three columns of table 1, and presented graphically in Figure 4. For the full network, about 37% of authors have only written with one other person, and about 22% have written with 2 others. The distribution of coauthors has a fatter tail in the 89–99 period compared to the 75–85 period, with fewer nodes having only one coauthor and more people with 3 or more coauthors.

The observed distribution does not fit a strict power law, having a curved, rather than linear shape, suggesting that the network was not

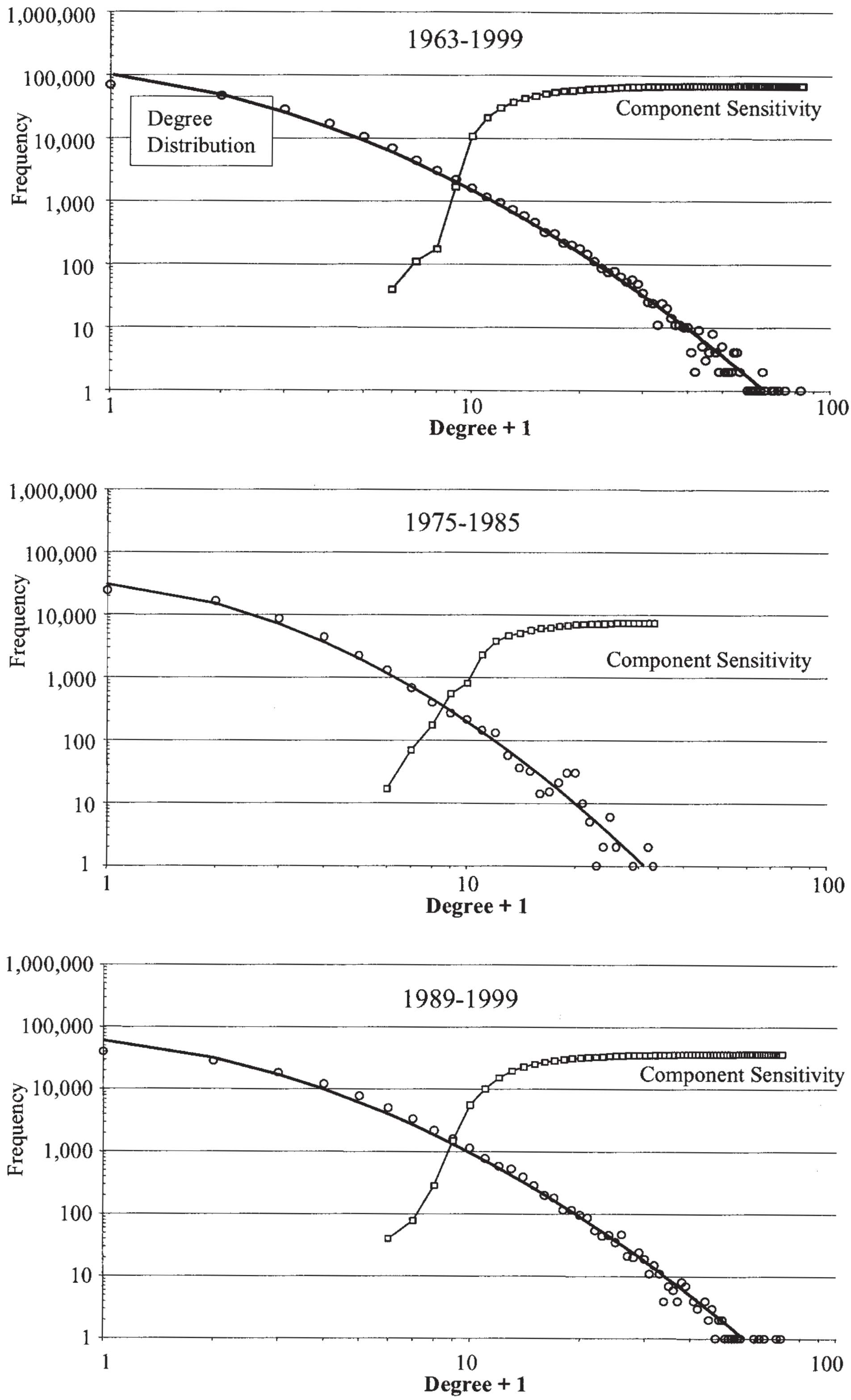


Figure 4. Scale-Free properties of Coauthorship Networks

generated solely by a preferential attachment process. Most prior research on power-law distributions tests for a power-law by fitting a regression line to the observed points in the log-log space. Using the simple regression approach, in each case the fit is improved by adding a squared term to the regression, indicating that the relation is not linear.[22]

Substantively, the key interest in scale-free networks lies in the extent to which a small number of high-degree actors are responsible for connecting the network, which puts them in a uniquely powerful position for influencing the direction of scientific practice. While the degree distribution does not conform to a strict power law, it is highly skewed, so it makes sense to test for this structural property directly. The second line in each panel of figure 3 (called *component sensitivity*) gives the size of the largest connected component when all actors with that degree or more are removed from the graph. It is clear from the figure, that the networks hold together well into the bulk of the tail, not disconnecting until we remove *all* actors with greater than 8–10 collaborators. While a minority of the total network has 8 or more collaborators, one would still have to remove over 500 of the most active nodes to disconnect the network. This network is *not* held together by a small number of network stars. Thus, while the network contains clear star actors – people with a disproportionate number of ties, and such actors are likely very influential within a local region of the network, information diffusion through the network does not depend on such actors.

[22] Jones and Handcock (2003) have recently critiqued much of this work. The regression method improperly weights cases and fails to account for autocorrelation between points. Their alternative maximum likelihood techniques also suggest that the observed data do not conform to a preferential attachment process. When viewed in the aggregate, the degree distribution is best fit with a variant of a negative binomial model, when viewed in the short run (75–85, 89–99) none of the standard models fit the observed distributions particularly well, suggesting that many alternative factors determine the number of collaborators. Thanks to Mark Handcock for providing these models.

