Heike Klüver, University of Konstanz
Christine Mahoney, University of Virginia*

FRAMING POLICY DEBATES: MEASURING FRAMING STRATEGIES OF INTEREST GROUPS

Abstract: Framing plays an important role in politics as interest groups strategically highlight some aspects of policy proposals while ignoring others to shape policy debates in their favor. However, due to methodological difficulties we have remarkably little systematic data about the effectiveness of framing strategies. This article therefore proposes a new methodological approach for measuring interest group framing that is based on a quantitative text analysis of interest group position papers and official policy documents. We are able to identify the frames employed by interest groups and assess their effectiveness by studying to what extent decision-makers are responsive to these frames. We test this new approach by comparing the results with a hand-coded content analysis and two established quantitative text analysis techniques.

Prepared for the 2012 APPAM (Association for Public Policy Analysis & Management) Fall Research Conference will be held November 8-10 in Baltimore, MD

*The authors follow alphabetical order. Both authors have contributed equally to the paper.

1. Introduction

Since the time of Aristotle's delineation of argumentation into the categories *Ethos*, *Pathos* and *Logos*, scholars have struggled to understand the power of persuasion. This is particularly true in the literature on public policy debates where participants in the policy process have incentives to frame policy issues in different ways in order gain an advantage. William Riker (1986; 1996) focused our attention on the ability of individual protagonists in the policy process to destabilize debates through strategic framing, or *heresthetics* -- that is structuring the debate so that you can win. Baumgartner and Jones (1993; 2009) document countless examples of this in the U.S. during the post-War period. These authors refer to "policy images" and note how shifts from positive to negative images can change venues and drastically alter interest group mobilization patterns, with significant consequences for public policy outcomes. Observers of public policymaking have been keenly aware of the impact of framing on public policy outcomes but quantifying and systematically documenting that influence has remained elusive.[1]

James Druckman (2001) was one of the first to study the limits to framing (e.g., the fact that many people may be strongly resistant to efforts by others to reframe a debate), but his study focused on framing processes at the mass level. Baumgartner et al. (2009) have conducted the most extensive study to date of collective framing processes, covering a random sample of 98 policy issues in the United States. Their results suggest that few issues are reframed, at least in the short run. Following each of the 98 issues over a four-year period, they found that fewer than five percent of the issues were significantly reframed. Much more common were stable frames understood by all members of the professional community of lobbyists and policy-makers surrounding the issue. The

[1] Framing studies are common in several other literatures within political science, including media-effects studies in the field of public opinion (see for example Gilliam and Iyengar, 2000; Berinski and Kinder, 2006). Similarly, Lakoff (2002) focuses on how members of the public respond to different types of emotional stimuli. These literatures, focused on individual psychological or cognitive responses to public policy stimuli among members of the mass public, are related to but different from the focus within the field of public policy. This literature deals with elites, not members of the mass public and is less interested in the individual-level cognitive response to frames than in the overall nature of the debate as reflected in publicly available documents. We follow the public policy tradition here.

individual lobbyists might seek to reframe an issue, but none had the power to unilaterally redirect *collective* attention favourable to their own position. Mahoney (2008) similarly found limits to the ability of lobbyists to reframe issues in her study comparing interest-group lobbying in the US and the EU. In contrast to members of the mass public, elite participants in a policy community have highly detailed understandings of the various elements of a debate, even those with which they disagree. As a result, it is not easy to introduce a "new" element of debate to a group of experts.

However, even though these studies have considerably enhanced our understanding of the role of framing in the policy process, we have little systematic knowledge about the determinants of successful framing. Significant reframing may be uncommon, but the question still remains: which types of frames are more successful in realizing public policy outcomes that are in line with an advocate's preferences?

The lack of empirical evidence is due to methodological difficulties in measuring interest group framing. Previous research has mainly focused on interviews with lobbyists, bureaucrats and politicians in order to shed light on political argumentation (e.g. Mahoney, 2008; Baumgartner et al., 2009). While interviews are an important channel through which information about framing and lobbying can be gathered, interviews are costly and can only be conducted for a limited number of cases. In addition, claims made in interviews might be subject to bias and it also may be difficult for interviewees to recollect past events. In this study, we therefore propose a new methodological approach to study interest group framing that is based on quantitative text analysis. Using this approach, it is possible to identify frames employed by interest groups and to assess their effectiveness across a large number of cases.

Considerable ambiguity characterizes the literature on framing, as scholars in different disciplines or subfields of political science have used a variety of terms and definitions. Entman (1991: 53) has described *framing* as "selecting and highlighting some features of reality while omitting others". We follow Entman (1991) and define a *frame* accordingly as a specific aspect of a policy proposal that is emphasized in a policy debate. A *dimension* by

contrast refers to the underlying structure of conflict according to which interest groups can be aligned in a legislative debate. Policy proposals often have differential effects on different social and professional groups, different geographic regions and can be justified in many ways. Accordingly, Baumgartner and Jones (1993; , 2009) refer to "non-contradictory argumentation" in which proponents of environmental protection may for example discuss negative implications for water quality whereas, in the same legislative debate, business interests may focus on the loss of jobs. These advocates use different *frames* to make their point – the environmental advocates discussing habitat preservation and possibly a second frame focusing on public health implications of water pollution, and the business advocates emphasizing that the cost to business from implementing new regulations will result in large-scale layoffs, but they usually oppose each other on the same *dimension*: for or against new regulations.

