Influence dynamics of adjacent academic disciplines.   
The case of economics and sociology.

Benjamin Rosche – CS6742

Economics and sociology are often described as two fundamentally different social sciences. Economics is characterized as the study of rational action, focusing on the efficient allocation of scarce resources. Sociology, by contrast, is characterized as the study of nonrational behavior, focusing on topics, such as norms and biases.

In this project, I use data from the digital library JSTOR to study how research topics in economics and sociology are co-evolving over time. The main language features used in this analysis is the co-occurrence of keywords in the research articles in economics and sociology. From Constellate, I download metadata of most articles that JSTOR offers in sociology and economics between 1900 to 2020. The metadata consists of the title, abstract, author, journal, and publication year of an article. For most articles, the full text is also available. Figure 1 plots the number of articles in each discipline by year. The figure shows that economics is the bigger discipline and that the JSTOR lags behind with adding the most recently published articles.

Figure 1: Number of articles by discipline

Chart, line chart

Description automatically generated

In this paper, I use the article titles to examine influence dynamics of economics and sociology. In next steps, however, I will switch to a full-text analysis in which I will draw on the introduction and conclusion of each article.

**Data cleaning**

From each article title, I extract terms (unigrams) that are longer than 3 characters, excluding stop-words. I end up with 45,289 distinct terms in economics and 29,830 distinct terms in sociology.

**The intersection of economics and sociology**

To find terms that are common to economics and sociology, I first take the top 1000 most frequent terms in each discipline and then get the intersection of those two sets. Here are examples from the resulting set:

['american', 'care', 'case', 'change', 'development', 'economic', 'education', 'effects', 'employment', 'family', 'health', 'impact', 'income', 'international', 'labor', 'life', 'market', 'model', 'national', 'policy', 'political', 'problems', 'public', 'role', 'rural', 'social', 'state', 'states', 'structure', 'theory', 'united', 'urban', 'welfare', 'women', 'work', 'world']

**Research objective and methodological strategy**

The objective of this project is to understand how terms in economics and sociology are co-evolving over time. While each discipline has domains of research with little overlap from the other one, they also share a several research areas. I pursue two strategies to examine who is influencing whom in those intersections over time.

1. *Fighting word analysis*

In an influential article, Monroe, Colaresi, and Quinn (2008) develop a method to identify terms that distinguish between two sets of documents. Like PMI and TF-IDF, their approach is based on term frequencies. I employ their method to analyze how whether terms at the intersection of economics and sociology change the relative frequency with which they appear in both disciplines over time. This relative frequency is expressed as a z-score. If this score is positive, the term appears more frequently in sociology (proportionally to the size of its corpus). If the score is negative, the term appears more frequently in economics.

1. *Co-occurrence networks*

My second strategy to understand the co-evolution of economics and sociology is based on co-occurrence networks. For each discipline, I identify term pairs (tuples) within the top 1000 most frequent words that appear together more often than we would expect by chance. I identify the pairs using pointwise mutual information (PMI), which measures the association between two terms. PMI is defined as , where P(A) refers to the frequency of term A and P(B) refers to the frequency of term B in the entire corpus. If PMI is larger (smaller) than 0, the terms A and B appear more (less) frequently together than we would expect by chance. For each term, I then create egonets of terms with which they frequently appear together (i.e., term + neighbors). I use a PMI of 0.5 as cutoff point. I use this procedure to create the co-occurrence networks by decade.

Using the created co-occurrence networks, I then analyze how the egonets of relevant terms at the intersection of sociology and economics have changed over time. In particular, I calculate the Jaccard similarity of egonets from time t to time t+1. More importantly, not only do I analyze the similarity of egonets over time separately by discipline, but I also analyze the similarity of a term’s egonet in economics at time t to its egonet in sociology at time t+1 (and vice versa). In doing so, I am able to assess who is influencing whom.

*Figure 2: Egonet for “inequality” in sociology and economics*

**Results**

1. **Fighting words analysis**

From the list of terms at the intersection of economics and sociology, I identify three term fields that would be interesting to analyze:

1. Subfields of sociology that have an equivalent in economics.

A fighting words analysis is indicative how strongly these subfields dominate the overall discourse in each discipline and, thus, how well they separate sociology from economics articles.

1. Words associated with certain research practices.

A fighting words analysis will give insight in how common such practices in each discipline are.

I analyze the following subfields: family sociology/economics (family, household), educational sociology/economics (education, school), political sociology/economics (policy, governance, institutional, capital), network sociology/economics (networks, graph), gender/race sociology/economics (women, men, gender, race), international sociology/economics (comparative, development.

Figure 3 displays the results of a fighting word analysis of subfields. We can see several patterns. First, economics is an imperialist discipline. Except for race and networks, all here considered subfields lean less toward sociology over time. Take family sociology for instance. Until 1980, the term “family” has become an ever-better separator of sociology and economics articles. In 1981, Gary Becker published his influential book “A treatise on the family”, which founded family and household economics, and won him a Nobel prize. Since then, “family”, while still being associated with sociology, has become more and more frequent in economics. This reversal is also visible for education (reversal started around 1950), for gender (reversal started around 2000), for organization (reversal started around 1970), and for migration (reversal started around 1940). Second, economics took the spot to inform policy as can be seen from the trajectory of “policy” and “government”. Third, race and networks are still firmly in the hand of sociologists.

*Figure 3: A fighting word analysis of subfields.*

A picture containing window, building

Description automatically generated

Regarding research practices, figure 4 shows that causal research is more commonly mentioned in economics and qualitative research more frequently mentioned in sociology. Experimental research, while more frequently mentioned in sociology until 1980, is now more common in economics. Finally, articles mentioning the word “theory” have been strongly associated with economics until about 1940, which around the time the foundations of neoclassical economics have been established. Since then, the term has become leaning more and more towards sociology.

*Figure 4: A fighting words analysis of research practices.*

A picture containing graphical user interface

Description automatically generated

1. **Co-occurrence analysis**

**Conclusion and next steps**