A Sociological Small World?

Is the social science collaboration network characterized by distinct clusters that are weakly connected to each other? Using formulas developed by Newman et al. (2001), we can test the observed graph properties relative to a random graph with a similar joint distribution of authors and papers. Any network that has significantly greater local clustering than expected by chance and average distances about equal to chance are considered small-world networks (Watts 1999; Watts and Strogatz 1998). Formally, one measures local clustering with the clustering coefficient, C, which is the proportion of all two-step contacts (collaborator's collaborators) that are also directly connected (called the transitivity index in prior work [Davis 1970; Davis and Leinhardt 1972; Harary et al. 1965; Holland and Leinhardt 1971]) and distance with l, the average path length between connected nodes. A *small-world* network has clustering that is higher than expected and average distances roughly equivalent to that expected in a random network of similar size and distribution of number of partners.

The top panel of Table 4 compares the clustering and path length statistics for the observed networks to the random expectation. For the first period, the observed clustering coefficient, C, is .194, which is not substantively different from the expected random value of .207. For the total network, the observed characteristic path length, l, is 9.81, which is significantly longer than the expected 7.57. Thus, distances are greater, and relations less clustered, than would be expected in a random graph with similar contribution structures, which means the graph does not have a small-world structure.

This result is largely replicated for the two period-specific networks. In each case, the clustering coefficient is a little smaller than random expectations, but distances are significantly greater than expected under random mixing, in direct contradiction to the small-world model. These findings suggest that the collaboration network is not composed of distinct, separate clusters. Instead, permeable theoretical boundaries likely result in a network that folds in on itself, connecting people at greater distances from widely different specialties.

Table 4. Comparison of Observed Coauthorship Structure to Equivalent Random Networks

	1963–1999	1975–1985	1989–1999
Nodes (n)[a]	128,151	35,109	87,731
Small-world Parameters			
Cluster Coefficient	.194	.306	.266
(Random expected)	(.207)	(.312)	(.302)
Average Path Length[b]	9.81	12.26	11.53
(Random expected)	(7.57)	(8.31)	(8.24)
Size of Largest Component			
Observed	68,285	7,492	36,772
Random Paper Assignment	95,078	16,736	59,736
(SD)	(169)	(131)	(145)
Ratio of Observed to Random	.72	.45	.62
Random Paper + One Publication	78,753	15,378	49,061
(SD)	(132)	(90)	(120)
Ratio of Observed to Random	.87	.49	.75
Size of Largest Bicomponent			
Observed	29,462	2,034	15,281
Ratio of Bicomponent to Component	.43	.27	.42
Random Paper Assignment	47,339	4,774	29,738
(SD)	(166)	(94)	(186)
Ratio of Bicomponent to Component	.50	.28	.50
Ratio of Observed to Random	.62	.43	.51
Random Paper + One Publication	48,769	7,882	31,806
(SD)	(153)	(78)	(123)
Ratio of Bicomponent to Component	.62	.51	.65
Ratio of Observed to Random	.60	.26	.48

[a] Excludes people without coauthors.

[b] Applies only within the largest connected component.

Structural Cohesion?

The final model based on commentaries of the discipline suggests a broad-based structurally cohesive collaboration network. The minimum requirement for cohesion is connectivity, and thus increases in the size of the largest connected component are a basic requirement for structural cohesion. While a necessary substrate, a component can be quite fragile, since removing a single person can disconnect the network. A stronger criterion for cohesion is the size of the largest bicomponent.[23]

We need a benchmark to meaningfully judge the size of a component or bicomponent in empirical networks. I construct comparison networks by randomly assigning the observed set of authors to the observed set of papers (which retains the observed publication volume distributions), then construct a random collaboration network from these randomized authorships. This is preferable to simply randomizing edges in the full network, since it maintains the necessary clustering that results from multiple authors on a single paper. Moreover, an authorship randomization approach allows me to control other mixing features, such as homophily on number of publications or the distribution of authors across specialties. If changes in the inclusion of a particular specialty with more coauthorships in the database were driving results, randomizing within specialty would effectively account for this bias. It should also be noted that components and bicomponents in random graphs effectively form an *upper bound* on component size, since under random mixing the components quickly converge to cover the entire graph (Palmer 1985). As such, the meaningful com-

[23] I focus here on the size of the *largest* bicomponent, but smaller bicomponents do exist in the network. However, they are usually many orders of magnitude smaller than the largest component. For example, the second largest bicomponent in the pooled network has fewer than 50 nodes.

parison over time is how much closer to the random value we get, conditional on the relevant mixing features of the network. These comparisons are given in the bottom panel of table 4.

Nearly half of all collaborating authors (68,285) are members of a single connected component, meaning that it is possible to trace a path from each to the other through coauthorship chains. This is about 72% of chance levels, if one simply assigned all authors to papers at random. In all years, the next-largest component is orders of magnitude smaller than the giant component. For the full network, nearly 60% of people who have coauthored but are not in the largest component are scattered across components of 2 or 3 people.

Simple random assignment, however, ignores the fact that many authors are only represented because they have coauthored a single paper that generates isolated dyads (cases such as {E and F} in figure 2). We can set the randomization process to match this parameter, ensuring that our simulated network contains as many necessarily isolated dyads as observed in the real graph (Random Paper + One Pub condition). Doing so lowers the expected size of the giant component. Compared to this more realistic simulation, the observed giant component is about 87% of the random expectation.[24] Looking over time, we see that the proportion of the population in the largest component has steadily increased relative to random expectation. Based on the one-pub restriction, the largest component in the early period was 49% of random expectation, rising to 75% of random expectation in the later period.

