UNIVERSITÉ DE LORRAINE

SUPERVISED PROJECT REPORT

Tracking Theoretical Shifts over Time: A Natural Language Processing Analysis of Cass R. Sunstein

Authors:
Omar CHERIF
Dalila LADLI
Maxime MÉLOUX

Supervisor:
Samuel FEREY

Reviewer:
Bart LAMIROY

A thesis submitted in fulfillment of the requirements for the degree of Master 1 in Natural Language Processing

in the

Bureau d'Économie Théorique et Appliquée
Institut du Digital, Management et Cognition

June 18, 2022

Declaration of Authorship

We, Omar CHERIF, Dalila LADLI and Maxime MÉLOUX, declare that this report titled, "Tracking Theoretical Shifts over Time: A Natural Language Processing Analysis of Cass R. Sunstein" and the work presented in it are our own. We confirm that:

- This work was done wholly or mainly while in candidature for a degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where we have consulted the published work of others, this is always clearly attributed.
- Where we have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely our own work.
- We have acknowledged all main sources of help.
- Where the thesis is based on work done by ourselves jointly with others, we have made clear exactly what was done by others and what we have contributed myself.

Signed: Omar CHERIF, Dalila LADLI and Maxime MÉLOUX

Date: June 18, 2022

UNIVERSITÉ DE LORRAINE

Abstract

Collégium Lorraine Management Innovation

Institut du Digital, Management et Cognition

Master 1 in Natural Language Processing

Tracking Theoretical Shifts over Time:
A Natural Language Processing Analysis of Cass R. Sunstein

by Omar CHERIF
Dalila LADLI
Maxime MÉLOUX

In this report, we compare multiple methods such as topic modelling, dynamic topic modelling, statistical analysis and semantic networks for the task of analyzing ideological evolution of an author on based on a given corpus. We then build, clean and process a corpus containing the publications of the American legal scholar Cass R. Sunstein between 1987 and 2005. We apply the aforementioned methods as well as custom-made ones on this corpus, and draw conclusions on how the author's approaches to economics and political theory have shifted over time.

Acknowledgements

We would like to thank our supervisor, prof. Samuel Ferey, for his valuable advice, time and expertise, as well as the head of the master program, prof. Maxime Amblard, for his help and support throughout the year.

Contents

Introduction

Preamble

In the history of the law and economics movement, the introduction of behavioral considerations into law and economics has been a major turning point (Ferey, 2009; Sheffrin, 2017). While the Posnerian and Beckerian paradigms dominated the law and economics movement until the mid-1980s, the Chicago School was criticized on its very foundations: the assumption of economic rationality applied to legal behavior. Cass Sunstein played a major role in these critiques of standard legal economics (Sheffrin, 2017). Born in 1954, a lawyer and political theorist like Posner and specialized in constitutional law, Cass Sunstein is widely identified with the behavioral economics of law movement. At the end of the 1990s, he published with Christine Jolls and Richard Thaler an important book entirely dedicated to behavioral approaches (Sunstein, 2000) and then went on to develop the famous model of libertarian paternalism. More broadly, he had a significant impact on President Obama's administration in developing the application of these behavioral policies[1].

But how did Sunstein shift from political theory and public law to behavioral economics? The main topic of this dissertation is to better understand the Sunstein changeover from Republicanism, inspired by a new thinking about political freedom, to libertarian paternalism, mainly inspired by cognitive psychology.

Working hypotheses

The main hypothesis endorsed by our supervisor, prof. Samuel Ferey, is that Sunstein expressed in the 1980s his commitment to participate in the renewal of republican[2] thought in political theory. As such, the behavioral turn in law and economics should not be understood from an economic theory standpoint alone, but needs to be embedded in a broader picture involving the evolution of legal theory and political theory. More precisely, by developing his conception of a republican theory of freedom and autonomy, Sunstein intended to oppose the traditional liberal model represented by the Chicago School politics. The discovery of experimental and behavioral economics - and in particular of the research program carried out by Kahneman and Tversky - allowed him to firmly establish a large part of his intuitions concerning the distinction to be made between private agents pursuing their own interests and citizens participating in public life. Behavioral economics is an answer to Sunstein questions about politics. But this theoretical and empirical deepening has an unexpected consequence: by focusing attention on individual cognitive problems, the most critical themes of the republican thought are greatly diminished. The model of liberal paternalism is, from this point of view, emblematic of the path taken: Sunstein's republican political ideas are still there but under the sign of a syncretic model with very little political stakes.

[1] "(a) Executive departments and agencies (agencies) are encouraged to: (i) identify policies, programs, and operations where applying behavioral science insights may yield substantial improvements in public welfare, program outcomes, and program cost effectiveness; (ii) develop strategies for applying behavioral science insights to programs and, where possible, rigorously test and evaluate the impact of these insights; (iii) recruit behavioral science experts to join the Federal Government as necessary to achieve the goals of this directive; and (iv) strengthen agency relationships with the research community to better use empirical findings from the behavioral sciences." (Office of the Federal Register, National Archives and Records Administration, 2016)

[2] Republicanism has to be understood in its philosophical meaning (the Republican tradition from Machiavelli to Pocock and Pettit, "Republican tradition") and not in its political meaning ("Republicans vs Democrats", "Donald Trump is a Republican").

However, this hypothesis lacks evidence; the main idea of the dissertation is to provide automatic tools and results in order to confirm of deny the "hermeneutic" hypothesis expressed by our supervisor. To carry out the work, we identified, in accordance with our supervisor, different interesting sub-questions as:

- How to consider the relevant corpus?
- Is there a changeover from Republicanism to Behavioral Economics? If so, how are the different topics dealt with by Sunstein over time related to this changeover? If such a changeover occurred, at what time?
- How are the main sub-topics of Republicanism (i.e. public deliberation, civic virtue, citizenship, autonomy, self-government, etc.) changed/modified through the use of behavioral law and economics? How to consider the place of sub-themes of economics and cognitive psychology over time?
- Are changes in the use of certain words and concepts (preference, autonomy, poverty) noticeable?

Methodology

To conduct this analysis, we will use various methods from Natural Language Processing (NLP), a field of computer science that focuses on language and its applications. In the recent years, the use of NLP in humanities has gathered significant momentum, especially for proving and developing hypotheses. This is evidenced by the growing number of conferences linking NLP to economics such as ECONLP (Hahn et al., 2018; Hahn et al., 2019) as well as the influx of literature using NLP methods for various applications in theoretical and applied economics (Gentzkow et al., 2019; Siegel, 2018).

For most NLP purposes, a corpus (plural *corpora*) is required to be able to infer or generate relevant information. Corpora can be extracted from a variety of sources, including online ones. They require cleaning and/or pre-processing in order to yield the best results.

Our work will therefore consist in three phases:

- Acquisition and pre-processing of the corpus. This phase is particularly important, as it will determine the accuracy and relevance of the analyses we will conduct.
- The first analysis phase will be conducted using mostly supervised and superficial statistical methods. They will serve as a way to evaluate the quality of our corpus and to get a first intuition of whether the working hypothesis is true.
- Finally, we will apply more complex and largely unsupervised methods, relying on semantics and modern NLP techniques, to the corpus. This more fine-grained analysis will allow us to validate or invalidate the working hypothesis, and can also be used to reveal novel information in the dataset.

Perspectives

The goal of our analysis is to ultimately build a way to process scientific texts and to be able to draw objective indicators from them, in addition to a more literary and classical interpretative approach. If done well, this program could be a valuable tool in analyzing the positioning and the ideological evolution of economists or political figures. One can also imagine applications in detection of contradictions or visualization of a given political landscape.

Chapter 1

Corpus acquisition and processing

1.1 Selection

Our hypothesis suggests that the shift in Sunstein's conception of political theory, if present, occurred around 1995. After discussion with our supervisor, it was established that the working corpus would be made of all journal articles written and published by Sunstein between 1987 and 2005 (included).

1.2 Retrieval

Two main data sources were used to obtain the selected articles.

- JSTOR [1] is a digital library of journal articles pertaining to social sciences and humanities. Using the advanced search function of the website, 215 articles were identified and downloaded using the `selenium` Python library.
- Chicago Unbound [2] is an online repository containing, amongst other, research articles published the University of Chicago Press. Since Sunstein spent a significant part of his career working at the University of Chicago, we were able to identify and download 415 articles using the advanced search function of the website and `selenium`.

The resulting corpus, consisting of 620 PDF files, contained many duplicates. In particular, Chicago Unbound contained several versions of some articles, including working papers. A manual selection was conducted, with the objective to keep digital documents rather than scanned ones whenever possible to facilitate text extraction down the line. This came with the expense of selecting working versions of certain articles, which may slightly differ from the published articles themselves. After filtering, we obtained a list of 231 PDF articles.

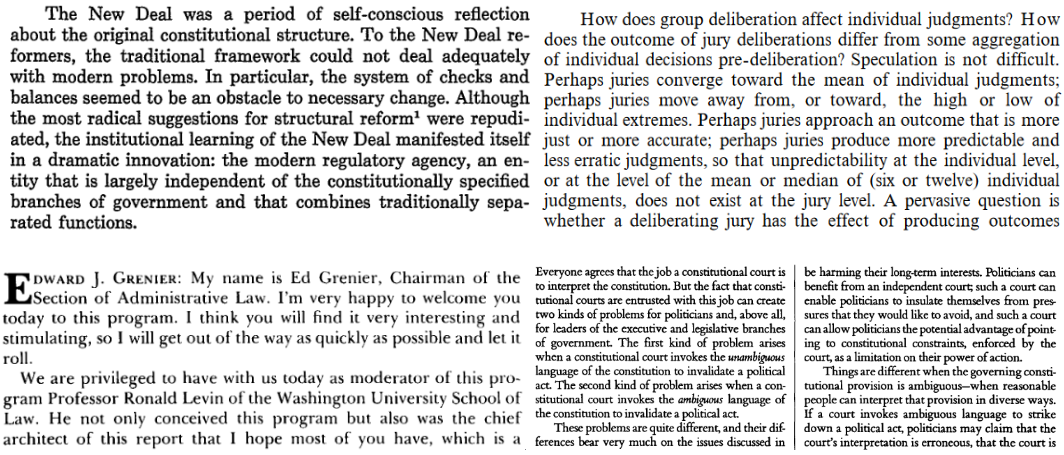
1.3 Text extraction

The resulting PDF files cannot be processed directly; text must first be extracted. In particular, we wished to extract titles and paragraphs, but also footnotes. However, footnotes are placed the bottom of pages, which makes text extraction likely to place them between pages, therefore breaking the continuity of the main text content. As a result, we decided to aggregate all the footnotes of each given article, and move them after the last page of the main text content.

Optical Character Recognition (OCR) is the most common way to extract text from scanned images. For computer-generated text, there is often a text layer in the PDF file that can be directly extracted or copy-pasted.

The main obstacle to OCR is that our corpus contains articles formatted with varying layouts, fonts and quality (see Fig. 1.1). As a result, multiple tools had to be used to obtain the

[1] http://www.jstor.org/
[2] https://chicagounbound.uchicago.edu/



The New Deal was a period of self-conscious reflection about the original constitutional structure. To the New Deal reformers, the traditional framework could not deal adequately with modern problems. In particular, the system of checks and balances seemed to be an obstacle to necessary change. Although the most radical suggestions for structural reform[1] were repudiated, the institutional learning of the New Deal manifested itself in a dramatic innovation: the modern regulatory agency, an entity that is largely independent of the constitutionally specified branches of government and that combines traditionally separated functions.

How does group deliberation affect individual judgments? How does the outcome of jury deliberations differ from some aggregation of individual decisions pre-deliberation? Speculation is not difficult. Perhaps juries converge toward the mean of individual judgments; perhaps juries move away from, or toward, the high or low of individual extremes. Perhaps juries approach an outcome that is more just or more accurate; perhaps juries produce more predictable and less erratic judgments, so that unpredictability at the individual level, or at the level of the mean or median of (six or twelve) individual judgments, does not exist at the jury level. A pervasive question is whether a deliberating jury has the effect of producing outcomes

EDWARD J. GRENIER: My name is Ed Grenier, Chairman of the Section of Administrative Law. I'm very happy to welcome you today to this program. I think you will find it very interesting and stimulating, so I will get out of the way as quickly as possible and let it roll.

We are privileged to have with us today as moderator of this program Professor Ronald Levin of the Washington University School of Law. He not only conceived this program but also was the chief architect of this report that I hope most of you have, which is a

Everyone agrees that the job a constitutional court is to interpret the constitution. But the fact that constitutional courts are entrusted with this job can create two kinds of problems for politicians and, above all, for leaders of the executive and legislative branches of government. The first kind of problem arises when a constitutional court invokes the *unambiguous* language of the constitution to invalidate a political act. The second kind of problem arises when a constitutional court invokes the *ambiguous* language of the constitution to invalidate a political act.

These problems are quite different, and their differences bear very much on the issues discussed in

be harming their long-term interests. Politicians can benefit from an independent court; such a court can enable politicians to insulate themselves from pressures that they would like to avoid, and such a court can allow politicians the potential advantage of pointing to constitutional constraints, enforced by the court, as a limitation on their power of action.

Things are different when the governing constitutional provision is ambiguous—when reasonable people can interpret that provision in diverse ways. If a court invokes ambiguous language to strike down a political act, politicians may claim that the court's interpretation is erroneous, that the court is

FIGURE 1.1: Some samples from our corpus. From left to right and top to bottom: High-quality scan (thick and clear font), computer-typed article, medium-poor-quality scan (thin font, initial capital), two-column scan.

best results, with some methods performing better on certain articles. We give below a short overview of those methods.

1.3.1 LAPDFText

LAPDFText (Ramakrishnan et al., 2012) is a Java tool made specifically for layout-aware extraction of scientific articles. It outputs an XML file, subdivided into pages, chunks (based on the font family and font size), and words. It is also able to output a text file, in which only chunks of the most common font size and family have been extracted, and the text put back together.

1.3.2 CERMine

CERMine (Tkaczyk et al., 2015) is an tool (available by default through a webservice) also aimed at performing layout-aware content extraction of scientific articles. It can output either a text file or an XML one, containing various sections corresponding to chapters and paragraphs in the text, as well as a specific section in which all footnotes have been put together.

1.3.3 Tesseract/OCRmyPDF

OCRmyPDF [3] is a Python package built on Google's Tesseract OCR library (Smith, 2007). It performs high-accuracy OCR on a PDF and can output the result in a text file, and is able to detect and extract multilingual text.

1.3.4 Selection

The pros and cons of all three methods are summarized in Table 1.1.

Each article was ran through all three tools independently, and a manual selection was performed to pick the best output. Text mode LAPDFText was selected for around 80% of the articles, due to its output usually being close to what was expected. Footnotes were then manually copied from the PDF and added at the end of the text file.

However, we found that LAPDFText was unable to deal with dashes and numbers for computer-generated articles. For these articles, the CERMine version was selected instead, but footnotes had to manually be removed from the middle of the text and moved to the end.

[3] `https://pypi.org/project/ocrmypdf/`

<table>
<tr><th>Tool</th><th>Pros</th><th>Cons</th></tr>
<tr><td>LAPDFText</td><td>• Very fast
• Very detailed results (in XML mode), from which footnotes can be extracted</td><td>• Hard to set-up initially
• Needs further processing for best results (in XML mode)
• Doesn't extract footnotes in text mode
• Cannot deal with some Unicode characters in computer-generated articles</td></tr>
<tr><td>CERMine</td><td>• Full Unicode support
• Able to parse metadata, sections and footnotes</td><td>• Slow, can only process one file at a time (webservice)
• Footnote detection can be inaccurate and result in them being left in the text</td></tr>
<tr><td>OCRmyPDF</td><td>• Extremely accurate
• Supports text in foreign language
• Extracts references</td><td>• Places footnotes in the middle of the text
• Does not perform de-hyphenation</td></tr>
</table>

TABLE 1.1: A comparison of the three main OCR tools used.

Finally, 4 articles yielded extremely poor results for both LAPDFText and CERMine, likely due to misalignment of the OCR layer onto the PDF file. This resulted into garbled output, in which words were not split properly and letter order was incorrect (see Fig. 1.2). In these cases, we resorted to using OCRmyPDF, and since it only affected a low number articles, we manually performed de-hyphenation on the resulting output. If more articles had been affected by this issue, we would have looked at any of the existing automated de-hyphenation methods, since manual de-hyphenation can be a very time-consuming task.

```
Neglect CASSR.SUNSTEIN csunstei@uchicago.edu
KarNl.LlewellDyinstS.ervicPer,ofo,fJurisprudLeanwcSechooDl,epartmoefnPtoliticSac
lience, UniversoiftCyhicagLoawSchoo1l,111East60thStreeCth,icagIoL,60637U, SA
Whenstronegmotioanrseinvolvepde,oplteendtofocuosnthebadnesosftheoutcomrea,thet
rhanonthe
probabiltihtyattheoutcomweilolccurT.heresulti"npgrobabilnietgylecht"elptsoexpl
aienxcessirveactions tolow-probabriilsitkoysfcatastropTheer.rorisshtoswa
workinkgnowledogfeprobabilniteyglecptr,oducing
publifcearthamtighgtreatelyxceetdhediscountheadrmA.sa
resulotfprobabilniteyglecpte,oploefteanre
farmorceoncernaebdouttheriskosfterroritshmanabousttatisticlaalrlygerrisktshat
thecyonfroinotrdinary
lifeI.ntheconteoxftterroriasnmdanalogoruisskts,helegaslystefmrequenrtelsypontd
osprobabilniteyglect,
resultiinngregulatitohnamtighbteunjustifoireedvencounterproduBcuttipvueb.lifc
earisitselafcosta,nd itisassociatwedithmanoythecrostsi,nthefor
mof"rippeleffectpsr"oducbedyfearA. sa normatmivaetter,
governmsehnotulrdeduceevenunjustiffieeadrif,thebenefoitftsheresponcsaenbeshow
tnooutweitghhecosts. JELClassificatioKnO:D,8 Terroristsshowa workingknowledgeof
threenoteworthpyointsaboutfear.Of these, thefirstwoare well-knownT.he
```

FIGURE 1.2: An example of poor OCR results (CERMine). LAPDFText produced a very similar output for this article.

1.4 Text processing

Finally, as mentioned earlier, the output of the OCR phase was not entirely accurate. Several issues were detected in the output files, and since our analysis partially relied on the presence of rather infrequent words in the corpus, cleaning was required before conducting any further analysis. Cleaning was performed in two stages.

1.4.1 Manual cleaning

A first stage of systematic, manual cleaning of the corpus, which took an estimated 40 to 60 man-hours, was performed in order to solve the following issues:

- Pages entirely missing from the output (usually 2-3 missing pages per article, but up to half of the pages for some articles)
- Page headers and footers (page numbers, article title) incorrectly added to the text
- Page and line breaks incorrectly resolved
- Output of highly structured data (tables, figures) added to the text, and often garbled

1.4.2 Automatic cleaning

The second cleaning phase was undertaken to solve the following issues:

- Footnote numbers appearing in the text, or incorrectly parsed as punctuation
- De-hyphenation incorrectly resolved, resulting in words being split
- Dashes removed from the text, resulting in words being concatenated
- Dashes and hyphens parsed as the same character
- Typographical errors or errors in character recognition
- Inconsistent spelling of some words (British/American spelling, optional hyphenation)
- Inconsistent capitalization resulting from OCR errors (e.g.: HowITzER)

In order to solve these issues, a large Python program, nicknamed the Monster, was created.

The Monster is an interactive program which conducts a sequential text analysis using a sliding window. The program starts with an initial wordlist, consisting of 58,000 American English words[4], two empty user-defined lists of substitution rules (a general one, and one for case), an automatically generated list of substitution rules, and an empty list of allowed case shapes.

The program runs through each file sequentially, detecting words as non-empty character sequences between whitespace. Each word is then stripped from punctuation and compared against the wordlist. If a word is not present in the wordlist or has inconsistent case (i.e. both capital and non-capital letters, with at least 2 of each), user action is required. Several options are possible, detailed in Tables A.1 and A.2.

Additionally, the program detects the most common types of character OCR errors based on user-defined substitution rules. Those rules are accounted for when computing the Levenshtein distance between words; common character substitutions are given a value of 0.3.

Finally, a few actions are performed automatically, notably the replacement of certain characters with some others: ellipsis to three periods, curly quotes to straight ones, etc. Some unknown words are also automatically added to the wordlist when detected, such as words of the form well-[adjective], anti-[noun] or [adverb]-[adjective] (in words such as *constitutionally-based*, *governmentally-funded*...)

Statistics

The automated cleaning took an estimated 15 to 20 man-hours. Table 1.2 sums up the data gathered by the Monster.

Type	Number of items learned
Unknown words	11,648
Word substitution rules	2,078
Unknown word cases	797
Case substitution rules	224

TABLE 1.2: Statistics obtained using the Monster.

[4]The full list can be found at http://www.mieliestronk.com/wordlist.html.

