Cross-Validating Measurement Techniques of Party Positioning*

Christine Arnold,†
Department of Political Science; University of Maastricht

Simon Hug‡
Département de science politique, Université de Genève

and Tobias Schulz§
Swiss Federal Research Institute for Forest, Snow and Landscape

Paper prepared for presentation at the Annual Meeting of the
American Political Science Association
Toronto (September 3-6, 2009)

First version: March 2009, this version: August 25, 2009

* This paper emanates from the DOSEI (Domestic Structures and European Integration) project funded by the European Union in its Fifth Framework program and directed by Thomas König (Universität Speyer). The Swiss part was financed by the *Bundesamt für Bildung und Wissenschaft* (Grant number BBW Nr 02.0313) and the *Grundlagenforschungsfonds* of the University of St. Gallen (Grant number G12161103). Thanks are especially due to the numerous people (listed in the footnote 5) who helped us locate and translate party manifestos.

†Department of Political Science; Faculty of Arts and Social Sciences; University of Maastricht; PO Box 616; 6200 MD Maastricht; The Netherlands; phone: +31 (0)43 388 3052; email: c.arnold@politics.unimaas.nl

‡ Département de science politique, Faculté des sciences économiques et sociales; Université de Genève; 40 Bd du Pont d'Arve; 1211 Genève 4; Switzerland; phone ++41 22 379 83 78; email: simon.hug@politic.unige.ch

§Swiss Federal Research Institute for Forest, Snow and Landscape (WSL) , Zürcherstrasse 111, CH-8903 Birmensdorf, Switzerland; phone +41 44 7392 477; email: +41 44 7392 215 tobias.schulz@wsl.ch

Abstract

The deepening and widening of the EU is turning out to be a force capable of rocking national governments, jeopardizing party cohesion, and pitting public opinion against the preferences of party elites. The accuracy of measurement tools to assess party positioning towards European issues increasingly has become a central concern. Only if we have valid and reliable techniques of estimating the position of parties can we begin to ask if parties indeed represent public opinion.

The purpose of this paper is to cross-validate the utility of manual content analysis and computer-based automatic methods. We will assess which technique provides better estimates of party positions on EU constitutional issues and which of those estimates in turn correlate most strongly with the preferences of the respective voters.

This paper will employ data from the European Election Survey 2004 and the national party manifestos that were issued for the European Parliament election of 2004.

1 Introduction

Studies on European integration rely increasingly on data measuring policy positions of political parties to address very diverse research questions. Researchers wishing to assess the positions of political parties toward European integration in general (e.g., Ray, 1999), the main political conflict lines in the political system of the European Union (EU) (e.g., Gabel and Anderson, 2004; Hooghe, Marks and Wilson, 2004), how members of the European Parliament represent national or party interests (e.g., Hix, 2002; Thomassen, Noury and Voeten, 2004) or how parties intervene or influence treaty negotiations (e.g., Benoit, Laver, Arnold, Pennings and Hosli, 2005; Hug, 2007; Hug and Schulz, 2007; König and Finke, 2007) all base their analyses on policy positions of political parties. These studies, taken as examples, all rely on different measurement strategies, using different types of data, be it party manifestos, expert surveys, or roll call votes, etc.. Increasingly, however, and mostly due to the development of new techniques to analyze textual data, party manifestos have become a more central source in such endeavors.

While there clearly is an increase in the use of textual data, the ways in which this data is used and analyzed differ considerably. A series of scholars have already compared these various approaches, for instance some of the contributions in Laver (2001) (see also Budge, 2001; Laver, Benoit and Garry, 2003; Ray, 2007; Slapin and Proksch, 2008; Benoit, Mikhaylov and Laver, 2009 (forthcoming); Helbling and Tresch, 2009; Monroe and Schrodt, 2009) come, however, often to quite different conclusions. Our goal in this paper is not to contribute directly to this debate, but to explore different ways in which textual data can be analyzed in a very specific context. The specific context is characterized by a research question requiring policy positions of political parties on scales that can be directly compared with policy positions of other actors obtained, possibly, through other means.

Research questions leading to such data requirements often appear in studies of representation,[1] but also appear in studies of international negotiations dealing with two-level games. In the latter context, and as the few graphical illustrations in Putnam's (1988) article show, preferences of different sets of actors on the same scale are necessary to test the various implications of such two-level games.

[1]Even though the methodological difficulties in such studies have been prominently highlighted by Achen (1977, 1978), his warnings seem largely forgotten in the literature.

Our empirical example stems from a research project[2] dealing with the negotiations leading up to the EU's "Treaty establishing a constitution for Europe."[3] In this project several studies relied on measures of preferences that the political parties had over various aspect of the "Treaty establishing a constitution for Europe." Textual data, in terms of party manifestos for the 2004 European Parliament election, provided an important source for such information. Analyses of this data, under the constraint that it should allow for comparable information as from other sources, were carried out under different assumptions and with different result. These differences will be discussed in detail in the present paper.

In the next section we will briefly review the various methods currently employed in analyzing textual data for the purposes of extracting policy positions. This review will not be exhaustive but focus on those methods that are easily amenable to a multilingual context (i.e., the manifestos of parties competing in EP elections come in almost 20 different languages) and to the possibility of linking it with other data, as discussed above. In section three we introduce our empirical work, first by discussing the data requirements, before turning to a description of how we analyzed the textual data. Following up on this subsection we offer first a preliminary plausibility check of the resulting data, before comparing our measures of policy preferences with those obtained by other methods. In section four we conclude and sketch our future research plans.

2 Techniques to derive Party and Government Policy Position

Normative democratic theory tells us that in a representative democracy there should be some match between the interests of the people and the policies that their representatives promote (Dahl, 1971; Wessels, 1999). The deepening and widening of the EU has pushed questions relating to democratic accountability and representation to the forefront of European studies. We are simultaneously observing a gap between citizens and elites when it comes to EU preferences

[2]See `http://www2.sowi.uni-mannheim.de/lspol2/dosei/` and, for instance, König and Hug (2006) for a sample publication stemming from this project.

[3]This treaty, after having foundered in the ratification referendums in France and the Netherlands has been resuscitated with slight modifications under the new name of the "Lisbon Treaty" and is in the process of clearing the last ratification hurdles.

and an enhanced politicization of EU matters in national contexts through the increased use of referendums as a legitimation-tool and the activities of political entrepreneurs. In this context, the accuracy of measurement tools to assess party and government positioning towards European issues increasingly has become a central concern. Only if we have valid and reliable sources and techniques of estimating the position of parties and can we begin to ask if parties indeed represent public opinion.

A variety of sources are used to measure the position of parties and governments. These sources include expert surveys where observers are asked about their perceptions of parties' policy position (Castles and Mair, 1984; Laver and Hunt, 1992; Huber and Inglehart, 1995; Ray, 1999; Pennings, 2002; Benoit and Laver, 2006; Steenbergen and Marks, 2007).

Also analyzed are party manifestos, legislative speeches, speeches of government representatives, legislative bills, and position papers submitted during deliberative processes such as the Constitutional Convention in the EU (Budge, Klingemann, Andrea and Bara, 2001; Laver, Benoit and Garry, 2003; Benoit, Laver, Arnold, Pennings and Hosli, 2005; Arnold and Franklin, 2006) Furthermore, roll-call data is used to derive issue preferences from legislative behavior (Hix, Noury and Roland, 2006). Less commonly, scholars have used the self-placement of voters in opinion surveys to derive parties' position (Gabel and Huber, 2000). Finally, also less frequently, scholars have used information on budget outlays to derive a policy position of governments.

The jury is still out on the question which source for positioning parties and governments would be most adequate. Some scholars find that expert surveys provide the most accurate data for party positioning on European integration (Marks, Hooghe, Steenbergen and Bakke, 2007). Others contend that party manifestos is a better source to capture the changing party position over time (McDonald, Mendes and Kim, 2007; Volkens, 2007).

Also a wide variety of techniques are employed to analyze these disparate sources of party positions. These techniques range from fully manual to fully automated; on the one hand we have studies relying on face-to-face expert interviews (e.g., König and Hug, 2006). On the other hand we have unsupervised topic classification with the help of computer algorithms (Hopkins and King, 2007). Especially advances in information retrieval, computational lin-

guistics and natural language processing have opened up new opportunities for research in political science (see for example Jurafsky and Martin, 2000; Laver, Benoit and Garry, 2003; Simon and Xenos, 2004; Diermeier, Godbout, Yu and Kaufmann, 2007; Evans, McIntosh, Lin and Cates, 2007; Yu, Kaufmann and Diermeier, 2008; Klebanov, Diermeier and Beigman, 2009). The ease with which one can now collect large volumes of political text, process and clean them has encouraged several innovative research projects. Examples include tracking legislative agendas and political topics and ideal-point estimation of ideological positioning (Laver, Benoit and Garry, 2003; Monroe and Ko, 2004; Quinn, Monroe, Colaresi, Crespin and Radev, 2006; Slapin and Proksch, 2008; Monroe, Colaresi and Quinn, 2009 (forthcoming)). Additionally, the utility of a range of sophisticated statistical algorithms has been established and increasingly these algorithms are being applied to the analysis of political texts (Manning and Schütze, 2002; McGuire and Vanberg, 2005). Thus, text in political science research is increasingly treated as data.

Among the techniques used for text categorization, one can differentiate four general approaches depending on the relative level of human involvement (Hillard, Purpura and Wilkerson, 2007; Cardie and Wilkerson, 2008). On the one had we have text categorizing techniques with relatively heavy involvement of the researcher, as is found in manual coding or to some degree in dictionary-based coding. On the other hand there are techniques with minimal and next to no human involvement, as is found in supervised learning and unsupervised learning techniques.

Text classification done through manual coding has traditionally been the most common methodology used to study political texts. In this approach, often called thematic content analysis, scholars manually code text units and then construct from the existence and frequency of the coded units the occurrence of concepts (Roberts, 1997; Popping, 2000). Prominent examples of this approach is the Comparative Manifesto Project and the Policy Agendas and Congressional Bills Project (Budge, 2001; Baumgartner and Jones, 2002; Adler and Wilkerson, 2005; Baumgartner and Jones, 2005; Klingemann, Volkens, Bara, McDonald and Budge, 2007). The advantage of this approach is that human coders are quite well capable of handling the idiosyncrasies of natural language and are thus capable to assign meaning to texts on a high cognitive level. However,

much evidence is also available on the disadvantages of this approach. Apart from the obvious factors such as the costs involved for the labor-intensive creation of the coding scheme and the coding of the documents, one of the key draw-backs of this approach is the question of reliability. Especially the Comparative Manifesto Project has received considerable criticism in the past few years and at the same time also benefited from proposals to improve its methodology (Pelizzo, 2003; Benoit and Laver, 2007; Hansen, 2008; Benoit, Mikhaylov and Laver, 2009 (forthcoming); Lowe, Benoit, Mikhaylov and Laver, 2009) A second approach to text classification with relatively high human involvement is dictionary-based content analysis. In this approach typically a subset of the documents is manually coded to develop a dictionary. The terms in the dictionary are then linked with broader categories. All of the documents are then run through this dictionary and the occurrence of the terms are tallied up (Krippendorff, 2004). In political science this approach has been applied to party manifestos (Laver and Garry, 2000; Pennings, 2002; Pennings and Keman, 2002), position papers submitted during deliberative processes such as the Constitutional Convention in the EU (Pennings, 2008 (forthcoming)), and to news wire feds by the Kansas Event Data project (Gerner, Schrodt, Francisco and Weddle, 1994). An advantage of this approach is that after the initial investment has been made of creating the dictionary the subsequent analysis even of very large data is relatively cheap. This is, of course, only provided the dictionary is indeed a valid measurement instrument.