We find a similar story with respect to the size of the largest bicomponent. Bicomponents are nested within components, and 43% (29,462 people) of the members of the largest component are also members of the largest bicomponent, or about 60% of the random expected size. Again, the relative proportion has increased over time, moving from 26% in the early period to 48% in the later period (based on the one-pub randomization model).[25]

Theoretical consensus should be higher among pairs of people embedded in higher-order k-components. While a complete cohesive blocking (Moody and White 2003) of the total coauthorship network is impossible because of its size, we can estimate the distribution of higher-order connectivity for the entire graph based on the connectivity distribution among a sample of dyads. By definition, every pair in the largest component (N = 68,285) has at least 1 path connecting them, and every pair within the largest bicomponent (N = 29,462) has at least two paths. Nested within this bicomponent, the largest tricomponent has approximately 14,627 (CI = 14,372–15,375) members, the largest 4-component has approximately 7,992 (CI = 6972–8068), and approximately 5,203 (CI = 4564–5667) are connected by 5 or more independent paths.[26]

Higher-order connectivity appears to have increased over time as well. The bicomponent for the 1975–1985 period is small enough to allow a complete enumeration of all nested con-

[24] Additional controls were checked, including conditioning on mixing by number of authors beyond isolated dyads (equivalent to fixing the diagonal of the mixing matrix, while the 'one-pub' condition only fixes the 1,1 cell), and constraining mixing within areas, which accounts for changes in specialty representation over time. Neither of these restrictions have as strong an effect on the expected values as the one-pub restriction. Tables available on request.

[25] Because bicomponents must be nested within components, and because our observed component is smaller than the random component, directly comparing the size of the observed bicomponent to the random graph somewhat underestimates the relative cohesion in the observed network. To account for this underestimate, I present the ratio of the size of the largest bicomponent to the size of the largest component for both the observed and the random graph. These figures also show that cohesion has increased over time, from .51 in the early period to .65 in the later period.

[26] Estimates are based on the number of node-independent paths connecting randomly sampled pairs of nodes. I then back-estimate the size of the *k*-component from the distribution of node-independent paths. I estimate confidence intervals by bootstrapping the resulting distribution, then using distribution means at 5% and 95%. Because this estimate is based on a sample, it is impossible to identify the sets of nodes that comprise the higher-order *k*-components. These should probably be taken as high-end estimates, since people can belong to different *k*-components, though my informal explorations of these data suggest that this is unlikely at these lower k-levels.

nectivity sets. Here we find that while there are many 3 and 4 components in the network, they are always very small (usually less than 10 members). These small higher-order k-components are linked together within the larger bicomponent in a manner that suggests a 'ridge-structure', where each group is partially embedded with other groups (Friedkin 1998). This is an image of a loose federation of coauthors linked within the wider cohesive set (though the core in the early period is only a small fraction of the total), with no significant schisms separating the network.

The 1989–1999 period admits to a greater fraction of the (much larger) bicomponent embedded in higher-order k-components. Approximately 5,023 (CI = 4,827–5,200) nodes are embedded in a 3-component, while 2,763 (CI = 2,559–2,885) are in a 4 component and 1,616 (CI = 1,368–1,703) are in components of $k > 4$. From these estimates, it appears that a substantial number of social scientists are deeply embedded within a highly cohesive coauthorship core, and that the size of this core has increased over time.

How are these cohesive sets related to each other in the network? The general shape of the network can be best represented with a *contour sociogram*.[27] In a traditional sociogram, points are arrayed spatially to minimize the distance between connected points and maximize the distance between disconnected points, as with the largest component in figure 2. For very large networks, a point-and-line sociogram is uninformative, because nodes simply crowd each other out, stacking on top of each other to reveal a largely uninformative cloud (just as a scatter plot of thousands of points often results in what appears to be an even spread over the entire space). However, we can use the bivariate distribution of points in this space to identify concentrations of nodes in the network. Any region of the graph with a comparatively high level of cohesion will have a larger number of nodes crowded together in that region, and thus a higher probability density value.[28] Figure 5 presents the contour plot for the largest bicomponent of the pooled network.

The general spherical distribution of nodes (there are no nodes outside the .21 contour) suggests that actors are spread relatively evenly across the possible space, showing no major divisions.[29] Within this comparatively flat terrain, there are two more prominent hills (seen here as those areas with a density of 2.74 and higher), connected by a high-density ridge (within the 2.11 contour). A primary subgroup analysis of the largest bicomponent reveals that the two peaks differ in the topics studied.[30] The highest hill in the "South-East" section corresponds to people writing in general sociology, while the "North-Central" hill has a concentration of people writing in applied health fields.

