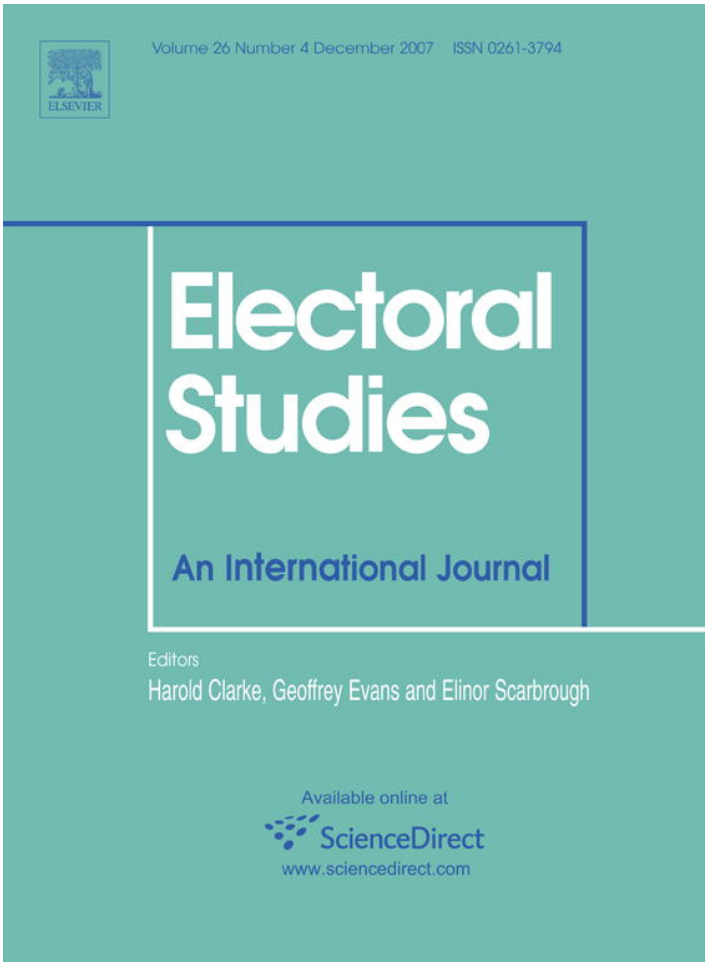
Provided for non-commercial research and education use.
Not for reproduction, distribution or commercial use.



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

http://www.elsevier.com/copyright

Author's personal copy

Available online at www.sciencedirect.com
ScienceDirect
Electoral Studies 26 (2007) 746–755
Electoral Studies
www.elsevier.com/locate/electstud

Estimating policy positions using political texts: An evaluation of the Wordscores approach

Robert Klemmensen [a], Sara Binzer Hobolt [b,*], Martin Ejnar Hansen [c]

[a] *Department of Political Science, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark*
[b] *Department of Politics and International Relations, University of Oxford, Manor Road, Oxford OX1 3UQ, United Kingdom*
[c] *Department of Political Science, University of Aarhus, Bartholins Allé, DK-8000 Aarhus C, Denmark*

Received 16 February 2007; revised 6 June 2007; accepted 28 July 2007

Abstract

This paper evaluates a recently developed method for extracting policy positions from political texts, known as *Wordscores*. This computerized content analysis technique is a potentially powerful tool for scholars interested in the study of political elites, since it promises an easy and efficient way of inferring policy position from texts and speeches. In this article, we provide a systematic evaluation of this promising method. Using Danish manifestos and government speeches from 1945 to 2005, we compare the policy positions extracted using Wordscores with measures of positions from the well-known Comparative Manifesto Project and cross-validate these with party expert surveys. Our analysis shows that the word scoring technique arrives at largely similar estimates to independently derived position measures and produces time series of government positions with high face validity.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Wordscores; Party positions; Comparative Manifesto Project; Measurement; Content analysis

1. Introduction

Studies of political competition often explicitly or implicitly refer to the positions of political actors on one or more dimensions. Deriving reliable and valid estimates of the policy positions of political actors is thus central to the study of political elites and electoral behaviour. It is therefore not surprising that there has been a surge in the literature on how to estimate policy positions. *Electoral Studies* recently devoted a special issue to the topic of how to estimate party positioning using expert, manifesto and survey data (see Marks, 2007). The hand-coded policy estimations of party manifestos provided by the Comparative Manifesto Project (CMP) are still the most widely used measure of party positions and the only data source that provides a long time series on these positions. Yet, while recent studies have highlighted problems with the CMP data, a new method for computerized content analysis, known as Wordscores, offers an alternative approach to extracting policy positions from political texts (see Laver et al., 2003). This has led to a lively debate on the pros and cons of the various methods (see Laver and Garry, 2000; Budge et al., 2001; Bakker et al., 2006; Budge and Pennings, 2007a,b; Benoit and Laver, 2007a,b). Advocates of the Wordscores technique have

* Corresponding author. Tel.: +44 1865 278830; fax: +44 1865 278725.
E-mail address: sara.hobolt@politics.ox.ac.uk (S.B. Hobolt).

0261-3794/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.electstud.2007.07.006

Author's personal copy

emphasized 'the potential for a huge increase in the scope and power of text analysis within political science' (Laver et al., 2003, p. 330), whereas its critics have argued that while the approach has great promise, 'the jury is still out on whether that will be fulfilled' (Budge and Pennings, 2007a, p. 129). Yet, the existing debate on Wordscores provides little clarity about the comparability, validity, and usefulness of this alternative method for estimating policy positions.

