

Helping Citizens to Locate Political Parties in the Policy Space: a Dataset for the 2014 elections to the European Parliament

Social and Behavioural Sciences

Kostas Gemenis

Max Planck Institute for the Study of Societies, Cologne, Germany
gemenis@mpifg.de

Fernando Mendez

University of Zurich, Switzerland
fernando.mendez@uzh.ch

Jonathan Wheatley

Oxford Brookes University, UK
jwheatley@brookes.ac.uk

Abstract

The authors present a dataset that contains the positions of 231 political parties across 28 countries on 30 policy issues that were considered salient for the 2014 elections to the European Parliament. The party position estimates were originally used in a voter information tool which compared the policy preferences of citizens to those of political parties. The paper discusses the estimation method in the context of the literature on estimating party positions, outlines the coding methodology, and introduces the value of the dataset for third-party users interested in studying political participation and representation.

Keywords

European Parliament elections – policy positions – political parties – Voting Advice Applications

Article by an MPIFG researcher

Kostas Gemenis, Fernando Mendez, Jonathan Wheatley: Helping Citizens to Locate Political Parties in the Policy Space: A Dataset for the 2014 Elections to the European Parliament. In: Research Data Journal for the Humanities and Social Sciences 1-14 (2019) (published online March 18). Brill
The original publication is available at the publisher's web site: <https://doi.org/10.1163/24523666-00401002>

© GEMENIS, MENDEZ AND WHEATLEY, 2019 | DOI:10.1163/24523666-00401002

This is an open access article distributed under the terms of the prevailing CC-BY license at the time of publication.

- Related data set “Political parties’ policy positions for the 2014 elections to the European Parliament” with DOI <http://doi.org/10.7802/1594> in repository “GESIS datorium”.

1. Introduction

Voting Advice Applications (VAAs) are contemporary technological solutions that aim to compare the positions of voters to the positions of candidates or political parties on a set of relevant policy issues and informing voters about how well their preferences match those of the parties or candidates.

Government-funded bodies originally developed VAAs as civic education tools with the intention of enhancing political knowledge and increasing electoral participation. The most popular VAA in the Netherlands, *StemWijzer*, was first developed in 1989 as an educational tool for high schools. Since the early 2000s, *StemWijzer* has served as a blueprint for many other VAAs, such as *Wahl-O-Mat* in Germany, as well being the precursor to other types of VAA designs, such as those of the Swiss-based *Smartvote*. The rise of the internet brought with it an ever-increasing number of users while also fuelling the spread of VAAs across Europe. VAAs have now become a more prominent and generalized feature of the European political landscape, often attracting millions of users during electoral contests in larger European countries (Garzia & Marschall, 2014).

In a similar way as for national elections, VAAs have been fielded during elections for the European Parliament. The decreasing levels of voter turnout in such elections have long alarmed those who hoped that the increase in the European Parliament’s powers under the co-decision procedure could dampen the indifference towards or disaffection with the European Union. Research has often indicated that VAA use during national election campaigns may boost turnout (Marschall & Schultze, 2012; Gemenis & Rosema, 2014; Garzia, Trechsel, & De Angelis, 2017; Germann & Gemenis, 2019), so it comes as no surprise that the 2014 elections to the European Parliament saw a proliferation of voter information tools such as *VoteMatch Europe* (<http://www.votematch.eu/>), *EUandI* (<http://euandi.eu/>), *EUvox* (<http://www.euvox2014.eu/>), and *Election2014* (<http://www.votewatch.eu/election/>). Starting with *EU Profiler* in 2009, most VAAs that were implemented at the supranational level (like the aforementioned tools) have allowed citizens to compare their policy preferences to those of political parties not only in their country of residence but in other EU member states as well. The central assumption of such VAAs, however, is that the positions of these parties can be estimated with considerable certainty. If they cannot, VAAs would provide citizens with misleading information.

In an attempt to contribute to the extensive academic debate on this subject (see Mair, 2001; Volkens, 2007; Laver, 2014), this paper describes the procedures and methods used to estimate party policy positions in the context of the *EUvox* VAA and presents the resulting database, which can be used by third-party researchers to investigate questions of broader interest in political science.

