How Disciplinarity Shapes Demography: 1956-2013

jimi adams, *Department of Health and Behavioral Sciences*

University of Colorado Denver

1. Abstract

Social network analysis provides strategies to examine scientific fields that can evaluate segmentation or consolidation of research(ers) within the field. Those structural patterns can then be overlaid with characteristics to identify what shapes the observed community structure. Demography, which draws from many academic disciplines to cover several well-bounded substantive areas, provides a case to **examine whether disciplinary boundaries or topical focus account for community structure in the field**. I use the complete histories of four general demography journals to examine the field's evolving community structure, the primary drivers of that structure, and to identify what topics bridge between those identified communities. **I find demography has a remarkably robust structural signature that is dominated more by topical organization than disciplinarity** – a pattern unique to demography in comparison with several similarly many disciplinary fields.

2. Research Questions

2.1 How clustered is the field of demography?

Social Network analysis identifies **communities** as groups marked by high levels of intra-group interaction with low levels of inter-group interaction. **Modularity** provides one strategy for measuring the segmentation between such communities. I estimate demography's community structure for the full period, and how it evolves across time.

2.2 What accounts for the modularity between demographic research communities?

Discipline-based: If research communities are primarily organized around disciplines, research problems will be widely distributed, with many disciplinary communities examining the same topics. This would be marked by **low correspondence** between the research communities & the topics covered within them. Such organization would indicate a general lack of problem-based coordination in the field (a "**multi-disciplinary**" approach).

Problem-based: If research communities are primarily organized around research problems, research topics will be narrowly distributed, with many disciplines contributing to single topic-based communities. This would be marked by **high correspondence** between research communities and the topics covered. Such organization would indicate a pattern of problem-based coordination in the field (an "**interdisciplinary**" approach).

2.3 How have these patterns changed over the life-course of demographic research?

How has demographic research matured over the last 5 decades? Does clustering become increasingly marked by disciplinary/problem orientation?

3. Data

All articles, notes, reviews and letters published in ***Demography, Demographic Research, PDR, PRPR,*** & ***Population Studies*** between **1956-2013**. Retrieved full bibliographic information (authors, cited references, etc.) from ISI, limited to papers with at least 1 shared reference.

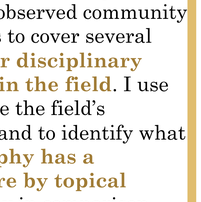
Journal	Total Articles	Non-Isolates
Demography	2,428	2,140
Demographic Research	515	481
Population & Development Review	2,927	1,315
Population Research & Policy Report	922	687
Population Studies	3,273	1,400
Total	10,065	6,023

4. Methods

4.1 Bibliographic Coupling Networks

a. **Network Construction** – weighted similarities of reference lists for all pairs of papers in the corpus. Network ties represent papers with stronger than expected similarities between their cited references.[1]

b. **Community Detection** – modularity-identified groups of papers that collectively rely upon similar unified bodies of literature. The modularity index summarizes the degree of segregation between those communities.[2] (see ***Figure 5.1 & 5.1a***)



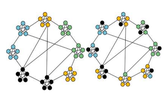
(Source 1, p. 3)

4.2 Topic Models

Frequencies of top **keywords** used per cluster (see ***Table 6.1*** & ***Figure 7***)

4.3 Correspondence between Communities & Topics

To examine how readily the identified communities (4.1b) align with topics (4.2) I rely on contingency analysis for categorical independence residuals, visualized with mosaic plots.[3] (see ***Figure 6.2***)



4.4 Evolution Over Observed Period

I present two analyses of temporal change for the patterns described above. First I examine changes in modularity (see ***Figure 5.2***). Then I examine how clustering of the identified topics evolves over adjacent 7-year periods (see ***Panel 7***), by identifying how topic/cluster patterns change across those slices.