In this paper, we aim to provide a systematic evaluation of the Wordscores method set out by Laver et al. (2003) (hereafter LBG). To conduct a thorough test of the method, we apply the technique to both Danish party election manifestos and government speeches in the post-war period. We thus evaluate the extent to which the Wordscores technique can produce valid time series data on policy positions in a volatile party system. The challenge of evaluating any method for measuring policy positions is, of course, that we have no knowledge of the ''true'' positions of parties and other political actors. We therefore compare the Wordscores estimates with both CMP data and party expert surveys. This cross-validation enables us to present a more thorough and nuanced evaluation of Wordscores, which can guide other scholars in their work.

2. Approaches to measuring policy positions of parties

Several data sources provide information about the positioning of political parties, including party expert surveys, public opinion surveys, roll-call data and party manifestos. Party expert surveys, in which national party experts are asked to place political parties relative to one another on one or more dimensions, have proliferated in recent decades (see Morgan, 1976; Janda, 1980; Castles and Mair, 1984; Laver and Hunt, 1992; Huber and Inglehart, 1995; Marks and Steenbergen, 2002; Benoit and Laver, 2006). Studies have shown that party expert surveys are among the most valid measures of party positions at our disposal (see Benoit and Laver, 2006; Marks et al., 2007). Yet despite its advantages, this method has the serious drawback that it cannot produce a time series of party positions, since it is difficult to measure party positions retrospectively using party experts (Mair, 2001). Moreover, it is questionable whether the within-country scores can be used to make cross-national comparisons, given that the experts have the national context in mind when ranking parties.

The best-known time series data on party positions are derived from party election manifestos. Party manifestos have become widely used by scholars due to the impressive cross-national and cross-temporal data set provided by the CMP (previously known as the Manifesto Research Group), which now comprises party manifestos from the main political parties in 51 parliamentary democracies covering all democratic elections since 1945 (Budge et al., 2001; Klingemann et al., 2007). The content analysis of these manifestos is carried out by human coders using a classification scheme which is made up of 56 standard categories, measuring parties' views on a large range of issues from market regulation to anti-imperialism. This data set is unique in that it allows researchers to track the policy preferences of political parties over time and across countries, using official documents of party positions (Budge et al., 1987; Budge and Klingemann, 2001; Volkens, 2001).

However, despite the obvious advantages of the CMP data, several scholars have criticized the data set in recent years. It has been argued that the CMP data are flawed because they do not provide us with a measure of uncertainty associated with the position estimates (see Armstrong and Bakker, 2006; Bakker et al., 2006; Benoit and Laver, 2007a,b). Moreover, the theoretical foundation of CMP, namely the *salience theory* of party competition, has come under attack. According to the salience theory, political parties do not seek direct confrontation on issues, but engage in selective emphasis of certain issues (Budge and Farlie, 1983; Klingemann et al., 1994; Riker, 1996; Budge et al., 2001). On this basis, the CMP coding measures the relative issue emphases of parties. A left–right position measure is calculated by grouping issue categories into ''right'' and ''left'' categories and subtracting one from the other. Some CMP categories are bipolar (pro/anti) and therefore more explicitly positional, for example ''social services expansion: positive'' and ''social services expansion: negative.'' Yet, other CMP coding categories deal in a unipolar way with positional issues: for example, ''nationalisation,'' which is classified as a left-wing issue, or ''law and order'', which is classified as a right-wing issue. The CMP left–right measure includes a mix of unipolar and bipolar categories. The logic is that the more a party manifesto emphasizes left-wing issues, such as state involvement in the economy, the more left-wing the party is. But several scholars have questioned whether it is possible to equate salience (or selective issue emphasis) with positioning. In certain policy areas, such as economic redistribution, moral issues and European

Author's personal copy

integration, it is difficult to deduce a party's position from the emphasis attached to the issue in the manifesto (see Laver and Garry, 2000; Laver, 2001; Bakker et al., 2006). It has thus been argued that to understand party competition, we need to distinguish between position and emphasis. Yet, ultimately, the validity of the different assumptions about the nature of party competition is an empirical question (Gabel and Huber, 2000, p. 96; Laver and Garry, 2000, p. 620), and in the analysis section below the CMP data are compared with expert data and Wordscores estimates. First, we turn to a more detailed description of the Wordscores approach.

3. The Wordscores approach

The Wordscores technique is a very different and innovative approach to extracting policy positions from political texts, such as manifestos and speeches, using computerized content analysis. This technique, developed by Michael Laver, Kenneth Benoit and John Garry (see Benoit and Laver, 2003; Laver and Benoit, 2002; Laver et al., 2003) distinguishes itself from previous methods of establishing the positions as the first probabilistic technique for coding political text using neither predefined coding dictionaries nor subjective judgement calls by human coders. Rather than looking for the meaning in the text, Wordscores treats words as ''data'', based on the assumption that the relative frequencies of the use of specific words by political actors provide manifestations of underlying political positions. The technique extracts positions by comparing the patterns of words in a set of *reference texts* to establish the positions of other texts, the *virgin texts*. The word scoring technique requires estimates of the positions of the reference texts on the policy dimensions under investigation (e.g. the left–right dimension) to be assigned by the researcher, based on information such as expert surveys, mass survey data or context-specific knowledge. The Wordscores procedure then generates a list of words from chosen reference texts, based on the relative occurrence of each word across and within texts, given the set of reference scores. Point estimates on the original policy dimension are then generated for virgin texts, computed as the mean of the scores of the words in the virgin text, weighted by their relative frequencies within those texts. In addition to yielding point estimates for virgin texts, the procedure also computes confidence intervals.

This novel approach to estimating policy positions has several advantages. First, by using computerized coding, the reliability of the content analysis is significantly enhanced since the human factor is removed from the coding process. Secondly, the method is far easier and more flexible to implement than existing methods, such as CMP and cross-national expert surveys, which are both expensive and labour intensive. Wordscores can easily be implemented using either a command line version for Stata or a graphical version written in Java.[1] Wordscores thus enables researchers to extract positions from documents that are rarely used, such as political speeches. Thirdly, since the positions obtained through Wordscores are only dependent on the appropriateness of the reference texts, researchers do not need to have any knowledge of the language of the texts, according to LBG. Finally, the word scoring approach gives us a measure of the uncertainty associated with each position score.