2. Research Problem

One of the biggest methodological challenges in the design of VAAs and voter information tools in general is to find an unbiased and reliable method of estimating the positions of political parties on the policy statements that are used to match voters to parties. For several decades, the primary reference source for party policy preferences was based on the laborious hand-coding of party manifestos using methods and codebooks developed in the 1970s (Budge, Klingemann, Volkens, Bara, & Tanenbaum, 2001). Such hand-coding, however, has been shown to be unreliable against conventional content analysis standards (see Mikhaylov, Laver, & Benoit, 2012; Gemenis, 2013a).

The advent of the internet brought with it a plethora of political documents in machine-readable format. Political scientists can now pre-process such documents and use them to estimate “quantities of interest” (Grimmer & Stewart, 2013). Supervised machine-learning algorithms appear to be particularly promising with regard to coding party manifestos. However, the outcomes of supervised machine learning to classify the content of party manifestos only moderately correspond to those by human coding (Bilbao-Jayo & Almeida, 2018; Wiedeman, 2018). Moreover, the applications are concerned with pre-defined policy categories that indicate the *relative emphasis* parties place on different issues instead of their *expressed position* on issues which is of interest in VAAs.

In contrast, the supervised computational scaling methods (Grimmer & Stewart, 2013, pp. 291–292) have been shown to lack validity, especially with regard to estimating positions on specific policy statements employed in VAAs (Bruinsma & Gemenis, forthcoming; Ruedin & Morales, 2017). Despite their promise, automated text analysis methods have not become a standard tool in political analysis due to the lack of rigorous studies demonstrating their widespread validity (Grimmer & Stewart, 2013). Moreover, the automated text analysis algorithms are not directly accessible or easily understandable by users, which makes the use of such automated methods problematic in terms of the standards set by VAA designers (e.g. the Lausanne Declaration on VAAs: Garzia & Marschall, 2014).

For these reasons, the VAAs operated by government-funded bodies rely on direct input from the parties or candidates, while the VAAs that are designed by political scientists rely on estimates given by other political scientists in so-called ‘expert surveys’ (Gemenis & Van Ham, 2014). The latter, however, have been criticized in terms of ideological bias, as experts have their own political preferences (Curini, 2010), and in terms of reliability, as experts often disagree with one another (Lindstädt, Proksch, & Slapin, 2018).

To address the problem of subjectiveness and unreliability in estimating parties’ positions on a wide variety of policy issues across EU member states, *EUvox* used an approach best described as an ‘iterative expert survey’ (Gemenis, 2015). The approach consists of an interactive forecasting technique that relies on the anonymous deliberations of a panel of coders and controlled feedback between multiple rounds of estimation, known as the Delphi method (Dalkey & Helmer, 1963). In many respects, this structured aggregation of responses has clear advantages over taking the mean (Steenbergen & Marks, 2007) or modal (Lindstädt, Proksch, & Slapin, 2018) response, or more sophisticated statistical aggregation approaches (Marquardt & Pemstein, 2018). A considerable body of evidence has shown that the Delphi method gives more accurate estimates compared to alternative methods (Rowe & Wright, 1999), and its application in estimating party policy positions has produced valid and reliable results (Gemenis, 2012; Gemenis & Van Ham, 2014). The premise of the Delphi method is that the iterative process induces coders to converge on an estimate by filtering out biased and/or uninformed estimates. Anonymity plays a crucial role, as it guarantees that the consensus reached among the coders is due to the quality of information and not to the coders’ personalities or political preferences. To achieve this, *EUvox* employed a web-based system that allowed panellists to estimate party positions in an anonymous, distributed, and asynchronous manner (Djouvas & Gemenis, 2019).

3. Methods

For the 2014 elections to the European Parliament, party positions were coded on a set of 30 questions representing policy issues that were considered salient for each country. For each party, the goal was to assign five coders to complete the task over at least two rounds of coding. In a few countries (Austria, Belgium/Flanders, France, Malta, Netherlands, Sweden), time constraints did not allow for a second round, while for a few parties the country teams of the project were unable to recruit five persons per party and the task was therefore carried out by fewer coders. Coding for *EUvox* took place on the *SmartCoding*

