

The State of Sociology: Evidence from Dissertation Abstracts

Micah Altman, Center for Equitable and Open Science, MIT Libraries, Massachusetts Institute of Technology

Philip N. Cohen, Department of Sociology, University of Maryland, College Park

(Authors are listed in alphabetical order. Contributions are described in detail in the Acknowledgements sections.)

Abstract: The academic discipline of sociology is divided methodologically and theoretically, which has long been a source of consternation as well as celebration. In this paper we describe the state of actually existing sociology as practiced in PhD-granting departments, by analyzing all dissertation abstracts from U.S. sociology departments archived at ProQuest from 2020 through early 2024 (N=1872) using natural language processing methods (Keyword Assisted Topic Models with Latent Dirichlet Allocation). Results show that more dissertations (64%) use qualitative than quantitative (47%) methods. Where geographic scope is identified, most focus on U.S. topics. The most common substantive topics are inequality, broadly defined (58%), followed by economic subjects (46%), race (37%), politics and health (35% each). The most pervasive difference concerns qualitative versus quantitative approaches. For example, qualitative methods are most common in organizations and movements (79%), and least common in life course (43%) and children and youth studies (54%). Quantitative methods are most associated with life course (61%) and ethnicity (59%), and least common in social movements (34%). Finally, using gender probabilities based on given names, we find that average female representation is somewhat higher in qualitative studies (61%) than quantitative (56%) studies, but subject of study matters more, with the average probability of female authorship highest in gender (72%) and family (66%) studies, and lowest in urban studies (52%). Women are more concentrated in qualitative than quantitative studies across almost all subject areas, but the differences are relatively small, so that gender divisions appear to be driven more by subjects of study than by methodological approaches. The data and methods developed in this paper may be extended to additional research questions.

Introduction

The academic discipline of sociology is divided methodologically and theoretically, which has long been a source of consternation as well as celebration. A variety of studies have assessed the methods and subject interests of sociology research, which are surely “entangled” (Schwemmer & Wieczorek 2020), as reflected in the long-standing finding that women are more likely than men are to do qualitative social science (Grant et al. 1987; Nunkoo et al. 2020; Plowman & Smith 2011). Others have focused on the frequency of studying of particular issues, such as climate change (Hiltner 2024; Scoville and McCumber 2023). There is less work assessing the epistemological or political perspectives of sociological research, which are more difficult to measure (Gross 2013). Elaborating on the nature of American sociological research is a particularly useful task at present, given the situation in which the discipline finds itself, described below.

This paper takes a descriptive approach to one aspect of this landscape: What are the major methodological approaches in American sociology today, and how are those associated with topics of study and author characteristics? We use dissertations for this question. Others have studied sociology journals (e.g., Schwemmer & Wieczorek 2020), which hinges on selection into publication, and includes many scholars who are not sociologists (in terms of degree field or employment status). A recent large survey of authors in sociology journals revealed that only half identified sociology as their primary discipline (Lockhart et al. 2024). And an analysis of dissertations going back to 1980 found less than half of those using the stem “sociol” in title, abstract, keywords, or department name were completed in sociology departments (Heiberger et al. 2021). Because we are interested in the workings of the discipline itself, we address our questions by analyzing dissertations produced in U.S. sociology departments, thereby avoiding publication bias, including only work from sociology departments, and providing an assessment of scholarship at a consistent state of development – the completion of a PhD degree.

Fault lines in sociology

The recognition of substantive and methodological divisions within sociology is longstanding (Abbott 2017; Ritzer 1975). Based on interviews with a handful of U.S. sociologists and his reading of the field, Gross (2013:199-200) described the field as, “a fractious, multiparadigm discipline composed of researchers focused on a wide array of problems and employing diverse theoretical approaches and methods.” He outlined some divisions this way:

“One important dividing line in the field is between sociologists who identify more with the humanities and those who view the discipline as a social science; another is between sociologists committed to an activist agenda and those who have more of a ‘professional’ orientation. Among sociology interviewees, nearly two-thirds said they did think critically about the notion of objectivity, but only one-third told us objectivity is an illusion. ... The epistemic culture of sociology is a hybrid, both in the sense that some champions of skepticism and subjectivism can be found in its ranks, and in that among

the rest acknowledgment of the difficulty of being objective — coupled with a commitment to grasp for it anyway — is regarded as *de rigueur*.”

We add that within sociology there are divergent approaches in the realm of operational science policy (Altman & Cohen 2023), such as in the areas of scholarly communication; and others that pertain to political stances taken by individual and organizational actors within the discipline. The divisions that center around questions of open science and open access have implications for the relationship of the discipline to other sciences – in ways that highlight or at least correlate with distinctions between quantitative and qualitative work, and that between activist-oriented and non-activist-oriented scholars.