Just like any other method, however, there are also potential pitfalls when using the Wordscores procedure. A common threat to the validity of the scores obtained using the text-based approaches is the lack of words. Regardless of whether we are interested in the position of political parties or the saliency of different issues, we are restricted by the number of words available in the texts we are analysing. The validity of the positions obtained by the Wordscores approach is furthermore dependent on the choice of reference text and the quality of the *a priori* scores attached to these reference texts. LBG emphasize that the choice of reference texts is crucial to the validity of the scores obtained and they point out that the reference texts have to share the properties of the virgin texts. Furthermore, LBG suggest that scholars use the so-called ''transformed scores'' when interpreting the raw estimates (''raw scores''). The transformed scores, which are produced along side the raw scores, are central because they permit a substantive interpretation of the results. However, in a recent paper Martin and Vanberg (in press) argue that the original transformation proposed by LBG suffers from two flaws. First, the LBG transformation is dependent on the combination of virgin texts which are included in the analysis. Second, Martin and Vanberg point out that whereas LBG do not score the reference texts, but assume them to be on the same metric as the virgin texts, it would be better to score both to make them truly comparable. To overcome these issues, Martin and Vanberg (in press) offer an alternative transformation of the raw Wordscores estimates. The main difference between the two transformation methods is that the Martin–Vanberg (MV) transformation provides scores for

[1] See http://www.wordscores.com/ for more details on how to download the software.

Author's personal copy

the reference text and places all texts on the original reference text scale, using the relative distances generated by the Wordscores raw scores. In this paper, we have used both the MV and the LBG transformations[2] (see Appendix 1).

The most critical assessment of the Wordscores approach was presented in a recent article by Budge and Pennings (2007a,b). In this article, they compare the Wordscores estimates unfavourably with their own CMP scores and argue that Wordscores ''flatten out'' party movements over time and that the Wordscores estimates are overtly sensitive to the choice of reference text. Yet, in response, Benoit and Laver (2007b) have argued that the assessment of Wordscores by Budge and Pennings is 'fundamentally flawed', because of their inaccurate implementation of the Wordscores approach and their treatment of CMP as the 'gold standard'.

In order to assess these rival claims, the remainder of this paper systematically applies the Wordscores technique to both manifestos and speeches, using different reference texts, and compares estimates from both of these text-based analyses, CMP and Wordscores, with independently derived party positions.

4. Estimating party positions

To evaluate these approaches empirically, we now turn to the analysis of the election manifestos of Danish political parties from 1945 to 2005; a total of 24 elections and 182 manifestos. The Danish party system provides an apposite case to test the Wordscores technique due to the many political parties and the radical transformation of the system in the ''earthquake election'' of 1973.[3] Moreover, the availability of expert data on Danish parties for the entire period allows us to cross-validate both the Wordscores estimates and the CMP estimates against the independently derived positions. Using expert data, we assess which of the two content analysis methods has the highest correspondence with the independently derived party positions. Substantially, we have restricted our analysis to the general left–right dimension, since this remains the dominant dimension in Danish politics (Andersen, 1984; Borre, 1995).

Table 1
Descriptive statistics for manifestos

	Virgin texts	Reference texts
Mean unique words (scored)	499 (280)	–
Mean total words (scored)	1809 (1946)	1726 (1273)
N	166	16

Note: standard deviations in brackets.

In Table 1, we show the descriptive statistics for the manifestos used in the analysis.[4] As shown, there is considerable variation in the length of the texts. The shortest manifesto of the Social Liberal Party in 1979 consists of only 178 words compared with the longest manifesto of 10,842 words by the Christian People's Party in 2001. This variation in word length obviously has consequences for the reliability of the estimates, yet, as LBG point out, one of the advantages of the Wordscores estimates is that they provides standard errors (Laver et al., 2003, p. 328). Hence, unlike most other methods, the estimates of uncertainty allow us to say whether differences between policy positions of texts are statistically significant, rather than product of measurement error or random variation.

To establish the position of each party on the left–right dimension by means of the Wordscores procedure, we have conducted the analysis using manifestos as reference texts. We have chosen the manifestos from two election years (1946 and 1975), before and after the 1973 upheaval in the Danish party system, to take into account changing issue emphases during the period. By using the same set of reference texts to estimate the positions of the manifestos from the entire post-war period, we can assess to what extent the Wordscores approach is capable of producing a genuine time series of the development in party positions.[5] The *a priori* codes of the reference texts were based on Damgaard's (2000) ordinal classification of the parties in these two elections.[6] Damgaard (2000) has placed the Danish political parties, using mainly party expert surveys and also roll-call analysis, in each election in the post-war period. These expert placements provide us

[2] All figures in this paper use MV transformations. See Appendix 1 for a comparison.

[3] The general election of December 1973 resulted in heavy losses for all the established parties represented in the Danish Parliament and several new parties entered Parliament, notably the centre-right Centre Democrats and the anti-tax Progress Party.

[4] We have only included parties which won representation in the Danish parliament in the analysis, since it was not possible to obtain independent position scores for all the parties actually competing in the elections. See Appendix 1.

[5] We have also run the same analysis using different sets of reference texts from other years, and the results are comparable, as long as we choose reference text from both before and after 1973.

[6] The exact codes we assigned to the parties in the 1947 election are: Com. 0, SD 3, SL 5, JP 7 Con. 8 Lib 10. In the 1975 election we used the following codes: Com 0, LS 1, SPP 2, SD 3 SL 5, CD 6, CPP 7 Lib 8, Con. 9, PP 10 (see Appendix 1 for party abbreviations).