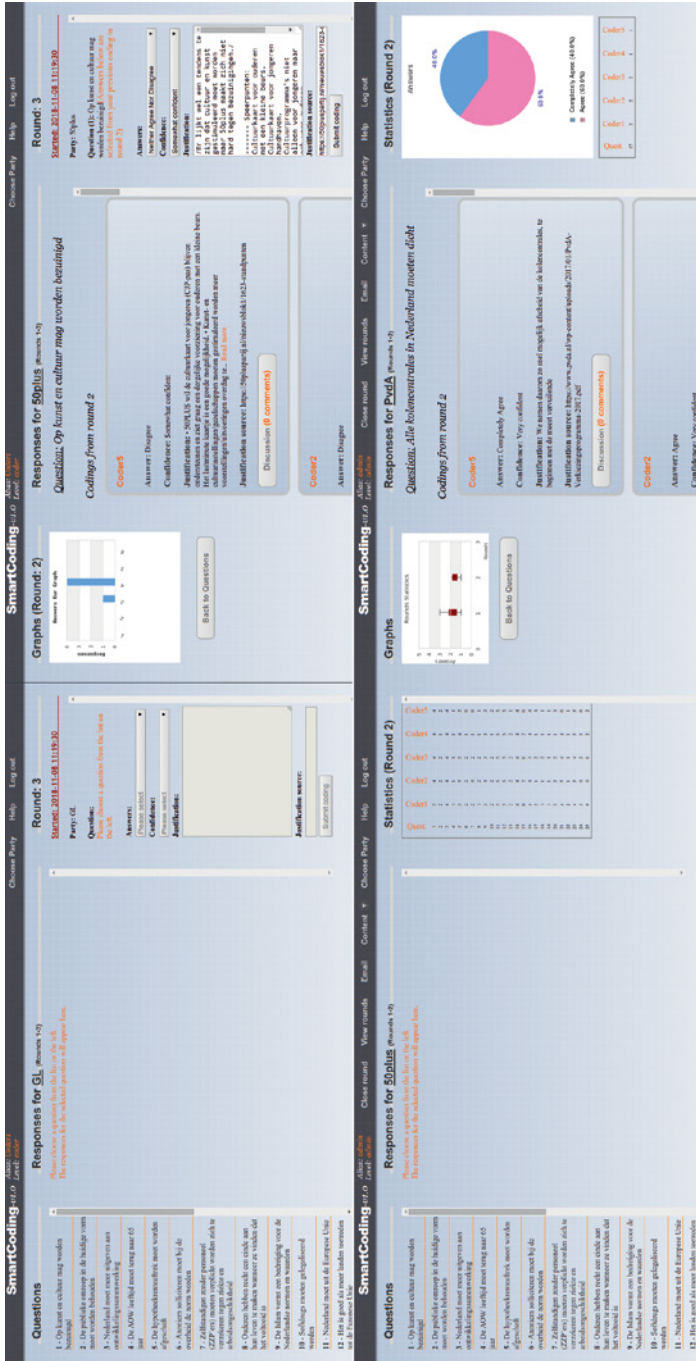


FIGURE 1 Screenshots from the SmartCoding v1.0 platform. The top panels show the coders' view during the coding process and the bottom panels show the moderator's view.

web platform, which was designed for use with the 'iterative expert survey' method as shown in Figure 1.

For each country, a separate instance of *SmartCoding* was initiated and the link to the instance was communicated along with the login details (username and password) to the moderator and coders of that particular country. The first time the coders entered the platform they answered a few personal background/demographic questions regarding their expertise. Once this information had been entered, they were able to see the parties that had been assigned to them for coding. Coders were able to select any of the available parties and code each party with respect to each issue statement. An example of an issue statement is:

Bulgaria should join the Eurozone

Statements such as the above represent an opinion. Coders had to decide whether or not the relevant party agreed with this opinion. They therefore needed to make a judgement using documented sources, their own expertise, or other contextual information. They had a five-point response scale with a 'no opinion' option for their judgement.

1. Completely agree;
2. Agree;
3. Neither agree nor disagree;
4. Disagree;
5. Completely disagree; or
6. No opinion.

Coders were asked to use the (6) 'No opinion' position for parties when the party's position on the issue was not mentioned in, and could not be inferred from, any relevant sources. Coders were instructed that coding results should contain as few as possible 'no opinion' positions.

Positions (2) 'Agree' and (4) 'Disagree' were used when a party took sides on the issue. For example, in the above statement, position (2) should be used if the party agreed that Bulgaria should join the Eurozone, while (4) should be used if the party disagreed and believed that Bulgaria should not join the Eurozone.

Positions (1) 'Completely agree' and (5) 'Completely disagree' were used when a party strongly emphasized the issue and did not place any restrictions on or have any doubts about the validity of the issue. In the example given, (1) should therefore be used when the party appeared to express a strong and unambiguous opinion that Bulgaria should join the Eurozone, while (5) should

be used when that party expressed an equally strong and unambiguous position that Bulgaria should not join the Eurozone.