Institutionally, the American Sociological Association (ASA) depends on subscription revenue and has lobbied against open access initiatives. Meanwhile, its elected leadership in recent years has leaned markedly left and activist-oriented, apparently reflecting the perspectives of the voting membership. Its positions – which reflect sometimes divergent interests of the elected academic leaders and the professional staff – may appear internally contradictory. Some sociologists have become disgruntled with ASA and publicly quit (including one of us),¹ partly because of opposition to open science² and open access.³ Others see ASA as too politicized or too leftist, too dominated by elite (or non-elite) institutions, too racist (or anti-racist), too intellectually diverse (or not diverse enough), and so on. The stakes have risen recently as some politicians and state governments have targeted sociology as part of the problem of “woke” / DEI (diversity, equity, and inclusion) domination in higher education (Cohen 2024). There are a raft of less dominant professional societies in sociology – some regional, and some focused on subject areas or types of research – that are in generally similar positions, and mostly have in common the dependence on revenue from journal subscriptions. Meanwhile, in addition to regional associations, some sociologists have fomented a new organization, the Society for Sociological Science, around the open-access journal *Sociological Science*. In the last year they held a conference that included as one of its defining issues, “ideas for improving sociological science.”⁴ The critique of ASA implied by the journal and its actions is not formally stated, but this may reflect divides in the discipline widening.

Recent political controversy outside academia regarding critical race theory and intersectionality, including politicians singling out sociology for public criticism and targeted policy, has increased scrutiny of these questions. Some question the legitimacy of the discipline by doubting its scientific nature, which is linked to the question of political bias, activist orientation, or exaggerated claims. In the Census Bureau coding, sociology is a “science, technology,

1

<https://familyinequality.wordpress.com/2021/11/06/why-im-leaving-the-american-sociological-association/>

2

<https://familyinequality.wordpress.com/2024/01/09/quick-update-on-american-sociological-reviews-failure-to-meet-current-social-science-standards/>

3

<https://familyinequality.wordpress.com/2023/08/11/american-sociological-association-in-absentia-but-not-silent-on-open-science/>

⁴ <https://socsciconf.org/>

engineering and math" (STEM) occupation⁵; at the National Science Foundation, sociology is funded under the Social & Economic Sciences division.⁶ However, although ASA regularly identifies sociology as a science in its public advocacy, the degree to which the discipline is science, or sciency, or scientific, or social scientific, is still contested (Cohen forthcoming). Whether this matters is debatable in many contexts, but in the case of scientific practices it is important because of its implications for the assessment of "reproducibility and replicability" in the discipline (National Academies 2019). Some sociologists favor this perspective while some oppose it, at least for some kinds of sociology (Breznau 2021; Khan et al. 2024). In both attitudes toward and practices related to open science, meanwhile, sociology lags behind economics, psychology, and political science (Ferguson et al. 2023).

Without surveying members of the discipline on their personal views, the political perspective of a discipline itself is difficult to assess. The ASA conference theme in 2024, selected by the current elected president, Joya Misra, was "Intersectional Solidarities," and she described it in terms that lean heavily left: "The 2024 theme emphasizes sociology as a form of liberatory praxis: an effort to not only understand structural inequities, but to intervene in socio-political struggles."⁷ On the other hand, in an op-ed opposing Florida state government's decision to remove sociology from a list of core college courses, Misra wrote:

"Sociology courses, by considering inequalities by wealth, income, race, gender, sexuality and age, may seem overly 'woke' to conservatives. But sociological findings are based on scientific analysis of data from objective sources, such as the U.S. Census Bureau."⁸

If Misra was elected, do her views reflect "the discipline" with regard to its members or its collective self-identity? These questions have been subject to debate within the discipline for decades, as recounted recently by a former association president Michael Burawoy with regard to Palestine.⁹ In one indication, in 2024 a member-initiated resolution in support of an Israeli ceasefire in Gaza passed with 59 percent of an election with 34 percent turnout (so, 20 percent of members voted in favor, 14 percent against).¹⁰

This review includes a wider array of subjects than we have thus far addressed with empirical analysis. For example, we have not yet developed methods for identifying political bias, activist orientation, or the strength of claims sociologists advance in their dissertations. We do not yet have race or ethnic identifiers for authors, and have not collected the data necessary to analyze change over time. In the meantime, however, the analysis below presents data and methods that may be helpful in launching such additional investigations.

⁵ <https://www.census.gov/library/stories/2022/11/stem-workers-under-age-25.html>

⁶ <https://www.nsf.gov/sbe/ses/about.jsp>

⁷ <https://www.asanet.org/2024-annual-meeting/annual-meeting-theme/>

⁸

<https://theconversation.com/whats-sociology-a-sociologist-explains-why-floridas-college-students-should-get-the-chance-to-learn-how-social-forces-affect-everyones-lives-222365>

⁹ <https://www.sociologistsforpalestine.org/blog/sociologyfacesthequestion>

¹⁰ <https://www.asanet.org/about/governance-and-leadership/election/>

Procedures

We analyze 1,873 English-language dissertation abstracts from graduate programs with “sociology” in the department name, published to the ProQuest database from the beginning of 2020 through the first quarter of 2024. Using this corpus, we developed a set of non-exclusive topics that reflect both substantive and methodological characteristics. First, in a close reading of 49 abstracts, we identified keywords in dissertation abstracts that indicated qualitative and quantitative methods, and those substantively related to inequality of various kinds.¹¹ These were treated as definitive indicators of these core concepts. Second, from a list of 1444 ngrams of two words or less (excluding stop words) that appear (at least once) in > 2.5% of dissertations, we individually tagged 524 terms associated with 26 research subjects and four methodological approaches. For example, ngrams tagged as *subject:health* included, “health behaviors” and “mental health”; and those tagged *subject:race* included, “racial composition” and “race ethnicity.” Those tagged *method:empirical* included, “data analysis”; *method:longitudinal* include, “longitudinal study”; and so on. (Our tags for *method:theory* are limited to those using the ngrams that include “theor/y/ies/etical.”¹²) Some ngrams were used more than once, such as “race class,” which was tagged for race as well as class. The number of ngrams per tag are presented in **Table 1**.