Author's personal copy

with an apposite 'benchmark' with which to compare the two content analysis methods. For the 1945–1998 elections, we use data collected by Damgaard (2000), whereas we rely on the expert survey conducted by Benoit and Laver (2006) for the two most recent elections.[7]

To compare the CMP and Wordscores placements, we use the Spearman's Rho as a measure of fit between the party expert scores and the two content analysis methods.[8] Spearman's Rho measures the correlation between two ordinal scales (similar to Pearson's correlation of two interval variables), and thus enables us to compare the extent to which the two methods produce the same rank order of the parties as the party experts. The null hypothesis is that the variables are independent – hence that there is no association between the expert rank order and the rank order produced by the content analyses. The results are shown in Table 2.

As Table 2 shows, the CMP measure has the highest correspondence with the expert data. It correlates significantly with the expert ordering in 21 out of 23 elections. The Wordscores estimates perform a little worse correlating significantly with the expert ordering in 17 out of 22 elections (note that 1947 and 1975 are not scored because they are reference texts). Appendix 1 shows the Wordscores estimates obtained for each party included in the analysis with confidence intervals. It shows that 89% of the manifestos have scores which are significantly different from 0 indicating that their score is reliable. Shorter manifestos are significantly more likely to yield unreliable scores, compared with longer texts.[9] The CMP placements of shorter manifestos may, of course, suffer from similar problems, yet we do not know, given that CMP provides no measure of uncertainty associated with the positions.[10]

Hence, this analysis shows that when comparing with independent party expert judgements, the party position estimates of the CMP project perform better than the comparable Wordscores estimates, but the analysis also reveals a high degree of similarity between the two. Hence, this suggests that the Wordscores technique could potentially be a valuable method when researchers want to code documents that have not already been hand-coded such as, for example, political speeches. This will be discussed in more detail below.

Table 2
Spearman correlations with expert party placements

Year	Wordscores	CMP
1945	0.89**	0.89**
1947	–	0.61
1950	0.71	0.94***
1953a	0.71	0.83**
1953b	0.94***	0.77*
1957	0.83**	0.94***
1960	0.31	0.83**
1964	0.83**	0.94***
1966	0.89**	0.94***
1968	0.54	0.88**
1971	0.6	0.80*
1973	0.67**	0.73**
1975	–	0.81***
1977	0.83***	0.63**
1979	0.76**	0.68**
1981	0.63*	0.73**
1984	0.73**	0.85***
1987	0.95***	0.88***
1988	0.90***	0.81**
1990	0.71**	0.81**
1994	0.62*	0.95***
1998	0.85***	0.79***
2001	0.96***	0.79**
2005	0.86**	–

Note: Wordscores and CMP placements of parties are correlated with party expert rankings, provided by Damgaard (2000) and Benoit and Laver (2006).
***$p < 0.01$, **$p < 0.05$, *$p < 0.10$.

In Fig. 1 we have plotted the development in the positions of the four old political parties in the Danish party system – the Liberal Party, the Conservative Party, the Social Liberal Party and the Social Democratic Party – using the same CMP scores and Wordscores estimates as in Table 2. As shown in the figure, the Wordscores and CMP placements of parties are very comparable. As we would expect, both measures place the Social Democrats to the left of the Conservative and Liberal parties with the Social Liberal Party positioned in the middle. However, the Wordscores approach seems to allow for larger swings in party positions than the CMP approach in this analysis.

In the next section, we apply the Wordscores technique to government speeches in order to evaluate whether this approach can produce valid time series data on government positions.

[7] We have also run the analysis using a Chapel Hill's 2002 survey for the 2001 election (see Marks and Steenbergen, 2002). Using these expert data gives the same results, but since this survey includes fewer parties, we have chosen the Benoit and Laver surveys for both the 2001 and the 2005 elections.

[8] The main reason for using Spearman's Rho is that Damgaard provides an ordinal placement of the parties.

[9] Our analysis shows that manifestos with unreliable scores on average contain 1164 words, whereas the manifestos with reliable scores contain on average 1893 words.

[10] In a recent paper, Armstrong and Bakker (2006) have developed a method by which standard errors can be associated to the CMP scores.