Coders were also instructed that position (3) 'Neither agree nor disagree' was an argued position, unlike (6) 'No opinion'. Position (3) was to be used either i) when the party only addressed an issue indirectly or vaguely and did not express an opinion clearly in favour or against; or ii) when the party clearly emphasized both sides of the issue (i.e. positive and negative aspects) and made a point of not taking sides; or iii) when a party was divided on an issue (i.e. equally prominent figures within the party took different sides of the argument).

Coders were instructed to reach their decisions on how to code each party on each statement anonymously and not to collaborate. In addition, they were asked to do all they could to be objective and not let their prejudices about certain parties get the better of them. For example, a coder who was sympathetic to pro-Euro party X should not code party Y 'Disagree' on whether Bulgaria should join the Eurozone based on a prejudice that party Y were pro-Russia communists who obviously opposed the Euro (even if they say they support it). If party Y's leaders said that they wanted Bulgaria to join the Eurozone, the party must be coded accordingly.

In addition, coders were instructed to indicate how they arrived at their judgement in each case by entering some information in a box entitled 'Justification'. Coders were told that they could make use of all possible sources of information on party positions (including manifestos, websites, news reports, and press releases). *EUvox* did not employ a set 'hierarchy of documents' (cf. Garzia, Trechsel, & De Sio, 2017), which meant that the coders were free to decide where to source the most relevant information. When these sources proved to be insufficient or otherwise unavailable, coders were asked to make a personal judgement based on their expertise or other contextual information. They were told that the information provided in the 'Justification' box should help convince other coders during the next round of coding that the judgement provided in that case was correct. Moreover, coders were instructed to provide the source of this evidence (e.g. a specific URL, page number of party manifesto, etc.) in the box entitled 'Source'.

Coders were also asked to indicate how confident they felt about their coding of a party for each statement. The options on a drop-down menu on the *SmartCoding* web platform were:

1. Very confident;
2. Somewhat confident;
3. Not confident at all.

Once coding for the first round was completed, a moderator selected for each country inspected the file with the aggregated results produced by the *SmartCoding* platform to check whether sufficient consensus had been reached among the panel. For each party/statement combination, *Smart Coding* had been programmed to export the median estimate as well as Van der Eijk's A coefficient of agreement (Van der Eijk, 2001). Coefficient A ranges from 1 to -1, where 1 indicates perfect agreement among the coders and -1 indicates polarization, with half of the coders placing the party at (1) 'Completely agree' and the other half at (5) 'Completely disagree'. Moderators were asked to check for instances where A was under 0.7 and to review the individual responses in these party/statement pair. It appeared to be often the case that coders mistake the direction of the response scale. For example, two coders might give exactly the same source and justification, but one will code the party as (1) 'Completely agree' and the other as (5) 'Completely disagree'. This probably means that one of the coders has mistaken the direction of the 1–5 response scale that was given for the coding. Moderators were instructed that, in such cases, it was advisable to enter a comment under the justification given by the coder in question to point out this inconsistency. It was at the moderator's discretion to enter as many such comments on estimates/justifications given by coders as necessary. The moderators could therefore point out possible mistakes, inconsistencies, biases, and misunderstandings of the meaning of the question or the response scale.

The moderator then initiated a second round of coding, in which coders were able to see the estimates, justifications, and level of confidence given by each of the other anonymous coders, as well as the comments made by the moderator. In the second (and any subsequent) round, coders were asked to review their initial (first round) estimates in the light of the information (estimates, confidence, justifications, and sources) given by all other coders and of the (possible) comments given by the moderator. Also, in the second (and any subsequent) round, *SmartCoding* allowed coders to comment on estimates given in previous rounds, and to respond to comments left by the moderator. *SmartCoding* allows this process of moderation and iteration to be repeated over several rounds until sufficient consensus is reached. In *EUvox*, two and (rarely) three rounds were usually sufficient, although, as previously noted, in a handful of countries time constraints did not allow for coding of parties over more than one round. When coding for the parties in a particular country ended, the moderator was asked to arbitrate over any unresolved cases where agreement in specific party/statement pairs was below 0.7 in terms of Van der Eijk's (2001) perceptual agreement coefficient, as well as in cases where the majority of coders gave a (3) 'Neither agree nor disagree' or

a (6) 'No opinion' position. To resolve these cases, moderators were instructed to take into account the justifications and levels of confidence given by the coders.