Many abstracts do not explicitly state the population under study, but we also identified terms associated with studies of either the United States or other societies (“world”), partly based on geographic terms and partly based on a list of common survey datasets with known populations.¹³ In addition, we added terms associated with analysis of US versus global contexts in the close reading of 49 abstracts (e.g., “Ivy League” and “Title IX” indicate US studies, while “low-income countries” and “Chinese government” indicate global topics).

¹¹ For example, terms that indicated qualitative dissertations include “ethnography,” “qualitative study,” and “participant observation”; those indicating quantitative work included statistical terms such as “bivariate regression,” “poisson,” and “moderation models”; and those indicating inequality included, “racism,” “social class,” and “stratification.”

¹² In hand coding, we attempted to identify dissertations that were primarily theory generating, rather than using existing theories. Finding only 1 such dissertation in the 50 we coded, we dropped that code from this analysis.

¹³ The list of surveys is here:

<https://familyinequality.wordpress.com/2022/10/15/these-are-the-datasets-of-our-lives-43-of-them-anyway/>.

Table 1. Ngram tagging categories

Subjects		Methods	
<u>Label</u>	<u># of tags</u>	<u>Label</u>	<u># of tags</u>
health	53	empirical	88
race	48	qualitative	39
gender	32	quantitative	24
education	29	longitudinal	5
family	29	theory	5
inequality	28		
econ	23		
crim	21		
politics	14		
class	12		
ethnicity	11		
immigration	10		
violence	9		
sexuality	9		
orgs	7		
children and youth	7		
movements	6		
urban	6		
environment	6		
networks	5		
lifecourse	5		
identity	4		
media	3		
rural	3		
religion	2		
stigma	1		

From these two sets of tags – high-confidence keywords and topic tags – we used Keyword Assisted Topic Models (KeyATM) with Latent Dirichelet Allocation (LDA) to produce term-topic clusters of dissertation abstracts (Eshima, Kosuke and Sasaki 2024). Dissertations assigned to each topic include those that had exact matches to associated ngrams, as well as those identified by the model to include related terms. To illustrate the topics identified, **Table 2** lists the most strongly identified dissertations in the top four substantive topics – inequality, race, politics and health – and those tagged as quantitative, qualitative, and theory.

Table 2. Most strongly identified dissertations in top four substantive topics and methods

<p><u>Inequality</u></p> <ol style="list-style-type: none">1. Income Inequality and Caste in India: Evidence from India Human Development Surveys2. Heeding the Call: An Empirical Evaluation of Gentrification Research3. Rural Oklahoma and the Nexus of Disaster Vulnerability, Risk, and Resilience4. The Social Etiology of Prescription Psychotherapeutic Misuse among Rural and Urban Adolescents and Emerging Adults: Concurrent and Prospective Stress-Process Analyses5. Rural/Urban Health Disparities among the Oldest Old in China
<p><u>Race</u></p> <ol style="list-style-type: none">1. Understanding “Our” Similarities and Differences Academically, Socially, and Psychologically: The Race-Gendered Experiences of Black Men and White Men Enrolled in STEM Doctoral Programs2. Social Stratification in the Mortgage Market Post the Great Recession3. The Macro and Micro Foundations of Racial Residential Segregation: A Contemporary Analysis4. Internalizing Achievement Inequality: The Development of Racial/Ethnic Differences in Mathematics Attitudes and Their Implications for Persistence in STEM5. Dancing While Black: Managing Racial Fatigue in Ballet
<p><u>Politics</u></p> <ol style="list-style-type: none">1. The Social Dynamics of Political and Economic Crises in the Twenty- First Century2. Unsettling Science: How Activists Deployed Science in the Conversion Therapy Debate3. Reading, Evaluating and Commemorating Feminism : Excluding and Reviving Dynamics of <i>L'Euguélonne's</i> Reception in Québec4. the Fraternal Twins: A Comparative Study of Hegemony Building, Political Mobilization and Crony Capitalism in Neo-Authoritarian Turkey and Russia (1991–2021)5. Militarism, Democracy, and Concordance: The Role of Citizenry in (Re)-Establishing Democratic Order in Argentina and Turkey
<p><u>Health</u></p> <ol style="list-style-type: none">1. Exploring Differential Exposure to Adverse Social Determinants of Health for Children with Developmental Disabilities and Their Related Health Outcomes and Resiliency2. Societal Shocks as Social Determinants of Health3. Exploring Multiple Stakeholder Experiences of Healthcare: Opportunities for Intervention and Change to Improve Care and Outcomes for People with Mental Illness4. Patient Engagement and the Epistemics of Medical Authority: Diagnosis Resistance in US Primary Care5. Educational Disparities in Chronic Pain and Life Expectancy: Gaps and Pathways
<p><u>Qualitative</u></p> <ol style="list-style-type: none">1. Culturing Evolution: A History and Anthropology of a Cognitive Science of Culture in Illiberal Hungary2. Creativity on Demand: Cognition, Materiality, and Sociality in Improvisation3. Sociological Silhouettes: The Archaeology and Demography of Knowledge4. The Ghost in the Machine: Organ Transplantation and the Phenomenology of Embodiment5. Observing Prestige: Visibility and Performance in the Sociology of Knowledge
<p><u>Quantitative</u></p> <ol style="list-style-type: none">1. Hurricanes and Heart Problems: Natural Disasters, Social Capital, and Cardiovascular Mortality2. Using Recurrent Neural Networks to Code Interviewer Question-Asking Behaviors: A Proof of Concept3. Assessing Area-Level Suicide and Overdose Rates via Google Search Term Data in the US4. Recidivism and Crime Specialization Among Male Inmates in Chinese Prisons5. Reputation and Generalization in Social Context: Turnout Reporting and Intergroup Relations
<p><u>Theory</u></p> <ol style="list-style-type: none">1. Group Cognition: Reclaiming C.S. Peirce's Specialized Theory of Cognition and Its Implications for Sociology and the Cognitive Sciences2. Peircean Pragmatism, Critical Realism, and Cultural Sociology3. Status, Influence, and Emotions: Emotional Contagion in Status Hierarchies4. "It Was Handed to Them": The Origins of Targeted Delivery and the Spirit of Nanomedicine5. Three Essays on Diversity Dynamics in Science and Culture