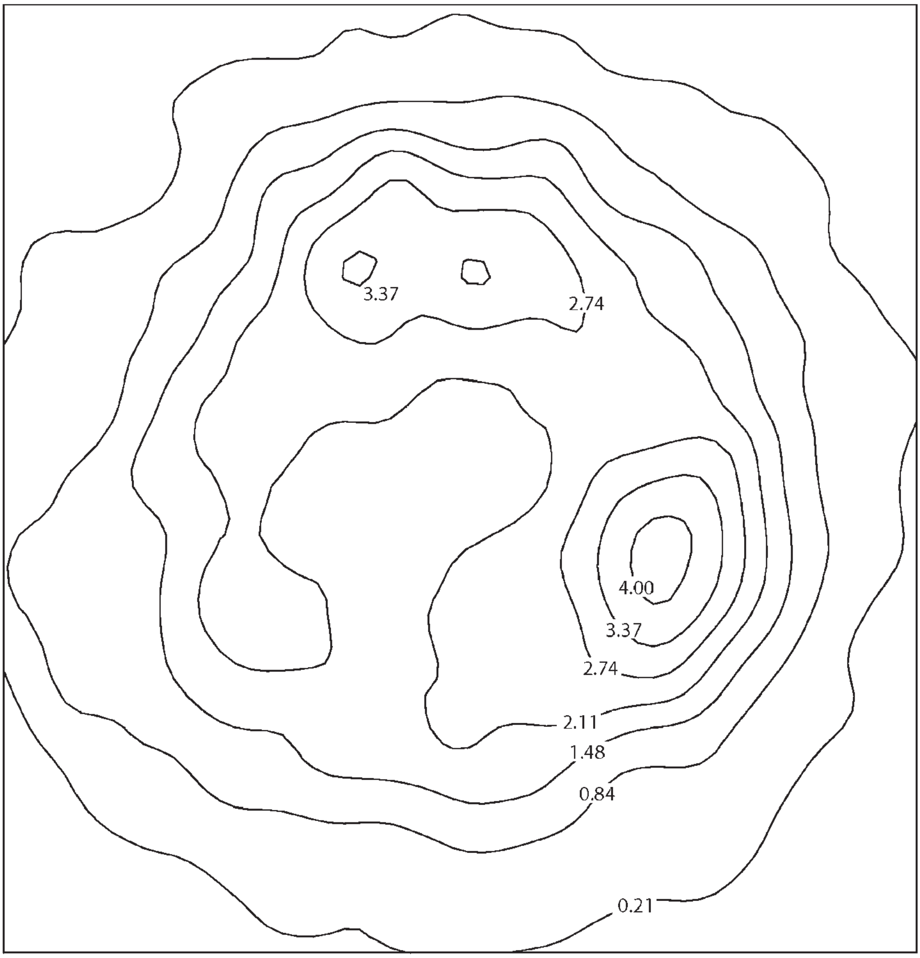
As is always the case with cluster analytic techniques, it is tempting to reify such clusters at the expense of the general topology. Friedkin (1998) notes that the cluster structure of a network can be compared to geographical topography, and describes "ridge structures" as "sequentially overlapping and densely occupied regions of the social space" (p.125). This ridge-structure image seems appropriate for the internal structure of the observed collaboration network. There are areas of relative concentration, but they overlap substantially. As suggested by previous work on the structure of sociology specialties (Cappell and Guterbock 1992; Daipha 2001; Ennis 1992; Richards 1984), high levels of intergroup contact, weak internal structure, and strong overall connectivity point toward a generalized cohesion *within* the sociology coauthorship core. This is a structure that should promote theoretical integration, since ideas and information can poten-

[27] To my knowledge, this is the first time networks have been represented with this type of figure.

[28] The concentration of points is often very uneven, resulting in very jagged contour plots. I have used a non-parametric kernal density estimation technique to smooth the resulting surface, with a bandwidth of .8. Alternative figures with less smoothing are available from the author on request.

[29] Contrast the relatively even distribution of nodes here with a large racially segregated network, such as http://www.sociology.ohio-state.edu/jwm/racel.gif.

[30] Prior versions of this paper included a detailed analysis of the subgroup structure of the core bicomponent. The two peaks correspond to clusters uncovered in the detailed cohesive group analysis. Combined they contain only 15.7% of the nodes in the graph. These results are available from the author on request.



Contour Sociogram. Contour values denote bivariate density estimates, indicating every 15th percent of the range between the 5th and 95th percentiles.

Figure 5. Social Science Coauthorship Network Largest Bicomponent (n = 29,462)

tially flow quite freely, seldom getting trapped on one side of a topological chasm.

Specialty Area and Structural Embeddedness

If a particular type of social science dominates the coauthorship core, then this connected group might have greater influence in shaping questions in the field (Crane 1972). If, instead, specialties are evenly distributed across the network, then we have greater evidence for functional integration (Hargens 1975). We can model individual embeddedness within the coauthorship network using a modeling strategy parallel to that used for network participation. The dependent variable now is *structural embeddedness* (Moody and White 2003) in the network: those in small isolated components are coded 1, those in the largest connected component are coded 2, and those in the largest bicomponent are coded 3. These results appear in Table 5.

In addition to measures defined for the network participation models, I add additional measures based on the local network patterns. I control for network volume measures, including the number of unique collaborators and the average number of authors per paper. *Mentorship* captures the preferential attachment hypothesis directly by looking at the relative publication frequency of coauthors. Mentorship is measured as the number of publications for the focal individual minus the number of publications for his/her coauthor, averaged over all

Table 5. Ordered Logistic Regression of Network Embeddedness on Network and Publication Characteristics