In the following section, we briefly discuss the state of the art in quantitative text analysis. Part three illustrates the text analysis approach employed in this study, followed by a description of the research design and the results of the quantitative text analysis. Part six presents a cross-validation of the new approach by comparing the results to independent estimates before the concluding section, which summarizes the overall findings. This article presents the method on an example case. The research is part of a large-scale project, which will apply the technique to 100 issues being debated in the European Union between 2008-2010[2].

2. Text analysis in political science

Textual data is arguably the most widely available data source in the study of politics. Political documents have a great potential to reveal information about the positions, attitudes and activities of their authors at a precise point in time. Unlike interviews, texts can be analyzed as many times as one wishes and the information extracted from texts does not get less reliable as time passes. Content analysis was developed to make

[2] We acknowledge the support of the U.S. National Science Foundation (Award ID: 1102978 "Framing Policy Debates in the European Union")

systematic use of this rich data source. It is a "research technique for making replicable and valid inferences from texts (...) to the context of their use" (Krippendorf, 2004: 19). Political scientists have therefore quickly begun to use quantitative text analysis to gather systematic information about political actors and processes. The most important applications of quantitative text analysis in political science fall into two different categories: First, starting with the well-known Comparative Manifesto Project (CMP) (Budge et al., 2001; Klingemann et al., 2006), scholars have employed text analysis techniques to measure the positions of political actors (e.g. Laver and Garry, 2000; Laver et al., 2003; Slapin and Proksch, 2008). Second, another important application of quantitative text analysis in political science is the classification of documents into policy areas (e.g. Hillard et al., 2008; Grimmer, 2010; Quinn et al., 2010). These techniques allow for automatically classifying thousands of documents such as press releases, bills and speeches into issue areas.

However, the systematic analysis of framing has only received little attention in political science. Cheryl Schonhardt-Bailey has successfully used quantitative text analysis to study framing and dimensionality of parliamentary debates, speeches of presidential candidates and transcripts of the Federal Reserve's Federal Open Market Committee (Schonhardt-Bailey, 2005, , 2006; Bailey and Schonhardt-Bailey, 2008; Schonhardt-Bailey, 2008). More specifically, she has applied a combination of cluster and correspondence analysis in order to identify frames and to assess the dimensionality of political debates. Similarly, Weale et al. (2012) have analyzed legislative debates on abortion in the House of Commons relying on the same technique. While this quantitative text analysis approach has been successfully applied to speeches and committee deliberations, there are, to our knowledge, no studies that employ quantitative text analysis to study framing and political argumentation of interest groups.

3. Studying interest group framing using quantitative text analysis

This article proposes a new methodological approach that allows for measuring framing across a large number of cases. Following Schonhardt-Bailey (2008), we combine a cluster and a correspondence analysis which we apply to a text corpus made up of interest group submissions to legislative consultations as well as official policy documents produced by the political institutions. First, a cluster analysis is conducted in order to identify the frames employed by interest groups and second, a correspondence analysis is carried out which allows for assessing the dimensionality of policy debates. We use the software package T-LAB to carry out this analysis which provides all the necessary functions (Lancia, 2009).[3]

T-LAB relies on co-occurrence analysis, which is the statistical analysis of words that appear together in designated spans of a text corpus (here: documents). The underlying assumption justifying the use of co-occurrence analysis is that words that co-occur "in similar contexts tend to have similar meaning" and "documents that contain similar word patterns tend to have similar topics" (Lancia, 2007: 25). Using the presence or absence of words in each document, the program generates a term-document-matrix on which to base the classification process. This matrix contains documents in rows and the occurrence of words in each text in columns. Based on an unsupervised bisecting K-means algorithm, T-LAB then identifies clusters of documents in a bottom-up process (Steinbach et al., 2000). The clusters can be interpreted as frames used by various actors in a policy debate (see also Miller, 1997; Schonhardt-Bailey, 2005, , 2006; Bailey and Schonhardt-Bailey, 2008; Schonhardt-Bailey, 2008).

In a second step, correspondence analysis is used to assess the dimensionality of these frames. Correspondence analysis allows for spatial representation of the relationship between the clusters as distances in dimensions (for further details, see Greenacre, 1984). T-LAB cross-tabulates document clusters and words in order to create a second matrix that can be used for factor correspondence analysis. Correspondence analysis provides a

[3] Alternative software packages are ALCESTE (Image, 2009) or DTM VIC (Lebart, 2012). We decided to use T-Lab for the text analysis given that it is more transparent about the text analysis procedure and given that it is accompanied by a detailed documentation. Schonhardt-Bailey (2012) has shown that all three techniques arrive at essentially the same findings.

measure that indicates the amount of variance explained by the dimensions. It aims to account for a maximum amount of variance along the first dimension. The second dimension then seeks to account for a maximum amount of remaining variance and so forth. The correspondence analysis provides coordinates for individual interest groups, the frames and decision-makers in the (potentially) multidimensional policy space.