More recently, with supervised learning approaches, a third approach of text classification has become possible which is characterized by much less human involvement (Sebastiani, 2002). In this approach, annotated text is used to train a classifier. This classifier is then tested on new text not contained in the training corpus. The performance of the classifier on this new text is then evaluated. There is a wide variety of algorithms that have been developed in recent years. Among the most popular classifiers in political science are Support Vector Machines which have been identified as one of the most efficient classification methods (Joachims, 1998) and naïve Bayes which also has been found to perform well. These methods have been applied to Senatorial speeches (Diermeier, Godbout, Yu and Kaufmann, 2007) and speeches from the House of Commons (Purpura and Hillard, 2006) and to blogs (Hopkins and King, 2007).

Another method, similar to a supervised learning approach and widely used in political science is the wordscore technique (Laver, Benoit and Garry, 2003). This method makes use of supervised word frequencies which are calculated both for reference texts and virgin texts. One of the key advantages of this method is that it produces interval-level scores for texts along a pre-defined dimension, with no need for the researcher to try to understand and subjectively interpret the meaning of words of the texts. This opens up the possibility to analyze large collections of texts independent of the language of the texts. This method has been applied to party manifestos (Kritzinger, Cavatorta and Chari, 2004; Hug and Schulz, 2007; Klemmensen, Hobolt and Hansen, 2007) advocacy briefs submitted to the US Supreme Court (Evans, McIntosh, Lin and Cates, 2007), debates in the Senate (Bertelli and Grose, 2006), government speeches (Giannetti and Laver, 2005), and interest group influence in the EU (Klüver, 2009 (forthcoming)). Recently alternative ways have been proposed how the wordscores technique can be improved (Lowe, 2008).

A fourth approach to text classification is unsupervised learning, much of it is advanced by computational linguistics. This approach has no predefined set of codes, dictionaries, or training set, but instead texts are statistically placed into categories to maximize internal coherence. In political science this approach has been used on Congressional bills (Hillard, Purpura and Wilkerson, 2007) and on blogs (Hopkins and King, 2007). Also, Quinn, Monroe, Colaresi, Crespin and Radev (2006), in the Dynamics of Political Rhetoric and Political Representation Project, used a multinomial mixture model to automatically sort Senatorial speeches into topic clusters. For very large corpora of texts this is a useful approach.

Finally, there are also hybrids of these approaches. For instance the wordfish approach applies a unsupervised scaling technique to text classification (Slapin and Proksch, 2008). But a necessary prior step before the technique can be applied is that a human coder needs to identify the segments in a document which are relevant to the dimension of interest. This approach has been applied to German legislation (**?**), and party manifestos (Proksch and Slapin, 2009 (forthcoming)). The usefulness of this approach hinges on the feasibility of slicing documents into relevant policy dimensions.

3 Empirics

Almost all of the techniques and approaches discussed above might, a-priori, be helpful to derive policy positions of political parties on such a specific topic as the EU's constitutional treaty. Given that the policy positions are to be used for studies of representation and tests of implications stemming from two-level games, it is important, however, to derive policy positions for parties on the same, or a closely related scale as the positions of other actors.

In the DOSEI-project the most detailed data on policy positions were collected through expert surveys on government positions.[4] Hence, most of the studies attempted to link information on other actors, like voters (Hug and Schulz, 2005) or parties (König and Finke, 2007), with the data from the expert survey. Here we will proceed similarly, given that the importance of identical or similar scales is of paramount importance.

3.1 Data used

For the European Parliament elections in 2004, we collected all national party manifestos we were able to find. For our analysis we selected parties that won at least one seat in the current Parliament. In terms of pre-processing the documents, we first converted all documents to ASCII format to extract the text. The documents originally were in pdf, doc or html formats. The manifestos of the old member states were then machine-translated using software for each language that produces the best results. For example the Systran Software was used in the cases of Spanish, Italian, German, French and Portuguese. The google translation tool was used in other cases and to correct translation omissions from the systran translation. Although the translations were often grammatically not quite correct, the meaning of the text was broadly clear and manual coding was possible. For the manifestos of the new member states human translators were hired and machine-translation was applied in the case of Cyprus only.[5] For many

[4] Table 1 below provides an overview over the questions employed.

[5] We would like to express our gratitude to the following persons who helped in finding translators, who actually translated texts or helped finding additional party manifestos: Agnes Lux and Kata Török (Hungary), Martina Chabreckova and Michaela Jacova (Slovakia), Visvaldis Valtenbergs (Latvia), Maarja Lühiste (Estonia), Mindaugas Jurkynas (Lithuania), Yiannos Katsourides (Cyprus), Hana Vykoupilova, Patricie Mertova and Michaela Jacova (Czech Republic), Lina Pavletic (Slovenia) and Silvia Makowski (Poland).

of the languages of the new Member States, however, machine-translation did not produce sufficiently good results. Finally, we then broke the document into units of one sentence each.

After the pre-processing and translation of the documents, we manually coded each document. Each sentence was assigned to a pre-defined category, based on the expert questionnaire developed for the measurement of actors' positions on the Constitution in the EU-funded DOSEI project (see table 1).

Each of these pre-defined categories was actually an answer category belonging to one of the items of the original Dosei expert survey. Hence, after having coded the manifestos this way, we had to recompile the information into a form that is equivalent to the ordinal variables of the original DOSEI expert survey. Since any category of any item could receive more than one tag, we applied the rule that the one category was finally measured that received most tags. In case of a tie, i.e. two or more categories received the same amount of tags, simply the first, i.e. the lowest ordinal category was chosen.

These same party manifestos have already been used in another context and analyzed with another tool. Hug and Schulz (2007) used these texts to determine the positions of the national parties in a negotiation and bargaining space defined by the policy preferences of national governments. More precisely Hug and Schulz (2007) first determined with the help of a factor analysis for ordinal variables (Martin and Quinn, 2004; Quinn, 2004) a two-dimensional bargaining space on the basis of data collected in an expert survey on the EU member states' governments positions on the various issues debated in relation with the EU's constitutional treaty. In this bargaining space were also located the status quo (i.e., the treaties up to the Nice treaty), the draft treaty proposed by the Convention, and the proposal adopted at the IGC. With each of these positions in the bargaining space is obviously a legal text related, and these text were employed as reference texts for a wordscores analyses of the party manifestos (Laver, Benoit and Garry, 2003).[6] We will use the results of these analyses in one of the analyses below, since they offer a further plausibility test of the codings we undertook.

[6]We report the positions of the political parties based on this approach in a table in the appendix.

3.2 A first plausibility test

In the expert survey carried out in the context of the DOSEI-project (König and Hug, 2006) the experts not only evaluated the government positions but also alerted the interviewers to the preferences of other important actors in the policy formation process at the domestic level. While most of these additional actors were ministries or particular organizations, some of them, namely seven, were political parties. For all these actors we have the assessment of experts regarding the former's position on the various issues debated at the IGC.[7] Hence, this information allows us a first plausibility check of our hand coding of party manifestos for the Belgian Socialist party, two Czech parties (ODS and CSSD), the French UMP, the Italian Lega Nord, the Slovak KDH and the Spanish Socialist party.

Before looking more closely at the degree of correspondence between evaluations of party positions by experts and the hand-codings of the party manifestos, an important element needs to be mentioned. While the experts were encouraged to give evaluations of the party positions on roughly 70 issues, it is very unlikely that all these issues were mentioned in party manifestos. Hence it cannot astonish, that the number of issue positions for the parties we will be able to compare varies between 10 (Italian Lega Nord and French UMP) and 30 (Belgian Socialists) for an average of close to 18 issues.

Taking this element into consideration we find that the codes from our hand coding correspond on average about half of the time (47.35%) with the assessment of the expert survey. Behind this average a large variance is hidden, however. While we find correspondences in only 14.29% of all codes for the Czech ODS (while for the Czech CSSD this increases to almost 54%) this raises to 70% for the Lega Nord. The remaining parties all hover around the average of about 50%.[8]

While for these comparisons the possible correspondence is direct, we can do similar analyses for three additional parties if we assume that in countries with single party governments the government position corresponds to the party position. While during the IGC in 2004 there were five one party governments,

[7] Strictly speaking, interviewed experts were asked whether inside the government there were actors with divergent views on the constitutional treaty. If they identified such actors, they were asked to indicate on which issues these actors differed. For all other issues it was assumed that they shared the government's view.

[8] The details of these analyses are reported in tables in the appendix.

i.e. in Greece, Malta, Spain, Sweden and United Kingdom, we can carry out the relevant comparisons only for the three latter countries.[9]

When carrying out the same comparisons as those above we find largely similar results. We find for the coding of the three governments on average a correspondence of roughly 50% with a low of 14% for the Swedish government and a high of 59% for the Spanish government.

3.3 A spatial comparison

As the comparisons with the data stemming from the expert survey suggest, one drawback of the party manifestos is obviously that they do not cover in as much detail the various issues that relate to the EU's constitutional treaty. This, however, might be a blessing in disguise, since it might well be, that parties do not have clearly defined positions on all these issues, and even if they have them, they might not want them to be as publicly visible as that.

Limiting ourselves to the purely textual data from the party manifestos, allows us to compare different ways to analyze the texts as discussed above. As most of the other techniques discussed above attempt to infer on the basis of the textual data policy positions in a space with reduced dimensionality, it is instrumental to reduce the information stemming from the hand coding to a limited set of dimensions.

[9] As mentioned above, the party manifestos from Malta were in such poor quality, that no electronic file could be generated, while translation problems occurred for the Greek party manifestos.

Table 1: Two-dimensional factoranalytic solution of hand coded data: items included and corresponding factor loadings