1.5 Resulting corpus

The final corpus contains 231 files in text format, summing up to over 2.7 million words and 17.4 million characters. Each file is associated with its publication year. It is worth noting that article publication may take several months or years, depending on the circumstances. Therefore, the date stored for each article might not always reflect the true chronology of Sunstein's thought. Furthermore, a lot of the analyses conducted in our study rely on dividing the data into multiple time splits. When this was the case, we chose to work with splits spanning one year, as this is the maximal granularity that could be reached in our corpus, and grouping multiple years together resulted in a low number (10 or fewer) of splits, which was not sufficient for our methods to yield relevant results.

Table 1.3 sums up general statistics about the resulting files.

Year	Paragraphs	Words	Characters
1987	1,595	98,544	627,995
1988	1,174	71,361	458,721
1989	1,278	75,197	483,221
1990	1,446	86,009	552,713
1991	1,584	93,466	602,503
1992	2,169	123,170	768,690
1993	1,926	126,923	788,603
1994	2,991	189,087	1,175,592
1995	3,702	186,976	1,171,579
1996	3,864	188,158	1,184,515
1997	1,795	93,778	588,914
1998	2,884	155,237	976,722
1999	5,878	316,009	1,986,588
2000	3,505	188,047	1,172,732
2001	1,618	151,953	946,594
2002	1,594	184,765	1,182,955
2003	939	105,085	667,539
2004	1,695	188,386	1,195,351
2005	1,297	141,532	946,278
Total	42,934	2,763,683	17,477,805

TABLE 1.3: Statistics on the Sunstein corpus

1.5.1 Evaluation

To measure the effect of the manual and automated cleaning steps, we can resort to measuring corpus perplexity. Perplexity is a metric used in information theory to measure how well a given model or probability distribution can predict a sample (Nicolov et al., 2009). When used on a corpus, it is a measure of how "likely" each sentence is. We expect that cleaning the corpus should lower the overall perplexity.

We compute perplexity using OpenAI's GPT-2 (Radford et al., 2019). Specifically, we choose the model `gpt2-large` and compute, for each article, the average sentence perplexity. Table 1.4 sums up the minimum, maximum, average and standard deviation of the perplexity obtained over all articles.

Corpus	Minimum	Maximum	Average	Std. deviation
CERMine	11.618	269.665	22.057	20.778
LAPDFText	11.018	39.360	18.418	3.315
Manually cleaned	10.995	23.275	17.085	2.268
Manually + automatically cleaned	10.495	22.130	16.705	2.151

TABLE 1.4: Statistics about the average sentence perplexity computed over all articles.

As expected, we observe that cleaning significantly lower the corpus perplexity. We first note that the CERMine corpus has very high average and maximum perplexity, which we attribute to the fact that some articles contain garbled output. LAPDFText produces better results overall for all metrics. Manual cleaning significantly brings down the metrics, which can be interpreted as a large improvement in overall corpus quality. Finally, automatic cleaning brings further improvements to the corpus quality, although the effects are less noticeable.

Furthermore, the effects of cleaning the corpus can be seen in the improvement of the results yielded by the analysis methods themselves. This has been observed throughout the research process carried out on the corpus, with our results gaining in coherence and relevance as the corpus improved in quality.

Chapter 2

NLP for validating hypotheses

This chapter focuses on approaches that were initially developed as the corpus was being cleaned. They rely on statistical and largely supervised approaches.

An important note is that we have little knowledge in the fields of economics and political theory. Therefore, for most of the analyses conducted in this chapter as well as the next one, the results were checked and validated by prof. Ferey using his expert knowledge. This has been our main evaluation method throughout this project.

2.1 TF-IDF

Our first approach was to measure the frequency of certain ideologically-heavy words over time. We initially attempted to detect the most important words in an unsupervised way, through a statistical method called Term Frequency–Inverse Document Frequency (TF-IDF).

2.1.1 Presentation

TF-IDF (Rajaraman and Ullman, 2011) is a statistical analysis technique for extracting the most relevant terms in a document. It first computes the frequency of each term (number of occurrences divided by the number of words in the document). This number is then multiplied by the inverse document frequency (a measure of how common each word is in the overall corpus). This allows for so-called stop words, i.e. words with no semantic content (such as *the* and *is*) as well as words that are frequent in all documents to be ignored.

We concatenated the articles of each year into a single document per year, computed the TF-IDF for all words in the corpus, and extracted the words with the highest score in each time slice. A sample of the results is given in table 2.1.

2.1.2 Results and analysis

Year	1987	1994	2001
Top terms	Deal	executive	academics
	Lochner	Framers	academic
	agency	President	cascade
	See	valuation	informational
	Chevron	unit	professors
	agencies	GDP	fields
	administrative	departments	signals
	checks	Congress	faculty
	inaction	supra	books
	guidelines	PACs	clerks

TABLE 2.1: The 10 words with the highest TF-IDF in the Sunstein corpus, for three selected years.

After checking the top terms for each year with prof. Ferey, we found that they did not produce results that could be used to either prove or disprove our working hypothesis. We therefore switched to a supervised approach, in which prof. Ferey gave us a list of words central to the working hypothesis. We then computed the TF-IDF of each of those words for each year, and normalized the scores, which are given in figure 2.1.

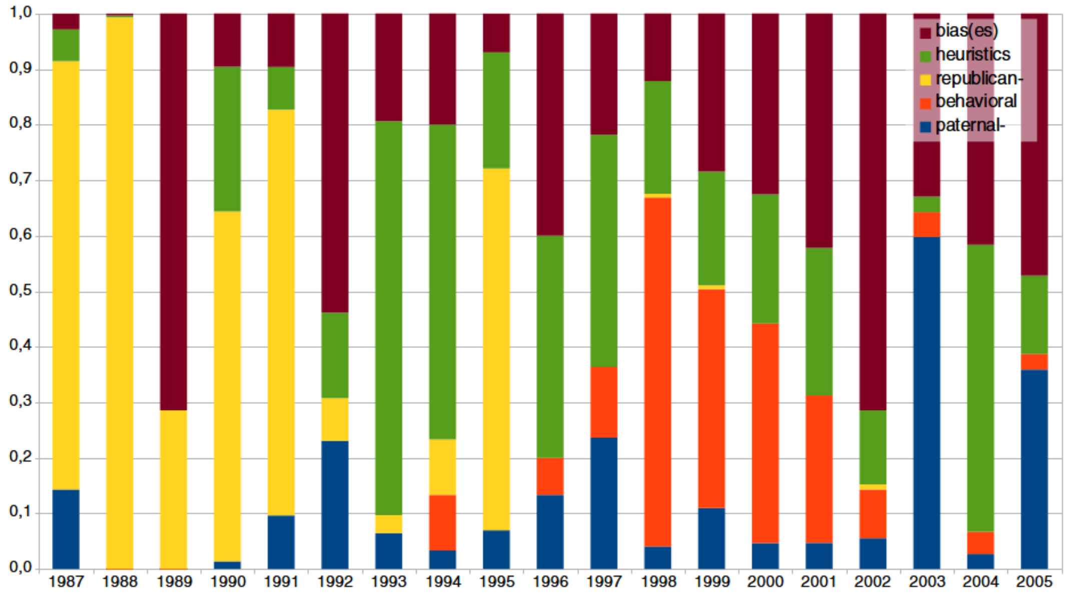


FIGURE 2.1: TF-IDF of selected words over time, normalized.

Several observations support the working hypothesis. In particular, we can clearly see that the word *republican*, which is very present at first, gradually disappears. In contrast, the words *bias* and *heuristics* are present in the early years of the corpus, but the frequency of *heuristics*, a specific term from behavioral economics, increases gradually, while *bias*, a relatively general term, remains constant. On the other hand, *behavioral* first appears in 1994 and clearly peaks in 1998, right as *republican* all but disappears from the corpus. Finally, *paternal-* (paternalism, paternalist) progressively replaces *behavioral*. We also verified that the appearances of *paternal-* before 1992 in the corpus refer to the general meaning of the term, which is not linked to libertarian paternalism.

However, prof. Ferey was more interested in an input-free analysis of the corpus, to avoid possible confirmation biases. We therefore had to find a different approach.

2.2 Transition modeling

A first, simple way to model the transition in Sunstein's thought would be to use linear regression on the frequency of terms. We could then try to infer an ideological shift from the words whose frequency have increased or decreased in frequency.

There are two problems with this approach. The first one is that only two parameters control the output of a linear regression: the slope and the intercept. We believe that this is too simple of a model to accurately model a transition. Furthermore, we believe that there is little reason why word frequencies should constantly increase or decrease throughout the corpus.

This led us to come up with a slightly more complex model for transitions, which we introduce in this section. This model relies on a specific mathematical tool: the sigmoid function.

2.2.1 The sigmoid function

The sigmoid function, whose graph is shown in figure 2.2, is defined as follows:

$$h(x) = \frac{1}{1+e^{-x}}$$

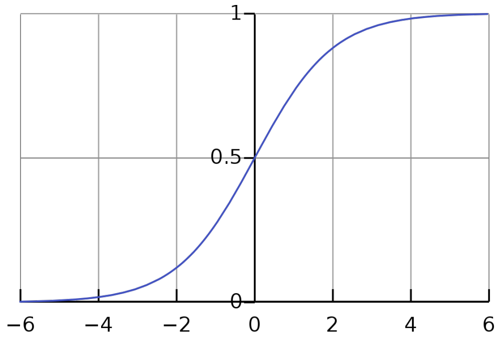


FIGURE 2.2: The sigmoid function.

The choice of the sigmoid comes from our hypothesis that an ideological shift can be represented by initially having a constant frequency for a given word, and then transitioning to a new frequency, which may be higher or lower. This new frequency then becomes constant, modeling the end of the transition. This represents the intuition that when first discovering a new idea or concept, an author will start using a specific set of words more and more. The sigmoid function will hence allow us to detect the appearance and disappearance of words in the vocabulary of an author.

More specifically, we can control several parameters of the sigmoid curve using the following version:

$$f(L, k, x, x_0, b) = \frac{L}{1+e^{-k(x-x_0)}} + b$$

The parameters are the following:

- L is the amplitude of the transition
- k is the speed of the transition
- x_0 is the center point (date) of the transition
- b is the initial frequency

We would like to point out that this function cannot approximate temporary transitions, such as the one for of the word *behavioral* in figure 2.1. For that, a different model would be required, such as a Gaussian distribution. However, the sigmoid can approximate constant frequencies (using $L = 0$) as well as linear transitions (using a low value of k).

2.2.2 Methodology

For every word in the vocabulary of the corpus, we compute the relative frequencies of the word in each year split. We then deduce the optimal parameters L, k, x_0 and b for that word using the `curve_fit` function from the `scipy`[1] library. We then compute the r^2 coefficient of determination, to see how well the curve approximates the word's frequency over time. We restrict ourselves to words having $r^2 > 0.4$, which yields a list of 560 words. Prof. Ferey then extracted a selection of 27 relevant words from that list. Those words were grouped into 4 categories (behavioral economics, general economics, republican thought and miscellaneous).

Our working hypothesis would predict the following:

[1] https://scipy.org/

- Words in the category of behavioral economics should increase in frequency, with a transition around 1995.
- Words from the economics category should also become more frequent. While Cass R. Sunstein mainly researches political theory, his interest in behavioral economics should lead him to talk more and more about economics in general. This transition should also occur around 1995.
- Words linked with the republican thought (republicanism) should become less and less frequent over time. This transition might occur around 1995 or later, as Sunstein found his previous beliefs to be incompatible with his new interest in behavioral economics.
- Words in the miscellaneous category could exhibit different levels of change. Amongst others, words linked with psychology and cognitive sciences should become more frequent.

2.2.3 Results and analysis

The obtained graphs are displayed in figure 2.3.

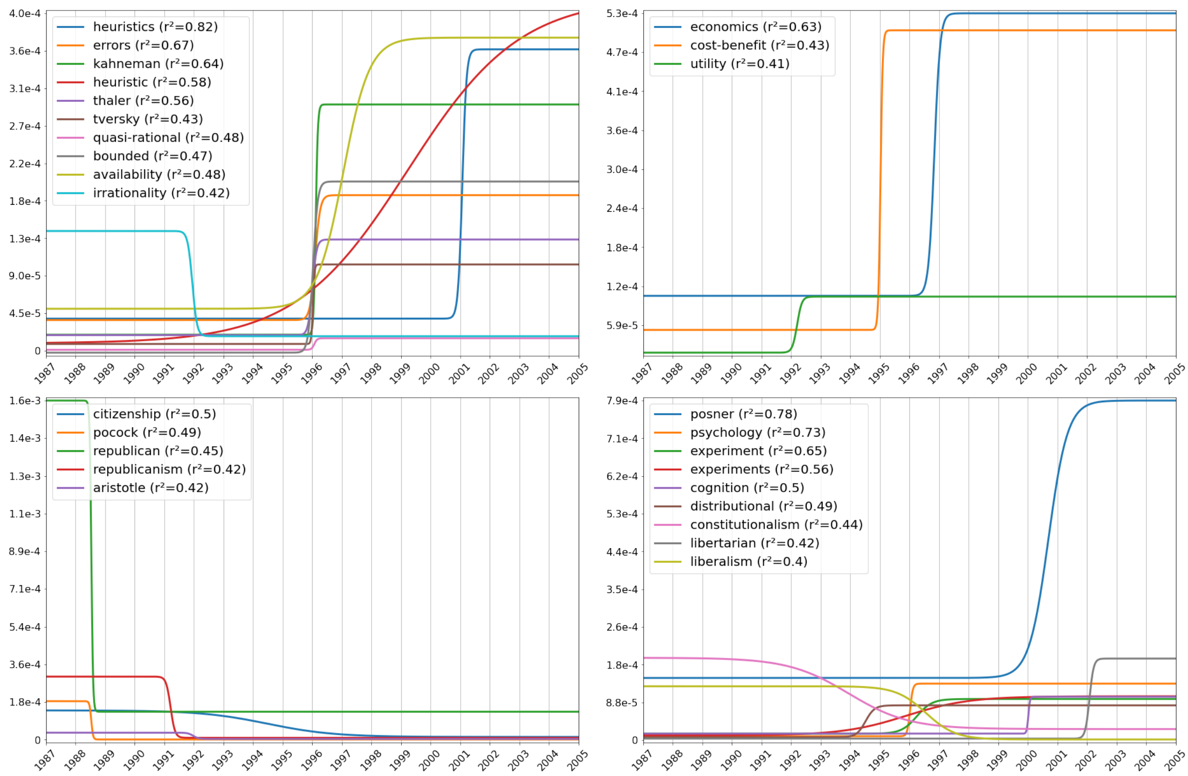


FIGURE 2.3: The fitted sigmoid curves for the four groups of words (from left to right and top to bottom: behavioral economics, general economics, republican thought and miscellaneous)

We observe that the results largely support our hypothesis. For behavioral economics, most of the words undergo a sharp transition in 1996. The words *heuristic* and *heuristics* experience respectively a slower and delayed transition, while *irrationality* shows a marked and early decline in frequency. We attribute the latter to the fact that *irrationality* is a generic word that can be used outside of the realm of behavioral economics. In the general economics group, all words experience a frequency increase between 1992 and 1995, which matches our initial expectations. Similarly, words from the republicanism group all but disappear; however, this transition occurs earlier than anticipated. Finally, the miscellaneous group is harder to interpret, but we notice a sharp rise in the frequency of words associated with psychology and cognitive sciences.

Despite those seemingly good results, the question of the accuracy of the model remains. While we have only selected curves for which $r^2 > 0.4$, this threshold might be insufficient to capture true transitions. A comparison of observed versus fitted data for two selected words is given in figure 2.4. It shows that the fitted curve is generally accurate when r^2 is low, and that the predicted transition date matches what is visually seen. When r^2 is lower, however, the overall shape of the transition is acceptable, but its date is more questionable. These limitations are therefore to be taken into account when interpreting the pertinence of our results.

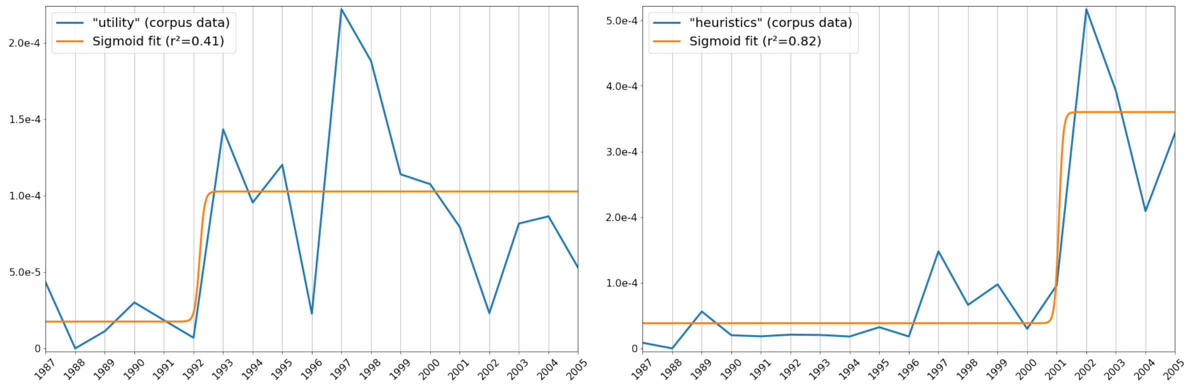


FIGURE 2.4: A comparison of observed versus fitted data for a low (left) and a high (right) value of r^2.

2.3 Semantic analysis with Tropes

2.3.1 Presentation

Tropes (Molette and Landré, 1994) is a semantic analysis software that performs text mining on a corpus. It features semantic classification as well as thesaurus generation based on Anglo-Saxon linguistics, particularly linked to John Lyons' work. To categorize data, it relies on morphosyntactic analysis, semantic networks and lexicons such as political dictionaries with a predefined semantic classification. The results are presented in the form of statistical reports or hypertext graphical representations. Among all the functions and tools that Tropes offers, we used the Scenario, Chronological Analysis and Actor Analysis tools.

2.3.2 Methodology

We chose Tropes, among all other semantic analysis software, because it is one of the only tools that allows the user to create their own dictionary (using the Scenario feature), unlike other software such as Iramuteq[2], which proposes only one method of classification.

We then had to choose between using an already existing classification or creating our own hierarchical classification. For the sake of precision and completeness, we created our own Scenario. This led us to a semi-supervised method of application, where Tropes automatically used the dictionary we implemented on the corpus.

A Scenario is made of a certain amount of semantic groups, namely groupings of words and/or equivalent classes, which can be nested up to 9 levels of depth.

Our Scenario was implemented using the words found on our corpus, with no prior dictionary, to avoid any external influence on the outcome. Afterwards, we filtered all obtained classes and kept only the ones likely to confirm or disprove our hypothesis. Some words were missing from the list of used references or proposed expressions by the terminology extractor. We completed our Scenario with a relevant subset of those words.

[2] http://www.iramuteq.org/

We then ran our created Scenario on our corpus of articles from 1987 to 2005. From these, Tropes extracted and returned the relationships between terms and their number of occurrences.

To visualize our results, we used two types of graphs:

- Graph – Star: This type of graph displays the relationship between a central word and its contextually surrounding ones, based on co-occurrence frequencies. Those words are displayed on the left, when they occur before the central element, and on the right, when they occur after the central element.
- Graph – Distribution: This type of graph shows a chronological histogram of a word's distribution. It is formed by splitting the text into fixed-length strings. The graph contains the frequency of occurrence of the subject within each string. For each string, a bar is displayed in a chronological order. The numbers represent the occurrences of the selected word within the string.

Although there are additional types of visualisations which we could extract from Tropes, we only focused on these two types of graphs due to the three-dimensional aspect and animated features encountered in other visualisations, which are difficult to include in a report such as this one.

2.3.3 Star graphs

To bring out the chronological aspect in our analysis, we decided to split the corpus into two sections: before and after 1994. This particular date was chosen based on prof. Ferey's expertise, and as it effectively splits the corpus in halves of near-equal sizes. We then compared the different relationships that several keywords, also given by prof. Ferey, have with each other.

Results

The resulting graphs are given in figures 2.5 to 2.9.

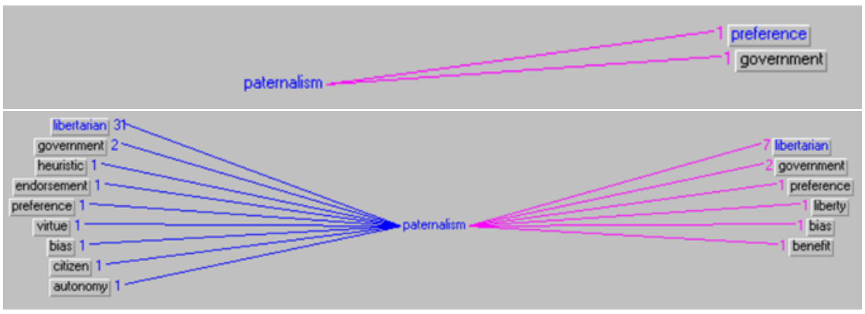


FIGURE 2.5: Star graph of the word "paternalism" before and after 1994.