4. Research design

In order to test our proposed framing measurement approach, we conducted a case study in which we provide a detailed application of the text analysis technique and test its validity by comparing the results to independent estimates. More specifically, we analyzed the policy debate surrounding the 2007 proposal of the European Commission on reduction of CO_2 emissions from cars. We selected this legislative debate as Klüver (2009) has already analyzed this policy debate by means of Wordfish, Wordscores and manual hand-coding. We are therefore able to compare the results obtained by our framing analysis with independent estimates obtained by three other established methods and to therefore cross-validate our findings.

In order to reduce CO_2 emissions from cars, the European Commission launched a legislative initiative in February 2007 by adopting a Communication which laid out the plans for an upcoming proposal on this issue. The European Commission suggested a variety of measures to reduce automobile CO_2 emissions to 120g/km until 2012 such as fiscal measures, an increase in the use of biofuels and a code of good practice on car advertising. On the basis of the Communication setting out the envisaged measures, the European Commission launched a public consultation in which stakeholders could submit position papers expressing their views on the proposed legislative framework. The consultation closed in July 2007 and the European Commission adopted its official legislative proposal in December 2007. By comparing the interest group position papers with the Communication and the final Commission proposal, we examine the framing strategies and their effectiveness during the policy formulation stage.

A variety of different interest groups participated in the consultation on the CO_2 emissions framework such as car manufacturers, environmental NGOs or representatives of the press industry. The policy debate on the reduction of CO_2 emissions is therefore not only an ideal case for testing the applicability of our framing approach as independent position estimates are available to cross-check our results, but also as the heterogeneity of interest groups participating in this debate implies that a wide variety of arguments are employed that can be captured using the framing analysis.

The frames employed by interest groups were extracted from their submissions in the public consultation. In order to assess which frames were most successful in shaping the official position of the European Commission, we analyzed its Communication released in February and the preamble of the proposal which was adopted in December 2007. We solely focused on the preamble rather than taking the entire legislative proposal due to considerations related to the comparability of texts. The Communication is a continuous text drafted without any restrictions. By contrast, the legislative proposal follows strict structural guidelines and employs highly technical legal terminology. Given that quantitative text requires that documents use a similar pool of words, we therefore cannot use the entire legislative proposal for the computerized content analysis (Laver et al., 2003: 315). However, the preamble of the proposal is suitable for quantitative text analysis as it summarizes the content of the proposal and at the same time is written like a continuous text comparable to the Communication.

5. Quantitative text analysis

In this section, we present the framing analysis of the CO_2 emissions debate conducted with the proposed quantitative text analysis approach. The framing analysis includes the following steps: First, preparation of the text corpus. Second, cluster analysis to identify frames and third, correspondence analysis to examine the dimensionality of the policy debate and to locate interest groups and the European Commission in the policy space.

5.1. Text preparation

The preparation of documents includes various manual and automated steps. At first, all the documents have to be transformed to machine-readable text files. Second, all the text passages not directly referring to the policy debate have to be manually removed from the documents, such as contact details, repetition of consultation questions or quotations of the Commission document. Third, spelling errors have to be corrected. While these modifications are standard practice in quantitative text analysis, several additional modifications have to be conducted that are specific to T-Lab, the text analysis program which we employ to conduct the framing analysis. First, the documents have to be collapsed into one single file with each original document tagged with identification variables. Second, a list of key terms is selected on which basis the cluster and correspondence analysis is carried out. T-LAB allows for an automated or a manual selection of key terms. In this analysis, key terms were automatically selected on the basis of the Chi^2 criterion after stopwords have been removed from the documents.

5.2. Cluster analysis

After the text corpus had been prepared, a cluster analysis was conducted in order to identify the frames used in the analysis. Table 1 reports the most typical words per cluster according to their Chi^2 value. Three document clusters could be identified: The first and smallest cluster (12% of the documents) comprises texts using words such as "advertising", "press" and "media". The list of typical words of this frame (cluster) clearly indicates its focus on the impact of the legislative proposal on the press and advertising industry. The following excerpt from FAEP whose contribution is grouped into this cluster

underlines that this frame deals with the implications for the press and advertising business:[4]

> "**Publishers** would strongly oppose any political measure that has the potential to create an imbalance in the **advertising** revenues of the **press** as this would have a severe impact on the independence and diversity of the **press**" (Source: FAEP).

TABLE 1 ABOUT HERE

The second cluster, which encompasses 28% of the documents, is marked by words such as "automotive", "segments" or "product". The key words show that this cluster comprises documents emphasizing the negative impact of the proposal on the automobile manufacturers. This frame is nicely illustrated in the following text passage taken from VDA which is the German automobile manufacturers association:

> "A policy discriminating against premium vehicles would damage a key area for generating **value** added and employment in the European **automotive** industry, and primarily in the German **automotive** industry" (Source: VDA).