To identify differences in the theoretical or methodological approaches across subject areas, in our analysis we present rates of co-occurrence of these methods and substantive topics. Finally, in a preliminary description of demographic patterns to dissertation topics, we show the predicted proportion female for each topic-by-methods group, using given-name-based gender imputation (Blevins and Mullen 2015) matched to the World Gender Name Dictionary (Martinez, et al. 2021).¹⁴

Results

The topics identified by the LDA model and their frequencies in the abstracts data are shown in **Figure 1**. This shows, for example, that 87% of dissertations were identified as empirical, 64% qualitative, 47% quantitative, and 38% referencing theories. U.S.-based dissertations (56%) were much more common than studies referencing other societies (21%). Among the 14 most common substantive topics, inequality was the most common (58%), followed by economic subjects (46%), race (37%), politics and health (35% each), family (32%), education (30%), and gender (30%).

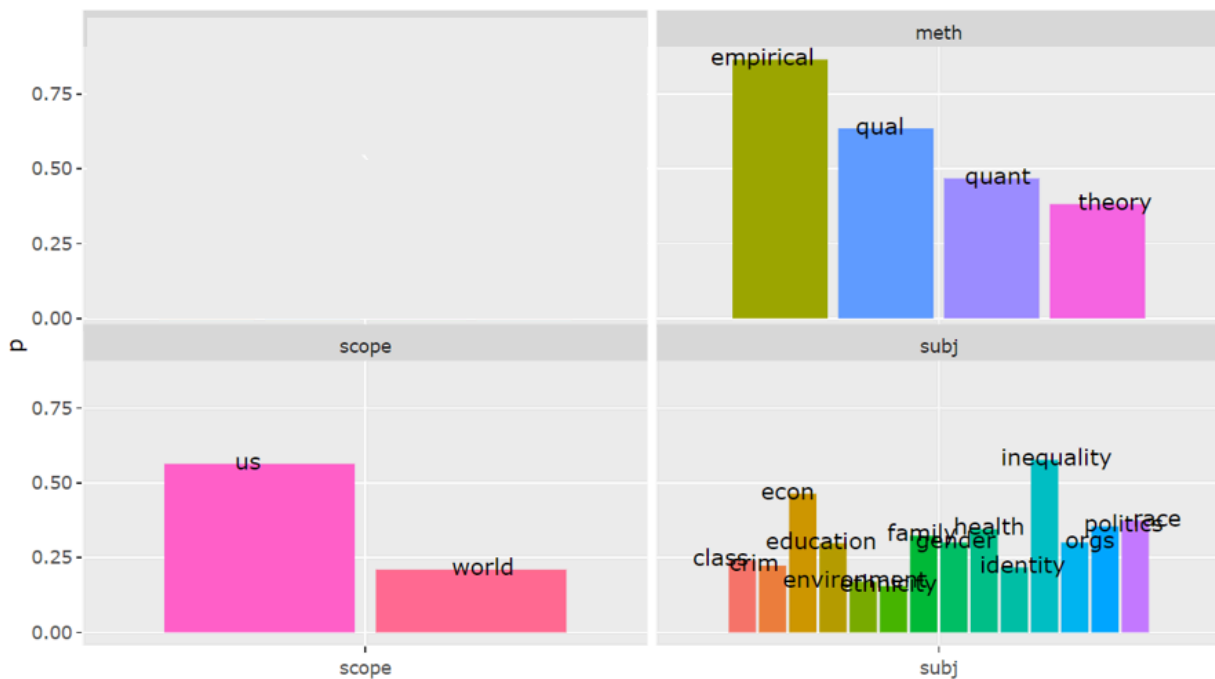


Figure 1. LDA topic model frequencies.

¹⁴ West et al. (2013) compute gender composition of authors across subfields of sociology using articles in the JSTOR corpus from 1990 to 2011, and report they “vary dramatically in gender composition,” but use a different algorithm for defining subfields, and do not provide detailed breakdowns.