References

[1] Garfield E. "From Bibliographic Coupling to Co-Citation Analysis via Algorithmic Historio-Bibliography: A Citationist's tribute to Belver C. Griffith." Presented at Drexel University. Philadelphia, PA; November 2001.

[2] Porter MA, Onnela J-P, Mucha PJ. "Communities in Networks." *Notices of the AMS* 2009;56(9):1082-1166.

[3] Friendly M. "Mosaic Displays for Loglinear Models." *Proceedings of the Statistical Graphics Section* 1992;61-68.

5. Findings I: Research Communities: Bibliographic Coupling Network

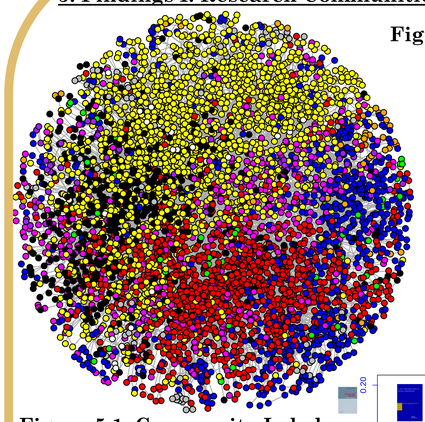
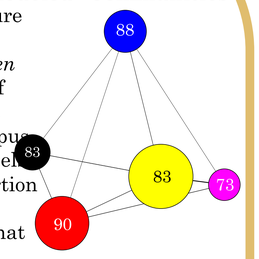


Figure 5.1. Community Labels The bibliographic coupling network's large component (N=4848). Ties are included between 2 papers if their cosine-weighted similarity score is ≥ sample mean + 2sd). Nodes are colored by communities as identified by the Fast Greedy algorithm implemented in igraph 0.7-1 for R 3.1.3.

Figure 5.1a. "Blockmodeled" Communities A reduction of Figure 5.1, summarizing connectivity *between* clusters (w/ > 5% of the corpus). Size is proportional to corpus representation; labels indicate the proportion of bibliographic coupling weights that fall *within* cluster).



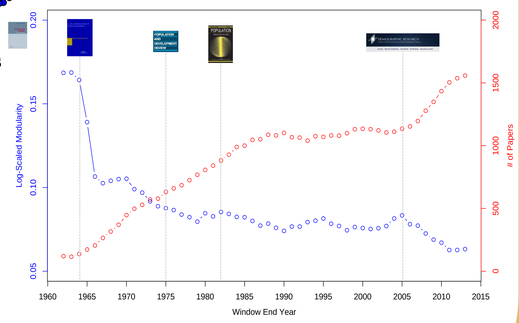


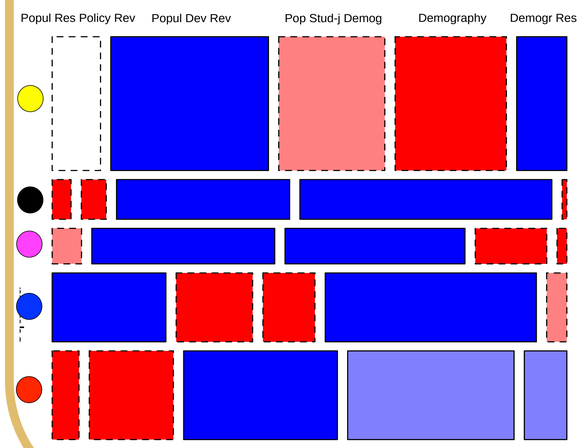
Figure 5.2. Modularity Temporal Change Networks comprising all papers published within time-varying moving windows, *ending* in the listed year. The general pattern is low and equilibrated modularity

6. Findings II: Community-Topic Correspondence

Table 6.1 – Cluster Topics. A label is applied to each cluster based on the dominant theme(s) as indicated by its top used keywords.