The third and largest cluster (60% of the documents) is represented by typical words such as "LPG", "biodiesel" and "natural". Further analysis using the keyword-in-context function of the open source text analysis program *Yoshikoder* (Lowe, 2009) reveals that these terms are used to discuss the negative effects of global warming on the environment and to highlight the environmental superiority of alternative technologies such as hybrid or electric cars as well as biofuels. The following excerpt from the contribution of Transport & Environment which is an environmental NGO highlights the nature of this frame:

> "Legislation on CO_2 from cars will oblige car makers to implement CO_2 saving technology on their vehicles. (...) They appear not to be willing to pay to avoid climate change, and do not even consider lifetime **fuel** savings, even if to do so would be in their own best interests. (...) CO_2 regulation will lead to a quicker

[4] In the following excerpts, only the 15 terms with the highest Chi2 value are marked in bold. Further terms related to these clusters are not marked.

> and more widespread adoption of **fuel** saving technology across Europe's car fleet. (...) Strong regulation will to slow climate change, strongly reduce our oil bill and bring high tech development to Europe" (Source: T&E).

In a second step, we compared a manual coding of actor type to the clusters in which the program classified the contributions of each interest group. The underlying assumption of comparing cluster membership and actor type is that interest groups of the same kind should use similar framing strategies. For instance, it is plausible to assume that environmental groups would rely on a similar framing strategy as they share the same fundamental policy goal which is the protection of the environment. Drawing on information gathered from the interest group submissions and websites concerning their interests and their organizational structure, we coded the interest groups into five different categories: Traditional automobile industry groups which represent the interests of the car manufacturers (n=6); Alternative automobile industry groups which amongst others represent the interests of manufacturers of electric cars and producers of biofuels (n=4); Environmental non-governmental organizations (NGOs) fighting for environmental protection (n=6); Press groups representing the interests of the print media and the advertising industry (n=2) and a variety of other groups (n=5).

Table 2 compares the clusters obtained by the text analysis with the coded group type in order to assess the validity of the measurement. Each row represents an interest group together with the cluster membership of the document it submitted to the consultation. Cluster scores represent the degree to which each document is a member of the various clusters as well as the best cluster solution according to these scores. The results show that the automated identification of clusters corresponds very strongly (though not perfectly) with a manual coding of group type. Of all the interest groups classified into substantial actor type categories, 89 per cent were grouped into the same clusters as their fellow groups.[5] For instance, all environmental NGOs except for FOE (Friends of the Earth) were

[5] Substantial categories refer to traditional automobile industry groups, alternative industry groups, environmental NGOs and press groups. All interest groups coded as "other" groups are excluded as this

grouped into the environmental cluster. The cluster analysis also classified the initial (Comm1) and final (Comm2) Commission document. The cluster membership scores of the Commission have changed over time. While the membership in the environment cluster has decreased from 0.42 to 0.38, the membership in the industry cluster has increased from 0.36 to 0.39. The cluster analysis therefore indicates that the European Commission was responsive to the industry frame at the expense of the environment frame.

TABLE 2 ABOUT HERE

5.3. Correspondence analysis

In a second step, the underlying dimensions of the frames are identified and the frames are mapped spatially using correspondence analysis. Correspondence analysis provides a measure which indicates the amount of association explained by the underlying dimensions. The correspondence analysis identifies a two-dimensional space in which the frames are located (see figure 1 and table 3). The first dimension displayed on the x-axis of figure 1 accounts for 58 per cent and the second dimension accounts for 42 per cent of the variance. Whereas the "environmental" and "industry" frames mainly oppose each other on the first dimension and hardly differ in respect to the second dimension, the "press" frame largely differs from the other two frames along the second dimension and is located more or less in the middle of the "environment" and "industry" frame on the first dimension. Hence, the debate surrounding the policy proposal on the reduction of CO_2 emissions from cars is characterized by a two-dimensional space.

FIGURE 1 ABOUT HERE

Finally, as we have two measures of the location of the European Commission, we can assess the direction of any movement in the official position. In this example, we compare the initial location (t_0) of the February 2007 Communication and the final location (t_1) of

category includes a wide variety of different groups and it is therefore not plausible that they use the same framing strategies.

the legislative proposal adopted by the Commission in December of that same year, after the consultation materials described above had been submitted and reviewed. These are indicated in figure 1 as COMM1 and COMM2 and the movement between the two time points is marked by an arrow. Figure 1 shows that the European Commission moved primarily on the first dimension away from the environmental and alternative industry groups towards the traditional automobile industry.

TABLE 3 ABOUT HERE

Using the coordinates in the multidimensional space (see table 4), one can furthermore compute the distance between the European Commission and the frames (clusters) or even the single interest groups to assess in which direction the Commission moved from t_0 to t_1. We calculate the distance between the European Commission and the three frames along the two dimensions of debate (see table 5). The analysis shows that the press frame was most distant from the European Commission at the beginning as well as the end of the policy formulation debate. The environmental frame is the closest to the European Commission, but the Commission increased its distance from 0.14 to 0.26 over time. By contrast, the industry frame is closer to the European Commission at time point t_1 than at t_0. In conclusion, the Commission moved over time towards the center of the political space, aligning more with the traditional car industry than in its initial position, but remaining closer to the position of the environment cluster than to the others even at the end of the policy formulation process.