	1963–1999			1975–1985			1989–1999		
Variable	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3	Mod 1	Mod 2	Mod 3
Exposure	1.01	1.01	1.01	1.02[a]	1.02[a]	1.02[a]	1.04	1.05	1.06
Number of Publications	1.17	1.12	1.13	1.36	1.32	1.34	1.33	1.14	1.16
Year of 1st Publication	1.01	1.01	1.01	1.02	1.03	1.02	1.02	1.02	1.02
Male Author (probability)	.81	.83	.83	.91	.93[a]	.93[a]	.82	.85	.85
Authors per Paper	.97	1.01[a]	1.00[a]	.82	.82	.81	1.00[a]	1.03	1.02[a]
Unique Coauthors	1.90	1.81	1.80	1.62	1.63	1.62	1.69	1.61	1.59
Mentorship	.85	.85	.85	.74	.74	.75	.72	.73	.73
Coauthor Diversity	.78	.86	.88	1.02[a]	1.0[a]	1.09[a]	.97[a]	1.08[a]	1.14
Complete Coverage	1.41	1.44	1.46	2.94	2.55	2.49	1.17	1.31	1.35
Priority Coverage	1.16	1.17	1.14	1.54	1.42	1.38	1.09	1.11	1.08
Quantitative Work	—	—	1.58	—	—	1.55	—	—	1.83
Specialty Area (code)									
Radical (25)	.70	.95[a]	.99[a]	.32	.43[a]	.47[a]	.49	.77[a]	.88[a]
Marxist (30)	.57	.76	.77	.56	.62	.65	.55[a]	.87[a]	1.04[a]
Knowledge (22)	1.20	.86	.87[a]	.75[a]	.89[a]	.90[a]	.59	.94[a]	1.06[a]
History & Theory (2)	1.05[a]	1.01[a]	1.02[a]	.98[a]	1.05[a]	1.07[a]	.84	1.06[a]	1.09[a]
Culture & Society (5)	.97[a]	.80	.82	.54	.55	.62	.64	.81	.88[a]
Visual (33)	.94[a]	1.02[a]	1.11[a]	1.01[a]	.97[a]	1.33[a]	.72[a]	1.30[a]	1.49[a]
Language & Arts (13)	.87	.80	.80	.61	.65	.69	.78	.95[a]	.95[a]
Political Sociology (9)	.88	.99[a]	.98[a]	.74	.81	.81	.80	1.05[a]	1.04[a]
Science (17)	.81	.89	.90	.89[a]	.94[a]	.94[a]	.68	.87	.92[a]
Social Change (7)	1.02[a]	.97[a]	.97[a]	.84[a]	.87[a]	.88[a]	.72	1.00[a]	.97[a]
Religion (15)	.95	1.01[a]	1.00[a]	.97[a]	1.02[a]	1.00[a]	1.00[a]	1.19	1.15
Group Interaction (4)	.97[a]	1.07[a]	1.07[a]	1.14[a]	1.17[a]	1.12[a]	.96[a]	1.21	1.21
Urban (12)	.87	.96	.95[a]	1.05[a]	1.05[a]	1.05[a]	.71	.95[a]	.92[a]
Community Development (23)	.68	1.15[a]	1.17[a]	1.52	1.50	1.51	.65	.89[a]	.88[a]
Female Gender (29)	.95[a]	.96[a]	.96[a]	1.15	1.12[a]	1.13[a]	.78	.94[a]	.95[a]
Social Development (83 = 36)	.74	.82	.81	.80	.80	.81	.73	.93[a]	.91[a]
Social Control (16)	.95	1.11	1.11	1.03[a]	1.04[a]	1.04[a]	1.01[a]	1.19	1.20
Policy & Plan (24)	.83	.91[a]	.94[a]	.94[a]	.96[a]	1.02[a]	.82	1.07[a]	1.11[a]
Clinical (31)	.57	1.15[a]	1.17[a]	1.13[a]	1.13[a]	1.12[a]	.04	.02	.03[a]
Mass Phenomena (8)	.81	1.09	1.09	.80	.83	.82	.92	1.15	1.14
Rural (11)	1.07	1.04[a]	1.03[a]	1.53	1.57	1.53	.92[a]	1.20	1.18
Education (14)	.92	—	—	.98[a]	—	—	.81	—	—
Methodology (1)	1.05[a]	1.14	1.13	1.08[a]	1.12[a]	1.10[a]	1.19	1.48	1.45
Environmental (26)	.92	1.03[a]	1.05[a]	1.41	1.41	1.43	.85	1.08[a]	1.06[a]
Violence (28)	.96[a]	.98[a]	.98[a]	1.15[a]	1.12[a]	1.10[a]	.98[a]	1.14[a]	1.13[a]
Demography (18)	.98[a]	1.06[a]	1.04[a]	1.30	1.30	1.21	1.00[a]	1.20	1.15
Social Differ. (10)	.88	1.14	1.12	1.13	1.15[a]	1.11[a]	.96[a]	1.21	1.16
Poverty (27)	.96[a]	1.03[a]	1.02[a]	1.08[a]	1.19[a]	1.15[a]	1.03[a]	1.26	1.25
Social Plan/Policy (72/35)	1.01[a]	1.10	1.10	.83	.86	.85	.95[a]	1.16	1.19
Complex Orgs (6)	.73	1.11	1.09	1.06[a]	1.08[a]	1.07[a]	.90	1.10	1.06[a]
Business (32)	1.12[a]	1.29[a]	1.28[a]	1.51[a]	1.46[a]	1.21[a]	.99[a]	1.38[a]	1.25[a]
Social Psychology (3)	.91	1.19	1.17	.96[a]	.97[a]	.94[a]	1.07	1.33	1.29
Social Problems (21)	1.20	1.23	1.21	1.23	1.25	1.24	1.13	1.31	1.28
Family (19)	1.12	1.18	1.16	1.07	1.07[a]	1.04[a]	1.18	1.40	1.35
Health (20)	1.13	1.19	1.18	1.12	1.13[a]	1.12[a]	1.21	1.43	1.40
Social Welfare (61/34)	1.08	1.13	1.14	.86	.88[a]	.88[a]	1.16	1.35	1.35
R-Square	.499[b]	.506	.511	.404[b]	.418	.422	.478[b]	.489	.497
N		86,498			24,897			56,632	

Note: Data shown as odds ratios. Unless otherwise noted, all cell values are significant at $p \leq .01$ (tables with detailed significance levels are available from the author). Mod = Model.

[a] Value is *not* significant at the $p < .01$

[b] Pertains only to the publication and demographic characteristics.

coauthors. Mentors will thus have large positive scores. *Coauthor diversity* measures the extent to which a person coauthors with different coauthors, relative to their opportunity to coauthor with others. This is calculated as the number of observed coauthors divided by the maximum possible number of coauthors given the number of papers published and the number of authors on each paper.[31] We would expect that those who have higher coauthor diversity would be more deeply embedded in the coauthorship network.