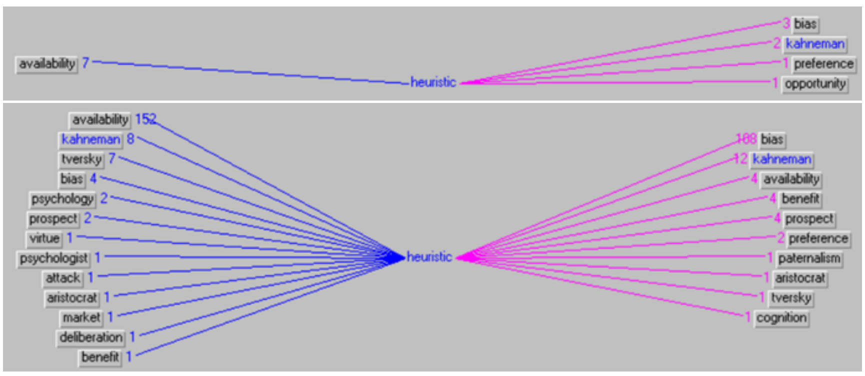


FIGURE 2.6: Star graph of the word "heuristic" before and after 1994.

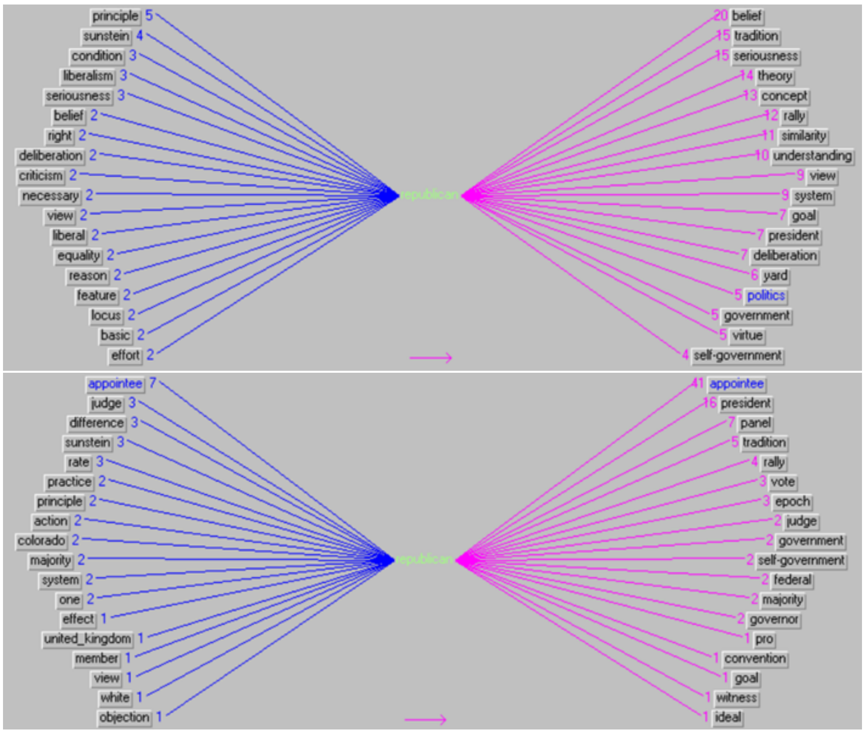


FIGURE 2.7: Star graph of the word "republican" before and after 1994.

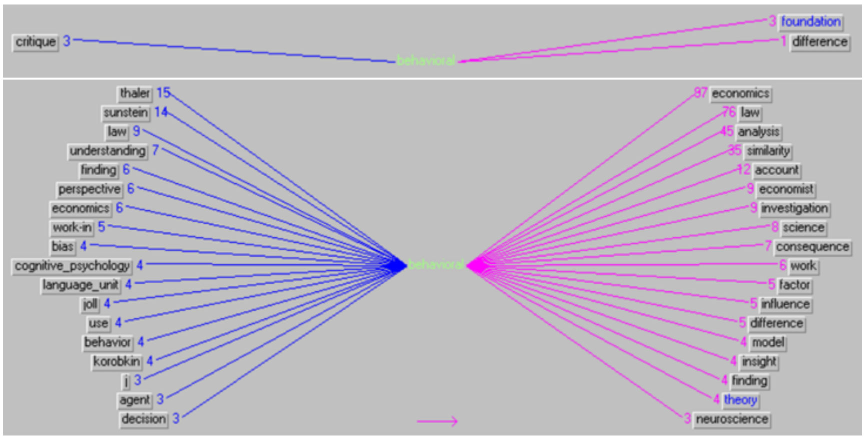


FIGURE 2.8: Star graph of the word "behavioral" before and after 1994.

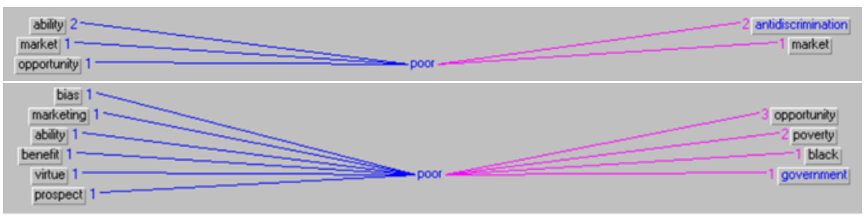


FIGURE 2.9: Star graph of the word "poor" before and after 1994.

Analysis

Figure 2.5 shows that while there was a sparse mention of *paternalism* in the corpus before 1994, its usage was very limited and not related to *libertarian*. After 1994, however, the term is linked to over a dozen new others. In particular, *libertarian* is the most commonly encountered term near it. We also see the appearance of a number of words from psychology and cognitive sciences, such as *bias* and *heuristic*.

Figure 2.6 displays the evolution of *heuristic*, a term describing the process through which humans reach decisions, often through the use of mental shortcuts. Before 1994, the word is linked to *availability* to form *availability heuristic*, a type of heuristics in which decisions are made based on immediately available information, such as memory. We also see a references to *bias* and *Kahneman*[3], but they are not many. After 1994, *prospect*, *Tversky* and *Kahneman* are directly linked to *heuristic*, and we see that *bias* is one of the most commonly encountered words in the context of *heuristic*, indicating the now preponderant place of psychology in Sunstein's thought. Furthermore, we now have references to *paternalism*, but also *market* and *deliberation*, indicating how behavioral economics are now influencing Sunstein's conception of both economics and politics.

Figure 2.7 shows that the term *republican* is associated with many concepts, both before and after 1994. One notable change is the disappearance of the term *virtue* from its context, a term referring to civic virtue, one of the key points treated in the republican tradition. This may suggest that Sunstein no longer agrees with the views of republicanism on the matter.

Figure 2.8 shows that *behavioral* undergoes an evolution similar to that of *heuristic*. Notably, we see the appearance of many terms coming from cognitive sciences and psychology, as well as terms such as *law*, *economics* and *work*, once more suggesting a spillover from behavioral science into other disciplines such as political theory.

Finally, 2.9 suggests that before 1994, Sunstein's conception of poverty is that it is related to a lack of *opportunities*, or an issue inherent to the *market*. After 1994, although those terms are still present, we also see the rise of *prospect* and *bias*, indicating that Sunstein's view of poverty is now influenced by prospect theory.

2.3.4 Tropes with Gephi

To get a more complete picture of the contextual data processed by Tropes rather than viewing one word at a time, we exported two star graphs (one for 1987-1993 and one for 1994-2005) computed using the scenario from Tropes. These graphs are essentially chained versions of the ones displayed in section 2.3.3, with vertices representing words and edges encoding the number of occurrences between two words. We generated visualizations and useful statistics from these graphs using Gephi[4].

We first visualized each graph using the Force Atlas 2 layout algorithm (Jacomy et al., 2014). This specific algorithm organizes the graph spatially by simulating attraction and repulsion forces between vertices. This implies that highly-connected vertices tend to become attractors and to stay close to the center of the graph. In our case, this method allows us to compare the semantic relationships between words before and after Sunstein's hypothesized shift.

We also used Gephi's built-in statistics module to compute the weighted degree of each node, which is a simple indicator of how connected a node is, as well as to subdivide the graph into clusters (modularity).

Results

We highlight a few examples of notable subgraphs in figures 2.10 to 2.13. Additional subgraphs can be found in figures C.9 to C.3. The full graphs, color-coded by clusters, can be found in figures C.10 (before 1994), C.11 (after 1994) and C.12 (full corpus).

[3]Daniel Kahneman and Amos Tversky are two psychologists and economists, which are mainly known for developing *prospect theory*, a theory of behavioral economics and finance using key findings from psychology, such as loss aversion (Kahneman and Tversky, 1979).

[4]`https://gephi.org/`

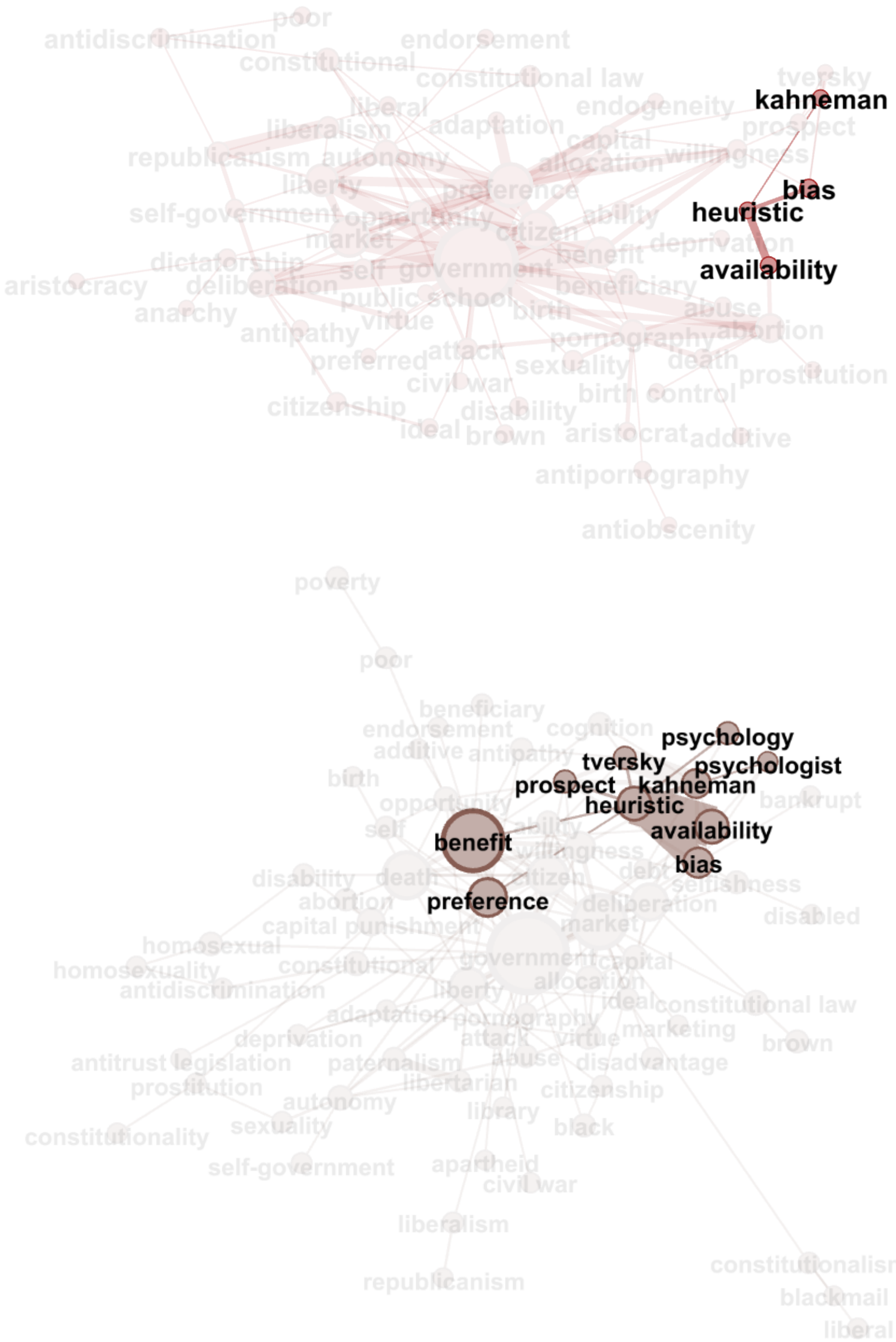


FIGURE 2.10: Subgraph of the word "hHeuristic" before and after 1994, visualized with Gephi.

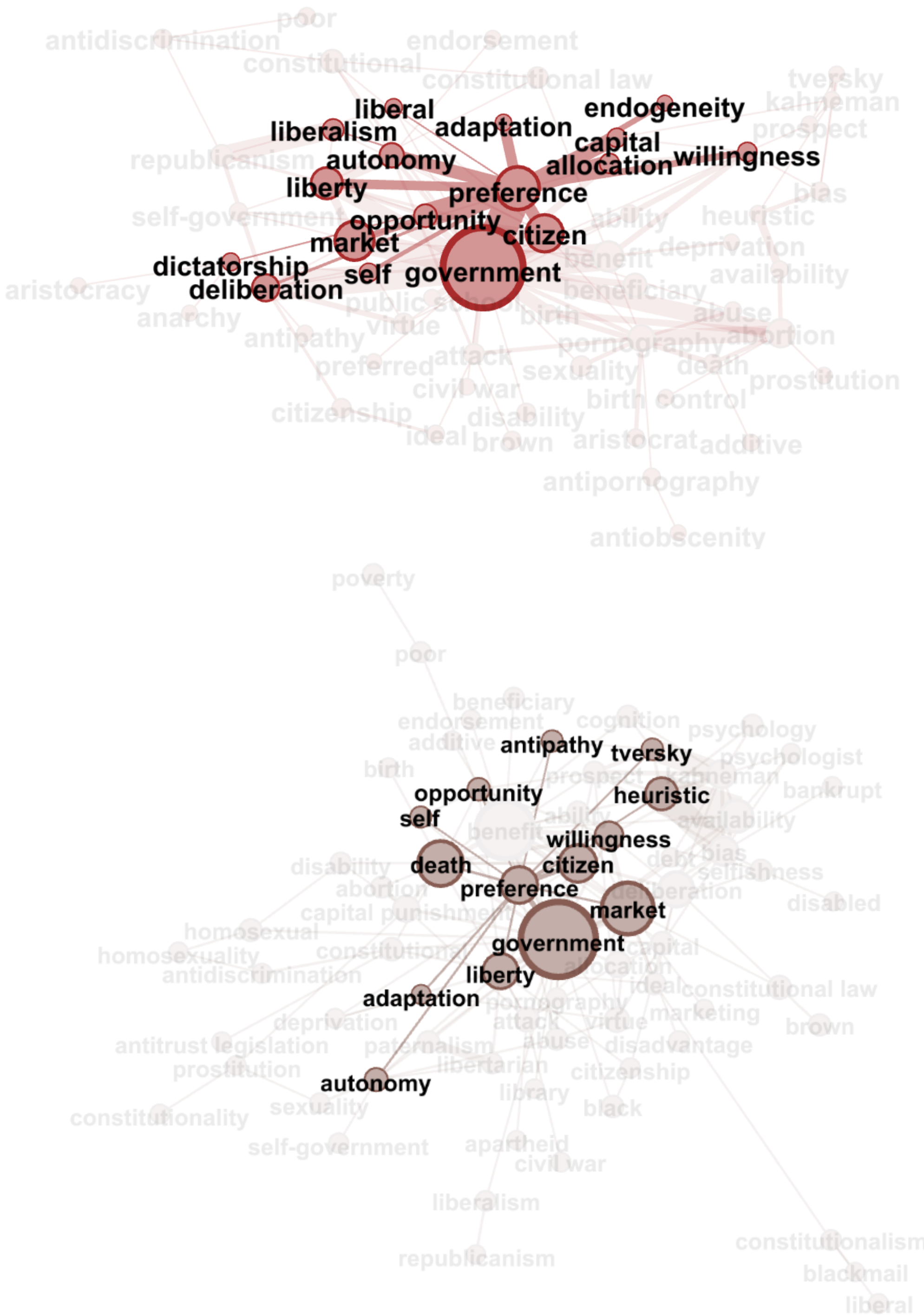


FIGURE 2.11: Subgraph of the word "preferences" before and after 1994, visualized with Gephi.

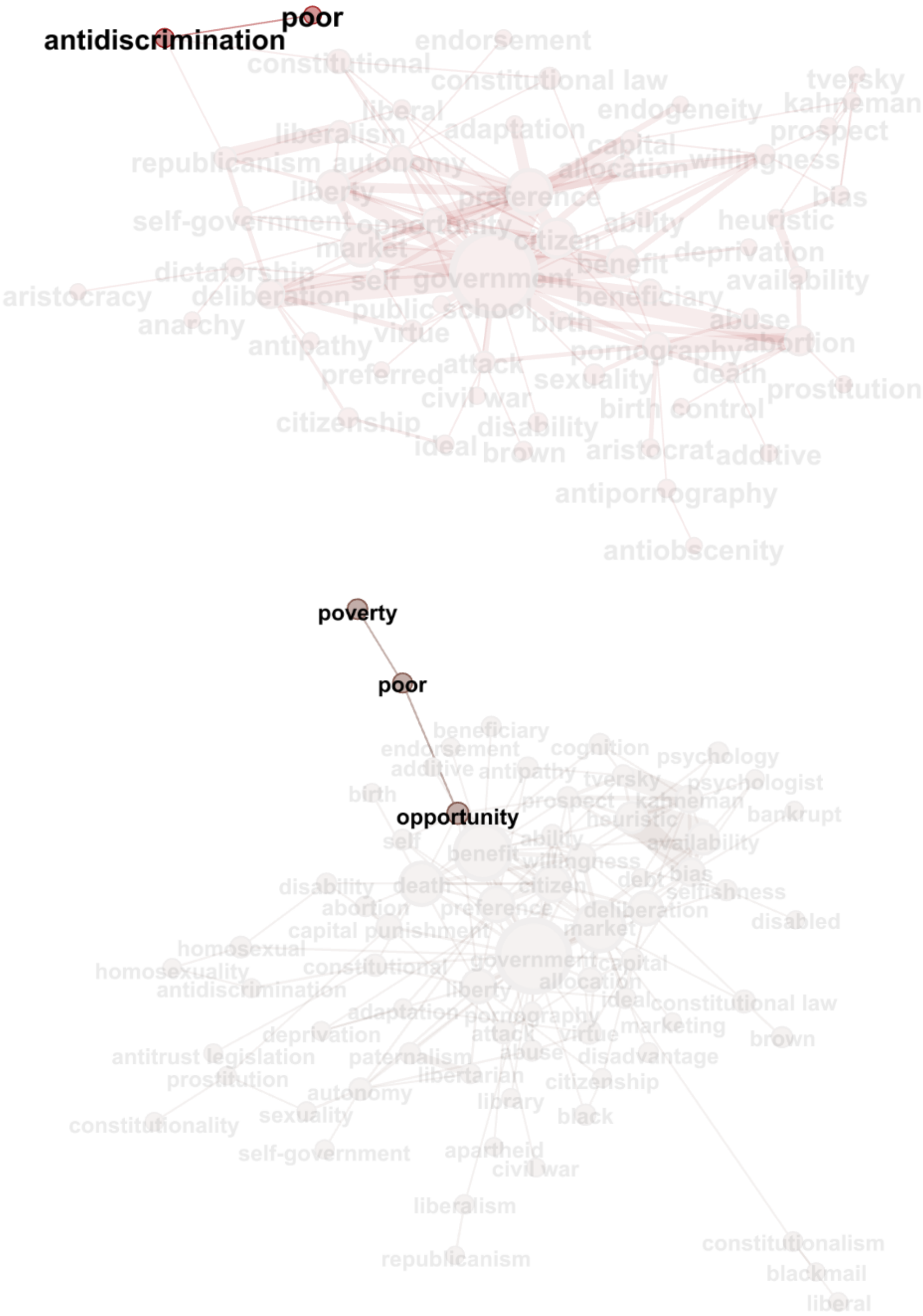


FIGURE 2.12: Subgraph of the word "poor" before and after 1994, visualized with Gephi.