Author's personal copy

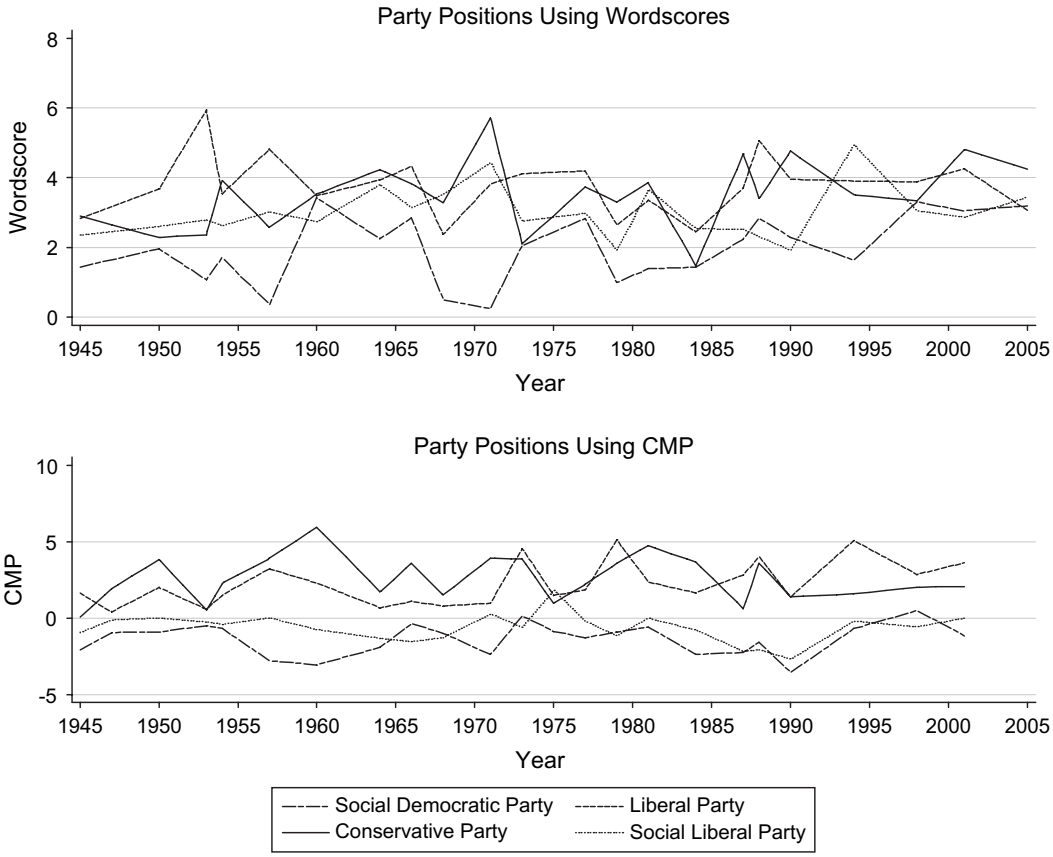


Fig. 1. Developments in party positions. Source: Danish Party Manifestos. Note: the Wordscores analysis uses the 1947 and 1975 manifestos as reference texts.

5. Locating the positions of governments over time

One of the major advantages of Wordscores is that this method can be used to extract policy positions from political documents other than party manifestos, such as, for example, political speeches. Political speeches have only recently been discovered as a valuable data source in the analysis of party positions (Laver and Benoit, 2002; Giannetti and Laver, 2005). The main benefit of using speeches to establish the position of political actors is that speeches are provided continuously over time, and not just prior to elections, and by a range of different political actors. By analysing speeches, we can examine a variety of research questions, such as how the government composition or the ideology of individual ministers affects public policy (see Giannetti and Laver, 2005; Laver et al., 2006). In this section, we apply the Wordscores procedure to the opening speeches in Parliament given by the Danish prime minister in order to evaluate whether this technique can generate meaningful estimates of changes in executive policy positions. In these speeches, the government's policy agenda for the forthcoming legislative session is outlined and this enables us to investigate the policy positions of governments (Hobolt and Klemmensen, 2005). First, we analyse the extent to which Wordscores estimates are robust when different reference texts are used. Second, we analyse whether the Wordscores approach is capable of capturing changes to the party composition and to the prime minister of governments over time. Budge and Pennings (2007a) argue that one of the major shortcomings of the word scoring approach is that it has not been successfully applied to longer time series of data. In our analysis, we intend to fill that gap by testing whether the Wordscores techniques provide valid measure of the policy positions of Danish governments, 1953–2006.

In Table 3, we show the descriptive statistics for the speeches used in the analysis. These texts are both

Table 3
Descriptive statistics for government speeches

	Virgin texts	Reference texts
Mean unique words (scored)	704 (90)	–
Mean total words (scored)	3709 (820)	3845 (847)
N	44	8

Note: standard deviations in brackets.

Author's personal copy

longer and more comparable than the manifestos and we would thus expect more precise estimates.

In order to assess the extent to which the choice of reference text influences the Wordscores estimates of the positions of the governments, we conduct five analyses with different reference texts. The main idea is to investigate whether using reference texts at different time points has consequences for the position estimates obtained. Changing political vocabulary over time could influence the validity of Wordscores estimates, and it is thus important to establish how sensitive the method is to the use of different texts.

In the first three analyses, we have used the 1953 speech as the (left-wing) anchor each time, but we changed the second reference text using the speeches made in 1968, 1983 and 2006 as the right-wing reference text. In the fourth analysis, we have chosen speeches from each end of the ideological spectrum as reference texts: the speech delivered in 1967 by Prime Minister Krag is chosen as the most left-wing speech in the post-war period and the speech given by Prime Minister Fogh Rasmussen in 2002 as the most right-wing speech.[11] The final analysis uses reference texts from two consecutive years (the speech by Social Democrat Anker Jørgensen in 1981 and Conservative Poul Schlüter made in 1982) in order to see whether the pattern in the previous analyses is just an artefact of using reference texts that are far removed from each other in time and thus a result of changing political vocabulary rather than changing positions.

Fig. 2 shows that the five analyses using different reference texts provide very similar conclusions about the policy shift in Danish politics over the past 50 years, namely that Danish governments have moved to the right. As the first three analyses illustrate, choosing reference texts further removed in time does not produce radically different results from the overall picture that Danish governments have moved to the right, although the degree to which recent governments have continued the shift to the right varies between the analyses. Equally, the analysis using reference text from consecutive year paints a very similar picture, albeit with a little more volatility in the positioning of governments.

Several studies of Danish party politics confirm this trend towards more right-wing positions. Green-Pedersen (2001) argues that the coalition governments in office from 1993 to 2001, led by the Social Democrats, moved to the right in order to satisfy the centre parties in the coalition. The historians Rasmussen and Rüdiger (1990) argue that the Conservative-led governments in office from 1982 also moved Danish politics considerable to the right. Hence the overall directional change in Danish politics detected by the Wordscores estimates is corroborated by qualitative studies of Danish politics.