As with network participation, time in the discipline (exposure) increases the likelihood of being in the core of the network, though more so in the later period than in the early period. As expected, the number of publications is also a strong predictor of being in the core, as is the number of unique coauthors. Relative diversity of coauthorship patterns results in a lower likelihood of being at the core overall, but the magnitude is small and inconsistent over time. In the later period, coauthorship diversity increases the likelihood of being at the core of the network, while it is nonsignificant in the early period. Mentorship is consistently negatively related to being at the heart of the network. To the extent that this type of publication pattern captures shop production, it may be that those shops are relatively isolated or that students do not then go on to write with new people.[32] Just as males were less likely to coauthor, conditional on coauthorship, males are less likely to occupy the core of the network suggesting that females are strongly integrated into the overall network core of the discipline.[33]

Looking at the multivariate models for specialty area (model 2), the general pattern of coefficients is similar to that for network participation. Those writing in areas such as theory and culture are slightly more likely to be at the periphery of the network and social psychology, sociology of business, family and social welfare more likely to be at the core, but the magnitudes are *much* smaller. Many of the areas are not statistically distinguishable from average, even with the large statistical power in these models, and the relative size of the odds ratios is much closer to 1 than in the coauthorship models. A clear summary of the weakness of specialty for predicting embeddedness is seen in the change in model fit with and without specialty fields. Across all three networks, the models only weakly improve by adding specialty area to the individual attributes. Turning to model 3, we find that quantitative work increases embeddedness, but has almost no effect on the value of the specialty area coefficients nor does it improve the model much.

Substantively, these models suggest that *once one enters the coauthorship network*, the key predictors of position are individual and publication characteristics. In contrast to network involvement, specialty area is a weak predictor of network embeddedness. Turning this finding around, it suggests that the cohesive core is spread relatively evenly across the specialty areas at risk to coauthorship.[34]

CONCLUSION AND DISCUSSION: SOCIAL INTEGRATION IN THE SOCIAL SCIENCES

Coauthorship is becoming increasingly more common in the social sciences. Nearly half of all papers and over two thirds of all papers in the

[31] Thanks to an anonymous reviewer for suggesting these measures.

[32] I have also tested a cohort based mentor measure, identifying those authors that are older than their coauthors (as measured by first publication appearance). The measure is either unrelated to core membership or negative, in much the same way as the publication volume measure.

[33] Prior analyses on the subgroup structure suggests that membership in smaller groups within the largest bicomponent differ by gender, with males more likely to be in the peak region at the southeast corner of figure 5. As with the coauthorship model, the embeddedness model benefits by controlling for database coverage. Those authors who publish in complete or priority journals are more likely to be in the core of the network.

[34] Previous readers suggested that this permeability might be due to the large number of people with few publications (such as graduate students who publish in disparate areas). Having published little, they may be unimportant to the general character of sociological production. In response, I have constructed a network of people with more than 3 publications and who have been in the discipline for more than 5 years. This significantly lowers the sample size but not its general topology. A greater proportion of people are in the largest bicomponent, but the bicomponent is no more fractured than when using the full sample.

ASR are coauthored. Coauthorship is not evenly distributed across sociological work. As predicted by others, coauthorship is more likely in specialties that admit to an easier division of labor. Research method seems particularly important, showing that quantitative work is more likely to be coauthored than non-quantitative work. While there is a specialty gap in network participation, among those who have participated, specialty area is only a weak predictor of network embeddedness. Thus, just as heterogeneity provides only limited information on social integration (Moody 2001), observations about fractionalization in the discipline based on increasing numbers of specialties might be misleading. The coauthorship pattern shows a steadily growing cohesive core, suggesting that while authors might specialize, their skills marry well with others creating an integrated collaboration network.

How do we account for the observed collaboration pattern? Two complimentary images of science production are suggestive. Abbott's (2001) description of social science as having permeable theoretical boundaries suggests that specialization within the social sciences does not necessarily generate divisions between specialists. Instead, competitors actively borrow ideas from each other (even if under new names), to cover the available idea space. This free mixing means that one's coauthors need not coauthor with each other, and thus the network as a whole admits to little clustering and few schisms, instead spreading quickly over the relevant idea spaces represented in the discipline. Friedkin's (1998) work suggests a specialty analogue to Abbott's competitive mixing model. Friedkin found that while contact clustered within specialties, these clusters where strongly connected to each other, creating network conduits through which ideas and information flow. Tie heterogeneity *within* groups means that groups can act as bridges between other groups but still maintain internal cohesion (Paxton and Moody 2002). Fleshing these hypotheses out will require examining the internal structure of the collaboration network in more detail, though preliminary work suggests that Friedkin's model fits for short-run images of later periods of the collaboration graph.

What do these findings suggest for the prospects of scientific consensus in sociology?[35] Data limitations demand cautious interpretations, but the structure is suggestive. First, the two most prominent models for consensus in an idea space are through references to recognized authority (Martin 2002; Crane 1972) or distributed interpersonal influence (Friedkin 1998). Both models suggest that systematic differences in network *participation* will generate an ideational gulf between those involved in the network and those without collaborations. In this case, the major divide centers largely on research method (quantitative or not) and theoretical focus (radical, cultural, and interpretive modes against largely positivist and empiricist modes). While this gulf appears substantial, increasing collaboration over-time suggests that it might be shrinking.

The two theoretical approaches offer slightly different predictions for future consensus *within* the connected collaboration network. The interpersonal influence model suggests that high overall cohesion will generate generalized consensus, as ideas circulate among the scientists connected in the network, though the long-term nature of the network suggests this might be a slow affair. A finer-level prediction is that consensus should be directly correlated with structural embeddedness, and those embedded in higher order k-components would be more similar to each other than those at the fringes of the network. However, the network does admit to a large inequality in numbers of collaborators, indicating clear stars even though these stars are not essential for connecting the entire network. Martin's authority-based perspective would suggest that those actors with many collaborators might have much more influence shaping ideas

[35] While suggestive, such proposals come with large caveats. Although coauthorship is a clear indicator of social connection, it is a stringent one. The trace of interaction can be found only after the collaboration is recorded through publication. It is likely that other types of social interaction are layered on top of the coauthorship network, which would likely lead to greater levels of cohesion than that observed through the coauthorship network. Second, the *Sociological Abstract* area codes may correlate only weakly with the humanist-positivist division that troubles many sociologists, building necessary ambiguity into these findings. Third, the database ignores book publications, which might systematically exclude some areas more than others.

than others, perhaps acting as "pumps" for ideas that are then quickly circulated through the well-connected regions of the collaboration graph.