TABLE 4 ABOUT HERE

TABLE 5 ABOUT HERE

6. Cross-validating the results

In order to check the validity of proposed framing analysis approach, we compare the results obtained by the cluster and correspondence analysis with position estimates

obtained by Klüver (2009) using three established text analysis techniques: manual hand-coding, Wordfish and Wordcores. Since these techniques were used to estimate the policy positions rather than the framing strategies of interest groups, we are not able to cross-check the identification of frames based on the cluster analysis. However, given that the correspondence analysis produces coordinates of interest groups and the European Commission in the policy space, we are able to cross-validate these estimates. In the following section, we briefly illustrate the three text analysis techniques before we compare the estimates to our results.

6.1. Hand-coding

First, Klüver (2009) conducted a manually coded content analysis of the interest group and Commission documents. She first developed a coding scheme that comprises 41 categories out of which 20 each are classified as either "pro environmental control" or "anti environmental control" categories (see table 6). An additional "others" category was reserved for all those units of analysis that could not be allocate to any of the 40 substantial categories. All interest group and Commission documents were then divided into natural sentences which constitute the units of analysis (see also Däubler et al., forthcoming). Finally, all sentences were assigned to at least one of the 41 categories by human coders. Klüver (2009) obtained policy position estimates for all interest groups as well as the European Commission on an a priori defined "environmental control" dimension by subtracting the percentage of "pro environmental control" sentences from the percentage of the "anti environmental control" sentences.

TABLE 6 ABOUT HERE

6.2. Wordscores

Klüver (2009) moreover analyzed the CO_2 emissions debate using Wordscores developed by Laver et al. (2003). Wordscores is a fully automated text analysis program for measuring policy positions of texts. It relies on the basic idea that ideal points can be estimated by comparing two sets of texts: reference texts and virgin texts. The ideal points

of the reference texts are known quantities while the policy positions of virgin texts are unknown to the researcher. Laver et al. (2003) suggested to estimate the position of virgin texts by combining the knowledge about the reference texts with information about the words included in the reference and virgin texts. The researcher has to assign reference values to the reference texts on a single policy dimension based on external information such as expert surveys or position estimates provided by the Comparative manifesto project. Wordscores then computes a list of words based on the assigned ideal points of the reference texts and the relative frequency of the words they contain within and across the set of analyzed documents. The location of the virgin texts in the policy space is estimated as the mean of all their scored words weighted by the relative frequencies of these words in those texts.

6.3. Wordfish

Finally, Klüver (2009) additionally analyzed the CO_2 emissions debate using Wordfish (Proksch and Slapin, 2008; Slapin and Proksch, 2008). Wordfish is the latest innovation in quantitative text analysis that allows for estimating the ideal points of texts on a single dimension. Based on the simplifying assumption that words are distributed according to a Poisson distribution, Wordfish estimates the policy positions of texts on the basis of the relative frequencies of words within and across texts. It is a model-based approach which includes several parameters including a text fixed effect controlling for the length of the documents, a word fixed effect acknowledging that some words are used more frequently than others, a word specific weight capturing the importance of words in discriminating between policy positions and finally the policy position estimate. Wordish estimates the parameters of the model using an expectation maximization (EM) algorithm (for further details see Slapin and Proksch, 2008). The model is identified by setting the mean of all estimated policy positions to 0 and their standard deviation to 1.

6.4. Comparison of the estimates

Since manual Hand-coding, Wordscores and Wordfish estimate policy positions only for a single dimension at a time, we compare them separately to the two dimensions identified by our framing analysis (see table 7).[6] The estimates obtained by T-LAB strongly correlate with the hand-coding (r=0.76), the Wordfish (r=0.74) and the Wordscores (r=0.50) estimates on the first dimension. However, on the second dimension, the correlation between our framing analysis estimates and the results obtained by hand-coding, Wordfish and Wordscores is only moderate. The comparison of the different estimates shows therefore strong correspondence for the first dimension, but little for the second one. In addition, according to all three approaches employed by Klüver (2009), the European Commission moved towards the traditional automobile industry over the course of the policy debate which corresponds to the Commission movement detected by the correspondence analysis.

TABLE 7 ABOUT HERE

In order to further investigate the nature of the second dimension, we examined the distribution of text across the different categories defined for the manual hand-coding analysis as this allows for a more detailed account of the actual content of the documents. As the conflict on the second dimension largely takes place between the press and advertising industry on the one hand and all the other groups on the other hand (see figure 1), we focused hereby on the share of sentences that are devoted to the category "Code of good practice on car advertising" which deals with all matters related to press and the advertising business (see table 6). Table 8 presents information about the relative number of sentences of each document that falls into this policy category. There are primarily two interest groups, AAUK (UK Advertising Association) and FAEP (European Federation of Magazine Publishers) which talk about this issue in their submissions. AAUK spends about 48 per cent of its submission on this policy category while 61 per cent of the FAEP submission deals with the code of good practice on car advertising. By contrast, most of the

[6] The estimates provided by Klüver (2009) were multiplied by -1 as they located the interest groups in the reverse order in the policy space. This mathematical transformation was solely performed for illustrative purposes and has no effect on the substantial results.

other interest groups do not even devote a single sentence to this issue. Only FOE (Friends of the earth) mentions the code of good practice in about 18 per cent of its position paper.