Question	First Dimension (FD)	Second Dimension (SE)	S.E. FD	S.E. SD	t-statistic FD	t-satistic SD
1. CHARTER OF FUNDAMENTAL RIGHTS	-0.045	-0.795	0.02	0.016	-2.315	-49.219
2. SUBSIDIARITY	0.071	-0.33	0.008	0.007	8.35	-47.933
3. RELIGIOUS REFERENCE IN THE PREAMBLE OF THE CONSTITUTION:	0.081	0.084	0.008	0.009	9.578	9.82
4. RIGHT TO WITHDRAW FROM THE UNION	0.845	0.391	0.021	0.027	40.046	14.636
5.1 OBJECTIVES: Market economy	-0.165	-0.478	0.013	0.012	-12.755	-39.175
5.2 OBJECTIVES: Employment	-0.235	-0.444	0.014	0.014	-16.497	-30.982
5.3 OBJECTIVES: High level of competitiveness	0.253	0.082	0.007	0.008	37.549	10.252
6. PRESIDENCY EUROPEAN COUNCIL (TERM)	0.033	-0.831	0.023	0.019	1.439	-44.427
7. ELECTION OF THE PRESIDENT EUROPEAN COUNCIL	-0.785	-1.216	0.035	0.036	-22.52	-34.232
8. QUALIFIED MAJORITY THRESHOLD	0.023	-0.088	0.013	0.01	1.838	-8.448
9. COMPOSITION COMMISSION	0.635	-0.047	0.013	0.017	47.059	-2.803
10. APPOINTMENT / ELECTION OF THE COMMISSION PRESIDENT	-1.179	-0.633	0.029	0.035	-40.218	-18.074
11. APPOINTMENT OF COMMISSIONERS	0.453	-0.933	0.029	0.024	15.549	-38.373
12. EXTERNAL REPRESENTATION OF THE UNION	-0.146	-0.643	0.019	0.017	-7.762	-38.344
14. JURISDICTION	-0.163	-0.486	0.015	0.014	-11.076	-34.412
15.2 RIGHT OF INITIATIVE OF LEGISLATIVE ACTS: EUROPEAN PARLIAMENT.	-0.887	-0.41	0.024	0.029	-36.418	-14.057
15.3 RIGHT OF INITIATIVE OF LEGISLATIVE ACTS: COUNCIL.	-0.869	-0.028	0.023	0.027	-38.288	-1.033
15.5 RIGHT OF INITIATIVE OF LEGISLATIVE ACTS: CITIZENS.	-0.167	-0.771	0.023	0.02	-7.326	-38.186
16. ENHANCED COOPERATION: SCOPE	0.201	-0.261	0.011	0.01	18.987	-25.306
17.1 DELEGATION OF COMPETENCIES: Agriculture	0.398	-0.093	0.012	0.012	33.926	-7.629
17.2 DELEGATION OF COMPETENCIES: Structural and cohesion policies	0.161	-0.713	0.019	0.016	8.32	-44.602
17.3 DELEGATION OF COMPETENCIES: Area of freedom, security and justice	0.777	-0.663	0.023	0.022	33.921	-29.618
17.4 DELEGATION OF COMPETENCIES: Foreign policy	1.009	-0.931	0.029	0.028	34.232	-33.126
17.5 DELEGATION OF COMPETENCIES: Economic Policy	0.674	-0.951	0.028	0.024	24.131	-39.021
17.6 DELEGATION OF COMPETENCIES: Tax harmoisation	0.148	-0.513	0.016	0.013	9.331	-40.885
17.7 DELEGATION OF COMPETENCIES: Employment policy	0.099	-0.505	0.017	0.014	6.016	-35.826
17.8 DELEGATION OF COMPETENCIES: Social policy	0.007	-0.545	0.015	0.012	0.434	-44.149
17.9 DELEGATION OF COMPETENCIES: Health	0.27	-0.977	0.025	0.02	10.924	-49.271
17.10 DELEGATION OF COMPETENCIES: Enviroment	0.192	-0.688	0.018	0.015	10.747	-47.052
17.11 DELEGATION OF COMPETENCIES: Education	0.077	-0.373	0.012	0.011	6.268	-33.082
17.12 DELEGATION OF COMPETENCIES: Reseach, technological development and space	0.639	-0.123	0.016	0.019	40.302	-6.651
18.A2 INVOLVEMENT OF THE PARLIAMENT: Structural and Cohesion Policies	0.96	-0.518	0.028	0.029	34.348	-17.584
18.A3 INVOLVEMENT OF THE PARLIAMENT: Area of Freedom, Security and Justice	0.606	-0.903	0.028	0.024	21.899	-37.169
18.A5 INVOLVEMENT OF THE PARLIAMENT: Tax Harmonisation	-0.126	-0.885	0.024	0.02	-5.366	-43.211
18.A6 INVOLVEMENT OF THE PARLIAMENT: Monetary Policy (for the Euro States)	0.446	-1.105	0.032	0.027	13.751	-40.612
18.A7 INVOLVEMENT OF THE PARLIAMENT: Economic Policy	0.504	-1.134	0.033	0.028	15.212	-39.98
18.A8 INVOLVEMENT OF THE PARLIAMENT: Employment Policy	0.979	-0.576	0.028	0.03	34.867	-19.205
18.A9 INVOLVEMENT OF THE PARLIAMENT: Social Policy	0.891	-0.816	0.028	0.026	31.452	-30.833
18.B1 COUNCIL VOTING RULE: Agriculture	-0.236	-0.926	0.026	0.023	-9.075	-40.004
18.B2 COUNCIL VOTING RULE: Structural and Cohesion Policies	-0.121	-0.048	0.016	0.017	-7.627	-2.872
18.B3 COUNCIL VOTING RULE: Area of Freedom, Security and Justice	-0.269	-0.802	0.024	0.022	-11.336	-37.263
18.B5 COUNCIL VOTING RULE: Tax Harmonisation	-0.996	-0.398	0.026	0.031	-38.634	-12.891
18.B6 COUNCIL VOTING RULE: Monetary Policy (for the Euro States)	-0.981	-0.417	0.026	0.031	-37.599	-13.46
18.B7 COUNCIL VOTING RULE: Economic Policy	-0.991	-0.397	0.026	0.031	-38.148	-12.769
18.B8 COUNCIL VOTING RULE: Employment Policy	-0.994	-0.424	0.025	0.031	-40.037	-13.547
18.B9 COUNCIL VOTING RULE: Social Policy	-0.116	-0.033	0.016	0.016	-7.268	-2.011
18.B10 COUNCIL VOTING RULE: Social Security Rights	-0.978	-0.429	0.026	0.032	-37.353	-13.62
18.B11 COUNCIL VOTING RULE: Common Foreign Policy	-0.892	-0.475	0.025	0.031	-35.662	-15.504
18.B12 COUNCIL VOTING RULE: Defence Policy	-0.907	-0.276	0.024	0.029	-37.432	-9.519
19. INVOLVEMENT OF THE PARLIAMENT IN THE ADOPTION OF THE BUDGET	0.315	-0.315	0.014	0.012	23.26	-25.91
20. STABILITY AND GROWTH PACT	-1.419	0.436	0.036	0.039	-39.727	11.26
21. STABILITY AND GROWTH PACT	-0.31	-0.448	0.021	0.022	-15.124	-20.612
22. COMMON SECURITY AND DEFENCE POLICY	0.013	-0.118	0.005	0.005	2.486	-24.198
23. MANAGEMENT SYSTEM FOR EXTERNAL BORDERS (SCOPE OF UNION ACTION)	0.013	-0.502	0.012	0.01	1.031	-50.509
24. MIGRATION AND ASYLUM POLICY: SCOPE OF UNION ACTION	0.428	-0.35	0.017	0.015	24.71	-22.594

Here we follow previous work which relied on factor analyses based on ordinal variables that were also carried out on identical data for government positions (e.g. Hug and Schulz, 2005, 2007) and the actors involved in the governmental consultations (e.g. Hug and König, 2006).[10] Table 1 reports the basic information of the ordinal factor analyses, for which we decided to extract two dimensions,[11] and also lists the questions used in the DOSEI-expert survey.

This factor-analytic solution obviously also allows for placing the various parties whose party manifestos we coded in a common space. Figure 1 depicts the distribution of points in this two-dimensional space.[12] Figure 3 provides the same information, however, for each country considered separately with the corresponding party labels.

Figures 2 and 4 provide exactly the same information based on the wordscores analyses used in Hug and Schulz (2007). As discussed by Hug and Schulz (2007) the wordscores analyses provides rather little variation in the positions of the political parties along the second dimension.[13]

[10] See Hix and Crombez (2005) and Finke (2009) for two different, though related, ways to reduce the dimensionality of the bargaining space.

[11] We chose two dimensions first of all for practical reasons, since the bargaining space defined by the governmental positions is best summarized in two dimensions (see Hug and Schulz, 2007; Finke, 2009), and second because such a solution seemed to fit the data best. Table 15 lists the PRE measures that result from a regression of the hand coded (ordinal or binary) variables on the different dimensional solutions of the ordinal factor analyses applied to this data. Since there are many missings in the hand coded variables, these ordinal regressions did not produce very reliable results in many cases. However, in the cases were a meaningful PRE measure could be computed for at least a model including the one-dimensional or two-dimensional solution, the PRE measures do actually reveal that a two-dimensional solution is superior to the one-dimensional solution for most items, since the PRE measure increases from the first to the second column of table 15.

[12] The policy positions of the political parties on which this figure relies appear in a table in the appendix.

[13] The reason for this reduced variation is due to the fact that the reference scores for the status quo, the Draft treaty of the Convention and the IGC outcome, were very similar and thus hardly discriminated on this second dimension.

Figure 1: Distribution of policy positions of political parties in 23 member states (factoranalytic solution of hand coded data)

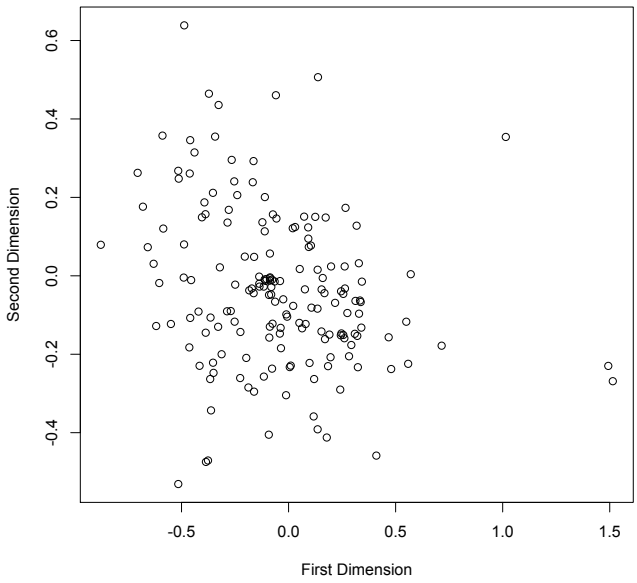


Figure 2: Distribution of policy positions of political parties in 23 member states (wordscored data)

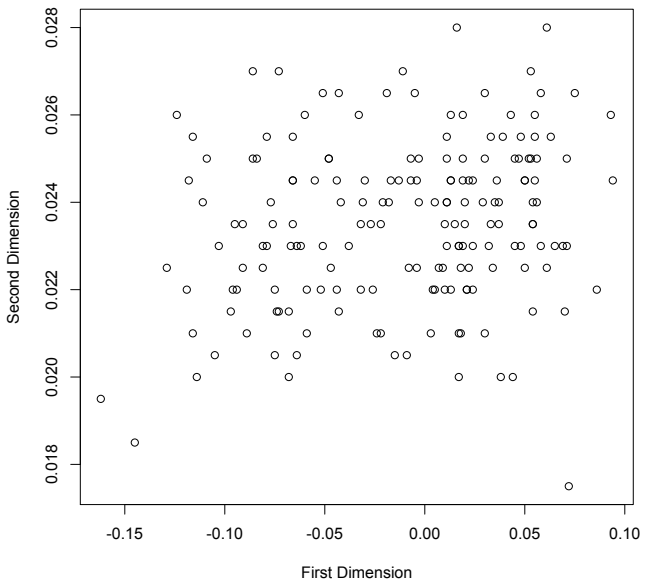


Figure 3: Distribution of policy positions per country (factoranalytic solution of hand coded data)