We can compare this pattern of government placements with the CMP data. In order to obtain a CMP score for the government, we have summed the CMP party position scores of the parties participating in the government weighted by the party's share of government seats. Hence, we expect that larger coalition parties are more responsible for the general political position of the government than smaller parties. Fig. 3 shows that the CMP and Wordscores[12] produce comparable time series of government positions, correlating at 0.59. However, whereas the two series look very similar until the early 1990s, they depart thereafter. According to Wordscores, the Danish centre-left government led by Poul Nyrup Rasmussen continued the general trend to the right in Danish politics, whereas as the CMP measure suggests that this was in fact the most left-wing government in the post-war period. Given the qualitative evidence provided by Green-Pedersen (2001) it seems unlikely that this social democratic government which came into office in 1993 was in fact the most left-wing government Denmark has seen since the Second World War. Furthermore, our previous analysis of the manifestos has shown that the Social Democratic Party, according to the Wordscores estimates, has moved to the right, and hence the Wordscores estimates of the speeches are consistent with the Wordscores analysis of the manifestos. It is worth noting, however, that when we use other sets of reference texts, as shown in Fig. 2, the Wordscores estimates also indicate that the right-wing shift in Danish politics plateaus already in the mid-1980s.

While the general right-wing trend in the positions of Danish executives is noteworthy, it is equally important to evaluate to what extent Wordscores estimates arrive at meaningful placements of individual governments that accord with our intuitive understanding of the Danish political system. One way of testing the extent to which Wordscores generates valid estimates is to assess

[11] The first speech was chosen because the Social Democrat-led government relied on the left-wing Socialist People's Party in order to stay in power (Mader, 1979). The 2002 speech was chosen because the Liberal-led government relied on the far-right Danish People's Party in order to get into office.

[12] In Fig. 3, we have used the Wordscores time series generated using the speeches from 1953 to 2006 as reference texts. As shown in Fig. 2, other sets of reference texts generate similar, though not identical, placements of governments.

Author's personal copy

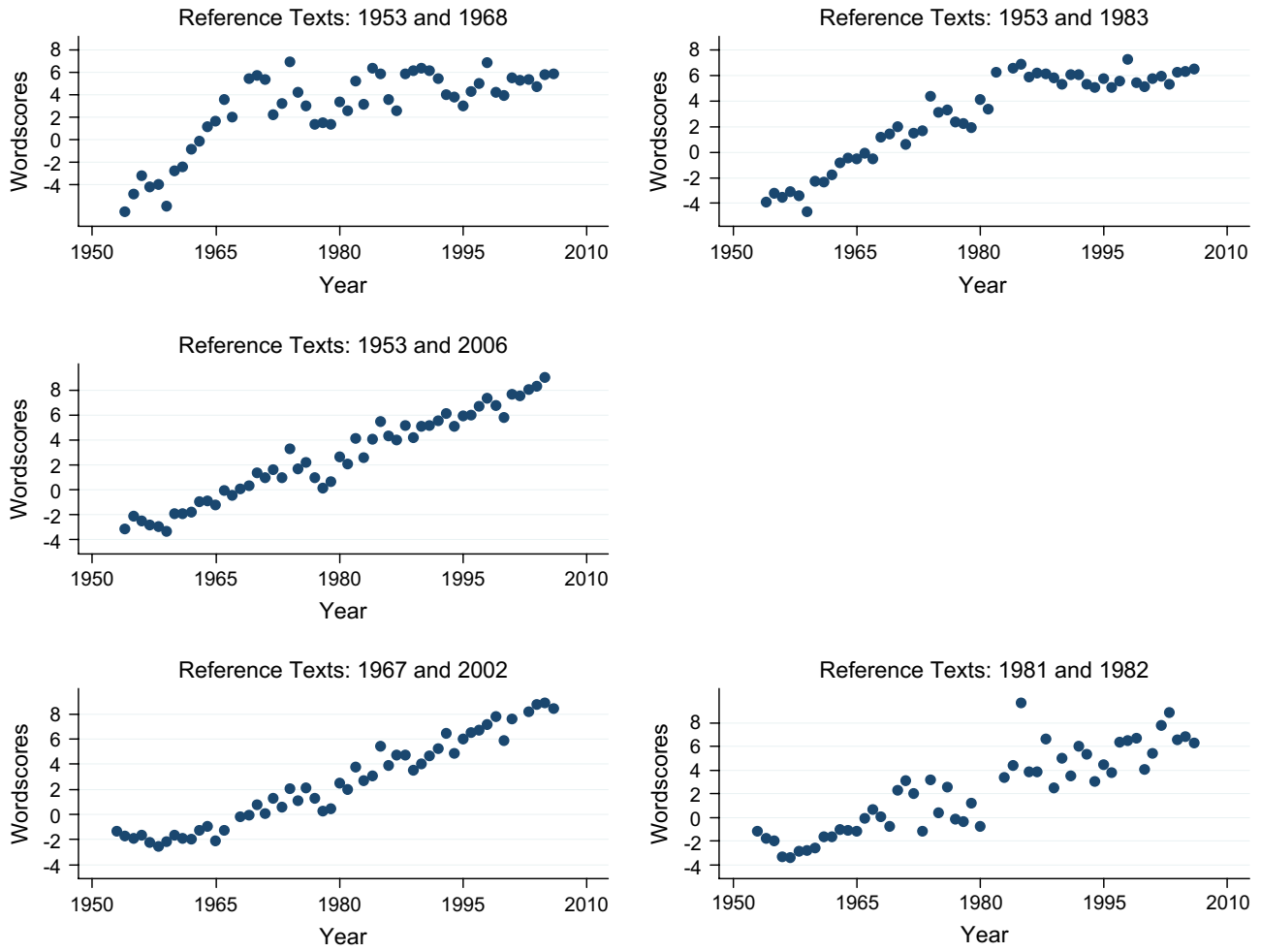


Fig. 2. Estimating changes government positions using different reference texts. Source: Danish prime ministers' opening speeches in Parliament, 1953–2006.

whether we can distinguish between left- and right-wing governments on the basis of these scores. Despite the general shift to the right in Danish politics, we would still expect that centre-left governments are, on average, to the left of centre-right governments. When we compare the means of Wordscores positions for governments headed by a Social Democratic prime minister to those led by a prime minister from a centre-right party we get a mean score for Social Democrat governments which is significantly lower than the mean for the group of governments with centre-right prime ministers.