TABLE 8 ABOUT HERE

The hand-coding analysis therefore indicates that the implications for the press is an important second dimension in the policy debate which is highly salient to two interest groups representing the press and advertising industry while most of the other interest groups do not even mention this issue in their submission. The classification of these two groups in the "press" cluster as well as the location of these groups in the policy space are therefore supported by the hand-coding analysis. While the conflict between FAEP and AAUK on the one hand and all the other groups on the other hand largely takes place on the second dimension, most other interest groups are indifferent to this dimension and primarily oppose each other on the first dimension which corresponds to the "environmental control dimension" identified by Klüver (2009). These results suggest that the T-LAB analysis accurately estimates the location of the interest groups and the European Commission on the main dimension, but that it additionally allows to detect cleavages not apparent to any technique assuming a single dimension.

7. Conclusion

The way political actors frame a policy initiative can have significant implications for the outcome of a legislative debate. Interest groups therefore strategically use political rhetoric to shape a political debate in a way that strengthens their position on a legislative proposal. Even though framing is at the core of understanding political outcomes, we have very little systematic data on framing strategies of interest groups and their impact on public policy. The lack of research and data on interest group framing is largely due to methodological difficulties in systematically studying framing.

In order to overcome the shortcomings of the literature, this study has introduced a new methodological approach to the analysis of interest group framing. Drawing on recent advances in quantitative text analysis, we have employed a combination of cluster and correspondence analysis to identify the frames put forward by interest groups and to assess the dimensionality of policy debates. By studying interest group position papers as well as official legislative documents adopted by decision-makers, we were furthermore able to empirically evaluate the success of different frames with regard to shaping the outcome of legislative debates. We have tested this new methodological approach on the basis of a policy debate launched by the European Commission in 2007 that dealt with the reduction of CO_2 emissions from cars.

As quantitative text analysis has not been applied before in the study of interest group framing, we have cross-validated our framing analysis by comparing our results with independent estimates obtained by Klüver (2009). Our results strongly correlate with estimates computed by manual hand-coding which is associated with a high degree of validity as well as with the two established automated text analysis techniques Wordfish and Wordscores. However, while the manual content analysis as well as Wordfish and Wordscores are restricted to examining one dimension, the methodological approach introduced in this study allows for multiple dimensions. As a result, while the T-LAB analysis presented in this article picked up two different dimensions, the estimates provided by Klüver (2009) neglect the second dimension.

The results of this study have major implications for the study of interest group framing. While methodological difficulties have largely prevented scholars from examining framing strategies and the impact of framing on public policy, the approach introduced in this study paves the way for the large-scale analysis of interest group framing.

REFERENCES

Bailey A and Schonhardt-Bailey C (2008) Does deliberation matter in FOMC monetary policymaking? The Volcker Revolution of 1979. *Political Analysis* 16(4): 404-427.

Baumgartner F R, Berry J M, Hojnacki M, Kimball D C and Leech B L (2009) *Lobbying and Policy Change: Who Wins, Who Loses, and Why*. Chicago: University of Chicago Press.

Baumgartner F R and Jones B D (1993) *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.

Baumgartner F R and Jones B D (2009) *Agendas and Instability in American Politics*. Chicago: University of Chicago Press.

Berinski A J and Kinder D R (2006) Making Sense of Issues Through Media Frames: Understanding the Kosovo Crisis. *Journal of Politics* 68(3): 640-656.

Budge I, Klingemann H-D, Volkens A, Bara J and Tanenbaum E (2001) *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945-1998*. Oxford: Oxford University Press.

Däubler T, Benoit K, Mikhaylov S and Laver M (forthcoming) Natural Sentences as Valid Units for Coded Political Texts. *British Journal of Political Science*

Druckman J N (2001) On the Limits of Framing Effects: Who Can Frame? *Journal of Politics* 63(4): 1041-1066.

Entman R M (1991) Framing US coverage of international news: Contrasts in narratives of the KAL and Iran air incidents. *Journal of Communication* 43(4): 51-58.

Gilliam F D J and Iyengar S I (2000) Prime Suspects: The Influence of Local Television News on the Viewing Public. *American Journal of Political Science* 44(3): 560-573.

Greenacre M J (1984) *Theory and application of correspondence analysis*. London: Academic Press.

Grimmer J (2010) A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Releases. *Political Analysis* 18(1): 1-35.

Hillard D, Purpura S and Wilkerson J (2008) Computer Assisted Topic Classification for Mixed Methods Social Science Research. *Journal of Information Technology and Politics* 4(4): 31-46.

Image (2009) ALCESTE, Logiciel d'Analyse de Données Textuelles, Version 4.7, http://www.image-zafar.com/, last accessed on 4 May 2009.

Klingemann H-D, Volkens A, Bara J L, Budge I and McDonald M D (2006) *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union, and OECD 1990-2003*. Oxford: Oxford University Press.

Klüver H (2009) Measuring interest group influence using quantitative text analysis. *European Union Politics* 10(4): 535-549.

Krippendorf K (2004) *Content Analysis: An introduction to its methodology* (Second Edition ed.). Thousand Oaks: Sage.