The co-occurrence proportions of substantive topics and methodological topics is shown in **Table 3**. The results show, for example, that between 85% (politics) and 94% (children and youth) of dissertations in substantive topics are identified as empirical; and between 3% (stigma) and 30% (life course) are associated with longitudinal methods. Qualitative methods are most common (79%) in organizations and movements; and least common in life course studies (43%). Quantitative methods are most associated with life course (61%), ethnicity (59%), and health studies (56%); and least common in social movements (34%), identity (36%), and politics (38%). (Quantitative studies are more common in subjects that are US rather than world focused, while qualitative studies are associated with non-US focus, which may partly result from the use of common secondary datasets with US respondents, as noted above.) Longitudinal dissertations are least common in stigma studies (3%) and most common in the life course area (30%) – representing a generally positive correlation between longitudinal and quantitative studies. Finally, theory was identified most commonly in the study of networks (57%) and stigma (46%), and least common in rural studies (32%). However, almost all subjects have between 34% and 40% theoretical identification. Every subject has fewer than one-third of abstracts associated with non-US contexts except immigration (33%), and all but four subjects have 25% or fewer world-focused studies. Thus, the most pervasive difference is that some subjects are much more associated with qualitative approaches while others are concentrated in quantitative work.

Table 3. Proportion of each subject by methodology and scope

Subject	(N)	Methodology				Scope		
		Empirical	Longitudinal	Qual	Quant	Theory	US	World
Class	450	0.90	0.14	0.69	0.45	0.35	0.62	0.21
Criminology	419	0.86	0.12	0.63	0.46	0.35	0.61	0.20
Economic	868	0.89	0.13	0.64	0.48	0.38	0.61	0.24
Education	558	0.88	0.14	0.60	0.55	0.39	0.59	0.17
Environment	317	0.90	0.12	0.70	0.41	0.39	0.56	0.29
Ethnicity	290	0.89	0.14	0.61	0.59	0.38	0.71	0.17
Family	606	0.90	0.17	0.58	0.50	0.35	0.59	0.17
Gender	563	0.90	0.10	0.64	0.50	0.39	0.56	0.20
Health	646	0.90	0.15	0.59	0.56	0.39	0.60	0.16
Identity	413	0.91	0.06	0.78	0.36	0.40	0.53	0.23
Immigration	271	0.90	0.13	0.68	0.45	0.35	0.71	0.33
Inequality	1,081	0.90	0.12	0.65	0.50	0.37	0.61	0.20
Life Course	202	0.93	0.30	0.43	0.61	0.38	0.66	0.13
Media	238	0.91	0.06	0.72	0.39	0.43	0.52	0.24
Movements	263	0.86	0.07	0.79	0.34	0.41	0.57	0.25
Networks	257	0.88	0.13	0.60	0.50	0.57	0.56	0.23
Organizations	563	0.88	0.08	0.79	0.39	0.38	0.60	0.22
Politics	642	0.85	0.11	0.71	0.38	0.37	0.61	0.28
Race	701	0.88	0.11	0.66	0.49	0.38	0.67	0.16
Religion	118	0.86	0.09	0.62	0.53	0.35	0.65	0.26
Rural	100	0.88	0.12	0.60	0.48	0.34	0.52	0.25
Sexuality	185	0.91	0.05	0.77	0.39	0.37	0.55	0.17
Stigma	90	0.93	0.03	0.71	0.38	0.46	0.48	0.13
Urban	280	0.87	0.14	0.64	0.50	0.35	0.72	0.22
Violence	202	0.86	0.10	0.65	0.45	0.38	0.62	0.21
children and youth	178	0.94	0.29	0.54	0.54	0.35	0.59	0.17

Categories are overlapping (percentages may exceed 100%).

Note: Excludes 3 “other” topics identified by the LDA model, and dissertations with no methodological topics (1-5% in each substantive area), or substantive topics (<5) identified. Numbers in parentheses are frequencies.

Next we turn to the gender composition of different substantive areas and methodological approaches. **Figure 2** shows the average probability of being female for dissertation authors by method, U.S. versus other studies, and among the 14 most common substantive topics. In the entire corpus, the average probability of being female, based on given names, is 57.3%. (In 2022, 62% of U.S. sociology doctorates were awarded to women [NCSES 2024].) The average female representation is highest in qualitative studies (61%) relative to empirical (59%), quantitative (56%), and dissertations discussing theory (54%). Studies of the U.S. show more female representation than those of other societies (56% versus 51%). In substantive areas, average probability of being female is highest in gender (72%) and family (66%) studies, and lowest in environment (54%).¹⁵

¹⁵ Krapf, Kreyenfeld, and Katharina (2016) also report female researchers are overrepresented in studies of family in the journal *Demography*.