Another way of evaluating the validity of the Wordscores estimates is to compare the variance of estimates within a government with the variance between governments. If the absolute government placements were valid, then we would expect greater variance between speeches made by different prime ministers compared with speeches given by the same prime minister at difference time points. In Table 4, we show the analysis of

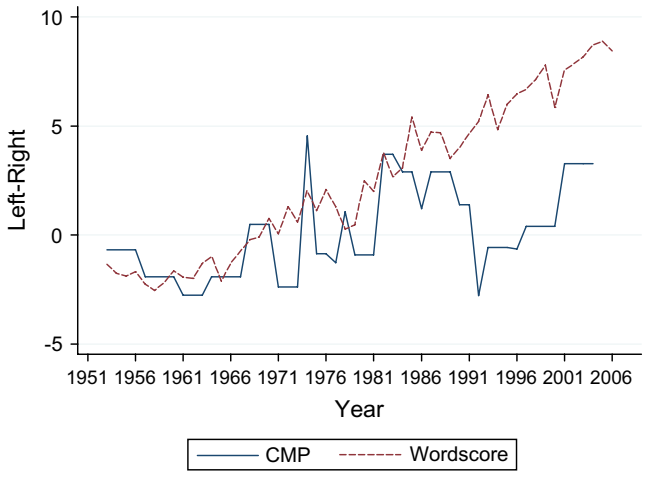


Fig. 3. Development in government positions, 1953–2006. Note: Wordscores estimates are generated using annual speeches by prime ministers' speeches (from 1953 to 2006 are used as reference texts). CMP estimates are calculated on the basis of the weighted mean position of parties in government.

Author's personal copy

Table 4
ANOVA of government positions

Reference texts	Between group	Within group	*F*-value
1953–1968	1.95	0.12	30.47***
1953–1983	7.97	0.17	103.41***
1953–2006	20.35	0.51	87.52***
1967–2002	16.66	0.41	89.58***
1981–1982	1.67	0.23	15.92***

***$p < 0.01$.

variance for each of the five time series presented in Fig. 2 (generated using different reference texts). For each of the five time series, the ANOVA confirms this expectation: there is significantly more variance between governments than within governments.[13]

These results indicate that the Wordscores placements of governments correspond with our intuitive understanding of politics, by capturing differences in the partisanship and prime minister of the governments.

6. Conclusion

The Wordscores approach is a pioneering new approach to the study of political texts. It offers a cheap, efficient and language blind technique for extracting policy positions from political texts. This approach can be extended not only to cross-national research of political actors, but also backwards in time. The only limitation is the availability of documents. It thus has several advantages over conventional methods of estimating policy positions. First, it is less costly and more flexible than any other method for estimating policy positions. Second, it can produce longer time series of policy positions. Currently the CMP data are the only time series on party positions and even these data are only collected at 4–5 year intervals.

But can we actually get reliable and valid estimates of policy positions on the basis of reference texts and matrices of word frequencies alone? This paper offers the first systematic and independent evaluation of the Wordscores approach, which cross-validates this technique with both CMP and expert data. In our investigation of party positions in Danish manifestos, we find that the CMP measure slightly outperforms the Wordscores estimates, when we compare with party expert surveys. However, the Wordscores estimates perform reasonable well, and they have the further advantage that we have a measure of uncertainty associated with each of our estimates. Moreover, our findings suggest that the Wordscores estimates provide time series of government positions with high face validity. The Wordscores estimates of government speeches identify a shift to the right in Danish politics and further captures changes from centre-left to centre-right governments (and vice versa). It is noteworthy, that these time series scores are very similar regardless of which set of reference texts is used. This suggests that the Wordscores technique is in fact less sensitive to the choice of reference text than has been suggested in articles by the inventors of this approach (Laver and Benoit, 2002; Laver et al., 2003).

Overall, our analyses indicate that the Wordscores approach does offer great promise indeed, but that scholars also need to be careful when relying on Wordscores estimates. Our analysis of party positions has shown that short documents are especially prone to generate erroneous or unreliable scores. This implies that the Wordscores approach can most safely be applied to relatively long texts with a well-known and dominant dimension of competition (e.g. left–right). Scholars should also choose reference texts of a similar nature to the texts under investigation and preferably use a set of reference texts representing each "extreme" on the dimension. These caveats suggest that although the Wordscores technique may be "language blind" in principle, the researcher cannot entirely do away with the need to know something about the texts under investigation and their political context. Mindful of these limitations, the Wordscores approach nevertheless allows scholars to analyse a vast number documents systematically and therefore begin to answer questions that have not yet been addressed empirically.

Acknowledgements

We thank Lars Bille for giving us full access to the Danish party manifestos collected by the late Tove Lise Schou. Moreover, we would like to thank Lanny W. Martin and Georg Vanberg for making their software programme available to us. Finally, we are grateful to Ken Benoit and Mogens N. Pedersen and the anonymous reviewers for extensive and very useful comments on earlier drafts of the paper.

Appendix 1. Supplementary data

Supplementary data associated with this article can be found in the online version, at doi:10.1016/j.electstud.2007.07.006.

[13] In this analysis we defined a government on the basis of parties participating in the governing coalition. The pair wise comparisons are too detailed to report but they are available upon request.

Author's personal copy

References

Armstrong, D., Bakker, R., 2006. Take that you Lousy dimension: a new model for manifesto-based left-right party placement. Unpublished working paper, University of Oxford.

