Extrating Policy Positions from Political Texts Using Words as Data - M. Laver & K. Benoit

Given a set of texts about which something is known (Referenztexte), the technique extracts data from these in the form of word frequencies and uses this information to estimate the policy positions of texts about which nothing is known (virgin Texte)

- cross-validation of the policy estimates it generates against existing published results
    - To do this we reanalyze the text data set used by Laver and Garry (2000) in their dictionary-based computer-coded content analysis of the manifestos of British and Irish political parties at the times of the 1992 and 1997 elections in each country
        - ♠ this used to compare the results with published estimates of the policy positions of the authors of these texts generated by dictionarybased computer-coding, hand-coded content analyses, and completely independent expert surveys
- extend the application of the technique beyond the analysis of party manifestos, to the estimation of legislator positions from parliamentary speeches

A model for locating political Texts on A PRIORI policy dimensions

A Priori or Inductive Analyses of Policy Positions

A priori: expert surveys, which offer policy scales with predetermined meanings to country experts who are asked to locate parties on them

- example of this way of doing things can be seen in the dictionary-based computer-coding technique applied by Laver and Garry (2000), which applies a predefined dictionary to each word in a political text, yielding estimated positions on predefined policy dimensions.

*Inductive: Using content analysis,for example, observed patterns in texts can be used to generate a matrix of similarities and dissimilarities between the texts under investigation. This matrix is then used in some form of dimensional analysis to provide a spatial representation of the texts. The analyst then provides substantive meanings for the underlying policy dimensions of this derived space, and these **a posteriori** dimensions form the basis of subsequent interpretations of policy positions. This is the approach used by the CMP in its hand-coded content analysis

- purely inductive spatial analysis of the policy positions of political texts is impossible
    - all spatial analyses boil down to the estimation of policy positions on a priori policy dimensions interpreted.* **IMO zu vernachlässigen**

Nachteil: a posteriori interpretation -> lack of any objective criterion for deciding between rival spatial interpretations, in situations in which the precise choice of interpretation can be critical to the purpose at hand

Nachteil a priori: The price for taking the a priori route, is the need to accept take-it-or-leave-it propositions about the number and substantive meaning of the policy dimensions under investigation. Using the introduced a priori method however, this price can be drastically reduced. This is because, once texts have been processed, it is very easy to reestimate their positions on a new a priori dimension in which the analyst might be interested.

- For this reason we concentrate here on estimating positions on a priori policy dimensions

The Essence of Our A PRiORI Approach

Estimating policy positions by comparing two sets of political texts.

- On one hand is a set of texts whose policy positions on well-defined a priori dimensions are "known" to the analyst We call these "reference" texts
- On the other hand is a set of texts whose policy positions we do not know but want to find out. We call these "virgin"texts
    - All we do know about the virgin texts is the words we find in them, which we compare to the words we have observed in reference texts with "known" policy positions.
- More specifically, we use the relative frequencies we observe for each of the different words in each of the reference texts to calculate the probability that we are reading a particular reference text, given that we are reading a particular word.
- For a particular a priori policy dimension, this allows us to generate a numerical "score" for each word.
    - This score is the expected policy position of any text, given only that we are reading the single word in question. Scoring words in this way replaces the predefined deterministic coding dictionary of traditional computer-coding techniques. It gives words policy scores, not having determined or even considered their meanings in advance but, instead, by treating words purely as data associated with a set of reference texts whose policy positions can be confidently estimated or assumed.

=> In this sense the set of real-world reference texts replaces the "artificial" coding dictionary used by traditional computer-coding techniques.

We use the word scores we generate from the reference texts to estimate the positions of virgin texts on the policy dimensions in which we are interested. Essentially, each word scored in a virgin text gives us a small amount of information about which of the reference texts the virgin text most closely resembles.

- This produces a conditional expectation of the virgin text's policy position, and each scored word in a virgin text adds to this information. Our procedure can thus be thought of as a type of Bayesian reading

FALLBEISPIEL

The reference texts are the 1992 manifestos of the British Labour, Liberal Democrat (LD), and Conservative parties. The research task is to estimate the unknown policy positions revealed by the 1997 manifestos of the same parties, which are thus treated as virgin texts.

When performed by computer, this procedure is entirely automatic, following two key decisions by the analyst: the choice of a particularset of reference texts and the identification of an estimated or assumed position for each reference text on each policy dimension of interest

Selection of Reference Texts

The hard-and-fast rule when selecting reference texts is that we must have access to confident estimates of, or assumptions about, their positions on the policy dimensions under investigation.