4. Data

- *EUvox* party positions dataset deposited at GESIS – DOI:<http://doi.org/10.7802/1594>
- **Temporal coverage:** 2014

The *EUvox* party positions dataset, available from the German social science data archive, GESIS (DOI:10.7802/1594), contains two main files: the party positions file, and the questionnaire file. In the party positions file, each row represents a political party, with the columns giving the name of the party in English and the original language, the number of rounds and coders that were employed in coding the positions of this party, and the positions of the party on each of the 30 policy statements. Out of these 30 statements, 21 were common across all countries (see Appendix). Among the common questions, there were seven for each of three broad dimensions: economic 'left-right', cultural 'liberal-conservative', based closely on the TAN/GAL scale proposed by the Chapel Hill Expert Survey team (Hooghe, Marks, & Wilson, 2002), and a pro/anti-EU dimension. A notable exception to the common questions was the wording of the statement about the Euro, which differed between Eurozone and non-Eurozone countries (see Appendix, Statement 1). Among the nine questions that differed among countries, some were common across several, but not all, countries. An example of such a statement was '*Protecting the environment is more important than fostering economic growth*', which was given in Finland, Ireland, Germany, Italy, Spain, Poland, Hungary, Estonia, and Bulgaria, but not in other countries. Other questions were unique to a single country, such as '*To combat unemployment, workers must accept signing individual contracts*', which was given only in Greece. The wording of all policy statements that were coded for each party can be found in the questionnaire file.

In total, the dataset contains the positions of 231 parties across 28 countries and on 30 policy issues, making it comparable to other major datasets in the field (e.g. the 2014 edition of the long-established Chapel Hill Expert Survey (<http://www.chesdata.eu>)). Moreover, the use of multiple coders and, in most cases, multiple rounds of coding offers increased confidence in the reliability of the estimates compared to other datasets with a comparable breadth of coverage (cf. Garzia, Trechsel & De Sio, 2017).

5. Concluding Remarks

Third-party users can employ the party position data offered in this dataset as is or scale the individual items into general dimensions of interest using data reduction techniques such as principal component analysis, multidimensional scaling, Mokken scale analysis, or Bayesian item response models (Louwerse & Otjes, 2012; Gemenis, 2013b; Germann, Mendez, Wheatley, & Serdült, 2015; Germann & Mendez, 2016; Linhart & Shikano, 2015; Dalton, 2016; Nyhuis & König, 2018) to study a wide variety of substantive questions in political science.

For instance, VAA data on party positions has been used to assess government coalitions (Linhart & Shikano, 2015), polarization across policy issues (Wall & Williams, 2018), the dimensionality of political space (e.g. Katsanidou & Otjes, 2016; Garry, Matthews, & Wheatley, 2017; Nyhuis & König, 2018), continuity and change in parties' ideological profiles (e.g. Dalton, 2016), and political representation (e.g. Rose & Borz, 2013), especially in combination with citizen data from election studies (e.g. Lefkofridi & Katsanidou, 2014). In addition, since VAAs employ the same questionnaire for parties and citizens, VAA datasets on party positions can be combined with datasets of VAA users (e.g. Wheatley, 2015; Wheatley & Mendez, forthcoming) in order to measure electoral utilities that can be used to test spatial models of voting (Mendez, 2017). Either way, VAA data on party positions offer unique opportunities to study political participation and representation beyond the VAA designers' original intention of helping citizens make informed choices during elections.

References

- Bilbao-Jayo, A., & Almeida, A. (2018). Automatic political discourse analysis with multi-scale convolutional neural networks and contextual data. *International Journal of Distributed Sensor Networks*, 14(11), doi:<http://doi.org/10.1177/155014771881827>.
- Bruinsma, B., & Gemenis, K. (forthcoming). Validating Wordscores: The promises and pitfalls of computational text scaling. *Communication Methods and Measures*.
- Budge, I., Klingemann, H. D., Volkens, A., Bara, J., & Tanenbaum, E. (Eds.). (2001). *Mapping Policy Preferences: Estimates for Parties, Governments and Electors, 1945–1998*. Oxford: Oxford University Press.
- Curini, L. (2010). Experts' political preferences and their impact on ideological bias: An unfolding analysis based on a Benoit-Laver expert survey. *Party Politics*, 16(3), 299–321.
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458–467.
- Dalton, R. J. (2016). Stability and change in party issue positions: The 2009 and 2014 European elections. *Electoral Studies*, 44, 525–534.