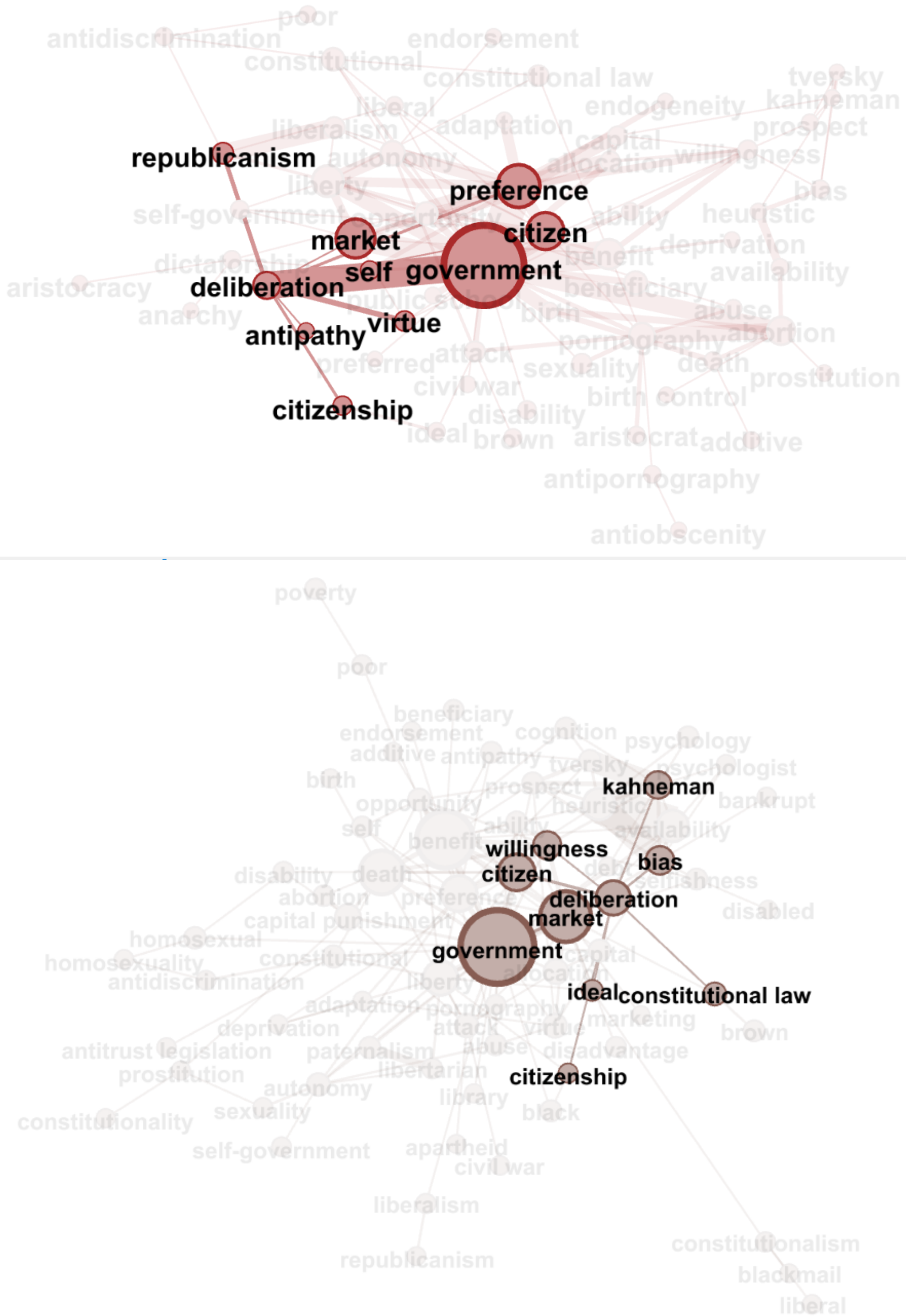


FIGURE 2.13: Subgraph of the word "deliberation" before and after 1994, visualized with Gephi.

1987-1993		1994-2005	
Word	Weighted degree	Word	Weighted degree
government	215	heuristic	393
preference	140	government	268
liberty	58	availability	247
citizen	53	bias	185
market	47	kahneman	182
pornography	47	market	155
abortion	46	tversky	177
deliberation	46	benefit	112
benefit	37	citizen	90
opportunity	36	willingness	70

TABLE 2.2: The 10 words with the highest weighted degree in both graphs.

Analysis

Figure 2.10 shows that before 1994, the words *heuristic*, *bias*, *Kahneman* and *Tversky* already form a notable subgraph, but are relatively disconnected from the rest of the graph. After 1994, this subgraph is much more noticeably into the main graphs, and forms connections with some of the main topics. This suggests that behavioral aspects were already a part of Sunstein's thought before 1994, but that it is only later on that he fully integrated their implications into his work.

In figure 2.11, we see that before 1994, *Preferences* is linked to words such as *autonomy*, *capital*, *allocation*, *government* and *market*, which is typical of the interpretation of preferences in republicanism. After 1994, *allocation* disappears and the link with *government* is significantly weaker. Whatsmore, *Tversky* and *heuristic* are now part of the subgraph. We believe this represents a shift in the way Sunstein approaches the question of preferences, only considering it from a cognitive and psychological perspective after 1994.

Figure 2.12 depicts the evolution of *poor* over time. It suggests a link between the question poverty and republicanism: following the link from *poor* to *antidiscrimination* then leads to *republicanism*. After 1994, however, the question of poverty becomes linked to *opportunity* and *market* (indirectly), which reflects a more economical and behavioral approach.

Finally, the evolution of the word *deliberation* in figure 2.13 also tends to confirm Sunstein's theoretical shift. Before 1994, it is physically located in a republican and political context, being linked to *government*, *constitutional* and *citizen*, while we later see the rise of elements of behavioral economics through the appearance of links with *Kahneman* and *Tversky*. We also observe that *deliberation* seems to act as a pivot points between the two psychologists and central political concepts.

From a quantitative standpoint, we observe a significant shift in the list of words with the highest weighted degree: While before 1994, words mostly refer to concepts of political theory (*government*, *liberty*, *citizen*, *deliberation*) as well as social questions (*pornography*, *abortion*), the second half of the corpus is dominated by terms from prospect theory (*heuristic*, *availability*, *bias*, *kahneman*, *tversky*, *willingness* [*to pay*]). This is a strong indicator of how behavioral economics has pervaded all aspects of Sunstein's reflection in the second half of the studied period.

To add on to this last point, we bring forward the idea that a theme can be important not because it is associated with many other themes (weighted degree), but because it pivots around distinct semantic universes, being an inescapable intermediary in the progression and argumentative coherence of the author's thoughts. In our case, this is illustrated by the fact that after 1994, *heuristic* is more central than *government* due to the high diversity of the subjects linked to it.

Chapter 3

NLP tools as a source of new hypotheses

In this chapter, we apply unsupervised methods to the corpus, with the goal of extracting information from the dataset while requiring no user input, save for validation of the results. In order to do that, we rely on the field of Topic Modeling.

Topic modeling, first introduced in Papadimitriou et al. (1998), is a branch of machine learning, the main focus of which is the extraction of the main themes and subjects from a text. Those themes are grouped into so-called topics, that is, list of words, which associated weights representing how strongly a word is associated with a topic.

3.1 Supervised approaches

Our initial perusing of the literature showed two promising topic modeling methods described in Gentzkow and Shapiro (2010) and Lin et al. (2008). However, those methods heavily rely on expert input. In our case, this approach would have required finding several large annotated corpora representing the four extreme points of the two axes between which Sunstein supposedly shifted (one for Keynesian economics, one for behavioral economics, one for republicanism and one for libertarian paternalism). This proved to be very hard to find, and would have raised the question of bias, since those corpora would have had to be constructed from a manual selection of texts.

We therefore focused on finding an entirely input-free solution.

3.2 Latent Dirichlet Allocation

3.2.1 Presentation

One of the first unsupervised methods for topic modeling is Latent Dirichlet Allocation (LDA) (Blei et al., 2003). The intuition behind this algorithm is the following: words that often appear in the same context are likely to be part of a common topic.

From there, for every document, we can deduce a list of n topics (a value that needs to be manually inputted), and the algorithm will generate a list of words that are theoretically part of each of the n topics. We therefore decided to apply LDA to each year split of our corpus, so as to see what the main topics of the author were, and how the words representing each topic evolved through time.

Our implementation uses the `gensim`[1] Python package. After pre-processing and tokenizing the text into documents, we ran the algorithm and extracted a total of 10 topics for each year split, with each containing 10 words.

Our results are summarized in table 3.1.

[1] https://radimrehurek.com/gensim/

3.2.2 Results

1987 (topics 1 and 2)	
.011*power	.017*state
.008*administrative	.014*court
.008*judicial	.012*statute
.008*understanding	.010*regulatory
.008*contract	.010*private
.006*state	.009*democracy
.006*agency	.009*protection
.006*justice	.009*provide
.006*lochner	.009*discrimination
.006*system	.009*agency

1995 (topics 1 and 2)	
.020*right	.009*lawyer
.011*people	.007*theory
.009*rule	.007*social
.008*category	.006*people
.008*constitutional	.006*constitutional
.007*political	.006*political
.007*particular	.006*general
.006*might	.005*right
.006*think	.005*rule
.006*change	.005*agreement

2005 (topics 1 and 2)	
.009*would	.014*precautionary
.008*people	.013*principle
.008*paper	.010*discount
.007*activity	.008*risk
.006*agency	.007*regulation
.006*perhaps	.007*supra
.006*engage	.006*benefit
.006*relatively	.020*sunstein
.006*procedure	.011*market
.006*likely	.011*posner

TABLE 3.1: Two of the main topics extracted from the corpus at different time points. Each word is associated with its relevance coefficient Φ.

In practice, those results are very hard to interpret, due to two main reasons:

- While the corpus is split into slices for each year, the topics are computed independently on each split. There is therefore no guarantee that the topics extracted from a given year match — or even overlap — with the topics extracted at a different time point. This makes it difficult to analyze the evolution of Sunstein's thought on a given concept.
- An expert is required to name and analyze these lists of words. As the results also depend on the value of n initially chosen, it would take a lot of time to apply this algorithm effectively within the problematics of ideological analysis.

3.3 Dynamic and Embedding Topic Modeling

As mentioned above, LDA has several downsides, mainly the impossibility of measuring the evolution of specific topics. Furthermore, it extracts topics through a purely syntactic analysis method.

Blei and Lafferty (2006) introduces an improved version of LDA called Dynamic Topic Modeling (DTM). This algorithm builds up on Blei et al. (2003) but takes into account time as part of the modeling. Specifically, it generates a fixed list of topics for all years, and then extracts the list of words and weighted associated with those topics for each time slice, therefore showcasing their evolution over time.

Blei and Lafferty (2006) is very technical and mathematically complex. While we found a Python implementation of the paper in the `gensim` library (using the `LDASeq` class), it turned out to use a lot of computational resources: a test on a small portion of our corpus took 6 hours to run. We estimate that running the program on the full corpus would take several days. Furthermore, we found that the quality of the LDA topics generated in section 3.2 was not satisfying, which led us to look for alternatives rather than spend a lot of time on this particular method.

Another kind of topic modeling algorithm that seemed promising was introduced in Dieng et al. (2020). This paper introduces a method called Embedding Topic Modeling (ETM), which the authors claim obtain significantly better results than the classic LDA. The main idea of the algorithm is to use word embeddings (Mikolov et al., 2013), which are a method of representing words as vectors in a high-dimensional space, with semantically similar words close to each other.

While we did not use ETM directly, it led us to the discovery of BERTopic, which combines ideas from both ETM and DTM.

3.4 BERTopic

3.4.1 Presentation

BERTopic (Grootendorst, 2022) is a recent, state-of-the-art topic modeling algorithm that relies on word embeddings (through BERT) to represent the evolution of topics in a corpus over time.

What is BERT ?

BERT, which stands for Bidirectional Encoder Representations from Transformers (Devlin et al., 2018; Turc et al., 2019)) is a series of very large language models that have been shown to perform very well in a large array of NLP tasks, including natural language generation and machine translation. BERT uses a transformer-type architecture (Vaswani et al., 2017) as well as contextual embeddings as part of its internal state.

3.4.2 Methodology

In the paper, Grootendorst also introduces the `BERTopic` Python library, which implements the algorithm described in the paper, but also includes lots of visualisation tools that proved to be very useful for our study.

Since BERTopic is optimized for corpora containing at least 1,000 short documents, we decided to split every article into paragraph-sized documents (one paragraph per document). As for the previous analyses, we chose to work with corpus slices containing one year per split.

Preprocessing was done by tokenizing the corpus as well as removing stop words. We then decided to mostly keep BERTopic's default settings, such as using 10 words per topic and only considering topics present in at least 35 documents. BERTopic can also consider n-grams as input instead of simply words. In our case, since stop words had already been removed, we chose to consider both unigrams and bigrams, since we assume most relevant phrases to contain either 1 or 2 non-stop words.

3.4.3 Results and analysis

Table 3.2 gives an overview of some of the extracted topics. The full list, containing a total of 160 topics, can be found in table B.1.

Topic 0	punitive	awards	damages	jury	juries
Topic 1	speech	amendment	free	first amendment	pornography
Topic 2	id	see id	see	id id	id 38
Topic 3	discrimination	affirmative	affirmative action	action	racial
Topic 4	president	executive	power	congress	presidential

TABLE 3.2: The top 5 words of the 5 first topics extracted by BERTopic on the overall corpus.

While it may not be obvious to a reader without an expert background in political theory, these topics are coherent and reflect actual themes of the underlying text (respectively: an experiment aiming at measuring how juries decide on the amount of punitive damages in a trial, a topic about free speech, a topic extracted from footnotes, a topic about discrimination, and a topic about the powers invested in the President of the United States as well as Congress).

3.4.4 Topic selection

Prof. Ferey discarded topics that weren't important in the context of our hypothesis, and formed a selection of 32 topics to be analyzed. We then used BERTopic's hierarchical clustering function to automatically group these topics into categories. The results are displayed in table 3.1.

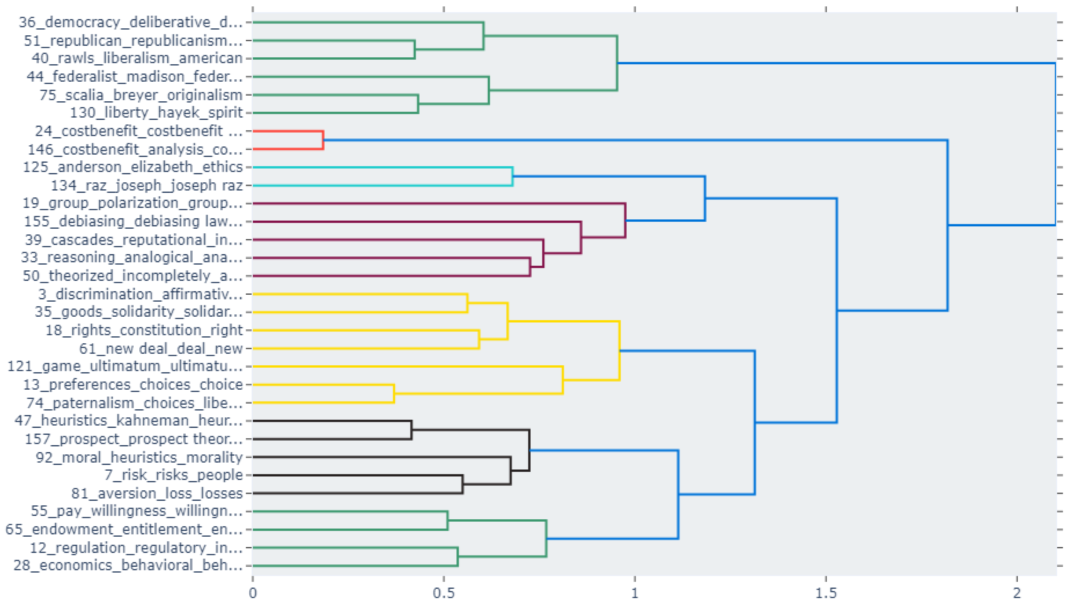


FIGURE 3.1: Hierarchical clustering of the selected topics

Thanks to this chart, we were able to automatically define several relevant categories of topics, described in table 3.3. Topic 65 was isolated for individual analysis, since it is central to prof. Ferey's theory.

Category	Topics
Political philosophy	36, 40, 44, 51, 75, 130
Economic approaches	24, 146
Ethics	125, 134
Judicial interpretations	19, 33, 39, 50, 155
Libertarian paternalism	13, 74
Moral philosophy	3, 13, 18, 35, 61, 74, 121
Behavioral economics	7, 47, 81, 92, 157
Endowment effect	65
Miscellaneous	12, 55

TABLE 3.3: Groups of topics extracted from BERTopic, labeled by prof. Ferey.

3.4.5 Evolution over time

Finally, we use BERTopic's time analysis to explore the frequency of topics over time. Figures 3.2 through 3.5 contain some of the key results, and the full list can be found in Appendix B.

Overall, those results provide strong evidence for the working hypothesis. Figure 3.2 shows an important rise in all topics related to behavioral economics, starting in 1994 and peaking

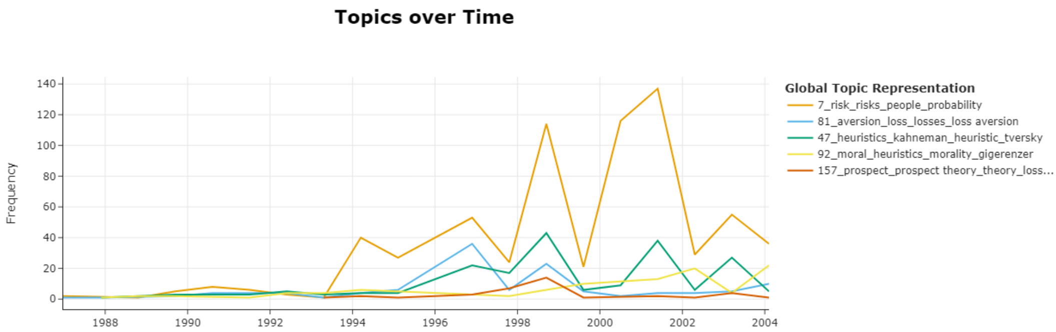


FIGURE 3.2: Evolution of the frequency of topics in the "Behavioral economics" category.

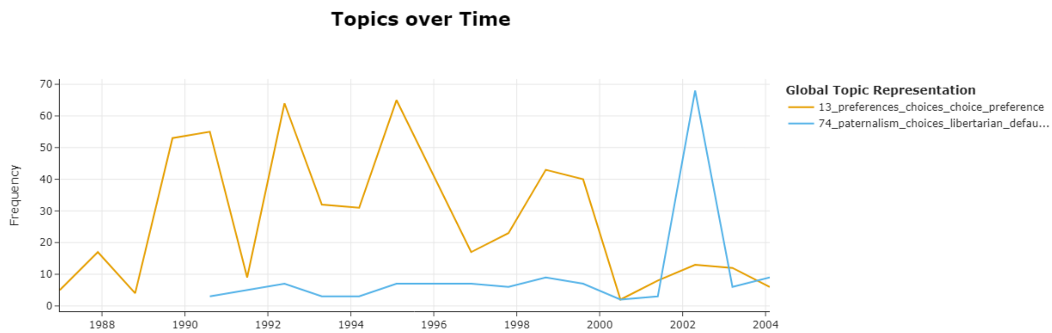


FIGURE 3.3: Evolution of the frequency of topics in the "Libertarian paternalism" category.

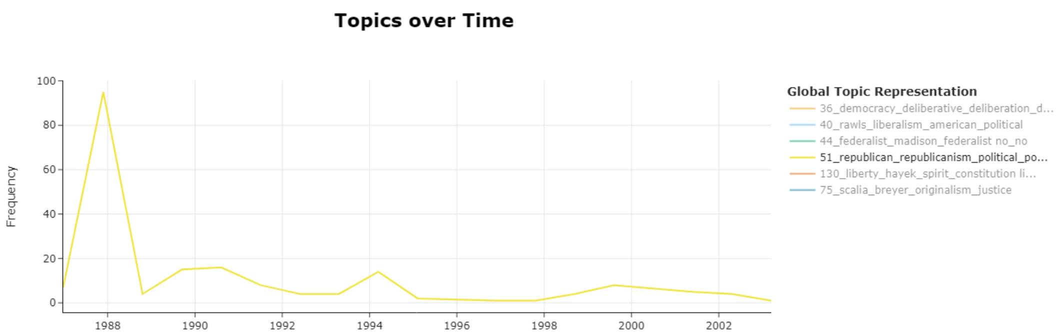


FIGURE 3.4: Evolution of topic 51 (republicanism) in the "Political philosophy" category.

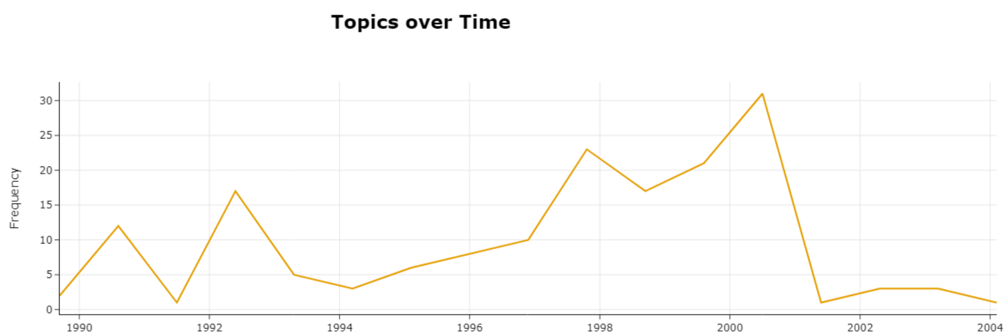


FIGURE 3.5: Evolution of topic 65 (endowment).

in 2001. Figure 3.3 displays Sunstein's thinking about a specific question: Should the state take into account individual preferences, or influence them? We first observe that topic 33, focused on collective preferences, dominates for most of the 1990's. Topic 74, representing libertarian paternalism, progressively appears during that period, and particularly comes into light in 2002, leading to a crash of topic 33. Figure 3.4 shows how topic 56 (republicanism) plummets in Sunstein's articles. After a peak in 1990, it decreases in importance up until 1995, where it all but dies out.