Lakoff G (2002) *Moral Politics: How Liberals and Conservatives Thin* (2nd ed.). Chicago: University of Chicago Press.

Lancia F (2007) *Word co-occurrence and similarity in meaning: Some methodological issues*. URL (consulted June 2012): http://www.mytlab.com/wcsmeaning.pdf.

Lancia F (2009) T-LAB Pro, Tools for Text Analysis, Version 6.1., http://www.tlab.it, last accessed on 4 May 2009.

Laver M, Benoit K and Garry J (2003) Extracting policy positions from political texts using word as data. *American Political Science Review* 97(2): 311-331.

Laver M and Garry J (2000) Estimating Policy Positions from Political Texts. *American Journal of Political Science* 44(3): 619-634.

Lebart L (2012) *Software DtmVic: Exploratory statistical processing of complex data sets comprising both numerical and textual data*. URL (consulted June 2012): http://www.dtmvic.com.

Lowe W (2009) Yoshikoder, Software for multilingual content analysis, Version 0.6.3., http://www.yoshikoder.org/, last accessed on 4 May 2009.

Mahoney C (2008) *Brussels versus the beltway: Advocacy in the United States and the European Union*. Washington, D.C.: Georgetown University Press.

Miller M M (1997) Frame mapping and analysis of news coverage of contentious issues. *Social Science Computer Review* 15(4): 367-378.

Proksch S-O and Slapin J B. (2008). WORDFISH: Scaling Software for Estimating Political Positions from Texts, Version 1.2.

Quinn K M, Monroe B, Colaresi M, Crespin M and Radev D (2010) How to analyze political attention with minimal assuptions and costs. *American Journal of Political Science* 54(1): 209-228.

Riker W (1986) *The Art of Political Manipulation*. New Haven: Yale University Press.

Riker W (1996) *The Strategy of Rhetoric: Campaigning for the American Constitution*. New Haven: Yale University Press.

Schattschneider E E (1960) *The semisovereign people: A realist's view of democracy in America*. New York: Holt.

Schonhardt-Bailey C (2005) Measuring ideas more effectively: An analysis of Bush and Kerry's national security speeches. *PS: Political Science & Politics* 38(4): 701-711.

Schonhardt-Bailey C (2006) *From the Corn Laws to Free Trade: Interests, Ideas, and Institutions in Historical Perspective*. Cambridge: MIT Press.

Schonhardt-Bailey C (2008) The Congressional Debate on Partial-Birth Abortion: Constitutional Gravitas and Moral Passion. *British Journal of Political Science* 38(3): 383-410.

Schonhardt-Bailey C (2012) *Looking at Congressional Committee Deliberations from Different Perspectives: Is the Added Effort Worth It?* Paper presented at the ESRC Methods Festival: Oxford, 5 July 2012

Slapin J B and Proksch S-O (2008) A Scaling Model for Estimating Time-Series Party Positions from Texts. *American Journal of Political Science* 52(3): 705–722.

Steinbach M, Karypis G and Kumar V (2000) *A comparison of document clustering techniques*. Technical Report #00-034.: Department of Computer Science and Engineering, University of Minnesota.

Weale A, Bicquelet A and Bara J (2012) Debating Abortion, Deliberative Reciprocity and Parliamentary Advocacy. *Political Studies* 60(3): 643-667.

TABLES

Table 1: Most prominent words distinguishing clusters of actors in the CO_2 emissions debate

Rank according to Chi^2 Value	Cluster 1: Press	Cluster 2: Industry	Cluster 3: Environment
1	advertising	target	lpg
2	press	political	energy
3	media	value	gas
4	promotional	function	fuel
5	print	approach	fuels
6	literature	automotive	biodiesel
7	publishers	models	oil
8	survey	segments	fuelled
9	believe	reduction	duty
10	restrictions	product	natural
11	marketing	complementary	light
12	information	system	methane
13	claim	technologies	biogas
14	freedom	N1	biomethane
15	penalties	rental	diesel
No of texts	3	7	15
% of texts	12%	28%	60%

Table 2: Comparison of actor type and cluster membership of interest groups

			Cluster membership scores		
Name	**Group type**	**Best solution**	**Press**	**Industry**	**Environment**
ADTS	Alt. Industry	Environment	0.11	0.31	0.58
AEGPL	Alt. Industry	Environment	0.10	0.19	0.72
EBB	Alt. Industry	Environment	0.13	0.23	0.64
ENGVA	Alt. Industry	Environment	0.09	0.19	0.73
COMM1	**Commission**	**Environment**	**0.22**	**0.36**	**0.42**
COMM2	**Commission**	**Industry**	**0.23**	**0.39**	**0.38**
FANC	Environ. Group	Environment	0.23	0.36	0.41
FOE	Environ. Group	Press	0.54	0.24	0.22
GREENPEACE	Environ. Group	Environment	0.23	0.35	0.43
RSPB	Environ. Group	Environment	0.25	0.35	0.41
TANDE	Environ. Group	Environment	0.27	0.31	0.43
WWF	Environ. Group	Environment	0.22	0.33	0.45
BEUC	Other	Industry	0.25	0.43	0.32
BVRLA	Other	Industry	0.19	0.54	0.27
ETRMA	Other	Environment	0.21	0.30	0.49
ETSC	Other	Environment	0.20	0.36	0.44
ETUC	Other	Industry	0.24	0.41	0.35
AAUK	Press	Press	0.68	0.16	0.16
FAEP	Press	Press	0.88	0.06	0.06
ACEA	Trad. Industry	Industry	0.18	0.56	0.26
JAMA	Trad. Industry	Industry	0.19	0.55	0.26
KAMA	Trad. Industry	Industry	0.19	0.53	0.28
RAI	Trad. Industry	Environment	0.22	0.36	0.43
SMMT	Trad. Industry	Industry	0.24	0.46	0.30
VDA	Trad. Industry	Industry	0.15	0.60	0.25

Note: Full names of the associations are available in the codebook accompanying the replication dataset.