Now this assumption is unlikely to be 100% correct, since the meaning and usage of words in party manifestos change over time

- In such instances it is possible to assume specific values for reference texts representing quintessential expressions of a view or policy whose position is known with a high degree of a priori confidence.

In other words, what we require for our set of reference texts is a set of estimates of, or assumptions about, policy positions that we are prepared to stand over and use as appropriate points of reference when analyzing the virgin texts in which we are ultimately interested.

We offer three further general **guidelines** in the selection of reference texts.

The first is that the reference texts should use the same lexicon, in the same context, as the virgin texts being analyzed.

The second guideline is that policy positions of the reference texts should "span" the dimensions in which we are interested.

- Trivially,if all reference texts have the same policy position on some dimension under investigation, then their content contains no information that can be used to distinguish between other texts on the same policy dimension.

The third general guideline is that the set of reference texts should contain as many different words as possible. The content of the virgin texts is analyzed in the context of the word universe of the reference texts. The more comprehensive this word universe, and thus the less often we find words in virgin texts that do not appear in any reference text, the better.

- The party manifestos that we analyze below are relatively long documents.

Generating Word Scores from Reference Texts

We begin with set R of reference texts, each having a policy position on dimension d that can be estimated or assumed with confidence.

We observe the relative frequency, as a proportion of the total number of words in the text, of each different word w used in reference text r.5 Let this be Fw

Swd is an average of the a priori reference text scores

Scoring Virgin Texts

Having calculated scores for all words in the word universe of the reference texts, the analysis of any set of virgin texts V of any size is very straightforward.

- First compute the relative frequency of each virgin text word, as a proportion of the total number of words in the virgin text. call this frequency F,V.
    - o This single numerical score represents the expected position of the virgin text on the a priori dimension under investigation. This inference is based on the assumption that the relative frequencies of word usage in the virgin texts arelinked to policy positions in the same way as the relative frequencies of word usage in the reference texts.

Interpreting Virgin Text Scores

To compare the virgin scores directly with the reference scores, therefore, we need to transform the scores of the virgin texts so that they have same dispersion metric as the reference texts. For each virgin text v on a dimension d

Estimating the Uncertainty of Text Scores

No previous political science work estimating policy positions using quantitative content analysis deals systematically with the uncertainty of any estimate generated.

This allows us for the first time to make systematic judgments about the extent to which differences between the estimated policy positions of two texts are in fact significant.

If we can compute a mean for any set of quantities, then we can also compute a variance

- The less this variance, the more the words in the text all correspond to the final score and hence the lower our uncertainty about that score.

Illustration Using a Sample Text

Estimating Economic Policy Positions of British and Irish Parties

British Party Positions on Economic Policy

The first stages in the analysis are to observe frequency counts for all words used in these reference texts12and to calculate relative word frequencies from these.

Using these relative frequencies and the reference text policy positions, we then calculated a word score on the economic policy dimension for every word used.

Having calculated word scores on the economic policy dimension for each of the 5,299 words used in the 1992 reference texts, we use these to estimate the positions of three "virgin" texts.

Bedeutung der Tabelle

The first row reports our estimated positions of the 1997 party manifestors, transformed to the same metric as the 1992 expert survey scores that were used as points of reference. Our firstpoint of comparison is with a set of 1997 expert survey scores, expressed in the same metric, highlighting the shift of the Labour Party to the center of this policy dimension

Table 2 also reports the standard errors associated with our raw estimates, from which we can conclude that differences among the estimated economic policy positions of the three manifestos are statistically significant.

To compare our results with those generated by other content analysis techniques, the last four rows in the top panel in Table2 report, in addition to our own estimates and those of the 1997 expert survey, two other textbased estimates of the 1997 economic policy positions of the British parties

Table 2 reports the mean absolute difference between the estimated positions of the parties on each standardized scale and the positions of the same parties in the expert survey.

=> This confirms our prima facie impression that our word-scored estimates are somewhat closer than the hand-coded content analysis to the expert survey estimate and are about as close to these as the more traditional dictionary-based computer-coded scale

Irish Party Positions on Economic Policy

Be that as it may, the results in Table 3 show that our approach, while generating results with good face validity in terms of subsequent coalition alignments, does not correspond as well as the other text-basedtechniques with expert survey.

- While this convergence is substantively plausible, an alternative possibility is that our estimates are less accurate than the others in this case.
- In other words, the standard errors generated by the word scoring technique are telling us that we should not feel as confident with its estimates for Ireland as we feel with those for Britain.

=> previous content analysis policy estimates of which we are aware report point estimates with no estimate whatsoever of associated error and, thus,are effectively blind to the potential problems arising from short texts we have diagnosed in the Irishcase.