cluster	label	% of corpus	% ties w/in cluster
yellow	fertility	34.9	83.1
black	family (planning)	10.5	83.3
magenta	population health	8.4	73.0
blue	migration	14.9	88.5
red	mortality	24.9	90.4

Figure 6.2 – Mosaic Plot, Correspondence between Clusters & Journals



Communities (color coded) correspond to Figures 5.1, 5.1a & Table 6.1. Height represents the relative contribution of the community to the complete corpus. The blue (red) bars denote those that are over- (under-) represented compared to random distribution of communities over journals. The mosaic plot is created with the vcn package in R 3.1.1. This suggests that certain literature clusters (and perhaps topics) are unevenly distributed across the examined journals.

7. Findings III: The Temporal Evolution of Blockmodeled Communities & (Bridging) Topics

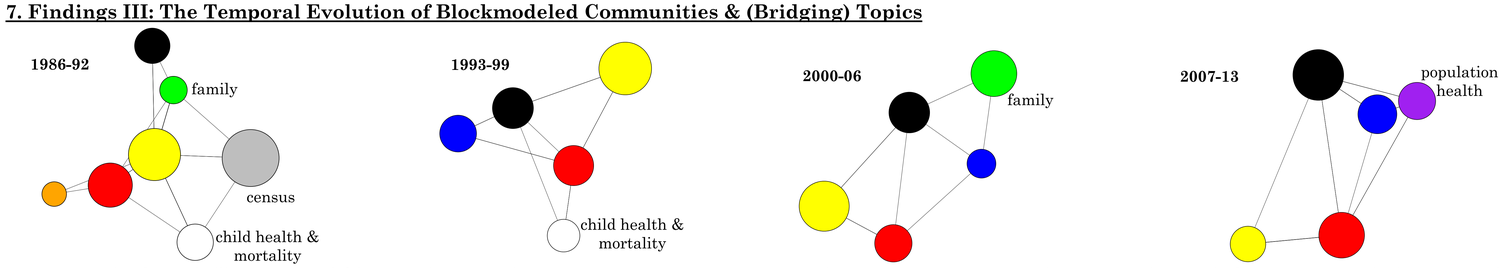


Figure 7. "Blockmodel" Temporal Slices. These blocks summarize the citation network segmentation across the observed period. Only blocks including >5% of all papers and ties including >1% of a blocks weighted bibliographic coupling scores are included. Moving windows are the width of the median citation age across the corpus (7y). These show the relative robustness of the core clusters—fertility (yellow), mortality (red) and migration (blue)—and highlight the evolution of other clusters.

8. Conclusions

Question 2.1 – The field of demography is **moderately clustered** (see Figure 5.1a & Table 6.1), and the dynamics of this bibliographic coupling network clustering appears relatively stable (see Figure 5.2).

Question 2.2 – These clusters appear to be much more dominated by research **content** (e.g., demographic processes) rather than disciplinary boundaries. However, in addition to that topical consolidation, there does appear to be some **journal-based clustering** across those topics (and within a few of them, see Figure 6.2).

Question 2.3 – The core topics are relatively **stable**. Most of the observed change is in their relative importance or the topics that comprise the other observed clusters.

In sum, demography is a clustered discipline, but one that largely clusters around topics, not disciplines. This pattern is consistent with interdisciplinary integration.

Next Steps – I am expanding these analyses to a more complete corpus of demography publications (N~166k). Also, the use and application of keywords is inconsistent across the corpus, so I plan to use a strategy known as "Topic Modeling" to more rigorously examine topic structure.

Acknowledgements:

This work is part of the collaborative IDEAS (Interdisciplinary Dynamics in Emerging Areas of Science) project w/ Ryan Light (Sociology, Oregon), which has received funding from:

- The Robert Wood Johnson Foundation, Health & Society Scholars Program at Columbia University
- A Junior Professorship Development Award from University of Oregon

Contact Information:

- jimi adams, Associate Professor, jimi.adams@ucdenver.edu