Table 3: Results of correspondence analysis

	Eigenvalue	Percentage	Cumul. Percentage
Factor 1 (x-axis)	0.19	58.00	58.00
Factor 2 (y-axis)	0.14	42.00	100.00

Table 4: Location of interest groups and the European Commission

Actor	Dimension 1	Dimension 2
PRESS CLUSTER	0.01	1.77
INDUSTRY CLUSTER	-0.60	-0.08
ENVIRONMENT CLUSTER	0.34	-0.08
COMMISSION t_0	0.21	-0.03
COMMISSION t_1	0.08	-0.05
AAUK	0.04	1.36
ACEA	-0.60	-0.08
ADTS	0.36	-0.20
AEGPL	0.38	-0.13
BEUC	-0.15	0.03
BVRLA	-0.34	0.03
EBB	0.40	-0.12
ENGVA	0.55	-0.19
ETRMA	0.37	-0.10
ETSC	0.20	-0.09
ETUC	-0.04	0.00
FAEP	-0.05	2.54
FANC	0.16	-0.05
FOE	0.02	0.81
GREENPEACE	0.20	-0.05
JAMA	-0.56	-0.06
KAMA	-0.52	-0.08
RAI	0.27	-0.08
RSPB	0.19	0.02
SMMT	-0.32	0.02
TANDE	0.26	0.05
VDA	-0.68	-0.11
WWF	0.24	-0.05

Table 5: Relative distance of the European Commission from frames in the debate, t_0 and t_1

Frames	Distance at t_0	Distance at t_1	Success
Press	1.81	1.82	-0.01
Industry	0.81	0.68	+0.13
Environment	0.14	0.26	-0.12

Table 6: Hand-coding classification scheme

Overall category	Environmental control	
	Pro	**Anti**
Reduction target	positive	negative
Appropriateness of measure	positive	negative
Inclusion of vans	positive	negative
Code of good practice on car advertising	positive	negative
Improved labeling to promote the purchase of fuel-efficient cars	positive	negative
Fiscal measures to promote the purchase of fuel-efficient cars	positive	negative
Penalties to enforce CO_2 reductions	positive	negative
Efficiency improvements of tyres	positive	negative
Efficiency improvements of air conditioning	positive	negative
Greater use of alternative fuels or automotive technology	positive	negative
Long-term reduction strategy	positive	negative
Averaging	negative	positive
Pooling	negative	positive
Banking	negative	positive
Individual targets for small-scale manufacturers	negative	positive
Exceptions for special purpose vehicles	negative	positive
Weight as parameter for calculating reduction targets	negative	positive
Inclusion of CO_2 reduction from cars into general Emissions Trading Regime	negative	positive
Monitoring	positive	negative
Crediting	negative	positive
Other	-	-

Source: Klüver (2009)

Table 7: Correlation of T-LAB coordinates with Hand-coding, Wordfish and Wordscores

		T-LAB
Dimension 1	**Hand-Coding**	0.76***
	Wordfish	0.74***
	Wordscores	0.50**
Dimension 2	**Hand-Coding**	0.34*
	Wordfish	0.05
	Wordscores	0.10

*p<0.1, **p<0.05, ***p<0.01. Source of Hand-Coding, Wordfish and Wordscores estimates: Klüver (2009)

Table 8: Investigation of second dimension

	Percentage of sentences devoted to policy category "Code of good practice on car advertising"	
Actor	**Positive**	**Negative**
AAUK	**0.00**	**48.28**
ACEA	0.00	0.00
ADTS	0.00	0.00
AEGPL	0.00	0.00
AVELE	0.00	0.00
AVERE	0.00	0.00
BEUC	2.67	0.00
BVRLA	0.00	0.00
COMM 1	6.52	0.00
COMM 2	0.00	0.00
EBB	0.00	0.00
ENGVA	0.00	0.00
ETRMA	0.00	0.00
ETSC	0.00	0.00
ETUC	0.00	0.00
FAEP	**0.00**	**60.71**
FANC	4.08	0.00
FOE	18.46	0.00
GREENPEACE	0.00	0.00
JAMA	0.00	1.96
KAMA	0.00	0.00
RAI	0.00	0.00
RSPB	0.00	0.00
SMMT	0.00	0.69
T&E	2.34	0.00
VDA	0.00	0.00
WWF	1.06	0.00

FIGURES

Figure 1: Two-dimensional policy space of the CO_2 emissions debate