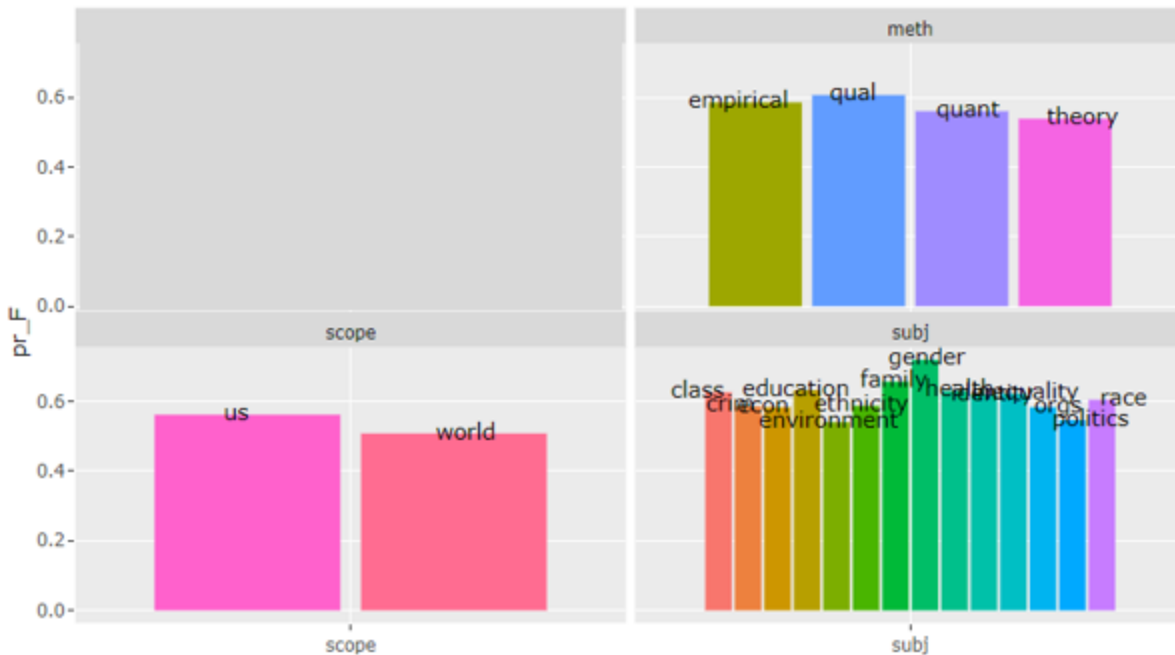


Figure 2. Gender ratio in each category

Finally, **Table 4** shows the average probability of being female by subject and method. The table is color coded for interpretation, so that those with the greatest female representation (77%) are blue, 50% is white, and the lowest value (15%) is red. The table includes some smaller fields than Figure 2, so female representation now ranges from a high of 72% in gender studies to lower levels in smaller fields: religion (48%), networks (51%), and urban studies (52%). In every subject except sexuality, qualitative dissertations are more likely to be written by women than are quantitative dissertations. However, despite that apparent consistency, the difference ranges from 0 points in sexuality to more than 10 points in class, rural, religion, media, and children and youth (a 17-point difference). There is also a wide range of gender composition in longitudinal studies, but these are often small cells. With regard to US versus world differences, in two areas (children and youth and life course) female authors are substantially more concentrated in non-US studies, but in most others female representation is higher in US-based studies, reaching 13 points higher in sexuality studies. Overall, the gender differences between subjects are greater than those within them, so that gender divisions in the discipline – at least as far as dissertation topics are concerned – appear to be driven more by subjects of study than by methodological approaches. That is, the gender differences between methods and scopes is about 5 points, while the range across subjects is 24 points (from the gender-balanced religion and network studies to the strongly female-leaning gender, family, and violence studies).

Table 4. Female proportion for each subject by method and scope

	(N)	Subject mean	Methodology					Scope	
			Empirical	Longitudinal	Qual	Quant	Theory	US	World
Class	450	0.63	0.64	0.54	0.67	0.57	0.59	0.61	0.56
Criminology	419	0.59	0.61	0.49	0.65	0.56	0.59	0.57	0.49
Economic	868	0.58	0.59	0.51	0.61	0.57	0.53	0.58	0.51
Education	558	0.63	0.64	0.55	0.67	0.62	0.58	0.61	0.59
Environment	317	0.54	0.55	0.57	0.55	0.53	0.46	0.53	0.55
Ethnicity	290	0.59	0.60	0.69	0.62	0.59	0.59	0.62	0.52
Family	606	0.66	0.67	0.58	0.70	0.62	0.61	0.64	0.64
Gender	563	0.72	0.73	0.70	0.77	0.70	0.72	0.73	0.70
Health	646	0.63	0.63	0.61	0.68	0.60	0.59	0.62	0.59
Identity	413	0.61	0.62	0.62	0.63	0.60	0.61	0.60	0.51
Immigration	271	0.60	0.61	0.61	0.64	0.56	0.57	0.57	0.55
Inequality	1,081	0.62	0.63	0.54	0.65	0.62	0.59	0.60	0.54
Lifecourse	202	0.64	0.66	0.62	0.69	0.62	0.56	0.62	0.74
Media	238	0.54	0.55	0.42	0.60	0.44	0.49	0.55	0.49
Movements	263	0.55	0.57	0.56	0.58	0.54	0.47	0.56	0.45
Networks	257	0.51	0.54	0.53	0.55	0.51	0.47	0.49	0.41
Organizations	563	0.59	0.60	0.42	0.63	0.56	0.52	0.58	0.49
Politics	642	0.54	0.57	0.46	0.58	0.52	0.51	0.51	0.49
Race	701	0.60	0.61	0.60	0.63	0.61	0.56	0.59	0.54
Religion	118	0.48	0.49	0.15	0.55	0.43	0.47	0.48	0.39
Rural	100	0.61	0.62	0.46	0.63	0.52	0.60	0.60	0.57
Sexuality	185	0.65	0.65	0.66	0.66	0.66	0.65	0.65	0.52
Stigma	90	0.63	0.61	0.33	0.64	0.59	0.62	0.65	0.65
Urban	280	0.52	0.52	0.53	0.53	0.52	0.43	0.49	0.45
Violence	202	0.67	0.69	0.65	0.72	0.63	0.68	0.65	0.61
children and youth	178	0.63	0.61	0.52	0.72	0.55	0.63	0.56	0.72
Category Means									
	–	–	0.58	0.53	0.61	0.56	0.54	0.56	0.51