While we lack the data necessary to answer this question directly, I suspect that the two approaches are both correct for different aspects of scientific consensus. I suspect that the high overall cohesion levels will generate consensus with respect to methods and rules of evidence, but that stars will act as "area authorities" with respect to particular theoretical or empirical claims. Thus, as methodological change continues to foster a division of labor based on how we *do* research, this will generate consensus on what counts as valid evidence for making scientific claims. However, competition for status within the discipline will likely revolve around stars who generate new ideas at the intersection of different research specialties. Classic treatments of the division of labor suggest that such integrated specialization should lead to organic solidarity, though this need not lead to a unification of particular ideas (Durkheim [1933] 1984; Hargens 1975; Hagstrom 1965; Whitley 2000). Perhaps, then, as Durkheim first suggested, cohesive collaboration networks will simultaneously allow for theoretical diversity and scientific consensus.

James Moody *is an Assistant Professor of Sociology at the Ohio State University. His research focuses broadly on social networks, with particular interest in linking individual action to global network dynamics. In addition to work on network methods, he has worked on questions related to the formation and dynamics of adolescent friendship networks, the diffusion of sexually transmitted diseases, and the network foundations for social solidarity.*

REFERENCES

Abbott, Andrew. 2000. "Reflections on the Future of Sociology." *Contemporary Sociology* 29:296–300.

———. 2001. *Chaos of Disciplines*. Chicago, IL: University of Chicago Press.

Allison, Paul D., Scott J. Long, and Tad K. Krauze. 1982. "Cumulative Advantage and Inequality in Science." *American Sociological Review* 47:615–25.

Babchuk, Nicholas, Keith Bruce, and Peters George. 1999. "Collaboration in Sociology and Other Scientific Disciplines: A Comparative Trend Analysis of Scholarship in the Social, Physical and Mathematical Sciences." *The American Sociologist* 30:5–21.

Barabasi, Albert-Laszlo and Reka Albert. 1999. "Emergence of Scaling in Random Networks." *Science* 286:509–12.

Bearman, Peter. 1993. *Relations into Rhetorics: Local Elite Social Structure in Norfolk, England: 1540–1640*. ASA Rose Monograph Series ed. New Brunswick, NJ: Rutgers University Press.

Bonacich, Phillip. 1987. "Power and Centrality: A Family of Measures." *American Journal of Sociology* 92:1170–1182.

Burt, Ronald S. 1987. "Social Contagion and Innovation: Cohesion Versus Structural Equivalence." *American Journal of Sociology* 92:1287–335.

Cappell, Charles L. and Thomas M. Guterbock. 1992. "Visible Colleges: The Social and Conceptual Structure of Sociology Specialties." *American Sociological Review* 57:266–73.

Cole, Jonathan R. and Stephen Cole. 1973. *Social Stratification in Science*. Chicago, IL: University of Chicago Press.

Collins, Harry M. 1998. "The Meaning of Data: Open and Closed Evidential Cultures in the Search for Gravitational Waves." *American Journal of Sociology* 104:293–338.

Collins, Randall. 2001. "Why the Social Sciences Won't Become High-Consensus, Rapid Discovery Sciences." Pp. 62–84 in *What's Wrong With Sociology*, edited by Stephen Cole. New Brunswick, NJ: Transaction Press.

———. 1986. "Is 1980s Sociology in the Doldrums?" *American Journal of Sociology* 91:1336–55.

Connell, R. W. 2000. "Sociology and World Market Society." *Contemporary Sociology* 29:291–96.

Crane, Diana. 1972. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago, IL: University of Chicago Press.

Crane, Diana and Henry Small. 1992. "American Sociology Since the Seventies: The Emerging Identity Crisis in the Discipline." Pp. 197–234 in *Sociology and Its Publics: The Forms and Fates of Disciplinary Organization*, edited by Terence C. Halliday and Janowitz Morris. Chicago, IL: University of Chicago Press.

Daipha, Phaedra. 2001. "The Intellectual and Social Organization of ASA 1990–1997." *The American Sociologist* 32:73–90.

Davis, James A. 1970. "Clustering and Hierarchy in Interpersonal Relations: Testing Two Graph Theoretical Models on 742 Sociomatrices." *American Sociological Review* 35:843–51.

———. 2001. "What's Wrong With Sociology?" Pp. 99–120 in *What's Wrong With Sociology?*, edited by Stephen Cole. New Brunswick, NJ: Transaction Press.

Davis, James A. and Samuel Leinhardt. 1972. "The Structure of Positive Relations in Small Groups." Pp. 218–51 in *Sociological Theories in Progress*, vol. 2, J. Berger, M. Zelditch, and B. Anderson. Boston, MA: Houghton Mifflin.

Durkheim, Emile. [1933] 1984. *The Division of Labor in Society,* translated by W. D. Halls. New York: The Free Press.

Endersby, J. W. 1996. "Collaborative Research in the Social Sciences: Multiple Authorship and Publication Credit." *Social Science Quarterly* 77:375–92.

Ennis, James G. 1992. "The Social Organization of Sociological Knowledge: Modeling the Intersection of Specialties." *American Sociological Review* 57:259–65.

Fisher, Bonnie S., Craig T. Cobane, Thomas M. Vander Ven, and Francis T. Cullen. 1998. "Trends and Patterns in Political Science." *Political Science Online*:847–56.

Friedkin, Noah E. 1998. *A Structural Theory of Social Influence*. Cambridge, England: Cambridge University Press.

Gould, Rodger. 2002. "The Origins of Status Hierarchies: A Formal Theory and Empirical Test." *American Journal of Sociology* 107:1143–78.

Hagstrom, Warren O. 1965. *The Scientific Community*. New York: Basic Books.