Estimating the Policy Positions of British and Irish Parties on the liberal Conservative Social Policy Dimension

Traditional techniques of content analysis have been very much less effective at providing reliable and stable estimates of policy positions on this dimension, a conclusion confirmed in a careful study by McDonald and Mendes

In applying our word scoring approach to a new policy dimension, we also reveal one of its chief advantages of flexibility,ease of use, and susceptibility to tests using different a priori conditions. Once the reference texts have been converted into the matrix of word probabilities Pwr,it is straightforward to compute word scores for a new dimension d' simply by changing the a priori set of reference scores to Ard'.

We can then very easily apply these new word scores to the virgin texts and thereby estimate their positions on d', which in most cases takes under one second of computing time.

British Party Positions on Social Policy

These results, summarized by the mean absolute differences, show that computer word scoring performs extraordinarily well in this previously troublesome area, far better than any other content analysis technique.

Irish Party Positions on Social Policy

The mean absolute differences again summarize the relative performance of the three content analysis techniques. These show that our word scoring technique, despite the fact that it uses no knowledge of the English language, performs strikingly better than the other content analysis techniques, performing remarkably well on a dimension that has previously presented content analysts with considerable problems.

Overall fit with Expert Surveys

Coding Non English Language Texts

our approach is language-blind it should work equally well in other languages

Overall, we take these results to show that our word scoring technique can migrate effectively into a non-English-language environment.

Using the Word Scoring Technique to analyze Legislative Speeches

Moving beyond party politics, there is no reason the technique should not be used to score texts generated by participants in any policy debate of interest, whether these are bureaucratic policy documents, the transcripts of speeches, court opinions, or international treaties and agreements

Use word scoring to analyze legislative speeches. Although most legislatures have long preserved such speeches as part of the written parliamentary record, these speeches have highly amenable to computerized analysis as they are increasingly published electronically.

- differ substantially from partymanifestos in several key respects
    - First, manifestos are typically comprehensive documents addressing a wide range of policy issues, while speeches tend to be muchmore restricted in focus.
    - Second, manifestos are published in a political context that is fairlywell defined. Greater care must be taken in establishing the political context of speeches if we are to justify the comparison of different speeches in the same analysis.
    - Third, because manifestos and speeches use different language registers and lexicons, the analysis of speeches requires types of reference text different from those used in the analysis of manifestos.
- Finally, political speeches tend to be much shorter than manifestos. With fewer words to analyze, statistical confidence in the results is likely to be reduced

if successful, however, we would consider it a major confirmation of the ability of our technique to extract political positions from texts using word frequencies as data.

FALLBEISPIEL

For reference positions in the debate, we postulated a priori the location of certainpartyleaders on the "proversus antigovernment" dimension.

The speech of the Taoiseach (prime minister) was assumed axiomatically to be progovernment and assigned a reference position of +1.0. The speech of the Fine Gael leader of the day and leader of the opposition, JohnBruton, was assumed axiomatically to be antigovernment and assigned a reference position of -1.0, as was the speech of Prionsias de Rossa, then leader of the opposition Workers' Party.

This allowed the calculation of word scores for all different words used in the debate in at least one of the reference texts-a total of 2,856 different words in all.

Having calculated word scores from the reference texts, it was then possible to estimate the positions of 55 other speakers on the pro- versus antigovernment

=> This is because our approach treats words as data and reflects the greater uncertainty that arises from having fewer data.

Overall we consider the use of word scoring beyond the analysis of party manifestos to be a considerable success, reproducing party positions in a no-confidence debate using no more than the relative word frequencies in speeches.

This also demonstrates three important features of the word scoring technique.

- First, in a context where independent estimates of reference scores are not available, assuming reference text positions using substantive local knowledge may yield promising and sensible results

- Second, we demonstrate that our method quickly and effortlessly handles a large number of texts that would have presented a daunting task using traditional methods. Third, we see that the method works even when texts are relatively short and provides estimates of the increased uncertainty arising from having less data

Conclusions and future Work

Our word scoring approach to placing political texts on policy dimensions has been demonstrated to be effective at replicating the results of content analysis techniques based on human- or computer-coding

Second, unlike traditional methods of content analysis, our technique provides quantitative measures of uncertainty for text-based estimates of policy positions.

These allow analysts to make informed judgments, when comparing two estimated policy positions, about whether differences between them can be viewed as significant or merely as products of chance or measurement errorsomething that has not been possible before.

Finally, because it treats words simply as data rather than requiring any knowledge of their meaning as used in the text, our word scoring method works irrespective of the language in which the texts are written