Finally, figure 3.5 shows the evolution of Sunstein's reflection on endowment (topic 65). This particular topic is relevant to our analysis because there exist two approaches to the endowment effect: a classical one, and one related to behavioral economics. While the graph of the topic does show a peak in the early 1990s and another one in 2001, the words associated with the topic did not really change between those two points.

We bring the reader's attention to the fact that along with the frequency of each topic, BERTopic outputs the list of word associated with each topic at each time point. We are not including this information here for brevity's sake, but we have generally noticed little evolution in each topic's words. One notable exception is that BERTopic outputs seemingly random words for time slices where the topic is not present (frequency of 1). Another exception is that in the case of topic 74 (paternalism), BERTopic gives a non-zero frequency in the early years of the corpus. However, the word list shows that before 1995, the topic contains words related to paternalism in general, and not the specific meaning of libertarian paternalism.

3.4.6 Sentiment Analysis

One problem with obtaining the frequency of topics is that it does not necessarily imply that the author agrees with or supports the topic. To take an extreme example, both a pacifist and a bellicist would have war as one of their main topics.

To clarify the author's position, we decided to use Sentiment Analysis, a branch of NLP focusing on detecting whether an author expresses a positive or negative sentiment in a text. For our analysis, we use Python's `NLTK`[2] package, which analyzes documents and returns for each one three scores and a final value:

- `Pos`: This value, between 0 and 1, estimates how positive the sentiment in a given text is.
- `Neg`: Similar to `Pos`, but for negativity.
- `Neu`: Similar to `Pos` and `Neg`, but for neutrality.
- `Compound`: This value, between -1 and 1, combines the last 3 values into a single metric, where a low value indicates an overall negative sentiment, and a high value a positive one.

We have combined sentiment analysis with the output of BERTopic, by collecting every document for a given topic, running sentiment analysis on it, grouping them by year and saving the average `Compound` value. The results for topic 18 (behavioral economics) is given in figure 3.6.

Unfortunately, after evaluation by prof. Ferey, those results do not seem to be meaningful, even when trying to analyze sentiment for a given word rather than a topic. We put forward the hypothesis that this may be due to the age and simplicity of the sentiment analysis methods used in `NLTK`, a package typically used for educative purposes than for state-of-the-art research. Other reasons include Sunstein's style of writing as an academician, which is not likely to use strongly loaded words, and the inclusion of unrelated words in the time slices for which the topic is not present (frequency of 1, as mentioned in subsection 3.4.5). While this algorithm can perform very well when sentiment can easily be extracted (I.V., 2016), academic writing might be too complex of an application for this tool.

[2] https://www.nltk.org/

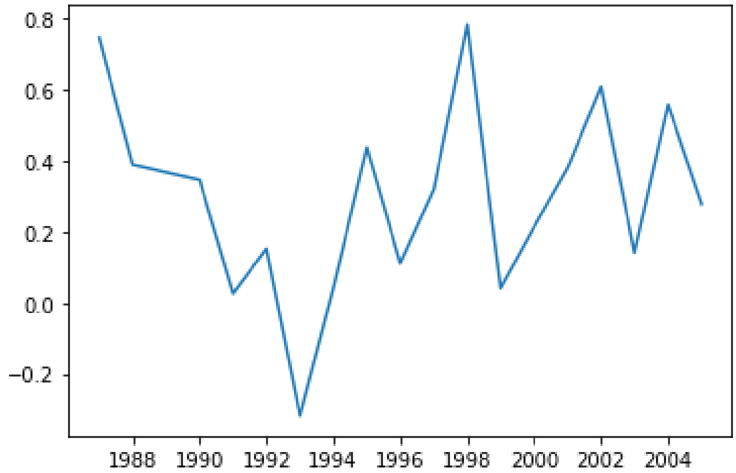


FIGURE 3.6: Sentiment analysis results for topic 28 (behavioral economics)
.

3.4.7 The endowment effect

Finally, we would like to give some insight on how the inconclusive results for topic 65 (endowment) could still be used. While we could not detect how Sunstein's approach to this topic has changed over time in section 3.4.5, we used BERTopic's topic similarity measure to identify the 5 closest topics, given in table 3.4. See figures B.9 to B.13 for a detailed breakdown of the topics.

Topic #	Main topic words	Similarity
74	endowment, effect, allocation, coase	0.787
92	valuation, goods, kinds, incommensurability	0.563
49	pay, willingness, relative, income	0.546
3	anderson, elizabeth, ethics, economics	0.484
18	prospect, theory, losses, tversky	0.473

TABLE 3.4: The top 5 closest topics to topic 65.

The qualitative interpretation of those results is as follows: The closest topic to topic 65 is topic 74, which contains words related to the classical interpretation of the endowment effect using the Coase theorem (Coase, 1960). This corresponds to what we expect from Sunstein's original positions. Topic 92 and 49 contain the interpretation of the endowment effect in the realm of economics. Topic 3 relates the topic to ethics, which is a first sign of a shift in Sunstein's reflection. Finally, topic 18 is about prospect theory, which is a direct result of behavioral economics. While no temporal information can be inferred from this list, it strongly hints that a shift may have occurred in Sunstein's conception of the endowment effect, and that this question is worth exploring further.

Conclusion

Results and interpretation

Through the use of multiple NLP methods, both unsupervised and supervised, we were able to bring forward data that supports prof. Ferey's hypothesis. There seems to be a shift in Sunstein's writing around the year 1995 to 2000, marked by the near-disappearance of republican thought and its gradual replacement by libertarian paternalism, influenced by behavioral economics and other fields such as psychology and cognitive sciences. Whether it is through semantic networks, transition modeling or topic modeling, our analysis tends to validate this theory through both qualitative and quantitative results. Semantic networks show that this changeover can further be found in specific sub-topics of Republicanism. Both semantic networks and topic modeling have also brought forward changes in the use of certain concepts such as the ones of preferences and poverty.

Future work

Generalization

In the future, we believe it would be interesting to test the methods developed in this report on other corpora, to try to ascertain how much of our results can generalize. One example of such corpora would be the Posner-Becker Blog.

Richard Posner and Gary Becker are leading modern figures in law and economics. From 2004 up to Becker's death in 2014, they both ran a blog where every week, one would comment recent events and the other would reply to that comment. Their editorial line is considered to be mostly libertarian. In total, over 700 articles were published in the blog.

One hypothesis that could be tested using our methods is to see whether the global financial crisis of 2007 has influenced Posner and Becker's conception of certain topics, in particular consumption and obesity. If such a shift occurred, it is expected that it would be tied to a single year (2007) and therefore much faster and more noticeable.

Methods

Furthermore, other methods, that have not been investigated in this report, could be applied for detecting ideological shift in a corpus such as ours.

Jelveh et al. (2014) and Jelveh et al. (2018) are two closely related papers that use both Pearson's χ^2 coefficient and LDA models for ideological detection, as in Gentzkow et al. (2019). The main interest of these papers is that they conduct analysis of political ideology on journal articles published by economists, which is very similar to our initial goal on Sunstein's corpus, and we believe the Posner-Becker corpus can be treated similarly. The results in this papers are all the more impressive because displays of ideology are typically discouraged in journal articles. We are confident that the methodology in those two articles could be replicated on our corpora to obtain initial results. However, those papers are highly technical, and implementing them would require a significant time investment, which was outside the scope of this study.

Ahmed and Xing (2010) describes a way to categorize ideological lexicon into 3 different subsets: ideological belief of the writer, topical content and topic-ideology interaction. The authors of the paper then introduce the multi-view topic model (mview-LDA), which is loosely based on that of Lin et al. (2008). This model is supposed to be able to differentiate and categorize lexicon into the aforementioned sections. We have found the content of this article seems to be complex in terms of mathematics, and decided to not implement it due to time constraints. However, it seems to be a significant upgrade of the model described in Lin et al. (2008) and is the most thorough ideological model we have found in the literature.

Closing words

In the opening section of this report, we mentioned the increase in the number of research articles applying NLP methods to various fields of humanities and social sciences, including economics and political theory. We hope that our analysis has further contributed to bridging the gap between those fields and that of computer science. It seems that both frequentist, statistical methods as well as more complex ones, such as topic modeling using word embeddings, can be used for various types of analyzes. Ideological analysis has yet to find an exhaustive method able to answer typical corpus questions, but we are confident that this will improve with time.

Bibliography

Amr Ahmed and Eric Xing. 2010. Staying Informed: Supervised and Semi-Supervised Multi-View Topical Analysis of Ideological Perspective. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1140–1150, Cambridge, MA. Association for Computational Linguistics.

David M. Blei and John D. Lafferty. 2006. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, ICML '06, pages 113–120, New York, NY, USA. Association for Computing Machinery.

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *The Journal of Machine Learning Research*, 3(null):993–1022.

R. H. Coase. 1960. The Problem of Social Cost. *The Journal of Law and Economics*, 3:1–44.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. Bert: Pre-training of deep bidirectional transformers for language understanding.

Adji B. Dieng, Francisco J. R. Ruiz, and David M. Blei. 2020. Topic Modeling in Embedding Spaces. *Transactions of the Association for Computational Linguistics*, 8:439–453.

Samuel Ferey. 2009. *Une histoire de l'analyse économique du droit: calcul rationnel et interprétation du droit*. Number 1 in Collection Droit et économie. Bruylant, Bruxelles.

Matthew Gentzkow, Bryan Kelly, and Matt Taddy. 2019. Text as Data. *Journal of Economic Literature*, 57(3):535–574.

Matthew Gentzkow and Jesse M. Shapiro. 2010. What Drives Media Slant? Evidence From U.S. Daily Newspapers. *Econometrica*, 78(1):35–71. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA7195.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

Udo Hahn, Véronique Hoste, and Ming-Feng Tsai, editors. 2018. *Proceedings of the First Workshop on Economics and Natural Language Processing*. Association for Computational Linguistics, Melbourne, Australia.

Udo Hahn, Véronique Hoste, and Zhu Zhang, editors. 2019. *Proceedings of the Second Workshop on Economics and Natural Language Processing*. Association for Computational Linguistics, Hong Kong.

Shravan I.V. 2016. Sentiment analysis in python using nltk. *OSFY - OpensourceForYou*.

Mathieu Jacomy, Tommaso Venturini, Sebastien Heymann, and Mathieu Bastian. 2014. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLOS ONE*, 9(6):1–12.

Zubin Jelveh, Bruce Kogut, and Suresh Naidu. 2014. Detecting Latent Ideology in Expert Text: Evidence From Academic Papers in Economics. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1804–1809, Doha, Qatar. Association for Computational Linguistics.

Zubin Jelveh, Bruce Kogut, and Suresh Naidu. 2018. Political Language in Economics. SSRN Scholarly Paper ID 2535453, Social Science Research Network, Rochester, NY.

Daniel Kahneman and Amos Tversky. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291.

Wei-Hao Lin, Eric Xing, and Alexander Hauptmann. 2008. A joint topic and perspective model for ideological discourse. In *Proceedings of the 2008th European Conference on Machine Learning and Knowledge Discovery in Databases - Volume Part II*, ECMLPKDD'08, pages 17–32, Berlin, Heidelberg. Springer-Verlag.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space.

Pierre Molette and Agnès Landré. 1994. Tropes.

N. Nicolov, G. Angelova, and R. Mitkov. 2009. *Recent Advances in Natural Language Processing V: Selected papers from RANLP 2007*. Current Issues in Linguistic Theory. John Benjamins Publishing Company.

Office of the Federal Register, National Archives and Records Administration. 2016. 3 CFR 13707 - Executive Order 13707 of September 15, 2015. Using Behavioral Science Insights To Better Serve the American People.

Christos H. Papadimitriou, Hisao Tamaki, Prabhakar Raghavan, and Santosh Vempala. 1998. Latent semantic indexing: A probabilistic analysis. In *Proceedings of the Seventeenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems*, PODS '98, page 159–168, New York, NY, USA. Association for Computing Machinery.

Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever, et al. 2019. Language models are unsupervised multitask learners. *OpenAI blog*, 1(8):9.

Anand Rajaraman and Jeffrey David Ullman. 2011. *Data Mining*, page 1–17. Cambridge University Press.

Cartic Ramakrishnan, Abhishek Patnia, Eduard Hovy, and Gully APC Burns. 2012. Layout-aware text extraction from full-text PDF of scientific articles. *Source Code for Biology and Medicine*, 7(1):7.

Steven M. Sheffrin. 2017. Behavioral Law and Economics Is Not Just a Refinement of Law and Economics. *Œconomia. History, Methodology, Philosophy*, pages 331–352.

Mélanie Siegel. 2018. Text Mining in Economics. *Semantic Applications*.

Ray Smith. 2007. An overview of the tesseract ocr engine. In *Proc. Ninth Int. Conference on Document Analysis and Recognition (ICDAR)*, pages 629–633.

Cass R. Sunstein, editor. 2000. *Behavioral Law and Economics*. Cambridge Series on Judgment and Decision Making. Cambridge University Press.

Dominika Tkaczyk, Paweł Szostek, Mateusz Fedoryszak, Piotr Jan Dendek, and Łukasz Bolikowski. 2015. CERMINE: automatic extraction of structured metadata from scientific literature. *International Journal on Document Analysis and Recognition (IJDAR)*, 18(4):317–335.

Iulia Turc, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. Well-read students learn better: On the importance of pre-training compact models. *arXiv preprint arXiv:1908.08962v2*.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is all you need.

Appendix A

Actions available in the Monster

Name	Action	Notes
1-10	Replace the word with one of 10 suggestions	Suggestions are generated by taking all words from the wordlist and outputting the ones with the lowest Levenshtein distance to the candidate word.
Add	Add the word to the wordlist	
Replace	Create a new substitution rule [candidate → replacement]	The replacement value is user-inputted.
Split	Split the word as suggested (maximizing the product of word probabilities)	The rule [candidate → split] is also created. **Ex:** *<toanswerfor> → <to> answer for*
De-hyphenize	Convert the hyphen into a dash	I.e. surround the dash with spaces. **Ex:** *<but-it> → <but> - it*
Context	Show more context around the word	Can be applied multiple times in a row.
Merge left	Merge the word with the closest one to the left	**Ex.:** *accept <ance> → <acceptance>*
Merge right	Merge the word with the closest one to the right	**Ex.:** *<enume> rate → <enumerate>*
One-time	Applies a regular expression substitution to the word's context	The regex is user-inputted. This allows substitutions outside of the sliding window.
Fast replace	Substitute the word for a value	Same as Replace, but does not add the rule to the global list.
Number removal	Remove digits from the word	**Ex.:** *<pr4ctice5> → <practice>*
Detach left	Perform a smart split based on punctuation within the word	**Ex.:** *<yet."He> → <yet>." He*
Detach right	Similar to Detach left, but aligns quotes to the right.	**Ex.:** *<it,said:"I> → <it>, said: "I*
New rule	Add a new substitution rule and applies it to the entire file	Both the search and replacement values are user-inputted.
Ignore	Move to the next word	Doesn't perform any changes.
Add hyphens	Similar to split, but also add hyphens with spaces (dashes) between the words	The rule [word → hyphenized] is also created. **Ex.:** *<truebut> → <true> — but*

TABLE A.1: A list of available options when encountering an unknown word. Angle brackets represent the position of the sliding window before and after the operation. Word probabilities are computed using the Python package `wordninja` [1].

Name	Action	Notes
Lower	Convert the word to lowercase	Also create the rule [wOrD $\rightarrow$ word].
Upper	Convert the word to uppercase	Also create the rule [wOrD $\rightarrow$ WORD].
Ignore	Move to the next word	Also adds the current word case to the allowed case list.
Sentence	Convert the word to sentence case	Also create the rule [wOrD $\rightarrow$ Word].
Custom	Create a new case substitution rule [wOrD $\rightarrow$ replacement]	The replacement value is user-inputted. Similar to Replace, but this rule is only applied for case substitution.
One-time	Applies a regular expression substitution to the word's context	The regex is user-inputted. Similar to the same command in non-case checking mode.

TABLE A.2: A list of available options when encountering inconsistent capitalization. Case check is not performed after certain regular operations such mas left and right merge. Note that the left part of substitution rules is case-sensitive.

Appendix B

Additional BERTopic results

TABLE B.1: The full list of topics extracted by BERTopic.

Topic #	Words
0	punitive, awards, damages, jury, juries, dollar, judgments, punishment, punitive damages, outrage
1	speech, amendment, free, first amendment, pornography, first, government, free speech, expression, viewpoint
2	id, see id, see, id id, id 38, 38, 28 id, id 535, id 67, 66 id
3	discrimination, affirmative, affirmative action, action, racial, race, caste, equality, black, principle
4	president, executive, power, congress, presidential, framers, powers, control, authority, independent
5	us, 1976, see, 1973, usc, united, cir, 424, eg, 410
6	epa, air, pollution, clean, environmental, act, standards, clean air, emissions, air act
7	risk, risks, people, probability, slovic, availability, heuristic, fear, availability heuristic, information
8	interpretation, statutory, interpretive, statutes, meaning, principles, courts, norms, legislative, text
9	standing, injury, plaintiff, article, injury fact, suit, court, congress, iii, article iii
10	risk, safety, health, osha, risks, regulation, occupational, workers, benefits, costs
11	vsl, lives, lifeyears, wtp, life, people, risk, risks, million, discount
12	regulation, regulatory, incentives, economic, efficiency, government, strategies, economic incentives, commandandcontrol, costs
13	preferences, choices, choice, preference, collective, consumption, social, people, private, may
14	chevron, agency, deference, interpretation, congress, agencies, courts, interpretations, court, administrative
15	administrative, judicial, review, agency, judicial review, courts, regulatory, agencies, law, administrative law
16	programming, broadcasters, television, code, public, public interest, broadcasting, interest, market, viewers
17	us, id, usc, ann, 478, see, 478 us, 448, see id, 1984
18	rights, constitution, right, property, constitutions, constitutional, economic, private, social, social economic
19	group, polarization, group polarization, members, groups, deliberation, extreme, direction, group members, shift
20	sunstein, cass, cass sunstein, rev, see cass, see, chi, regulatory state, see sunstein, forthcoming
21	precautionary, precautionary principle, principle, risks, warming, global, risk, global warming, environmental, catastrophic
22	university, chicago, professor, university chicago, karl, school, llewellyn, law school, karl llewellyn, department political

Continued on next page

Table B.1 – continued from previous page

Topic #	Words
23	supra, note, supra note, see, see supra, kuran supra, sullivan supra, kuran, wildavsky supra, note 19
24	costbenefit, costbenefit analysis, analysis, costs, benefits, regulation, balancing, agencies, costs benefits, regulatory
25	norms, social, social norms, norm, roles, meanings, expressive, law, function, people
26	nondelegation, doctrine, nondelegation doctrine, canons, constitutional, congress, court, nondelegation canons, judicial, courts
27	supra, note, supra note, al supra, al, et, et al, see, note see, gunningham
28	economics, behavioral, behavior, law economics, economic, analysis, law, bounded, predictions, behavioral economics
29	iii, ii, puzzles, iv, practice, conclusion, ii iii, preliminaries, mechanisms, implications
30	supra, accompanying, text, see supra, notes, note, supra note, supra notes, accompanying text, text see
31	usc, 1988, f2d, cir, us, 42, 42 usc, 29 usc, supp, 1982
32	omb, oira, order, executive, executive order, regulatory, agency, review, president, agencies
33	reasoning, analogical, analogical reasoning, equilibrium, reflective, analogy, reflective equilibrium, analogies, search, search reflective
34	abortion, roe, women, right, surrogacy, abortions, roe wade, wade, reproductive, sex
35	goods, solidarity, solidarity goods, value, good, people, others, social, enjoying, may
36	democracy, deliberative, deliberation, deliberative democracy, see, elster, ely, 1980, democratic, democracy distrust
37	introduction, august 2000, position august, 2000, position, introduction introduction, august, qualifications, landis administrative, 2000 position
38	fda, tobacco, drug, products, food, tobacco products, food drug, fda authority, authority, drugs
39	cascades, reputational, informational, cascade, others, people, academic, informational cascades, cascade effects, social
40	rawls, liberalism, american, political, john, john rawls, republicanism, liberal, see, political liberalism
41	valuation, goods, kinds, kinds valuation, incommensurability, metric, value, valued, diverse, single
42	rules, rule, factors, ex, decisions, cases, law, system, may, advance
43	sunstein supra, sunstein, supra, note, see sunstein, supra note, see, kahneman schkade, pildes sunstein, schkade
44	federalist, madison, federalist no, no, hamilton, ed, see federalist, rossiter, no 10, 1961
45	id id, id, , , , , , , ,
46	impeachment, president, misdemeanors, office, crimes, impeached, clinton, crimes misdemeanors, high crimes, president clinton
47	heuristics, kahneman, heuristic, tversky, biases, daniel, heuristics biases, availability, amos, daniel kahneman
48	accompanying, notes, text, text accompanying, accompanying notes, infra, see text, see infra, infra text, cited
49	secondorder, delegate, delegation, decision, decisions, burdens, may, secondorder decisions, strategies, costs
50	theorized, incompletely, agreements, incompletely theorized, theorized agreements, agreement, theory, law, disagree, theories
51	republican, republicanism, political, politics, republicans, liberal, deliberation, liberalism, private, republican thought