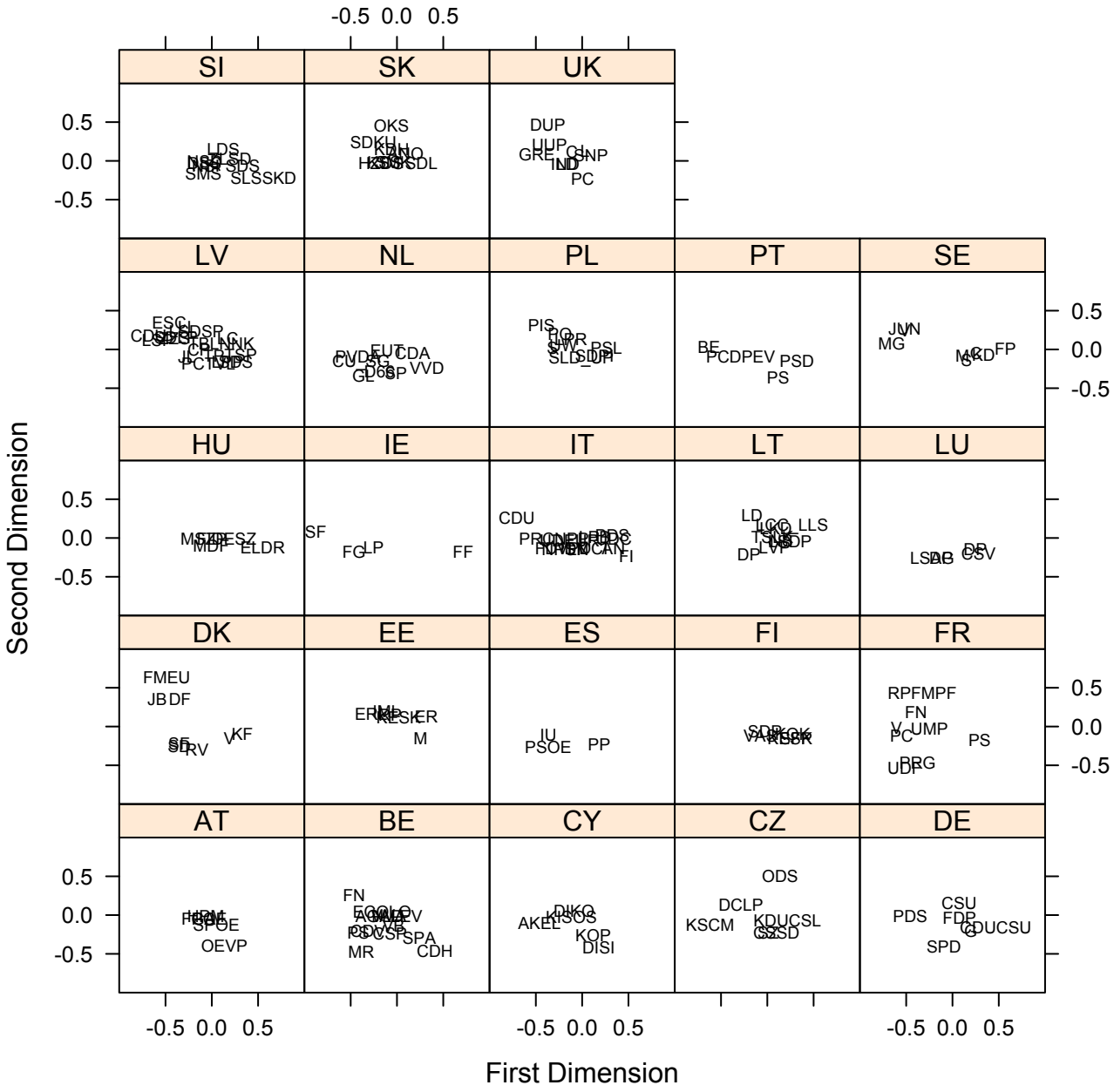
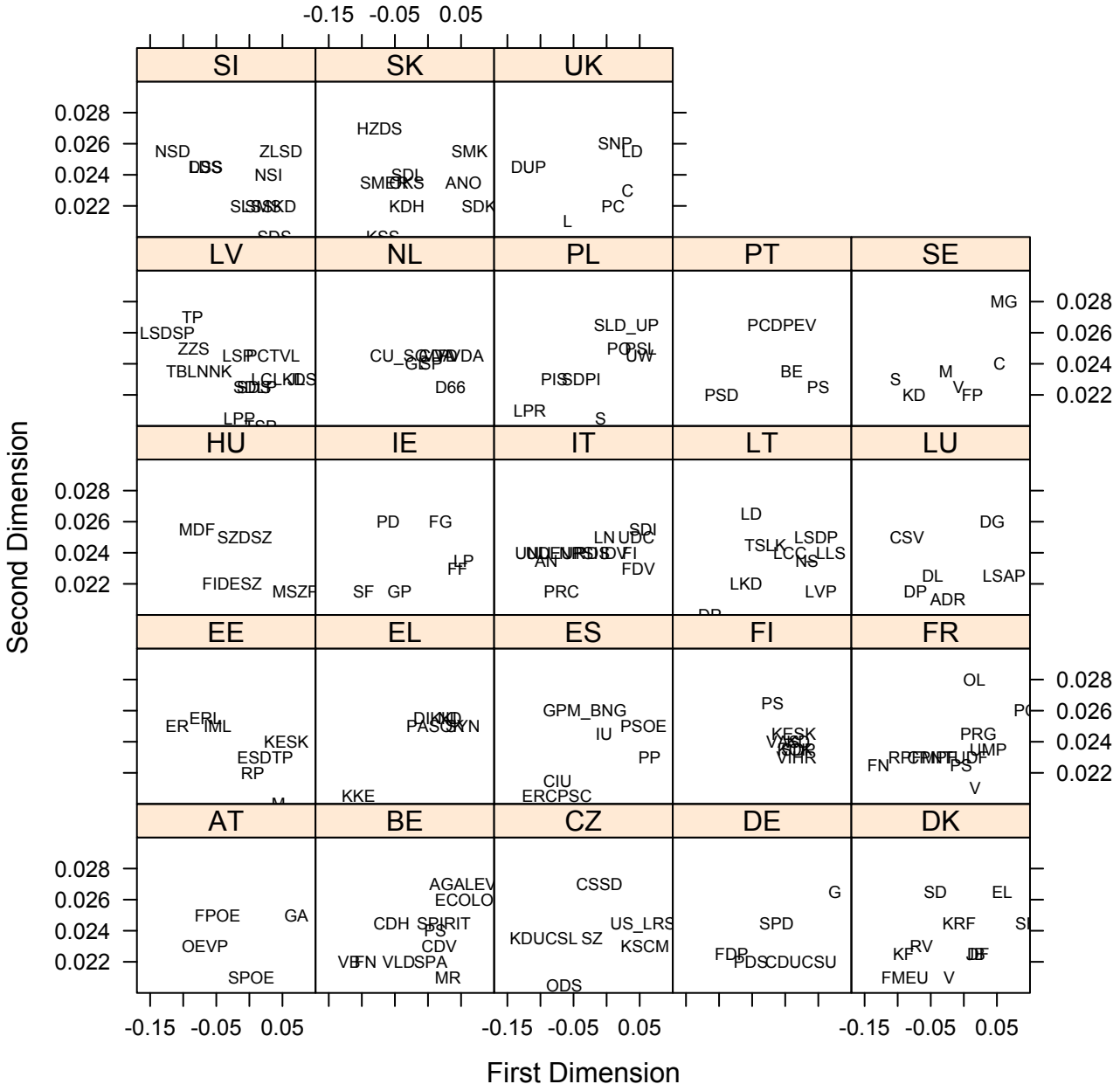


Figure 4: Distribution of policy positions per country (wordscored data)



To more systematically compare the derived policy positions for the parties from the hand coding of the party manifestos and the wordscores analyses we performed two simple linear regressions where one of the two factor analytic dimensions of the former analysis was the dependent variable, while the two wordscores dimensions were the two independent variables.[14]

Table 2: Explaining the two dimensions of the hand coded data

	Dependent variable hand coding dimension:			
	1	2	1	2
wordscores 1^{st} dimension	0.030	-1.658		
	(0.435)	(0.567)		
wordscores 2^{nd} dimension	-9.338	0.169		
	(6.854)	(8.930)		
1^{st} dimension Eurobarometer data			0.038	0.028
			(0.053)	(0.061)
2^{nd} dimension Eurobarometer data			0.078	-0.112
			(0.046)	(0.054)
constant	0.382	-0.117	-0.074	-0.070
	(0.323)	(0.421)	(0.034)	(0.039)
N	144	144	114	114
R^2	0.013	0.058	0.038	0.038
adj. R^2	-0.001	0.045	0.020	0.021
Resid. sd	0.298	0.388	0.310	0.358

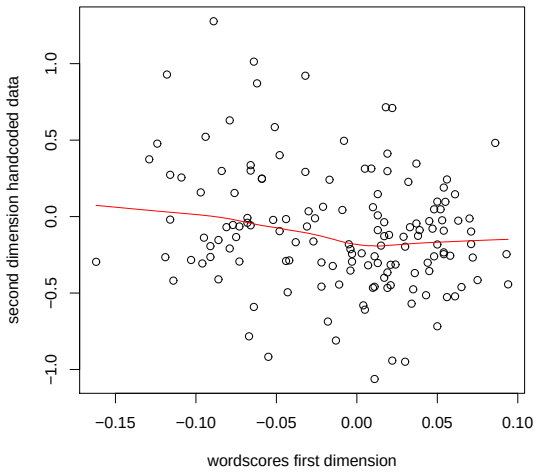
Standard errors in parentheses

The results reported in the first two columns of table 2 show that the first dimension of the factor analysis based on the hand coded data is negatively (though not statistically significantly) related to the second dimension of the wordscores analysis. The second dimension based on the hand coded data is also negatively and quite strongly (and statistically significantly) related to the first dimension from the wordscores analyses. We illustrate the former relationship graphically in figure 5 with the help of a lowess-curve.

In the last two columns of table 2 we report the results of two similar regression models that are based on differently derived policy positions for the political parties. More precisely Hug (2007) uses the data from the expert survey of government positions and combines it with the help of a bridging observation with aggregated data from Eurobarometer data on the preferences of party sympa-

[14]Proceeding this way allows for possible rotations of the two factor-analytic spaces.

Figure 5: Relationship between hand coded data and wordscore analysis



thizers.[15] As the results show, the first dimension based on the hand coded data is slightly related to the second dimension based on the Eurobarometer data, while this latter variable is strongly and negativelly related to the second dimension based on the hand coded data. We depict again this latter relationship graphically in figure 6.

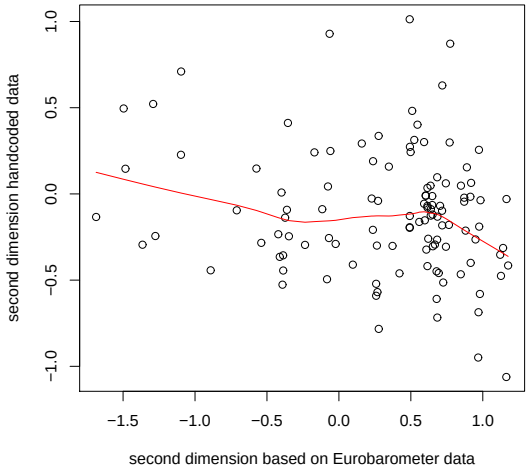
4 Conclusion

In this paper we have analyzed whether two fundamentally different techniques of deriving political positions from the same source of text produces similar results. The data source were the party manifestos of the European Party Election of 2004, and we analyzed them by applying the wordscores method on the one hand and a hand coding on the other hand. The variables resulting from the hand coding were brought in correspondence with the variables of the original DOSEI expert surveys, on which the coding scheme of the hand coding is based and which is also the basis of the reference scores that were applied in the wordscores analysis.

Given that the techniques applied are fundamentally different, the correspon-

[15]Finke (2009) proceeds similarly to link various treaty negotiations, while Shor, Berry and McCarty (2008) discuss some of the methodological issues related to this technique.

Figure 6: Relationship between hand coded data and Eurobarometer data



dence between the results is astonishingly high not only if the variables are compared directly (as done in a first plausibility test) but also if the wordscores dimensions are compared with the dimensions of a factor analysis of the hand coded data. This leads us to conclude that even fundamentally different methods lead to qualitatively similar results if they are applied to the same data and which one to choose is thus more a question of applicability and personal taste than a question of finding the true positions. This result is all the more interesting as in the current setting it seems that the hand coding method was clearly disadvantaged: given the detailed coding scheme derived from the questionnaire of an expert survey, only little information on these items could be found in the different texts. Nevertheless, the results are quite clearly correlated. Hence, the hand coding in the end was able to extract, more or less, the information that was "extractable" on the dimensions applied and it seems plausible that even the wordscores method using all words available in the texts, is not able to extract much more and much more different information, simply because it is not available in these texts.