- Djouvas, C. & Gemenis, K. (2019). *SmartCoding: An online platform for estimating political parties' policy positions*.
- Garry, J., Matthews, N., & Wheatley, J. (2017). Dimensionality of policy space in consociational Northern Ireland. *Political Studies*, 65(2), 493–511.
- Garzia, D. & Marschall, S. (Eds.). (2014). *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective*. Colchester: ECPR Press.
- Garzia, D., Trechsel, A. H., & De Angelis, A. (2017). Voting advice applications and electoral participation: A multi-method study. *Political Communication*, 34(3), 424–443.
- Garzia, D., Trechsel, A., & De Sio, L. (2017). Party placement in supranational elections: An introduction to the euandi 2014 dataset. *Party Politics*, 23(4), 333–341.
- Hooghe, L., Marks, G., & Wilson, C. J. (2002). Does left/right structure party positions on European integration? *Comparative Political Studies*, 35(8), 965–989.
- Gemenis, K. (2012, November). *A new approach for estimating parties' positions in voting advice applications*. Paper presented at the Interdisciplinary Perspectives on Voting Advice Applications Workshop, Limassol, Cyprus.
- Gemenis, K. (2013a). What to do (and not to do) with the comparative manifestos project data. *Political Studies*, 61(Suppl), 3–23.
- Gemenis, K. (2013b). Estimating parties' policy positions through voting advice applications: Some methodological considerations. *Acta Politica*, 48(3), 268–295.
- Gemenis, K. (2015). An iterative expert survey approach for estimating parties' policy positions. *Quality & Quantity*, 49(6), 2291–2306.
- Gemenis, K., & Rosema, M. (2014). Voting advice applications and electoral turnout. *Electoral Studies*, 36, 281–289.
- Gemenis, K., & Van Ham, C. T. (2014). Comparing methods for estimating parties' positions in Voting Advice Applications. In Garzia, D. & Marschall, S. (Eds.), *Matching Voters with Parties and Candidates: Voting Advice Applications in Comparative Perspective* (pp. 33–47). Colchester: ECPR Press.
- Germann, M., & Gemenis, K. (2019). Getting out the vote with Voting Advice Applications. *Political Communication*, 36(1), 149–170. doi:<http://doi.org/10.1080/10584609.2018.1526237>.
- Germann, M., & Mendez, F. (2016). Dynamic scale validation reloaded. *Quality & Quantity*, 50(3), 981–1007.
- Germann, M., Mendez, F., Wheatley, J., & Serdült, U. (2015). Spatial maps in voting advice applications: The case for dynamic scale validation. *Acta Politica*, 50(2), 214–238.
- Grimmer, J., & Stewart, B. M. (2013). Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3), 267–297.
- Katsanidou, A., & Otjes, S. (2016). How the European debt crisis reshaped national political space: The case of Greece. *European Union Politics*, 17(2), 262–284.
- Laver, M. (2014). Measuring policy positions in political space. *Annual Review of Political Science*, 17, 207–223.