Continued on next page

Table B.1 – continued from previous page

Topic #	Words
52	workers, employees, employers, waivable, waivers, employment, waiver, employee, employer, rights
53	scalia, dissenting, scalia dissenting, scalia supra, morrison, ct, concurring, justice, see scalia, see
54	cong, 1995, sess, rec, 141, 1st sess, cong rec, cong 1st, 1st, 141 cong
55	pay, willingness, willingness pay, relative, income, position, value, rank, relative position, valuation
56	military, war, security, national security, president, bush, maximalism, national, commissions, court
57	gdp, income, workers, poverty, unemployment, economic, wellbeing, growth, wage, relative
58	vi, summary, vi conclusion, widely publicized, background, conclusion, background background, iv disclosure, little ex, familiarity new
59	capital, punishment, capital punishment, death, penalty, moral, steiker, death penalty, murders, morally
60	lochner, unconstitutional, unconstitutional conditions, doctrine, common, lochnerlike, conditions, conditions doctrine, common law, constitutional
61	new deal, deal, new, checks, balances, checks balances, system, constitutional, government, national
62	us, inc, chevron, 467, 467 us, usa, 837, usa inc, chevron usa, us 837
63	montana, speed, highway, basic rule, patrol, limit, speed limit, mph, highway patrol, basic
64	suicide, physicianassisted, physicianassisted suicide, right, patient, right physicianassisted, death, patients, euthanasia, cases
65	endowment, entitlement, endowment effect, effect, allocation, coase, theorem, initial, default, initial allocation
66	animals, animal, human, animal rights, rights, beings, suffering, human beings, cruelty, animal welfare
67	cable, access, fcc, television, turner, broadcast, broadcasting, fairness, fairness doctrine, programming
68	ct, 112, 112 ct, kennedy concurring, lujan 112, 118, rav, 113, 118 ct, kennedy
69	experts, expert, data, antenatal, medical, care, statistical, doctors, standard, corticosteroids
70	marriage, marry, marriages, right marry, right, state, institution, marital, married, benefits
71	roosevelt, rights, bill, second bill, bill rights, property, second, constitutive, right, security
72	supra, see, see supra, note, accompanying, supra note, notes, text, id, accompanying notes
73	id see, see id, id, see, eg id, eg, see eg, , ,
74	paternalism, choices, libertarian, default, choice, employees, planner, preferences, choose, plan
75	scalia, breyer, originalism, justice, justice scalia, active liberty, originalist, liberty, constitution, active
76	waste, hazardous, canal, hazardous waste, love, love canal, disposal, contamination, mail, waste sites
77	smoking, tobacco, cigarettes, fda, health, viscusi, litigation, msa, addiction, taxes
78	dissenting, us, holmes, stevens, marshall, concurring, marshall dissenting, holmes dissenting, see, justice
79	senate, nominee, nominees, appointments, president, role, appointment, confirmation, court, advice
80	legal, computational, models, computer, program, reasoning, legal reasoning, legal expert, artificial intelligence, expert

Continued on next page

Table B.1 – continued from previous page

Topic #	Words
81	aversion, loss, losses, loss aversion, extremeness, option, people, extremeness aversion, choice, reference point
82	secession, secede, right secede, subunit, right, subunits, nation, civil, soviet, liberties
83	supra, note, supra note, brown supra, miller supra, see, mackinnon supra, brown, stewart supra, miller
84	usc, seq, et seq, 42, tbl, et, 42 usc, id, 16 usc, seq 16
85	stock, company, company stock, employees, employer, percent, employee, tax, contributions, employers
86	delaney, delaney clause, clause, cancer, food, additives, minimis, de, de minimis, substances
87	lawrence, sexual, sodomy, court, samesex, adultery, criminal, texas, desuetude, consensual
88	information, disclosure, informational, disclosure requirements, market, may, strategies, info, information disclosure, pl
89	tort, loss, damages, dependents, victim, tort law, death, losses, income, lost
92	moral, heuristics, morality, gigerenzer, moral heuristics, intuitions, trolley, judgments, brain, moral judgments
91	news, media, coverage, stories, attention, sensational, story, issues, minutes, newspapers
90	auto, nhtsa, standards, sideimpact, autos, trucks, buses, american trucking, trucking, sideimpact standards
93	sunstein cass, cass, sunstein, cass sunstein, sunstein edna, edna ullmannmargalit, edna, ullmannmargalit, strauss cass, ullmannmargalit cass
94	institutions, rules, institutional, ii, ramseyer eric, japan january, econometrics, econometrics japan, independence civil, rasmusen judicial
95	rules, adaptable, rule, privately, adaptable rules, privately adaptable, costs, may, private, endstates
96	ozone, epa, reg, fed, fed reg, pg, standard, standards, ppm, pm
97	florida, gore, bush, recount, bush gore, manual, court, vote, votes, manual recount
98	professors, school, schools, students, law, rankings, devins, academics, law school, law professors
99	animals, standing, animal, species, injury, wildlife, defenders wildlife, suit, aesthetic, animal welfare
100	mackinnon, feminism, women, sexuality, unmodified, sexual, pornography, feminism unmodified, sex, catharine
101	campaign, finance, pacs, campaign finance, buckley, limits, money, contributions, expenditures, political
102	const, art, const art, us const, art ii, cl, us, ii, ii us, art cl
103	voluntarily, risk, run, risks, involuntary, voluntary, involuntarily, incurred, deaths, voluntariness
104	bill, senate, house, reform, 1995, dole, act, moratorium, dole bill, glenn
105	rosen, murphy, laughter, car, frank, faster, faster car, laughter murphy, laughter rosen, speed
106	supra, note, supra note, see, glendon supra, glendon, eg, see eg, winter supra, note 29
107	program, allocations, programs, lives, allocation, condition, conditions, lives saved, equity, million
108	sunstein, democracy, cass, problem free, democracy problem, free, sunstein democracy, free speech, speech, cass sunstein
109	occupational exposure, exposure limit, exposure, occupational, asbestos, limit, oshah, limit oshah, asbestos occupational, fittings
110	speech, university, hate, universities, hate speech, restrictions, educational, students, viewpoint, campus
	Continued on next page

Table B.1 – continued from previous page

Topic #	Words
111	appointees, republican, panel, republican appointees, democratic appointees, democratic, judges, party, sitting, panels
112	counsel, independent counsel, independent, attorney, act, investigation, counsel act, attorney general, criminal, charges
113	formalism, bentham, judges, interpretation, rules, law, england, cases, formalist, hart
114	infra, see infra, infra part, app, infra app, section, part, see, infra section, see section
115	fcc, f2d, cir, dc, dc cir, syracuse, peace council, syracuse peace, broadcasting co, peace
116	sunstein cass, cass, sunstein, cass sunstein, sunstein thank, thank, thank thanks, born slightly, getting time, daughter born
117	environmental, stewart, environmental law, richard stewart, reforming, stewart reforming, reforming environmental, richard, ackerman richard, 13 colum
118	usc, clean, act, 42, air, air act, clean air, 42 usc, act 42, supp
119	richard, posner, cass, 1999, sunstein, cass sunstein, september, 2003, eric, august
120	cir, dc, dc cir, f2d, auto, curiam, center auto, en, en banc, banc
121	game, ultimatum, ultimatum game, responders, fairness, player, proposers, responder, proposer, offer
122	baker, edwin, edwin baker, baker advertising, advertising democratic, democratic press, frank choosing, right pond, choosing right, pond
123	hdi, development, human development, nations, report, development report, human, united nations, united, income
124	november, november 2000, internet, 2000 system, system november, 2000, internet november, 1995, dd, dam
125	anderson, elizabeth, ethics, elizabeth anderson, ethics economics, value ethics, anderson value, value, see elizabeth, economics
126	vicious, circle, breaking, epstein, vicious circle, breaking vicious, breyer breaking, breyer, richard epstein, cited
127	cir, theory perspective, perspective april, f2d, f3d, 2d, april 2001, political theory, 2001 political, april
128	cited, cited note, rev, chi, note, chi rev, stone, 57 chi, see, stone 54
129	conclusion conclusion, conclusion, ideas conclusion, provocative, learned, book8, principal opinion, conclusion rest, conclusion marketplace, theory conclusion
130	liberty, hayek, spirit, constitution liberty, constitution, hayek constitution, ed, marshall, 1960, baltimore
131	sex, orientation, sexual, discrimination, sexual orientation, discrimination basis, basis, samesex, basis sexual, homosexuality
132	exec, exec order, cfr, order, 12866, 12866 cfr, order no, order 12866, no 12, cited
133	infra, accompanying, see infra, text see, notes, text, accompanying text, infra notes, see, document drafting
134	raz, joseph, joseph raz, ethics, see joseph, morality, morality freedom, raz morality, ethics public, public domain
135	risk, students, alcohol drug, violence, risks, prevention programs, programs, fear, children, hans
136	oaths, compulsory, declarations, attachment, oath, compulsory declarations, heterogeneity, unity, compulsory oaths, pledge
137	minimalism, minimalists, minimalist, court, siegel, rulings, justice connor, connor, justice, war
138	baseball, beane, lewis, team, player, players, billy, teams, game, james
139	economics, posner, richard posner, richard, landes, september 1993, landes richard, law economics, working, 1993
140	mines, uranium, tailings, equipment, mill tailings, uranium mill, standards, electrical equipment, move uranium, equipment standards

Continued on next page

Table B.1 – continued from previous page

Topic #	Words
141	viscusi, kip, kip viscusi, fatal, viscusi fatal, fatal tradeoffs, tradeoffs, see kip, tradeoffs public, private responsibilities
142	supra part, see supra, part, supra, ib, see part, part ib, ib see, see, supra see
143	czech, republic, charter, soviet, independence, czech republic, ny, ny times, times, dobbs
144	rationality, bounded, bounded rationality, elster, conlisk bounded, john conlisk, jon elster, conlisk, irrationality, jon
145	posner supra, posner, supra, note, supra note, see posner, see, richard posner, posner sex, sex reason
146	costbenefit, analysis, costbenefit analysis, costs, benefits, costs benefits, effects, regulatory, disaggregated, bottom line
147	id id, id, invalid, invalid invalid, id invalid, , , , ,
148	problem judg, judg, iv problem, problem problem, iv, problem, , , ,
149	sunstein, cass, cass sunstein, see sunstein, legacy 87, sunstein lochner, lochner legacy, rev, 87 colum, laughter sunstein
151	see supra, supra pp, pp, supra, see, 46768, 46768 see, supra 423, 423 see, pp 46768
150	see supra, supra, accompanying, text, note, accompanying text, supra note, see, notes, supra text
152	fittings, corrosion, 947, 947 f2d, proof fittings, corrosion proof, epa 947, fittings epa, f2d 1201, 1201
153	february 1992, posner, 1992, richard posner, richard, february, posner sex, sex reason, sunstein, september
154	2005, 1996, may 2005, 60th, 1111, 2005 july, 2005 may, 1111 60th, july, july 2005
155	debiasing, debiasing law, bounded, selfserving, rationality, bounded rationality, law, boundedly, loewenstein, boundedly rational
156	cloning, clone, therapeutic, genetic, reproductive, therapeutic cloning, ban, reproductive cloning, nonreproductive cloning, human
157	prospect, prospect theory, theory, losses, tversky, gains, amos, kahneman amos, kahneman, amos tversky
158	see dietrich, dorner logic, logic failure, dietrich, dietrich dorner, dorner, markets, market, failure 1994, logic
159	breyer, breaking, vicious, circle, breyer breaking, breaking vicious, vicious circle, risk regulation, effective risk, stephen breyer

FIGURE B.1: Intertopic distance after topic selection. Labels are not displayed due to the interactive nature of the visualization. Generally, we observe little coherence between the clusters on this map and groups created automatically by BERTopic. This might be due to the loss of dimensionality when projecting the data onto a two-dimensional plane.

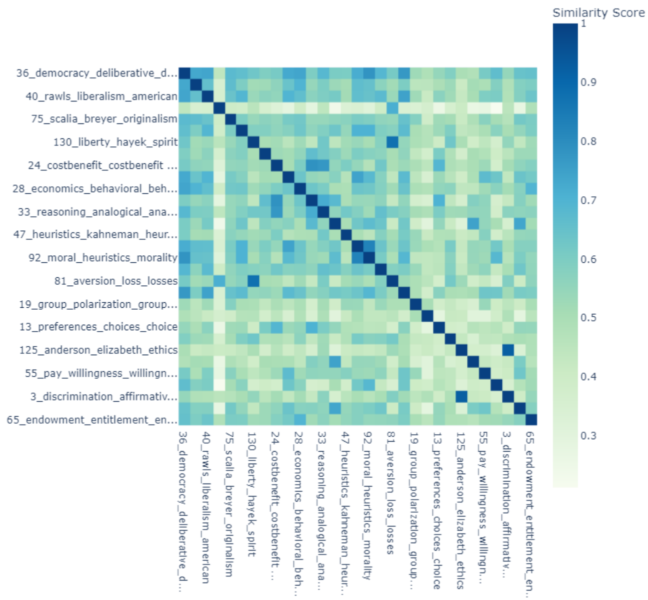


FIGURE B.2: A heatmap of topic correlation between selected topics. Please note that only half of the topic names are displayed in the x- and y-axis due to space limitations.

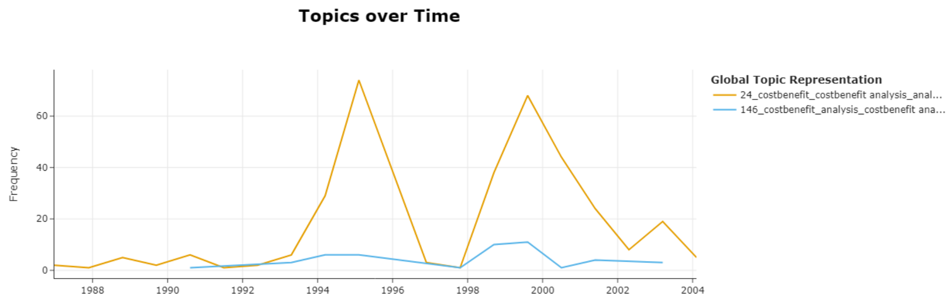


FIGURE B.3: Topic evolution over time for the "Economic Approaches" category, showing an increase in the frequency of economic topics during the transition to behavioral economics.

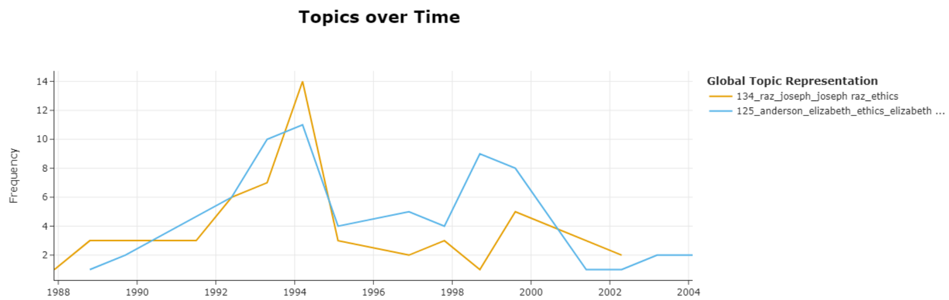


FIGURE B.4: Topic evolution over time for the "Ethics" category, also showing an increase in the frequency of topics at the beginning of the transition to behavioral economics.

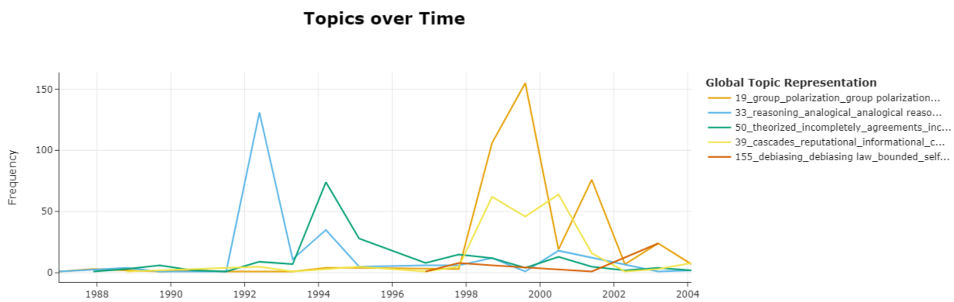


FIGURE B.5: Topic evolution over time for the "Judicial interpretation" category. We note the rise of topics 39 and 155, linked with prospect theory and behavioral economics.

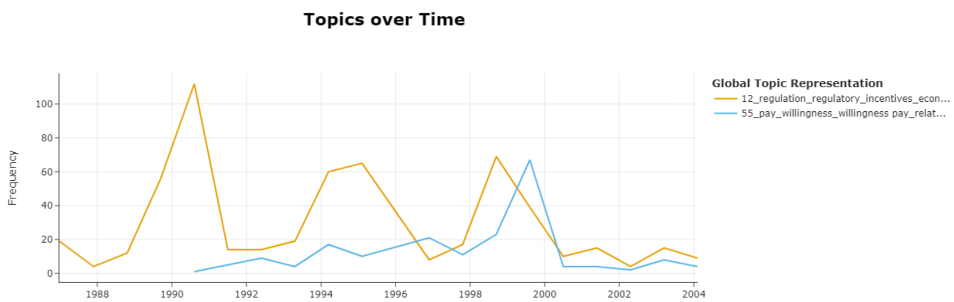


FIGURE B.6: Topic evolution over time for the "Miscellaneous" category.

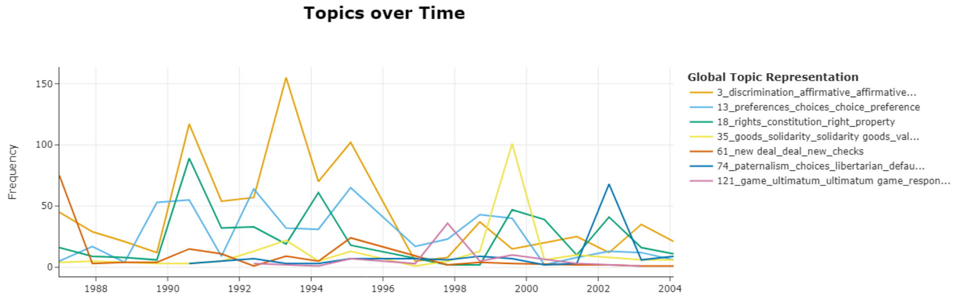


FIGURE B.7: Topic evolution over time for the "Moral philosophy" category.

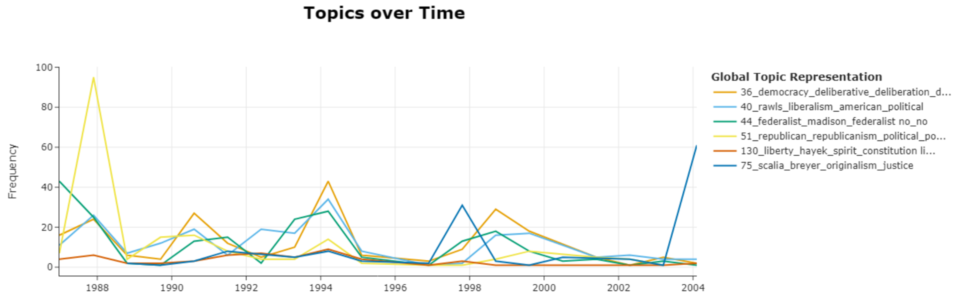


FIGURE B.8: Topic evolution over time for the "Political philosophy" category.

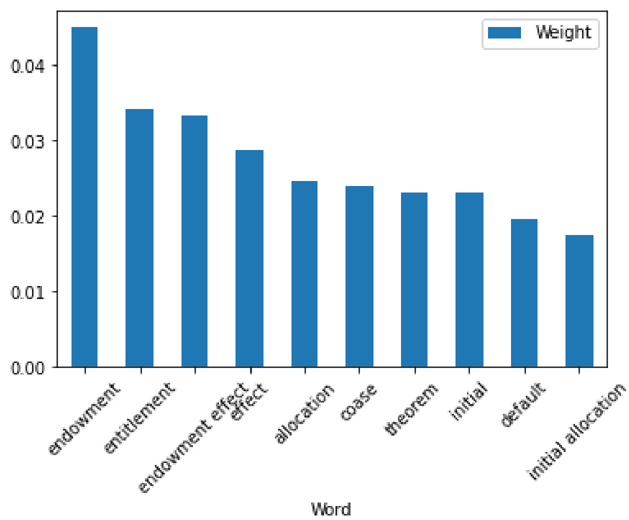


FIGURE B.9: The top words in topic 74.