Appendix

Tables 3-12 provide information on the correspondence of the hand codings and the expert survey responses on the positions of political parties. Table 13 provides information on the factor analysis of the hand coded party manifestos while table 14 reports the results of a wordscores analysis of the party manifestos for the 2004 European parliament election (*Source:* Hug and Schulz, 2007). Finally. table 15 provides information on how well the various dimensions of the factoranalytic solutions explained the underlying variables.

Table 3: Belgium: The Socialists

expert survey	party manifesto			
	1	2	3	4
1	0	2	1	0
2	0	14	0	1
3	1	7	2	1
4	0	0	0	1
5	0	0	0	0

Table 4: Czech ODS

expert survey	party manifesto		
	1	2	3
1	2	3	
1	2	1	1
2	8	0	0
3	1	1	0
4	0	0	0

Table 5: Czech CSSD

expert survey	party manifesto		
	2	3	4
1	2	0	1
2	6	0	1
3	1	1	1

Table 6: French UMP

expert survey	party manifesto			
	1	2	3	4
1	1	1	0	0
2	0	2	1	3
3	0	1	0	0
4	1	0	0	0
5	0	0	0	0

Table 7: Italian Lega Nord

expert survey	party manifesto			
	1	2	3	4
1	2	0	1	0
2	0	4	0	0
3	0	1	0	0
4	1	0	0	1

Table 8: Slobak KDH

expert survey	party manifesto			
	1	2	3	4
1	7	0	2	0
2	2	2	1	1
3	2	6	0	1
4	0	0	0	1

Table 9: Spanish PSOE

expert survey	party manifesto				
	1	2	3	4	5
1	0	0	0	0	0
2	0	13	0	0	0
3	0	9	3	0	0
4	0	0	0	1	0
5	0	0	0	0	1

Table 10: UK Government Labour

expert survey	party manifesto	
	1	2
1	0	0
2	1	3
3	2	1
4	0	0

Table 11: Spanish government PSOE

expert survey	party manifesto				
	1	2	3	4	5
1	0	5	0	0	0
2	0	14	2	0	0
3	0	3 1	1	0	
4	0	0	0	0	0
5	0	0	0	0	1

Table 12: Swedish Government Socialdemocrats

expert survey	party manifesto	
	2	3
1	0	1
2	1	2
3	3	0
4	0	0
5	0	0

Table 13: Factor Analysis of handcoded 2004 European Parliament Election Party Manifestos

Party	First Dimension (FD)	Second Dimension (SE)	S.E. FD	S.E. SD	t-statistic FD	t-satistic SD
SE C	0.256	-0.092	0.017	0.016	14.77	-5.795
SE FP	0.571	0.008	0.016	0.018	35.324	0.45
SE JUN	-0.515	0.536	0.022	0.022	-23.167	24.793
SE KD	0.337	-0.134	0.016	0.016	20.984	-8.205
SE MG	-0.658	0.146	0.02	0.022	-32.437	6.548
SE M	0.107	-0.162	0.015	0.015	7.24	-10.701
SE S	0.154	-0.284	0.016	0.016	9.328	-18.272
SE V	-0.513	0.496	0.018	0.018	-28.51	28.305
DK DF	-0.343	0.71	0.021	0.018	-16.195	38.483
DK FMEU	-0.487	1.277	0.032	0.025	-15.23	50.741
DK JB	-0.589	0.715	0.022	0.021	-26.363	34.772
DK KF	0.329	-0.193	0.016	0.016	21.062	-12.447
DK RV	-0.161	-0.591	0.019	0.017	-8.696	-34.132
DK SD	-0.351	-0.495	0.017	0.016	-21.179	-31.28
DK SF	-0.353	-0.443	0.016	0.016	-22.571	-27.602
DK V	0.191	-0.3	0.016	0.016	11.66	-19.316
FI VAS	-0.075	-0.244	0.018	0.017	-4.185	-14.238
FI KOK	0.275	-0.19	0.018	0.018	15.171	-10.739
FI SDP	-0.025	-0.12	0.015	0.015	-1.629	-8.019
FI KESK	0.244	-0.304	0.016	0.016	14.869	-18.634
FI SFP	0.309	-0.296	0.018	0.017	17.38	-16.932
BE CDH	0.41	-0.917	0.024	0.019	17.397	-47.284
BE CDV	-0.313	-0.4	0.014	0.014	-21.72	-27.699
BE ECOLO	-0.161	0.097	0.013	0.013	-12.381	7.638
BE MR	-0.386	-0.949	0.026	0.024	-14.607	-39.847
BE N-VA	-0.087	-0.008	0.02	0.02	-4.355	-0.396
BE PS	-0.416	-0.459	0.016	0.017	-25.997	-27.63
BE SPA	0.241	-0.58	0.018	0.016	13.189	-35.784
BE VB	-0.037	-0.265	0.019	0.019	-1.929	-13.996
BE FN	-0.462	0.522	0.02	0.019	-23.055	27.379
BE CSP	-0.077	-0.473	0.016	0.015	-4.691	-31.303
BE AGALEV	-0.087	-0.024	0.019	0.019	-4.503	-1.252
BE VLD	-0.103	-0.016	0.02	0.02	-5.15	-0.839
NL CDA	0.168	-0.088	0.014	0.013	12.407	-6.827
NL CU SG	-0.387	-0.29	0.014	0.016	-27.122	-18.45
NL ET	-0.108	-0.026	0.019	0.02	-5.547	-1.306
NL GL	-0.362	-0.686	0.019	0.018	-19.267	-38.962
NL SP	-0.012	-0.609	0.018	0.016	-0.644	-38.622
NL D66	-0.188	-0.57	0.018	0.018	-10.235	-32.236
NL PVDA	-0.421	-0.182	0.015	0.017	-27.641	-10.625
NL VVD	0.324	-0.466	0.019	0.018	17.144	-25.619
LU CSV	0.282	-0.41	0.017	0.016	16.716	-25.332
LU DG	-0.116	-0.514	0.018	0.018	-6.34	-29.243
LU DP	0.247	-0.294	0.017	0.017	14.552	-17.368
LU LSAP	-0.226	-0.521	0.018	0.017	-12.666	-31.129
FR FN	-0.393	0.374	0.018	0.018	-21.396	20.502
FR V	-0.604	-0.037	0.019	0.022	-31.494	-1.688
FR RPFMPF	-0.327	0.871	0.024	0.02	-13.732	44.513
FR PC	-0.55	-0.246	0.017	0.019	-33.146	-12.753
FR PRG	-0.376	-0.942	0.028	0.026	-13.59	-36.423
FR PS	0.293	-0.353	0.016	0.015	18.763	-23.703
FR UDF	-0.515	-1.062	0.027	0.025	-18.762	-41.943
FR UMP	-0.249	-0.045	0.011	0.012	-22.085	-3.811
IT AN	0.34	-0.264	0.018	0.018	18.822	-14.306
IT PDS	0.328	0.064	0.015	0.016	21.454	3.956
IT FDV	-0.329	-0.26	0.019	0.019	-17.687	-13.662
IT FI	0.479	-0.475	0.021	0.02	23.209	-24.255
IT LN	-0.041	-0.295	0.017	0.016	-2.424	-18.056
IT IDV	-0.088	-0.259	0.018	0.018	-4.823	-14.679
IT LPB	0.136	0.031	0.016	0.017	8.355	1.856
IT NPP	-0.112	-0.019	0.02	0.02	-5.612	-0.956

Table 13: *(continued)*

Party	First Dimension (FD)	Second Dimension (SE)	S.E. FD	S.E. SD	t-statistic FD	t-satistic SD
IT NPSI	-0.225	-0.286	0.019	0.019	-11.824	-14.986
IT PDCI	0.063	-0.268	0.017	0.016	3.773	-16.617
IT PRC	-0.49	-0.009	0.017	0.019	-29.184	-0.479
IT UDC	0.342	-0.03	0.017	0.018	19.701	-1.605
IT UDEUR	-0.136	-0.055	0.011	0.011	-12.379	-4.857
ES IU	-0.364	-0.213	0.013	0.014	-28.994	-14.971
ES PP	0.184	-0.461	0.015	0.012	12.704	-36.977
ES PSOE	-0.366	-0.527	0.018	0.018	-20.658	-28.636
PT BE	-0.63	0.061	0.017	0.019	-37.513	3.209
PT CDU	-0.705	0.525	0.021	0.021	-33.123	25.026
PT PCDPEV	-0.287	-0.181	0.012	0.012	-24.164	-14.628
PT PSD	0.321	-0.306	0.018	0.018	17.864	-17.193
PT PS	0.117	-0.717	0.019	0.015	6.114	-47.429
DE CSU	0.073	0.302	0.014	0.013	5.246	22.389
DE FDP	0.076	-0.069	0.011	0.01	7.041	-7.161
DE PDS	-0.455	-0.022	0.013	0.015	-34.045	-1.424
DE SPD	-0.092	-0.81	0.023	0.02	-3.969	-41.024
DE G	0.196	-0.415	0.013	0.011	15.178	-38.183
DE CDU	0.467	-0.314	0.018	0.018	25.922	-17.141
AT HPM	-0.065	-0.029	0.019	0.02	-3.405	-1.444
AT OEVP	0.136	-0.783	0.023	0.02	5.948	-39.404
AT FPOE	-0.081	-0.095	0.018	0.019	-4.431	-5.055
AT GA	-0.091	-0.098	0.019	0.018	-4.934	-5.286
AT SPOE	0.05	-0.24	0.017	0.016	2.91	-15.076
UK C	-0.111	0.227	0.009	0.009	-12.053	26.024
UK DUP	-0.372	0.929	0.026	0.022	-14.424	42.989
UK GP	-0.488	0.16	0.014	0.016	-35.219	10.059
UK LD	-0.163	-0.089	0.009	0.009	-18.247	-9.638
UK L	0.03	0.249	0.015	0.014	2.078	17.804
UK SNP	0.094	0.147	0.015	0.014	6.188	10.352
UK PC	0.006	-0.465	0.019	0.018	0.288	-26.193
UK IND	-0.184	-0.074	0.019	0.02	-9.918	-3.752
UK UUP	-0.353	0.423	0.017	0.016	-20.58	25.969
IE SF	-0.877	0.158	0.02	0.023	-44.809	6.786
IE FF	0.715	-0.356	0.026	0.028	27.246	-12.8
IE FG	-0.463	-0.365	0.016	0.017	-29.416	-21.192
IE LP	-0.252	-0.234	0.012	0.012	-21.523	-19.341
CO EFA	-0.107	-0.268	0.015	0.016	-6.98	-17.135
CO ELDR	-0.073	-0.225	0.013	0.011	-5.815	-20.434
CO EPPE	0.356	-0.177	0.017	0.017	21.31	-10.517
CO GM	-0.188	-0.667	0.02	0.019	-9.26	-34.826
CO PES	-0.401	-0.623	0.02	0.02	-19.702	-31.53
CO EUL	-0.1	-0.201	0.019	0.019	-5.399	-10.774
CZ CSSD	0.098	-0.444	0.018	0.016	5.581	-27.971
CZ DCLP	-0.285	0.272	0.013	0.013	-22.197	21.153
CZ KDUCSL	0.217	-0.138	0.015	0.013	14.871	-10.225
CZ KSCM	-0.618	-0.256	0.018	0.021	-34.01	-12.425
CZ ODS	0.137	1.013	0.026	0.021	5.375	47.808
CZ SZ	0.01	-0.458	0.019	0.018	0.521	-26.133
CY akel	-0.459	-0.215	0.016	0.018	-28.009	-12.083
CY diko	-0.087	0.114	0.018	0.017	-4.724	6.604
CY disi	0.178	-0.825	0.022	0.018	8.102	-46.899
CY kisos	-0.116	-0.055	0.019	0.02	-6.048	-2.75
CY kop	0.119	-0.526	0.02	0.017	6.038	-30.782
EE ER	0.318	0.256	0.017	0.017	18.737	15.1
EE ERL	-0.279	0.336	0.014	0.013	-20.496	26.712
EE IML	-0.111	0.402	0.013	0.011	-8.251	35.143
EE KESK	0.019	0.243	0.01	0.009	1.993	25.888
EE M	0.256	-0.302	0.016	0.015	15.725	-20.208
EE RP	-0.073	0.313	0.01	0.009	-7.445	34.797
HU ELDR	0.55	-0.234	0.016	0.016	33.872	-14.208