Categories are overlapping -- percentages may exceed 100%. Number of dissertations in category in parentheses.

Discussion and next steps

Analyzing 1,873 sociology dissertations completed between 2020 and 2024, we find broad differences in methods across subfields of the discipline, so that some are substantially more likely to use quantitative, qualitative, or longitudinal methods; rely on explicit theories; or focus on the US versus non-US contexts. Further, female authors clearly are more concentrated in some subfields than others.¹⁶ However, the differences could be interpreted as marginal. That is, no subfield is more than 79% or less than 43% qualitative, and none is more than 61% or less than 34% quantitative. Similarly, predicted gender composition varies, but only between relative parity (48% female) and strongly female-leaning (72%). In addition, some gender patterns persist across almost all areas, especially the greater tendency of women to use qualitative

¹⁶ These differences across sociology are comparable to those found in some disciplines, such as psychology (Odic and Wojcik 2020) land use science (Chen and Seto 2022), but larger than those reported in some other disciplines, for example oncology (Yalamanjali, Zhang, and Jagsi 2021) and geosciences (Pico et al. 2020).

(61%) versus quantitative (56%) methods. With regard to gender, differences in approach are more related to subjects of study than to qualitative versus quantitative methods.

We have established a method and data source that will allow the investigation to progress further. Specifically, using additional textual analysis, we may develop an assessment of the strength of claims made with regard to results and implications. With further author name analysis, we may assess race/ethnic composition probabilities. Using information on departmental affiliation, we may add a measure of departmental research productivity or centrality (based on citation metrics). Finally, we may expand the chronological range of the data incorporated, to allow investigation of changing patterns over time; and we may include additional disciplines for comparison. All of these can be done within the scope of methods we have thus far developed. With these steps taken, for example, we will be able to model the method of study, and strength of claims, as a function of subject of study, author characteristics, and departmental status.

Acknowledgements

We describe contributions to the paper using a standard taxonomy (see Allen *et al.* 2014). All authors take equal responsibility for the article in its current form. PC lead the writing of the original draft of the manuscript; MA lead data curation, methodology, formal analysis and visualization. All authors contributed to review and revision. All authors contributed to the conception of the article (including core ideas, analytical framework, and statement of research questions). All authors contributed to the project administration and to the writing process through direct writing, critical review, and commentary.

Data and code

Available from <https://github.com/MIT-Informatics/socabs/releases/tag/v-pre1.0>. Authors will deposit a copy of final replication code and data to Zenodo on final publication.

Competing Interests

The authors have no substantial conflicts of interest but wish to thank IMLS (award LG-250130-OLS-21) for their research support.

References

Abbott, Andrew. 2017. *Department and Discipline: Chicago Sociology at One Hundred*. University of Chicago Press.

Allen L, Scott J, Brand A, Hlava M, Altman M. Publishing: Credit where credit is due. *Nature*. 2014;508(7496):312-3.

Blevins, Cameron, and Lincoln Mullen. "Jane, John... Leslie? A Historical Method for Algorithmic Gender Prediction." *DHQ: Digital Humanities Quarterly* 9, no. 3 (2015).

Altman, Micah, and Philip N. Cohen. "We Are in a Period of Science Policy Innovation, yet There Are Major Evidence Gaps in Evaluating Their Effectiveness." *Impact of Social Sciences* (blog), July 27, 2023.

<https://blogs.lse.ac.uk/impactofsocialsciences/2023/07/27/we-are-in-a-period-of-science-policy-innovation-yet-there-are-major-evidence-gaps-in-evaluating-their-effectiveness/>.

Breznau, Nate. 2021. "Does Sociology Need Open Science?" *Societies* 11(1):9. Doi: [10.3390/soc11010009](https://doi.org/10.3390/soc11010009).

Chen, Tzu-Hsin Karen, and Karen C. Seto. 2022. "Gender and Authorship Patterns in Urban Land Science." *Journal of Land Use Science* 17(1):245–61. doi: 10.1080/1747423X.2021.2018515.

Cohen, Philip N. 2024. "How Sociology Can Save Itself." *Chronicle of Higher Education*, February 7.

Cohen, Philip N. Forthcoming. *Citizen Scholar: Public Engagement for Social Scientists*. New York: Columbia University Press.

Shusei Eshima, Kosuke Imai, and Tomoya Sasaki. 2024. "Keyword Assisted Topic Models." *American Journal of Political Science* 68(2).