Andersen, J.G., 1984. Decline of class voting or change in class voting – social class and party choice in Denmark in the 1970s. European Journal of Political Research 12 (3), 43–259.

Bakker, R., Edwards, E., de Vries, C., 2006. Fickle parties of changing dimensions? Testing the Comparability of the Party Manifesto Data Across Time and Space, Paper presented at the MPSA conference, Chicago.

Benoit, K., Laver, M., 2007a. Estimating party policy positions: comparing expert surveys and hand-coded content analysis. Electoral Studies 26 (1), 90–107.

Benoit, K., Laver, M., 2007b. Benchmarks for text analysis: a response to Budge and Pennings. Electoral Studies 26 (1), 130–135.

Benoit, K., Laver, M., 2006. Party Policy in Modern Democracies. Routledge, London.

Benoit, K., Laver, M., 2003. Estimating Irish party positions using computer wordscoring: the 2002 elections. Irish Political Studies 18 (1), 97–107.

Borre, O., 1995. Old and new politics in Denmark. Scandinavian Political Studies 18 (3), 187–195.

Budge, I., Farlie, D.J., 1983. Explaining and predicting elections. Issue Effect and Party Strategies in Twenty-Three Democracies. Allen and Unwin, London.

Budge, I., Klingemann, H.D., 2001. Finally! Comparative over-time mapping of party policy movement. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tannenbaum, E. (Eds.), Mapping Policy Preferences: Parties, Electors and Governments: 1945–1998. Oxford University Press, Oxford.

Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tannenbaum, E., 2001. Mapping Policy Preferences: Parties, Electors and Governments: 1945–1998. Oxford University Press, Oxford.

Budge, I., Pennings, P., 2007a. Do they work? Validating computerised word frequency estimates against long policy series. Electoral Studies 26 (1), 121–129.

Budge, I., Pennings, P., 2007b. Missing the message and shooting the messenger: Benoit and Laver's ''response''. Electoral Studies 26 (1), 136–141.

Budge, I., Robertson, D., Hearl, D. (Eds.), 1987. Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies. Cambridge University Press, Cambridge.

Castles, F., Mair, P., 1984. Left–right political scales – some expert judgement. European Journal of Political Research 12 (1), 73–88.

Damgaard, E., 2000. Denmark: the life and death of government coaltions'. In: Müller, W.C., Strøm, K. (Eds.), Coalition Governments in Western Europe. Oxford University Press, Oxford.

Gabel, M., Huber, J., 2000. Putting parties in their place: inferring ideological positions from party manifesto data. American Journal of Political Science 44 (1), 94–103.

Giannetti, D., Laver, M., 2005. Policy position and jobs in the government. European Journal of Political Research 44 (1), 91–120.

Green-Pedersen, C., 2001. Minority governments and party politics: the political and institutional background to the Danish miracle. Journal of Public Policy 21 (1), 53–70.

Hobolt, S.B., Klemmensen, R., 2005. Responsive government? Public opinion and policy preferences in Britain and Denmark. Political Studies 53, 379–402.

Huber, J.D., Inglehart, R., 1995. Expert interpretations of party space and party locations in 42 countries. Party Politics 1 (1), 73–111.

Janda, K., 1980. Political Parties. A Cross-national Survey. Free Press, New York.

Klingemann, H.-D., Volkens, A., Bara, J., Budge, I., McDonald, M.D., 2007. Mapping policy preferences II. Estimates for Parties, Electors, and Governments in Eastern Europe, European Union, and OECD 1990–2003. Oxford University Press, Oxford.

Klingemann, H.-D., Hofferbert, R., Budge, I., 1994. Parties, Policies and Democraies. Westview Press, Boulder.

Laver, M., 2001. Why should we estimate the policy position of political actors? In: Laver, M. (Ed.), Estimating the Policy Position of Political Actors. Routledge, London.

Laver, M., Benoit, K., 2002. Locating TDs in policy spaces: wordscoring dáil speeches. Irish Political Studies 17 (1), 59–73.

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

Laver, M., Benoit, K., Sauger, N., 2006. Policy competition in the 2002 French Legislative and Presidential Elections. European Journal of Political Research 45 (4), 667–697.

Laver, M., Garry, J., 2000. Estimating policy positions from political texts. American Journal of Political Science 44, 619–634.

Laver, M., Hunt, W.B., 1992. Policy and Party Competition. Routledge, New York.

Mader, E., 1979. SF under det ''røde'' cabinet. Odense University Press, Odense.

Mair, P., 2001. Searching for the positions of political actors: a review of approaches and a critical evaluation of expert surveys. In: Laver, M. (Ed.), Estimating the policy positions of political actors. Routledge, London.

Marks, G. (Ed.), 2007. Estimating error in measures of party positioning: expert, manifesto, and survey data compared. Electoral Studies 26 (1), 1–141 (Special issue).

Marks, G., Steenbergen, M., 2002. Chapel Hill Party Dataset 2002. University of North Carolina, Chapel Hill. Available from: http://www.unc.edu/~gwmarks/data.htm.

Marks, G., Hooghe, L., Steenbergen, M., Bakker, B., 2007. Cross-validating data on party positioning on European integration. Electoral Studies 26 (1), 23–38.

Martin, L.W., Vanberg, G. A robust transformation procedure for interpreting political text. Political Analysis, in press.

Morgan, M.-J., 1976. The modelling of governmental coalition formation: a policy-based approach with interval measurement. Ph.D. dissertation, University of Michigan.

Rasmussen, H., Rüdiger, M., 1990. Tiden efter 1945. Gyldendal, Copenhagen.

Riker, W.B., 1996. The Strategy of Rhetoric. Princeton University Press, New Haven.

Volkens, A., 2001. Manifesto research from 1979: from reliability to validity. In: Laver, M. (Ed.), Estimating the Policy Positions of Political Actors. Routledge, London.