Table 13: *(continued)*

Party	First Dimension (FD)	Second Dimension (SE)	S.E. FD	S.E. SD	t-statistic FD	t-satistic SD
HU FIDESZ	0.159	-0.011	0.011	0.011	14.012	-1.023
HU MDF	-0.007	-0.208	0.016	0.015	-0.417	-13.93
HU MSZP	-0.087	-0.012	0.02	0.019	-4.434	-0.642
LT DP	-0.198	-0.419	0.018	0.017	-10.765	-24.216
LT LCC LLS	0.266	0.347	0.019	0.017	14.289	19.895
LT LD	-0.164	0.585	0.021	0.019	-7.93	31.031
LT LKD	0.092	0.247	0.016	0.016	5.721	15.359
LT LSDP	0.245	-0.079	0.016	0.016	14.878	-4.79
LT LVP	0.079	-0.246	0.018	0.018	4.45	-13.905
LT NS	0.154	-0.069	0.014	0.013	11.216	-5.166
LT TSLK	0.052	0.035	0.012	0.012	4.32	2.973
LV ESC	-0.46	0.692	0.024	0.021	-19.308	32.408
LV LKDS	-0.68	0.352	0.019	0.02	-35.887	17.267
LV CONS	-0.137	-0.004	0.019	0.019	-7.226	-0.214
LV TBLNNK	0.104	0.154	0.015	0.014	7.15	11.24
LV LPP	0.17	-0.323	0.017	0.016	9.928	-20.649
LV PCTVL	-0.036	-0.369	0.019	0.017	-1.859	-21.151
LV LC	0.173	0.297	0.018	0.018	9.854	16.867
LV LL	-0.266	0.591	0.018	0.015	-14.975	39.29
LV LSDSP	-0.388	0.314	0.014	0.014	-28.004	22.911
LV LSP	-0.585	0.241	0.019	0.02	-31.623	12.278
LV JL	-0.27	-0.18	0.018	0.018	-15.311	-10.242
LV TSP	0.312	-0.128	0.013	0.014	23.693	-9.318
LV TP	0.022	-0.153	0.018	0.017	1.233	-8.772
LV SDS	0.261	-0.318	0.014	0.013	18.473	-23.893
LV ZZS	-0.405	0.298	0.017	0.017	-23.306	17.325
LV SDLP	-0.167	0.477	0.02	0.017	-8.413	27.304
PL LPR	-0.122	0.273	0.019	0.018	-6.396	15.172
PL PIS	-0.439	0.629	0.021	0.019	-21.426	33.834
PL PO	-0.24	0.412	0.015	0.013	-16.017	30.725
PL PSL	0.262	0.048	0.016	0.016	16.447	2.917
PL S	-0.321	0.043	0.019	0.02	-17.119	2.193
PL SDPI	0.135	-0.167	0.014	0.014	9.385	-11.956
PL SLD UP	-0.01	-0.197	0.018	0.019	-0.571	-10.53
PL UW	-0.204	0.098	0.019	0.019	-10.75	5.197
SI DSS	-0.081	-0.057	0.019	0.02	-4.153	-2.927
SI LDS	0.125	0.301	0.017	0.015	7.399	20.321
SI LDST	-0.079	0.004	0.019	0.019	-4.08	0.215
SI NSD	-0.077	-0.02	0.02	0.02	-3.945	-1.014
SI NSDT	-0.144	-0.023	0.02	0.02	-7.297	-1.189
SI NSI	-0.063	-0.132	0.014	0.014	-4.48	-9.483
SI SDS	0.334	-0.126	0.015	0.016	22.141	-8.011
SI SLSSKD	0.559	-0.449	0.02	0.019	28.52	-24.124
SI SMS	-0.09	-0.315	0.018	0.017	-4.93	-18.033
SI ZLSD	0.198	0.048	0.014	0.014	14.48	3.421
SK ANO	0.092	0.19	0.015	0.014	6.024	13.422
SK HZDS	-0.172	-0.065	0.012	0.012	-14.006	-5.363
SK KDH	-0.057	0.292	0.012	0.011	-4.723	26.747
SK KSS	-0.137	-0.04	0.015	0.014	-9.388	-2.889
SK OKS	-0.059	0.921	0.025	0.019	-2.405	49.407
SK SDKU	-0.254	0.482	0.016	0.014	-16.275	34.803
SK SDL	0.263	-0.065	0.015	0.016	17.527	-4.163
SK SMK	-0.04	-0.027	0.02	0.02	-2.044	-1.354
IGC	1.493	-0.459	0.031	0.037	48.875	-12.442
DFT	1.514	-0.537	0.032	0.038	46.998	-14.18
STQ	1.015	0.708	0.026	0.03	39.087	23.759

Table 14: Wordscores of 2004 European Parliament Election Party Manifestos

	First Dimension		Second Dimension		Words Scored	
Party	Score	SE	Score	SE	Total	Unique
Status Quo	-0.081		0.046			
Draft Const.	0.030		0.050			
IGC Outcome	0.018		0.042			
AT FPOE	-0.048	0.041	0.050	0.017	201	93
AT GA	0.071	0.027	0.050	0.013	318	144
AT OEVP	-0.067	0.010	0.046	0.004	3236	689
AT SPOE	0.003	0.016	0.042	0.008	1188	338
BE AGALEV	0.053	0.017	0.054	0.002	5665	753
BE CDV	0.017	0.016	0.046	0.002	6798	941
BE SPA	0.004	0.014	0.044	0.002	8523	994
BE SPIRIT	0.024	0.028	0.049	0.004	2404	558
BE VB	-0.119	0.011	0.044	0.001	7632	1343
BE VLD	-0.044	0.019	0.044	0.003	5151	737
BE CDH	-0.055	0.029	0.049	0.001	2108	1400
BE ECOLO	0.055	0.041	0.052	0.002	5338	878
BE FN	-0.094	0.044	0.044	0.002	4965	912
BE MR	0.030	0.104	0.042	0.004	901	339
BE PS	0.011	0.022	0.048	0.001	9410	1624
CZ CSSD	-0.011	0.062	0.054	0.003	1058	415
CZ KDUCSL	-0.095	0.037	0.047	0.002	2191	696
CZ KSCM	0.058	0.049	0.046	0.003	1365	495
CZ ODS	-0.064	0.052	0.041	0.003	1182	468
CZ SZ	-0.022	0.023	0.047	0.001	6287	1279
CZ US lrs	0.055	0.042	0.049	0.003	1844	669
DE CDUCSU	0.024	0.039	0.044	0.005	4323	737
DE FDP	-0.081	0.031	0.045	0.003	7309	1028
DE G	0.075	0.021	0.053	0.003	4542	1379
DE PDS	-0.052	0.021	0.044	0.003	3038	1390
DE SPD	-0.013	0.055	0.049	0.006	2196	501
DK DF	0.022	0.042	0.045	0.003	2133	511
DK EL	0.058	0.059	0.053	0.004	1154	326
DK FMEU	-0.089	0.063	0.042	0.005	188	36
DK JB	0.018	0.043	0.045	0.003	2467	494
DK KF	-0.091	0.042	0.045	0.003	2250	450
DK KRF	-0.007	0.063	0.049	0.004	1012	313
DK RV	-0.064	0.031	0.046	0.002	892	73
DK SD	-0.043	0.039	0.053	0.002	2869	606
DK SF	0.094	0.027	0.049	0.002	4755	782
DK V	-0.022	0.036	0.042	0.002	3022	625
EE ER	-0.109	0.033	0.050	0.004	481	153
EE ERL	-0.066	0.028	0.051	0.003	638	319
EE ESDTP	0.024	0.017	0.046	0.002	1039	361
EE IML	-0.048	0.025	0.050	0.003	741	299
EE KESK	0.056	0.026	0.048	0.003	535	277
EE M	0.044	0.025	0.040	0.003	644	274
EE RP	0.005	0.011	0.044	0.001	3224	905
EL DIKKI	0.011	0.022	0.051	0.001	3759	516
EL KKE	-0.105	0.021	0.041	0.001	3524	497
EL ND	0.033	0.025	0.051	0.001	2858	498
EL PASOK	0.011	0.030	0.050	0.001	2124	388
EL SYN	0.053	0.037	0.050	0.002	1220	291
ES CIU	-0.074	0.013	0.043	0.001	13239	393
ES ERCPSC	-0.075	0.016	0.041	0.001	8988	333
ES GPM BNG	-0.033	0.019	0.052	0.001	11835	934
ES IU	-0.004	0.014	0.049	0.001	6467	1471
ES PP	0.065	0.019	0.046	0.001	9760	1142
ES PSOE	0.056	0.025	0.050	0.001	5727	905
FI KD	0.020	0.016	0.048	0.001	1451	628
FI KESK	0.013	0.015	0.049	0.001	1664	611
FI KOK	0.015	0.018	0.047	0.001	1177	480