Harary, Frank, Robert Z. Norman, and Dorwin Cartwright. 1965. *Structural Models: An Introduction to the Theory of Directed Graphs*. New York: John Wiley and Sons.

Hargens, Lowell. 1975. *Patterns of Scientific Research*. Washington D.C.: The American Sociological Association.

Holland, Paul W. and Samuel Leinhardt. 1971. "Transitivity in Structural Models of Small Groups." *Comparative Groups Studies* 2:107–24.

Hudson, John. 1996. "Trends in Multi-authored Papers in Economics." *Journal of Economic Perspectives* 10:153–58.

Jones, James and Mark Handcock. 2003. "Sexual Contacts and Epidemic Thresholds." *Nature* 423:605–6.

Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*. Chicago, IL: University of Chicago Press.

Laband, D. N. and R. D. Tollison. 2000. "Intellectual Collaboration." *Journal of Political Economy* 108:632–62.

Leifer, Eric M. 1988. "Interaction Preludes to Role Setting: Exploratory Local Action." *American Sociological Review* 53:865–78.

Lieberson, Stanley and Freda B. Lynn. 2002. "Barking Up the Wrong Branch: Scientific Alternatives to the Current Model of Sociological Science." *Annual Review of Sociology* 28:1–19.

Mannheim, Karl. 1936. *Ideology and Utopia*. New York: Harcourt, Brace & Company.

Markovsky, Barry. 1998. "Social Network Conceptions of Solidarity." Pp. 343–72 in *The Problem of Solidarity: Theories and Models*, edited by Patrick Doreian and Thomas Fararo. Amsterdam, The Netherlands: Gordon and Breach.

Martin, John-Levi. 2002. "Power, Authority, and the Constraint of Belief Systems." *American Journal of Sociology* 107:861–904.

McDowell, John M. and Melvin Michael. 1983. "The Determinants of Co-Authorship: an Analysis of the Economics Literature." *Review of Economics and Statistics* 65:155–60.

Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159:56–63.

Milgram, S. 1969. "The Small World Problem." *Psychology Today* 22:61–7.

Moody, James. 2001. "Race, School Integration, and Friendship Segregation in America." *American Journal of Sociology* 107:679–716.

Moody, James and Douglas R. White. 2003. "Social Cohesion and Embeddedness: A Hierarchical Conception of Social Groups." *American Sociological Review* 68:103–27.

Newman, M. E. J. 2000. *Models of the Small World*. Sante Fe Institute Technical Paper.

———. 2001. "The Structure of Scientific Collaboration Networks." *Proceedings of the National Academy of Sciences* 98:404–9.

Newman, M. E. J., S. J. Strogatz, and D. J. Watts. 2001. "Random Graphs with Arbitrary Degree Distributions and Their Applications." *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* 64:026118.

Owen-Smith, Jason. 2001. "Managing Laboratory Work Through Skepticism: Processes of Evaluation and Control." *American Sociological Review* 66:427–52.

Palmer, Edger N. 1985. *Graphical Evolution: An Introduction to the Theory of Random Graphs*. New York: John Wiley and Sons.

Paxton, Pamela and James Moody. 2002. "Structure and Sentiment: Explaining Emotional Attachment to Group." *Social Psychology Quarterly* 66:34–47.

Richards, Jim. 1984. "Structure of Specialization Among American Population Scientists." *Scientometrics* 6:425–32.

Simmel, Georg. 1950. *The Sociology of Georg Simmel*, edited by Kurt H. Wolf. New York: The Free Press.

Simpson, Ida H. and Richard L. Simpson. 2001. "The Transformation of the American Sociological Association." Pp. 271–89 in *What's Wrong With Sociology*, edited by Stephen Cole. New Brunswick, NJ: Transaction Press.

Skvoretz, John. 1998. "Theoretical Models: Sociology's Missing Links." Pp. 238–52 in *What is Social Theory? The Philosophical Debates*, edited by Alan Sica. Oxford, UK: Blackwell.

Stinchcombe, Arthur L. 1994. "Disintegrated Disciplines and the Future of Sociology." *Sociological Forum* 9: 279–91; reprinted 2001.

Pp 85–94 in *What's Wrong With Sociology*, edited by Stephen Cole. Transaction Press.

Swidler, Ann and Jorge Arditi. 1994. "The New Sociology of Knowledge." *Annual Review of Sociology* 20:305–29.

Watts, Duncan J. 1999. "Networks, Dynamics, and the Small-World Phenomenon." *American Journal of Sociology* 105:493–527.

Watts, Duncan J. and Steven H. Strogatz. 1998. "Collective Dynamics of 'Small-World' Networks." *Nature* 393:440–442.

White, Douglas R. and Frank Harary. 2001. "The Cohesiveness of Blocks in Social Networks: Node Connectivity and Conditional Density." *Sociological Methodology* 31:305–59.

White, Douglas R., Michael Schnegg, Lilyan A. Brudner, and Hugo G. 2002. "Conectividad Múltiple, Fronteras e Integración: Compadrazgo y Parentesco en Tlaxcala Rural." Pp. 41–94, *Análisis de Redes: Aplicaciones en Ciencias Sociales*, Eds. Jorge Gil-Mendieta y Samuel Schmidt. Mexico, D.F.: IIMAS-UNAM. (Instituto de Investigaciones en Mathemáticas Aplicadas y en Systemas- Universidad Nacional Autónoma de México).

Whitley, Richard. 2000. *The Intellectual and Social Organization of the Sciences*. New York: Oxford University Press.

Zuckerman, Harriet. 1977. *Scientific Elite: Nobel Laureates in The United States*. New York: Free Press.

Zuckerman, Harriet and Robert K. Merton. 1973. "Age, Aging, and Age Structure in Science." Pp 497–560 in *The Sociology of Science: Theoretical and Empirical Investigations*, edited by Norman W. Storer. Chicago, IL: University of Chicago Press.