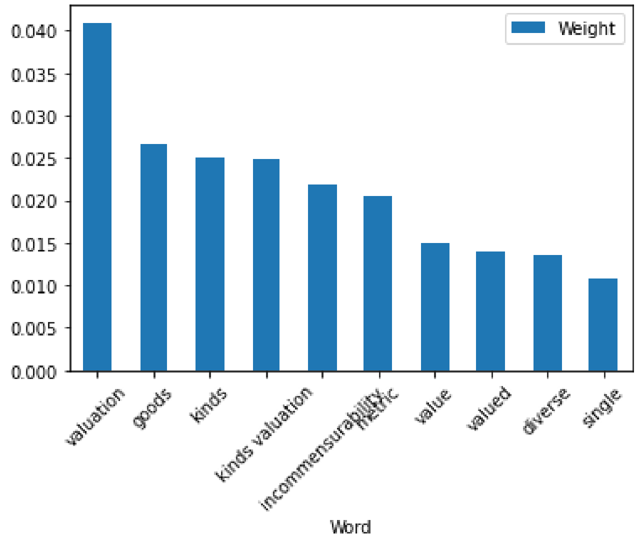


FIGURE B.10: The top words in topic 92.

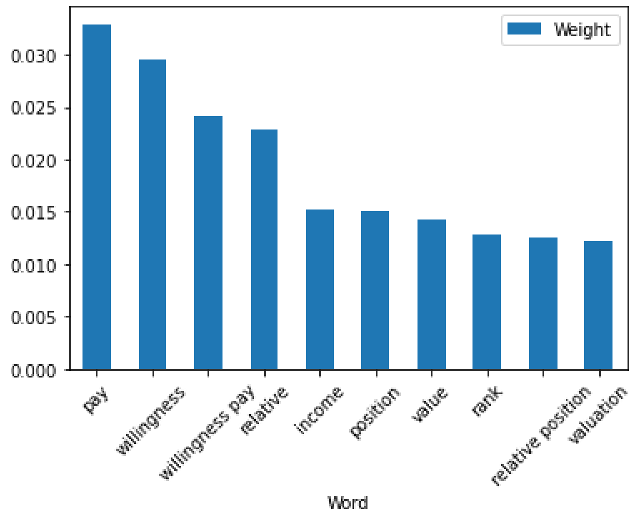


FIGURE B.11: The top words in topic 49.

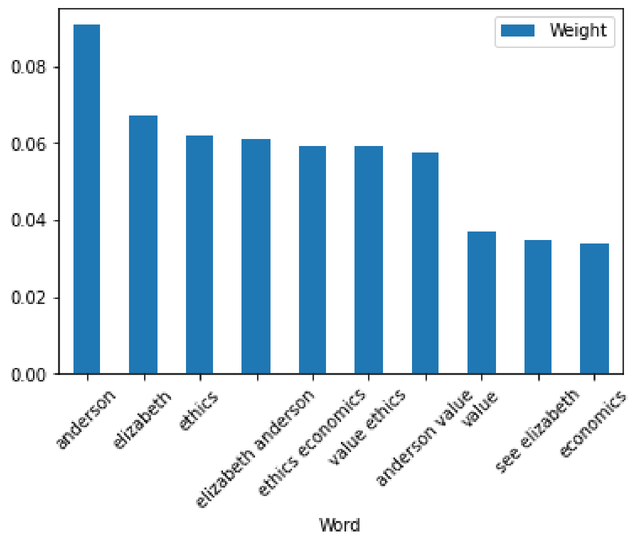


FIGURE B.12: The top words in topic 3.

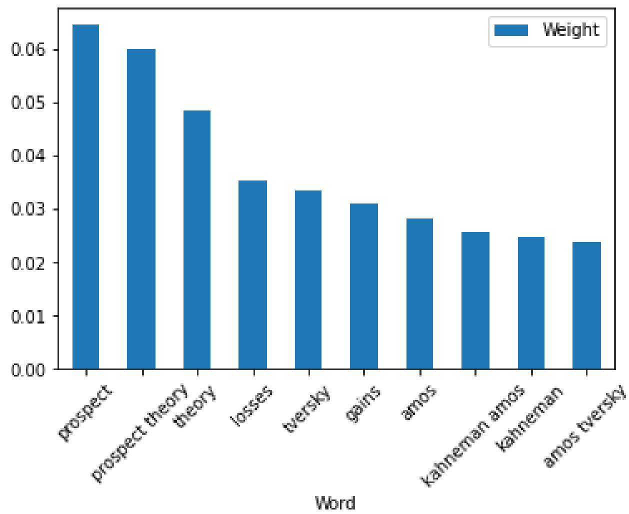


FIGURE B.13: The top words in topic 18.

Appendix C

Additional Tropes results

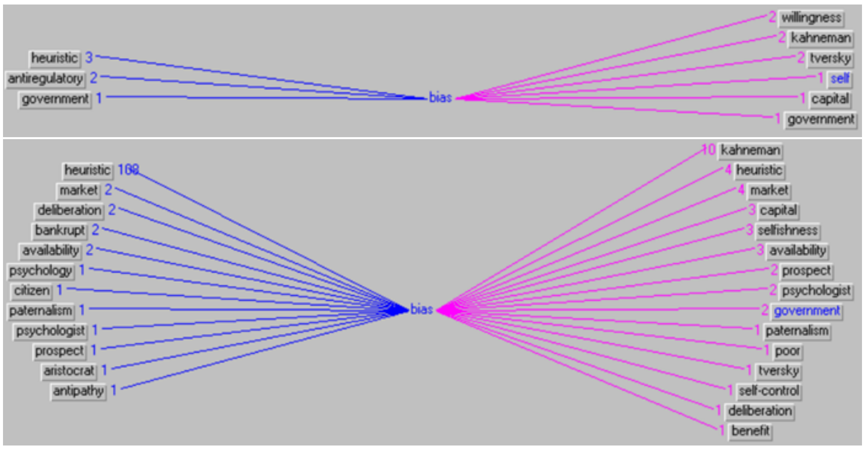


FIGURE C.1: Star graph of the word "bias" before and after 1994. The interpretation is very similar to that of the word "heuristic" in figure 2.6.

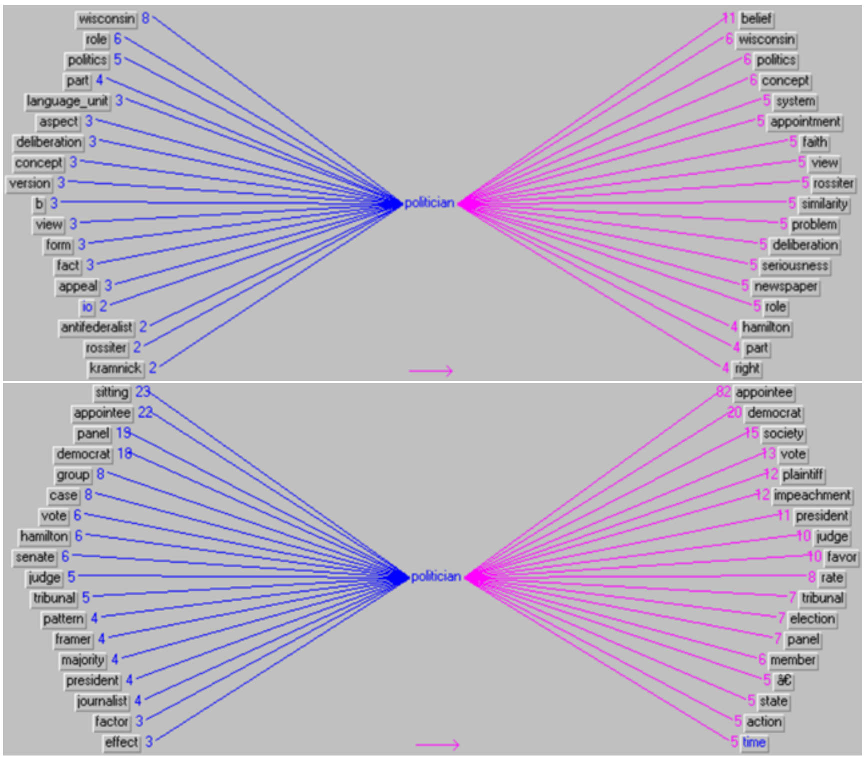


FIGURE C.2: Star graph of the word "politician" before and after 1994.

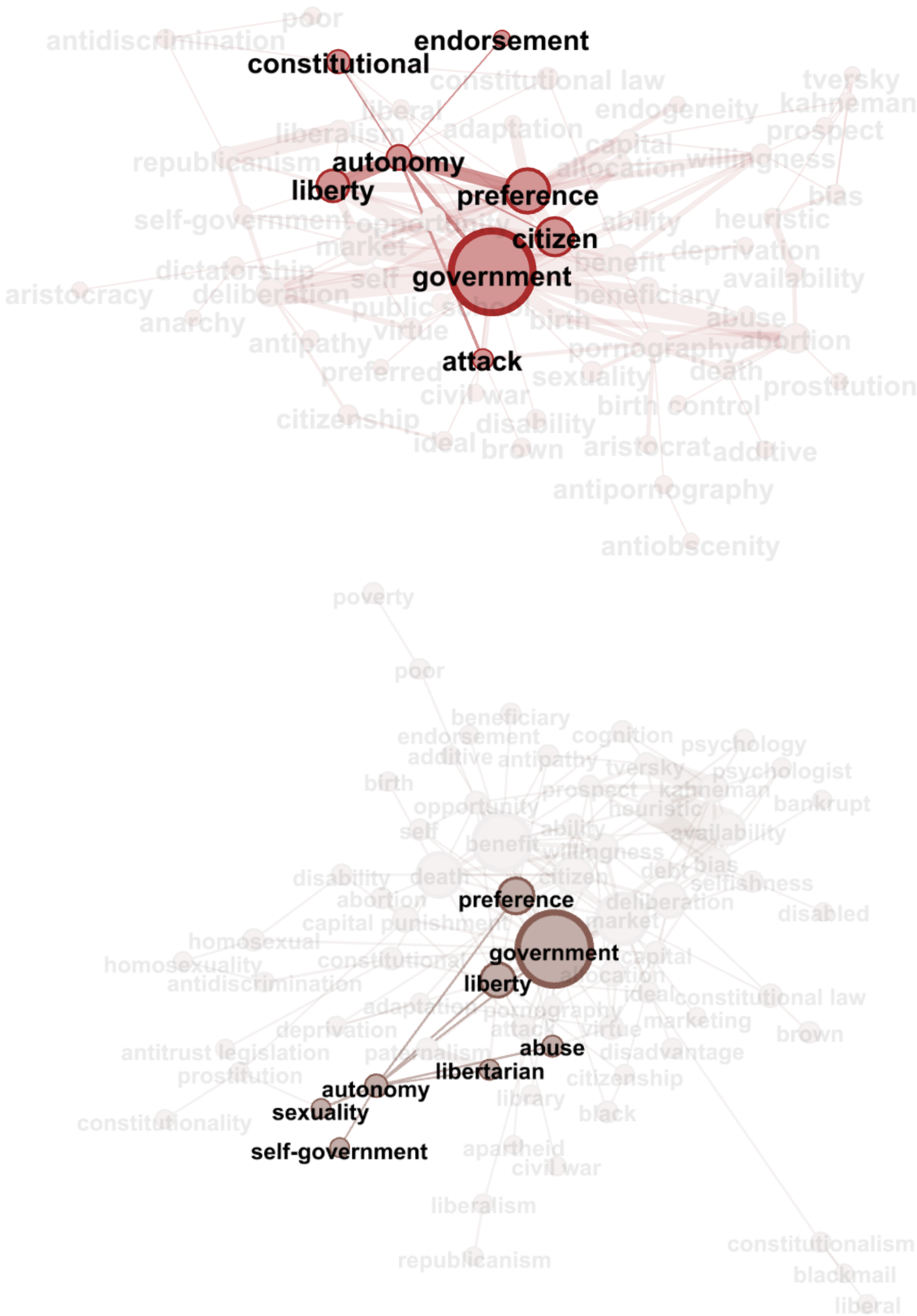


FIGURE C.3: Subgraph of the word "autonomy" before and after 1994, visualized with Gephi. We observe the weakening of the link with "preference", and the appearance of "libertarian".

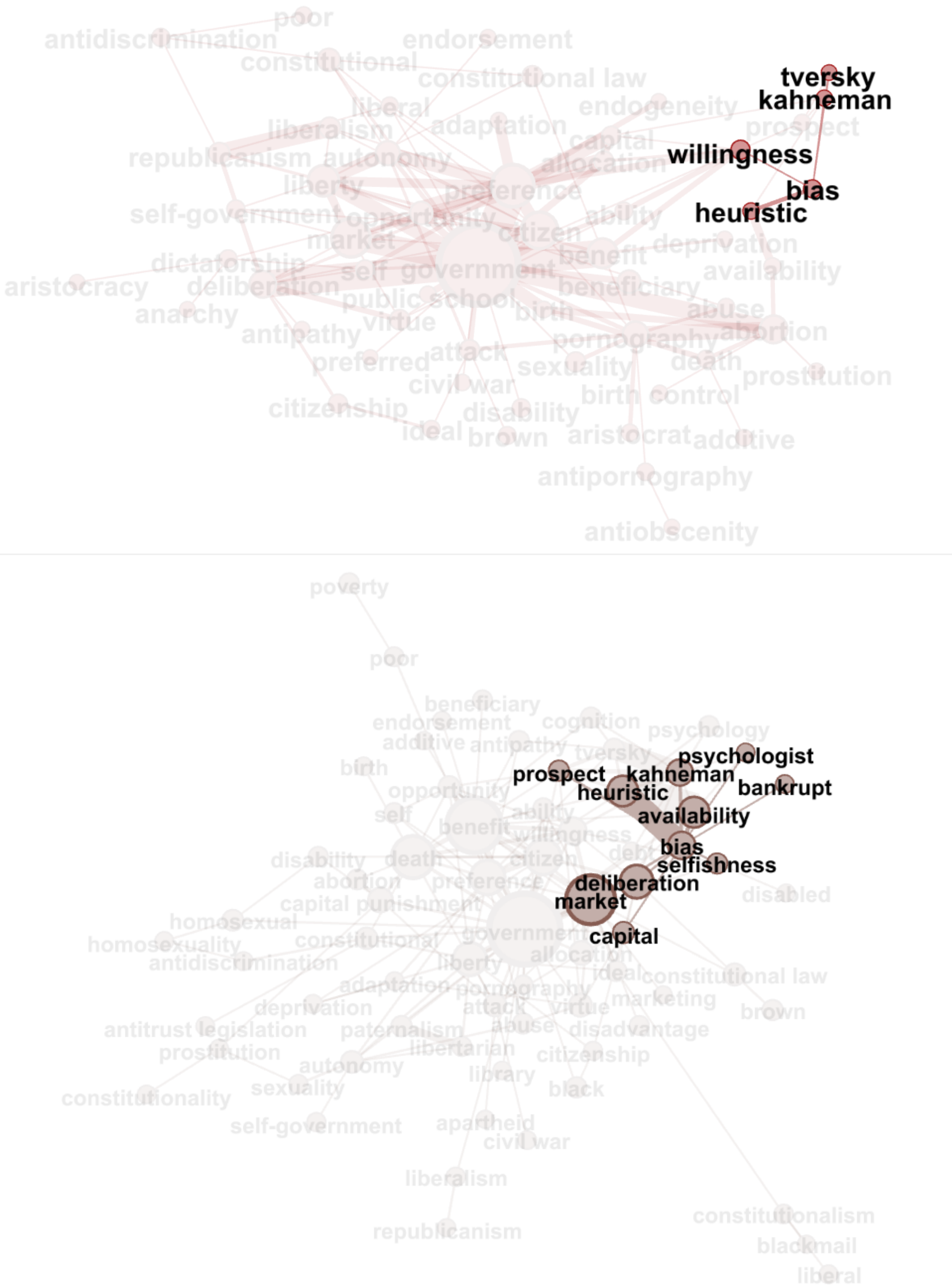


FIGURE C.4: Subgraph of the word "bias" before and after 1994, visualized with Gephi. The evolution is similar to that of "heuristic" in figure 2.10, with the appearance of links to "deliberation" and "market", which we interpret as a connection between the newly-discovered behavioral conception of economics and existing topics of political theory.

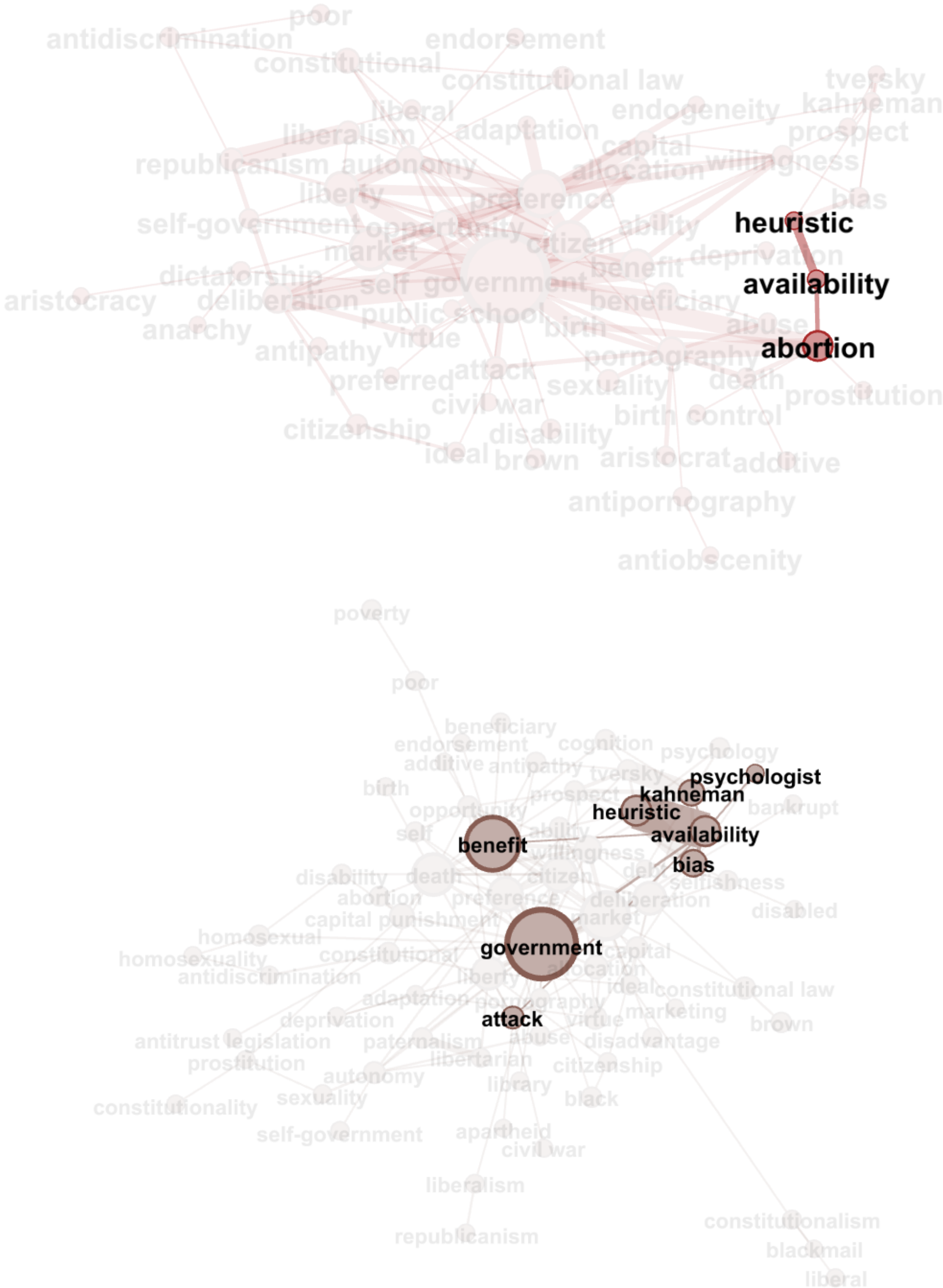


FIGURE C.5: Subgraph of the word "availability" before and after 1994, visualized with Gephi. We note the large development of the word's subgraph, as well as the establishing of a new connection with the central element "government".