Table 14: *(continued)*

	First Dimension		Second Dimension		Words Scored	
Party	Score	SE	Score	SE	Total	Unique
FI PS	-0.019	0.051	0.053	0.003	155	102
FI SDP	0.020	0.020	0.047	0.001	956	414
FI SFP	-0.162	0.060	0.039	0.001	149	15
FI VAS	-0.003	0.029	0.048	0.002	374	193
FI VIHR	0.017	0.012	0.046	0.001	2570	838
FR CPNT	-0.051	0.015	0.046	0.001	7814	728
FR FN	-0.129	0.017	0.045	0.002	6197	967
FR OL	0.016	0.041	0.056	0.005	875	249
FR PC	0.093	0.021	0.052	0.003	2308	403
FR PRG	0.022	0.034	0.049	0.004	1300	378
FR PS	-0.004	0.016	0.045	0.002	5676	861
FR RPFMPF	-0.062	0.008	0.046	0.001	4858	1844
FR UDF	0.011	0.012	0.046	0.001	9889	1094
FR UMP	0.037	0.030	0.047	0.003	1830	489
FR V	0.017	0.008	0.042	0.001	1922	1591
HU SZDSZ	-0.007	0.015	0.050	0.003	2475	537
HU FIDESZ	-0.026	0.008	0.044	0.002	3577	376
HU MDF	-0.079	0.027	0.051	0.006	83	2071
HU MSZP	0.070	0.046	0.043	0.009	984	205
IE FF	0.045	0.050	0.046	0.002	586	186
IE FG	0.019	0.045	0.052	0.002	6054	844
IE GP	-0.043	0.058	0.043	0.003	7428	1023
IE LP	0.054	0.059	0.047	0.003	4290	772
IE PD	-0.060	0.049	0.052	0.002	3818	693
IE SF	-0.097	0.021	0.043	0.001	6775	883
IT AN	-0.091	0.028	0.047	0.002	2196	1252
IT FDV	0.048	0.028	0.046	0.002	1757	469
IT FI	0.035	0.017	0.048	0.001	1411	425
IT IDV	0.011	0.019	0.048	0.001	3934	733
IT LN	-0.003	0.014	0.050	0.001	2878	717
IT NPSI	-0.042	0.011	0.048	0.001	6379	1042
IT PDCI	0.072	0.044	0.035	0.005	741	1143
IT PDS	-0.021	0.013	0.048	0.001	540	204
IT PRC	-0.068	0.019	0.043	0.001	6732	949
IT SDI	0.055	0.022	0.051	0.002	3004	603
IT UDC	0.045	0.040	0.050	0.003	2246	548
IT UDEUR	-0.077	0.009	0.048	0.001	911	345
IT UNU	-0.111	0.046	0.048	0.003	5562	1481
LT DP	-0.114	0.053	0.040	0.003	609	212
LT LCC LLS	0.037	0.047	0.048	0.003	458	261
LT LD	-0.051	0.106	0.053	0.005	720	369
LT LKD	-0.059	0.046	0.044	0.003	194	94
LT LSDP	0.047	0.051	0.050	0.003	785	449
LT LVP	0.054	0.048	0.043	0.003	541	303
LT NS	0.033	0.043	0.047	0.003	560	294
LT TSLK	-0.030	0.019	0.049	0.001	859	433
LU ADR	-0.024	0.027	0.042	0.002	4184	1178
LU CSV	-0.086	0.013	0.050	0.001	4962	891
LU DG	0.043	0.021	0.052	0.002	4271	1723
LU DL	-0.047	0.020	0.045	0.001	8935	1052
LU DP	-0.073	0.025	0.043	0.002	8167	1018
LU LSAP	0.061	0.051	0.045	0.003	1547	430
LV JL	0.071	0.024	0.046	0.004	286	159
LV LC	0.019	0.023	0.046	0.004	297	165
LV LKDS	0.069	0.021	0.046	0.004	333	176
LV LPP	-0.015	0.023	0.041	0.004	302	168
LV LSDSP	-0.124	0.028	0.052	0.003	328	184
LV LSP	-0.017	0.025	0.049	0.004	302	172
LV PCTVL	0.036	0.023	0.049	0.004	302	157
LV SDLP	0.009	0.031	0.045	0.005	188	128

Table 14: *(continued)*

	First Dimension		Second Dimension		Words Scored	
Party	Score	SE	Score	SE	Total	Unique
LV SDS	0.007	0.022	0.045	0.003	355	210
LV TBLNNK	-0.076	0.027	0.047	0.003	327	172
LV TP	-0.086	0.032	0.054	0.004	240	142
LV TSP	0.017	0.023	0.040	0.004	339	201
LV ZZS	-0.084	0.033	0.050	0.004	211	144
NL CDA	0.013	0.018	0.049	0.001	6349	792
NL CU SG	-0.044	0.012	0.049	0.001	6798	1166
NL D66	0.034	0.022	0.045	0.001	4355	728
NL GL	-0.018	0.011	0.048	0.001	7260	1437
NL LPF	-0.145	0.045	0.037	0.002	1390	361
NL PVDA	0.050	0.016	0.049	0.001	7235	959
NL SP	0.005	0.015	0.048	0.001	646	1078
NL VVD	0.019	0.050	0.049	0.003	958	321
PL LPR	-0.116	0.009	0.042	0.001	6136	1043
PL PIS	-0.079	0.044	0.046	0.006	401	210
PL PO	0.019	0.027	0.050	0.004	805	352
PL PSL	0.052	0.028	0.050	0.004	688	289
PL S	-0.009	0.061	0.041	0.007	170	106
PL SDPI	-0.038	0.059	0.046	0.009	154	93
PL SLD UP	0.030	0.034	0.053	0.006	452	252
PL UW	0.050	0.056	0.049	0.010	162	82
PT BE	0.010	0.029	0.047	0.003	2260	479
PT PCDPEV	-0.005	0.007	0.053	0.001	5658	1792
PT PS	0.050	0.017	0.045	0.002	6943	1016
PT PSD	-0.096	0.029	0.044	0.003	2893	658
SE C	0.054	0.013	0.048	0.001	3273	640
SE FP	0.013	0.011	0.044	0.001	4875	795
SE KD	-0.075	0.024	0.044	0.001	1274	362
SE M	-0.027	0.019	0.047	0.001	1611	388
SE MG	0.061	0.049	0.056	0.004	177	104
SE S	-0.103	0.038	0.046	0.002	660	247
SE V	-0.008	0.020	0.045	0.001	1571	433
SI DSS	-0.066	0.037	0.049	0.003	324	123
SI LDS	-0.066	0.037	0.049	0.003	324	123
SI NSD	-0.116	0.048	0.051	0.003	239	103
SI NSI	0.029	0.032	0.048	0.003	464	192
SI SDS	0.038	0.018	0.040	0.002	1481	543
SI SLSSKD	0.021	0.016	0.044	0.001	2169	609
SI SMS	0.021	0.016	0.044	0.001	2169	609
SI ZLSD	0.048	0.017	0.051	0.002	1822	484
SK ANO	0.054	0.138	0.047	0.005	79	59
SK HZDS	-0.073	0.040	0.054	0.002	529	129
SK KDH	-0.032	0.030	0.044	0.001	1270	575
SK KSS	-0.068	0.100	0.040	0.004	154	102
SK OKS	-0.032	0.017	0.047	0.001	4048	1033
SK SDKU	0.086	0.032	0.044	0.001	1208	463
SK SDL	-0.031	0.037	0.048	0.002	807	357
SK SMER	-0.066	0.026	0.047	0.001	1555	521
SK SMK	0.063	0.099	0.051	0.004	146	99
UK C	0.032	0.018	0.046	0.004	8176	964
UK DUP	-0.118	0.025	0.049	0.005	5006	757
UK L	-0.059	0.024	0.042	0.006	4989	680
UK LD	0.039	0.017	0.051	0.004	9395	1055
UK PC	0.010	0.033	0.044	0.007	2467	521
UK SNP	0.013	0.022	0.052	0.006	5921	765

The 'Score" is the *transformed* score calculated by the Wordscores program and consequently "SE" is the standard error of the transformed scores. "Unique" is the number of unique scored words and 'Total" the total of the words contained in a text.

Table 15: PRE Measures for Regressions of handcoded Items on different dimensional factoranalytic fits of the data

Item	1.dim.	2.dim.	3.dim.	4.dim.
q1	0.111	0.444	0.556	0.444
q2	0.152	0.261	0.261	0.457
q3	0	0	0	0.455
q4	0	1	1	1
q5a	0	0	0.467	0.133
q5b	0.074	0.222	0.667	0.889
q5c	0.077	0.128	0.154	0.077
q6	0.273	0.455	0.455	0.636
q7	0.091	1	1	0.273
q8	0	0	NA	0.25
q9	0	0.4	0.6	NA
q10	-0.2	1	1	1
q11	0.2	1	NA	1
q12	0	0.375	0.125	0.375
q14	-0.1	0.3	0.7	1
q15b	1	1	1	1
q15c	1	1	1	NA
q15e	0	1	1	1
q16	0	0.143	0.429	0.429
q171	0.355	0.452	0.613	0.548
q172	-0.2	0.6	0.4	0.6
q173	0.375	0.688	0.25	0.562
q174	0.706	0.824	NA	NA
q175	0.727	0.727	1	NA
q176	0.269	0.692	1	0.577
q177	0	1	1	1
q178	0	0.222	NA	0.167
q179	0	1	0.5	NA
q1710	0	0.167	0.167	NA
q1711	0	0	0.4	0
q1712	1	1	1	1
q18a2	1	1	1	1
q18a3	0.5	0.667	0.5	NA
q18a5	-0.143	0.714	0.714	1
q18a6	0.333	1	1	1
q18a7	0.8	1	1	NA
q18a8	1	1	1	1
q18a9	0.625	1	1	NA
q18b1	0	1	1	NA
q18b2	0	0	0	1
q18b3	0	1	1	1
q18b4	NA	NA	NA	NA
q18b5	1	1	1	1
q18b6	1	1	1	1
q18b7	1	1	1	1
q18b8	1	1	1	1
q18b9	0	0	0	1
q18b10	1	1	1	1
q18b11	1	1	1	1
q18b12	1	1	1	1
q19	0.1	0	0	0
q20	0.903	1	0.903	0.839
q21	0	1	1	NA
q22	0	0	0.227	0.182
q23	0.143	0.371	0.4	0.343
q24	0.839	0.903	0.871	0.839

References

Achen, Christopher H. 1977. "Measuring Representation: Perils of the Correlation Coefficient." *American Journal of Political Science* 21(4):805–815.

Achen, Christopher H. 1982. *Interpreting and Using Regression*. Beverly Hills: Sage Publications.

Adler, E. Scott and John Wilkerson. 2005. "The Scope and Urgency of Legislation." APSA.

Arnold, Christine and Mark Franklin. 2006. "Heterogeneity and Level of Public Support for European Integration: Exploring the Dynamics of European Policy-Making." Paper presented at the annual meeting of the American Political Science Association, Marriott, Loews Philadelphia, and the Pennsylvania Convention Center, Philadelphia, PA, Aug 31, 2006.

Baumgartner, Frank R. and Bryan D. Jones. 2005. *The Politics of Attention. How Government Prioritizes Problems*. Chicago: University of Chicago Press.

Baumgartner, Frank R. and Bryan D. Jones, eds. 2002. *Policy Dynamics*. Chicago: University of Chicago Press.

Benoit, Kenneth and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.

Benoit, Kenneth and Michael Laver. 2007. "Estimating Party Policy Positions: Comparing Expert Surveys and Hand-Coded Content Analysis." *Electoral Studies* 26(1):90–107.

Benoit, Kenneth, Michael Laver, Christine Arnold, Paul Pennings and Madeleine O. Hosli. 2005. "Measuring National Delegate Positions at the Convention on the Future of Europe Using Computerized Wordscoring." *European Union Politics* 6(3):291–313.

Benoit, Kenneth, Slava Mikhaylov and Michael Laver. 2009 (forthcoming). "Treating Words as Data with Error: Estimating Uncertainty in Text Statements of Policy Positions." *American Journal of Political Science* 53(2).

Bertelli, Anthony M. and Christian R. Grose. 2006. "The Spatial Model and the Senate Trial of President Clinton." *American Politics Research* 34:535–559.

Budge, Ian. 2001. Validating the Manifesto Research Group Approach: Theoretical Assumptions and Empirical Confirmations. In *Estimating the Policy Position of Political Actors*, ed. Michael Laver. London: Routledge pp. 50–65.

Budge, Ian, Hans-Dieter Klingemann, Volkens Andrea and Judi Bara, eds. 2001. *Mapping Policy Preferences : Estimates for Parties, Electors, and Governments, 1945-1998*. Oxford: Oxford University Press.

Cardie, Claire and John Wilkerson. 2008. "Guest Editors' Introduction: Text Annotation for Political Science Research." *Journal of Information Technology & Politics* 5(1):1–6.

Castles, Francis G. and Peter Mair. 1984. "Left-Right Political Scales: Some 'Expert' Judgements." *European Journal of Political Research* 12(1):73–88.

Dahl, Robert A. 1971. *Polyarchy. Participation and Opposition*. New Haven: Yale University Press.