- Lefkofridi, Z., & Katsanidou, A. (2014). Multilevel representation in the European Parliament. *European Union Politics*, 15(1), 108–131.
- Lindstädt, R., Proksch, S. O., & Slapin, J. (2018). When experts disagree: Response aggregation and its consequences in expert surveys. *Political Science Research and Methods*, doi:<http://doi.org/10.1017/psrm.2018.52>.
- Linhart, E., & Shikano, S. (2015). Koalitionsbildung nach der Bundestagswahl 2013: Parteien im Spannungsfeld zwischen Ämter-, Politik- und Stimmenmotivation. In Korte, K.-R. (Ed.), *Die Bundestagswahl 2013* (pp. 457–484). Wiesbaden: Springer.
- Louwerse, T., & Otjes, S. (2012). Design challenges in cross-national VAAs: the case of the EU Profiler. *International Journal of Electronic Governance*, 5(3–4), 279–297.
- 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* (pp. 10–30). London: Routledge.
- Marquardt, K. L., & Pemstein, D. (2018). IRT models for expert-coded panel data. *Political Analysis*, 26(4), 431–456.
- Marschall, S., & Schultze, M. (2012). Voting Advice Applications and their effect on voter turnout: the case of the German Wahl-O-Mat. *International Journal of Electronic Governance*, 5(3–4), 349–366.
- Mendez, F. (2017). Modeling proximity and directional decisional logic: What can we learn from applying statistical learning techniques to VAA-generated data? *Journal of Elections, Public Opinion and Parties*, 27(1), 31–55.
- Mikhaylov, S., Laver, M., & Benoit, K. R. (2012). Coder reliability and misclassification in the human coding of party manifestos. *Political Analysis*, 20(1), 78–91.
- Nyhuis, D., & König, P. (2018). Estimating the conflict dimensionality in the German Länder from Vote Advice Applications, 2014–2017. *Statistics, Politics and Policy*, doi:<http://doi.org/10.1515/spp-2018-0007>.
- Rose, R., & Borz, G. (2013). Aggregation and representation in European Parliament party groups. *West European Politics*, 36(3), 474–497.
- Rowe, G., & Wright, G. (1999). The Delphi technique as a forecasting tool: issues and analysis. *International Journal of Forecasting*, 15(4), 353–375.
- Ruedin, D., & Morales, L. (2017). Estimating party positions on immigration: Assessing the reliability and validity of different methods. *Party Politics*, doi:<http://doi.org/10.1177/1354068817713122>.
- Steenbergen, M. R., & Marks, G. (2007). Evaluating expert judgments. *European Journal of Political Research*, 46(3), 347–366.
- Van der Eijk, C. (2001). Measuring agreement in ordered rating scales. *Quality & Quantity*, 35(3), 325–341.
- Volkens, A. (2007). Strengths and weaknesses of approaches to measuring policy positions of parties. *Electoral Studies*, 26(1), 108–120.

- Wall, M., & Williams, S. (2018). Seeking evidence for a Welsh progressive consensus: Party positioning in the 2016 National Assembly for Wales election. *Parliamentary Affairs*, 71(4), 820–844.
- Wheatley, J. (2015). Identifying latent policy dimensions from public opinion data: An inductive approach. *Journal of Elections, Public Opinion & Parties*, 25(2), 215–233.
- Wheatley, J., & Mendez, F. (forthcoming). Reconceptualising dimensions of political competition in Europe: A demand side approach. *British Journal of Political Science*.
- Wiedemann, G. (2018). Proportional classification revisited: Automatic content analysis of political manifestos using active learning. *Social Science Computer Review*, doi:<http://doi.org/10.1177/0894439318758389>.

Appendix

TABLE A1 *Statements common to all countries in the EUvox dataset*

#	Dimension	Statement
1	EU	[Name of country] should exit the Euro (Eurozone countries) / [Name of country] should never adopt the Euro (non-Eurozone countries).
2	EU	A single member state should be able to block a treaty change, even if all the other member states agree to it.
3	EU	The right of EU citizens to work in [name of country] should be restricted.
4	EU	There should be a common EU foreign policy even if this limits the capacity of [name of country] to act independently.
5	EU	The EU should redistribute resources from richer to poorer EU regions.
6	EU	Overall, EU membership has been a bad thing for the [name of country].
7	EU	EU treaties should be decided by [name of national parliament] rather than by citizens in a referendum.
8	Economic	Free market competition makes the health care system function better.*
9	Economic	The number of public sector employees should be reduced.
10	Economic	The state should intervene as little as possible in the economy.
11	Economic	Wealth should be redistributed from the richest people to the poorest.

TABLE A1 *Statements common to all countries in the EUvox dataset (cont.)*

#	Dimension	Statement
12	Economic	Cutting government spending is a good way to solve the economic crisis.
13	Economic	It should be easy for companies to fire people
14	Economic	External loans from institutions such as the IMF are a good solution to crisis situations.
15	Cultural	Immigrants must adapt to the values and culture of [name of country].
16	Cultural	Restrictions on citizen privacy are acceptable in order to combat crime.
17	Cultural	To maintain public order, governments should be able to restrict demonstrations.
18	Cultural	Less serious crimes should be punished with community service, not imprisonment.
19	Cultural	Same sex couples should enjoy the same rights as heterosexual couples to marry.
20	Cultural	Women should be free to decide on matters of abortion.
21	Cultural	The recreational use of cannabis should be legal.

* In France, the question wording was 'Face au déficit de la sécurité sociale, les malades devraient les malades devraient payer une plus grande part de leurs dépenses médicaux.'