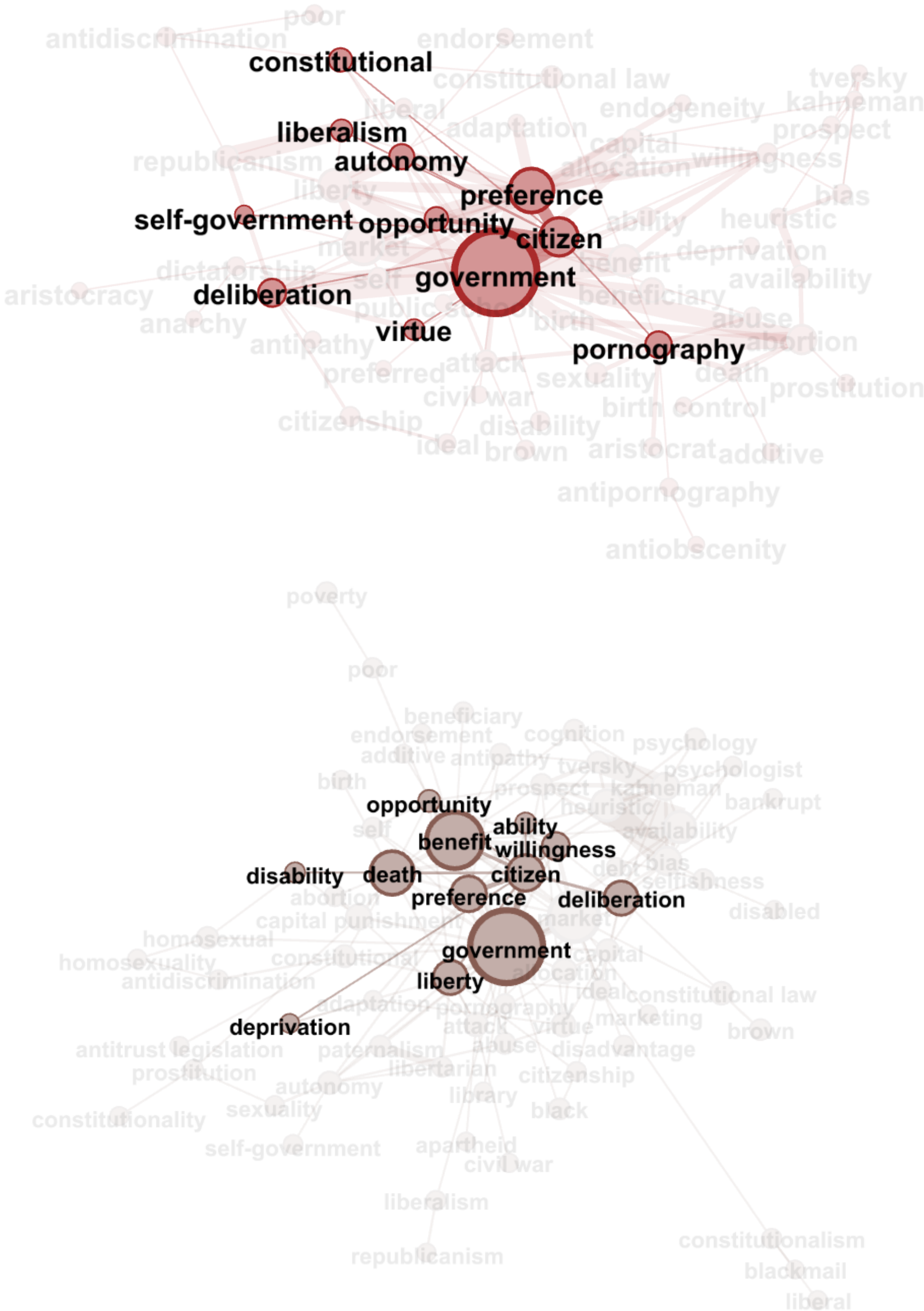


FIGURE C.6: Subgraph of the word "citizen" before and after 1994, visualized with Gephi. A noticeable evolution is the weakening of the link with "preference".

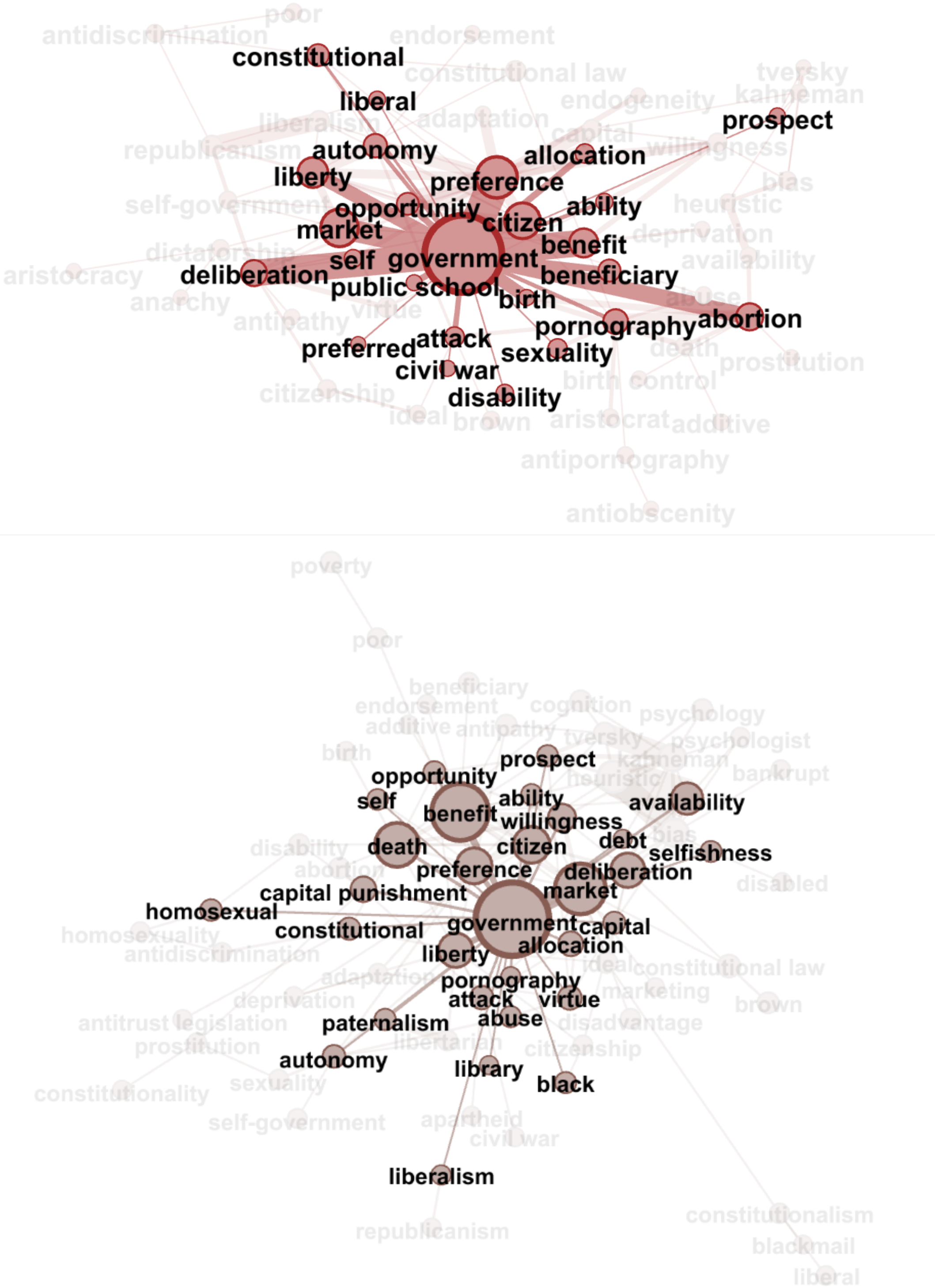


FIGURE C.7: Subgraph of the word "government" before and after 1994, visualized with Gephi. This central element is connected to several other terms related to prospect theory and libertarian paternalism after 1994.

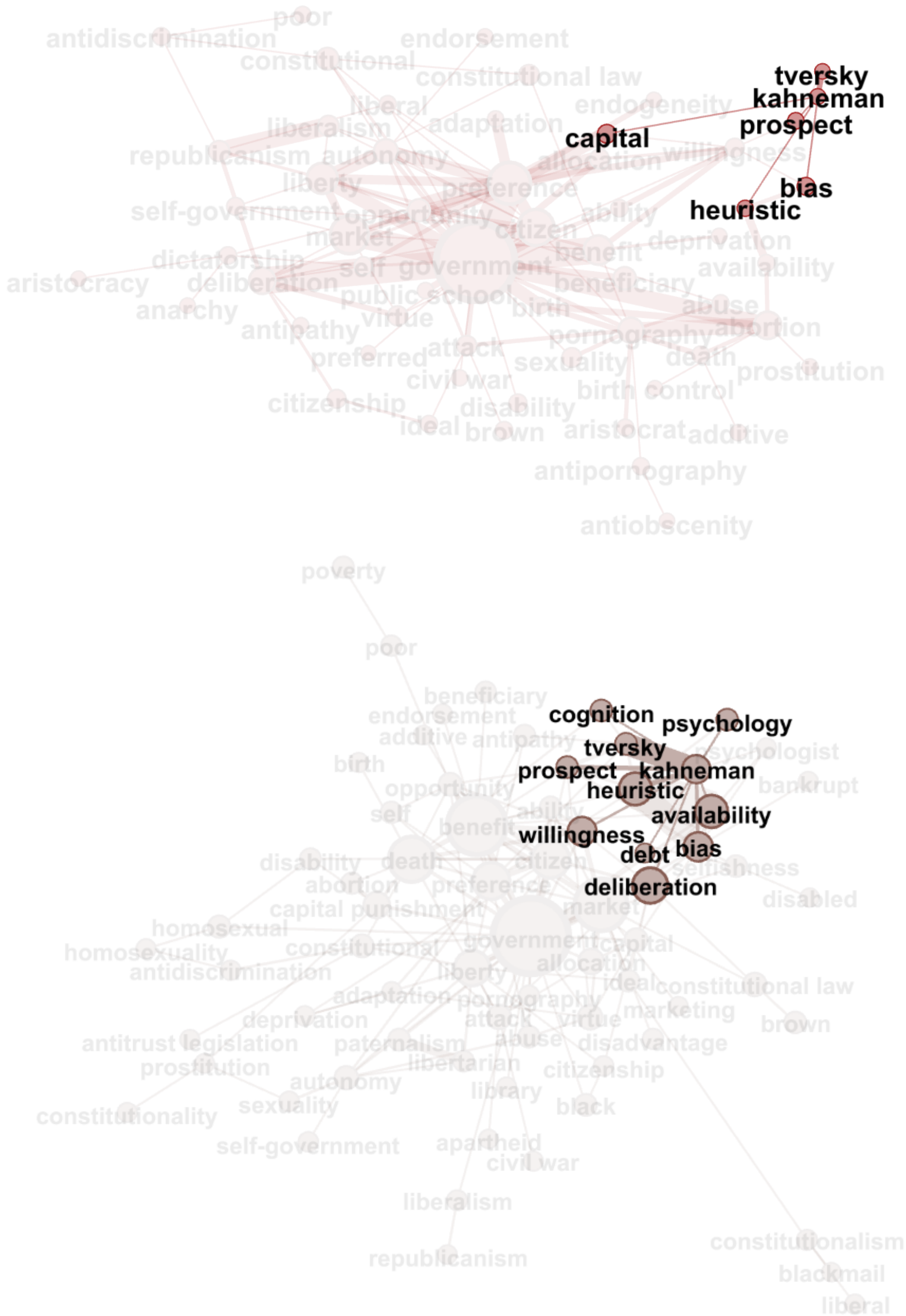


FIGURE C.8: Subgraph of the word "Kahneman" before and after 1994, visualized with Gephi. Note how the initial subgraph is weakly connected to the main structure, and becomes an integral part of it later on.

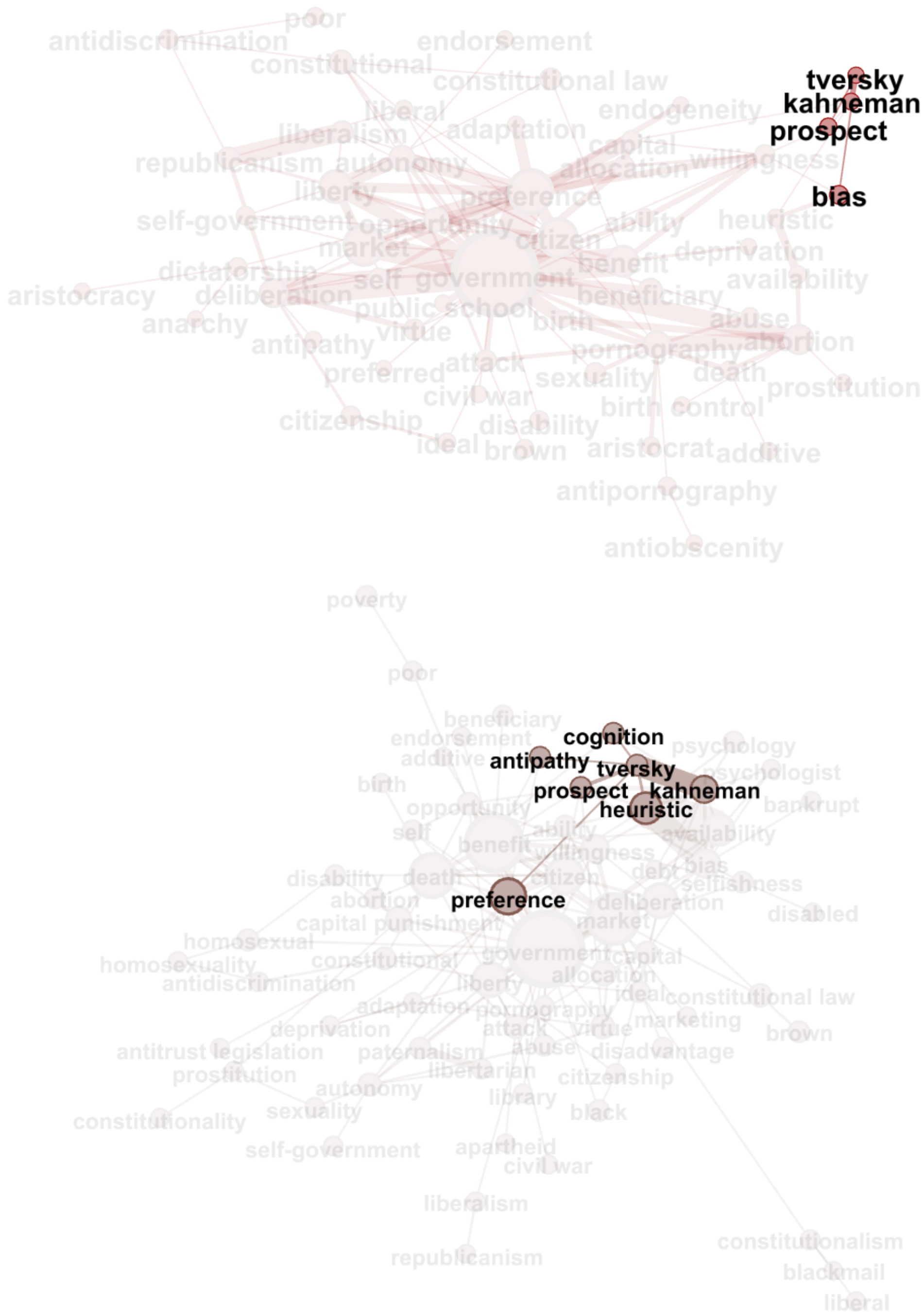


FIGURE C.9: Subgraph of the word "Tversky" before and after 1994, visualized with Gephi. The analysis is similar to that of "Kahneman" in figure C.8.

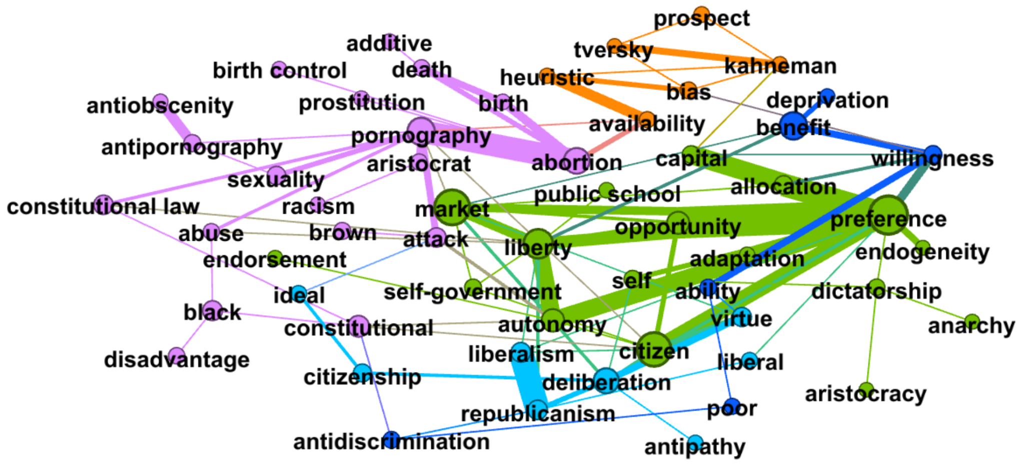


FIGURE C.10: Subgraph of the corpus before 1994, visualized with Gephi, with color-coded clusters.

FIGURE C.11: Subgraph of the corpus after 1994, visualized with Gephi, with color-coded clusters.

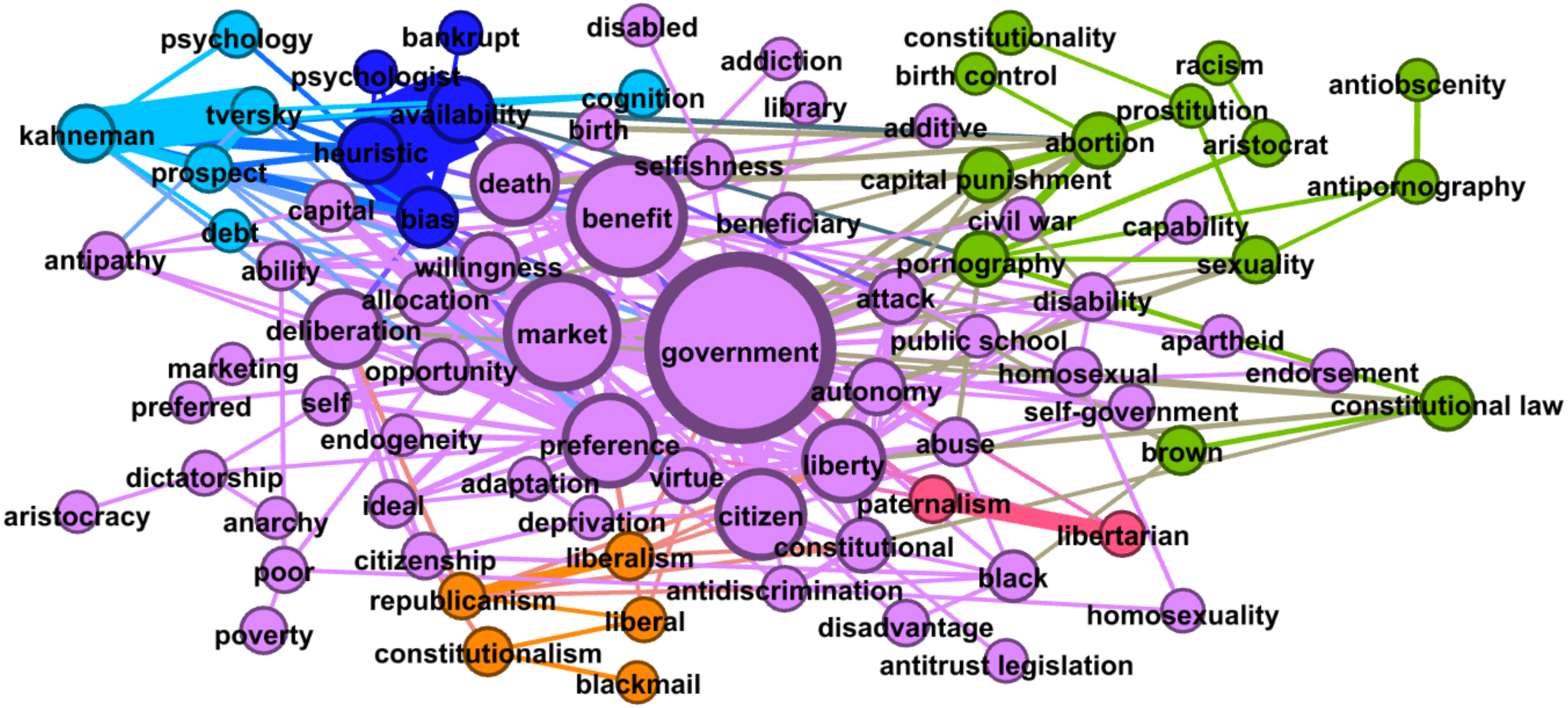


FIGURE C.12: Subgraph of the whole corpus, visualized with Gephi, with color-coded clusters.