Ferguson, Joel, Rebecca Littman, Garret Christensen, Elizabeth Levy Paluck, Nicholas Swanson, Zenan Wang, Edward Miguel, David Birke, and John-Henry Pezzuto. 2023. "Survey of Open Science Practices and Attitudes in the Social Sciences." *Nature Communications* 14(1):5401. doi: [10.1038/s41467-023-41111-1](https://doi.org/10.1038/s41467-023-41111-1).

Grant, Linda, Kathryn B. Ward, and Xue Lan Rong. 1987. "Is There An Association between Gender and Methods in Sociological Research?" *American Sociological Review* 52(6):856–62. doi: 10.2307/2095839.

Gross, Neil. 2013. *Why Are Professors Liberal and Why Do Conservatives Care?* Harvard University Press.

Haber, Noah, Emily R. Smith, Ellen Moscoe, Kathryn Andrews, Robin Audy, Winnie Bell, Alana T. Brennan, et al. "Causal Language and Strength of Inference in Academic and Media Articles

Shared in Social Media (CLAIMS): A Systematic Review.” *PLOS ONE* 13, no. 5 (May 30, 2018): e0196346. <https://doi.org/10.1371/journal.pone.0196346>.

Heiberger, Raphael H., Sebastian Munoz-Najar Galvez, and Daniel A. McFarland. 2021. “Facets of Specialization and Its Relation to Career Success: An Analysis of U.S. Sociology, 1980 to 2015.” *American Sociological Review* 86(6):1164–92. doi: 10.1177/00031224211056267.

Hiltner, Sofia. 2024. “Limited Attention to Climate Change in U.S. Sociology.” *The American Sociologist*. doi: 10.1007/s12108-024-09624-4.

Khan, Shamus, Jennifer S. Hirsch, and Ohad Zeltzer-Zubida. 2024. “A Dataset without a Code Book: Ethnography and Open Science.” *Frontiers in Sociology* 9. Doi: [10.3389/fsoc.2024.1308029](https://doi.org/10.3389/fsoc.2024.1308029).

Krapf, Sandra, Michaela Kreyenfeld, and Katharina Wolf. 2016. “Gendered Authorship and Demographic Research: An Analysis of 50 Years of Demography.” *Demography* 53(4):1169–84. doi: 10.1007/s13524-016-0482-x.

Lockhart, Jeffrey W., Molly M. King, and Christin Munsch. 2024. “Who Authors Social Science? Demographics and the Production of Knowledge.” *Social Currents* 23294965241246805. doi: [10.1177/23294965241246805](https://doi.org/10.1177/23294965241246805).

Martínez, Gema Lax, et al. Expanding the World Gender-Name Dictionary: WGND 2.0. No. 64. World Intellectual Property Organization-Economics and Statistics Division, 2021.

National Academies of Sciences, Engineering, and Medicine. 2019. Reproducibility and Replicability in Science. Washington, D.C.: National Academies Press.

National Center for Science and Engineering Statistics, Survey of Earned Doctorates. 2024. Table 1-5. <https://nces.nsf.gov/pubs/nsf24300/data-tables> (accessed July 7, 2024).

Nunkoo, Robin, Michael Thelwall, Jeynakshi Ladsawut, and Sandhiya Goolaup. 2020. “Three Decades of Tourism Scholarship: Gender, Collaboration and Research Methods.” *Tourism Management* 78:104056. doi: 10.1016/j.tourman.2019.104056.

Odic, Darko, and Erica H. Wojcik. 2020. “The Publication Gender Gap in Psychology.” *American Psychologist* 75(1):92–103. doi: 10.1037/amp0000480.

Pico, T., P. Bierman, K. Doyle, and S. Richardson. 2020. “First Authorship Gender Gap in the Geosciences.” *Earth and Space Science* 7(8):e2020EA001203. doi: 10.1029/2020EA001203.

Plowman, Donde Ashmos, and Anne D. Smith. 2011. "The Gendering of Organizational Research Methods: Evidence of Gender Patterns in Qualitative Research." *Qualitative Research in Organizations and Management: An International Journal* 6(1):64–82. doi: 10.1108/17465641111129399.

Ritzer, George. 1975. "Sociology: A Multiple Paradigm Science." *The American Sociologist* 10(3):156–67.

Schwemmer, Carsten, and Oliver Wieczorek. 2020. "The Methodological Divide of Sociology: Evidence from Two Decades of Journal Publications." *Sociology* 54(1):3–21. doi: 10.1177/0038038519853146.

Scoville, Caleb, and Andrew McCumber. 2023. "Climate Silence in Sociology? How Elite American Sociology, Environmental Sociology, and Science and Technology Studies Treat Climate Change." *Sociological Perspectives* 66(5):888–913. doi: 10.1177/07311214231180554.

West, Jevin D., Jennifer Jacquet, Molly M. King, Shelley J. Correll, and Carl T. Bergstrom. 2013. "The Role of Gender in Scholarly Authorship." *PLOS ONE* 8(7):e66212. doi: 10.1371/journal.pone.0066212.

Yalamanchali, Anirudh, Emily S. Zhang, and Reshma Jagsi. 2021. "Trends in Female Authorship in Major Journals of 3 Oncology Disciplines, 2002-2018." *JAMA Network Open* 4(4):e212252. doi: 10.1001/jamanetworkopen.2021.2252.