Diermeier, Daniel, J.-F. Godbout, B. Yu and S. Kaufmann. 2007. "Language and ideology in Congress." MPSA 2007, Chicago.

Evans, Michael, Wayne McIntosh, Jimmy Lin and Cynthia Cates. 2007. "Recounting the Courts? Applying Automated Content Analysis to Enhance Empirical Legal Research." *Journal of Empirical Legal Studies* 4(4):1007–1039.

Finke, Daniel. 2009. "Challenges to Intergovernmentalism: An Empirical Analysis of EU Treaty Negotiations since Maastricht." *West European Politics* 32(3):466 – 495.

Gabel, Matthew J. and Christopher J. Anderson. 2004. The Structure of Citizen Attitudes and the European Political Space. In *European Integration and Political Conflict*, ed. Gary Marks and Marco R. Steenbergen. Cambridge: Cambridge University Press chapter 1, pp. 13–31.

Gabel, Matthew J. and John D. Huber. 2000. "Putting Parties in their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data." *American Journal of Political Science* 44(1):94–103.

Gerner, Deborah J., Philip A. Schrodt, Ronald Francisco and Judith L. Weddle. 1994. "The Analysis of Political Events Using Machine Coded Data." *International Studies Quarterly* 38(march):91–119.

Giannetti, Danielle and Michael Laver. 2005. "Policy Position and Jobs in the Governemnt." *European Journal of Political Research* 44(1):91–120.

Hansen, Martin Ejnar. 2008. "Back to the Archives? A critique of the Danish part of the manifesto dataset." *Scandinavian Political Studies* 31(2):201–216.

Helbling, Marc and Anke Tresch. 2009. "Measuring party positions and issue salience from media coverage. Discussing and cross-validating new indicators." Paper prepared for presentation at the Anual Meeting of the American Political Science Association, Toronto, September 3-6, 2009.

Hillard, Dustin, Stephen Purpura and John Wilkerson. 2007. "An Active Learning Framework for Classifying Political Text." Paper presented at the annual meeting of the Midwest Political Science Association, Palmer House Hotel, Chicago, IL, Apr 12, 2007.

Hix, Simon. 2002. "Parliamentary Behavior with Two Principals: Preferences, Parties, and Voting in the European Parliament." *American Journal of Political Science* 46(3):688–698.

Hix, Simon, Abdul Noury and Gérard Roland. 2006. *Democracy in the European Parliament*. Cambridge: Cambridge University Press.

Hix, Simon and Christophe Crombez. 2005. "Extracting Ideal Point Estimates from Actors' Preferences in the EU Constitutional Negotiations." *European Union Politics* 6(3):353–376.

Hooghe, Liesbet, Gary Marks and Carole J. Wilson. 2004. Does Left/Right Structure Party Positions on European Integration? In *European Integration and Political Conflict*, ed. Gary Marks and Marco R. Steenbergen. Cambridge: Cambridge University Press chapter 6, pp. 120–140.

Hopkins, Daniel and Gary King. 2007. "Extracting Systematic Social Science Meaning from Text." Department of Government, Harvard University andInstitute for Quantitative Social Science.

Huber, John and Ronald Inglehart. 1995. "Expert Interpretations of Party Space and Party Locations in 42 Societies." *Party Politics* 1(1 (January)):73–111.

Hug, Simon. 2007. "Multilateral Bargaining at the 2004 IGC: An Empirical Assessment." EUSA Conference 2007: Montreal.

Hug, Simon and Thomas König. 2006. Conclusion. In *Policy-Making Processes and the European Constitution: A Comparative Study of Member States and Accession Countries*, ed. Thomas König and Simon Hug. London: Routledge pp. 260–278.

Hug, Simon and Tobias Schulz. 2005. "Using Mass Survey Data to Infer Political Positions." *European Union Politics* 6(3):339–352.

Hug, Simon and Tobias Schulz. 2007. "Referendums in the EU's Constitution Building Process." *Review of International Organizations* 2(2):177–218.

Joachims, T. 1998. "Text categorization with support vector machines: learning with many relevant features." *European conference on machine learning, ECML-98* p. 137-142.

Jurafsky, Daniel and James H. Martin. 2000. *Speech and Language Processing: An Introduction to Natural Language Processing, Speech Recognition, and Computational Linguistics*. Prentice-Hall.

Klebanov, Beata Beigman, Daniel Diermeier and Eyal Beigman. 2009. "Lexical Cohesion Analysis of Political Speech." *Political Analysis* .

Klemmensen, Robert, Sara B. Hobolt and Martin Ejnar Hansen. 2007. "Estimating policy positions using political texts: An evaluation of the Wordscores approach." *Electoral Studies* 26:746–755.

Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Michael D. McDonald and Ian Budge. 2007. *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union, and OECD 1990-2003*. Oxford: Oxford University Press,.

Klüver, Heike. 2009 (forthcoming). "Measuring interest group influence using quantitative text analysis." *European Union Politics.* .

König, Thomas and Daniel Finke. 2007. "Reforming the Equilibrium? Veto Players and Policy Change in the European Constitution-Building Process." *Review of International Organizations* 2(2).

König, Thomas and Simon Hug, eds. 2006. *Policy-Making Processes and the European Constitution: A Comparative Study of Member States and Accession Countries*. London: Routledge.

Krippendorff, Klaus. 2004. *Content Analysis: An Introduction to Its Methodology (2nd edition)*. Thousand Oaks, CA: Sage Publications.

Kritzinger, Sylvia, Francesco Cavatorta and Raj S. Chari. 2004. "Continuity and change in party positions towards Europe in Italian parties: an examination of parties' manifestos." *Journal of European Public Policy* 11(6):954-974.

Laver, Michael and Ben Hunt. 1992. *Policy and Party Competition*. New York: Routledge.

Laver, Michael, ed. 2001. *Estimating the Policy Position of Political Actors*. London: Routledge.

Laver, Michael and John Garry. 2000. "Estimating Positions from Political Texts." *American Journal of Political Science* 424:619–634.

Laver, Michael, Kenneth Benoit and John Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97(2):311–331.

Lowe, Will. 2008. "Understanding Wordscores." *Political Analysis* 16(4):356–371.

Lowe, Will, Kenneth Benoit, Slava Mikhaylov and Michael Laver. 2009. "Scaling Policy Positions From Hand-Coded Political Texts." Prepared for the 2009 Annual Meeting of the Midwest Political Science Association, Palmer House Hilton and Towers, Chicago, April 2-5.

Manning, Christopher D and Hinrich Schütze. 2002. *Foundations of statistical natural language processing (5th ed.)*. Cambridge: MIT Press.

Marks, Gary, Liesbet Hooghe, Marco R. Steenbergen and Ryan Bakke. 2007. "Crossvalidating Data on Party Positioning on European Integration." *Electoral Studies* 26:23–38.

Martin, Andrew D. and Kevin M. Quinn. 2004. "MCMCpack 0.4-8." (available from http://mcmcpack.wustl.edu).

McDonald, Michael D., Silvia M. Mendes and Myunghee Kim. 2007. "Cross-Temporal and Cross-National Comparisons of Party Left-Right Positions." *Electoral Studies* 26(1):62–75.

McGuire, Kevin T. and Georg Vanberg. 2005. "Mapping the Policies of the U.S. Supreme Court: Data, Opinions, and Constitutional Law." Presented at the annual meeting of the American Political Science Association, Washington, DC, 1-5 September.

Monroe, Burt L. and Maeda Ko. 2004. "Talk's cheap: Text-based ideal point estimation." Paper presented to the Political Methodology Society, Palo Alto, July 29ÛÒ31, 2004.

Monroe, Burt L. and Philip A. Schrodt. 2009. "Introduction to the Special Issue: The Statistical Analysis of Political Text." *Political Analysis* .

Monroe, Burt, Michael Colaresi and Kevin Quinn. 2009 (forthcoming). "Fightin' Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict." *Political Analysis* .

Pelizzo, Riccardo. 2003. "Party Positions or Party Direction? An Analysis of Party Manifesto Data (2003)." *West European Politics* 26:67–89.

Pennings, Paul. 2002. "The Dimensionality of the EU Policy Space. The European Elections of 1999." *European Union Politics* 3(1):59–79.

Pennings, Paul Christine Arnold. 2008 (forthcoming). "'Is constitutional politics like 'normal politics'? The case of the EU Constitution." *Political Studies* .

Pennings, Paul and Han Keman. 2002. "Towards a new methodology of estimating party policy positions." *Quality & Quantity* 36:55ÛÒ79.

Popping, Roel. 2000. *Computer-assisted text analysis*. London: Sage.

Proksch, Sven-Oliver and Jonathan B. Slapin. 2009 (forthcoming). "How to Avoid Pitfalls in Statistical Analysis of Political Texts: The Case of Germany." *German Politics* 18(3).

Purpura, Stephen and D. Hillard. 2006. "Automated Classification of Congressional Legislation." Proceedings of the Seventh International Conference on Digital Government Research. San Diego, CA.

Putnam, Robert D. 1988. "Diplomacy and Domestic Politics: The Logic of Two-Level Games." *International Organization* 42(3):427–460.

Quinn, Kevin M. 2004. "Bayesian Factor Analysis for Mixed Ordinal and Continuous Responses." *Political Analysis* 12(4):338–353.

Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin and Dragomir R. Radev. 2006. "An automated method of topic-coding legislative speech over time with application to the 105th-108th U.S. Senate." Midwest Political Science Association Meeting, 2006.

Ray, Leonard. 1999. "Measuring Party Orientations Towards European Integration: Results From an Expert Survey." *European Journal of Political Research* 36(October 2):283–306.

Ray, Leonard. 2007. "Validity of measured party positions on European integration: Assumptions, approaches, and a comparison of alternative measures." *Electoral Studies* 26(1):11–22.

Roberts, Carl W. 1997. *Text analysis for the social sciences: Methods for drawing inferences from texts and transcripts*. Mahwah, NJ: Lawrence Erlbaum.

Sebastiani, Fabrizio. 2002. "Machine Learning in Automated Text Categorization." *ACM Computing Surveys* 34:1–47.

Shor, Boris, Christopher Berry and Nolan McCarty. 2008. "Methodological Issues in Bridging Legislative Chambers to Create Common Space Ideal Points." University of Chicago and Princeton University.

Simon, A.F. and Michael A. Xenos. 2004. "Dimensional reduction of word-frequency data as a substitute for intersubjective content analysis." *Political Analysis* 12:63–75.

Slapin, Jonathan B. and Sven-Oliver Proksch. 2008. "A Poisson Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52(3):705–722.

Steenbergen, Marco and Gary Marks. 2007. "Evaluating Expert Surveys." *European Journal of Political Research* 46(3):346–366.

Thomassen, Jacques J.A., Abdul G. Noury and Erik Voeten. 2004. Political Competition in the European Parliament: Evidence from Roll Call and Survey Analyses. In *European Integration and Political Conflict*, ed. Gary Marks and Marco R. Steenbergen. Cambridge: Cambridge University Press chapter 7, pp. 141–164.

Volkens, Andrea. 2007. "Strengths and weaknesses of approaches to measuring policy positions of parties." *Electoral Studies* 26(1):108ÛÒ120.

Wessels, Bernhard. 1999. Whom to Represent? Role Orientations of Legislators in Europe. In *Political Representation and Legitimacy in the European Union*, ed. Hermann Schmitt and Jacques Thomassen. Oxford: Oxford University Press pp. 209–234.

Yu, Bei, Stefan Kaufmann and Daniel Diermeier. 2008. "Classifying Party Affiliation from Political Speech." *Journal of Information Technology & Politics* 5(1):33 – 48.