



## **Scaling hand-coded political texts to learn more about left-right policy content**

**LSE Research Online URL for this paper:** <http://eprints.lse.ac.uk/114447/>

Version: Published Version

---

### **Article:**

Däubler, Thomas and Benoit, Kenneth ORCID: 0000-0002-0797-564X (2022)  
Scaling hand-coded political texts to learn more about left-right policy content.  
Party Politics, 28 (5). 834 - 844. ISSN 1354-0688

<https://doi.org/10.1177/13540688211026076>

---

### **Reuse**

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial (CC BY-NC) licence. This licence allows you to remix, tweak, and build upon this work non-commercially, and any new works must also acknowledge the authors and be non-commercial. You don't have to license any derivative works on the same terms. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

# Scaling hand-coded political texts to learn more about left-right policy content

Party Politics

1–11

© The Author(s) 2021



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/13540688211026076

[journals.sagepub.com/home/ppq](https://journals.sagepub.com/home/ppq)**Thomas Däubler**

University College Dublin, Ireland

**Kenneth Benoit**

London School of Economics, UK

## Abstract

Manual annotation of the policy content of political texts forms the basis for one of the most widely used empirical measures in comparative politics: left-right policy positions. Bridging automated “text as data” approaches and qualitative content analysis, we apply statistical scaling to this data to learn more about the association of specific policy dimensions to the left-right super-dimension, in a way that minimizes *ex ante* assumptions about the substantive content of left-right policy. We apply a Bayesian negative binomial variant of Slapin and Proksch’s (2008) “wordfish” model to category counts from party manifestos coded by the Manifesto Project, providing a data-driven approach that offers new insights into the policy content of left and right. We demonstrate how this method also works with content not originally designed for measuring positions. In addition, we show how the approach can be extended to measure the policy content of two latent dimensions, with some categories contributing to both.

## Keywords

Comparative Agendas Project, IRT, Manifesto Project, measurement, party manifestos, policy positions

Configurations of party positioning can be complex if the policy space includes a multitude of specific dimensions or issues (e.g., Bakker et al., 2015; Benoit and Laver, 2006). Both politicians and scholars therefore often use low-dimensional representations (e.g., Gabel and Huber, 2000; Laver and Budge, 1992) to characterize the policy preferences of political actors. The most basic of such models refers to “left” and “right,” ordering party and voter positions on a single dimension (Downs, 1957). Two-dimensional representations provide a slightly more complex but richer model of the policy space, for example by distinguishing an economic and a “socio-cultural” axis of political contestation (e.g., Bornschier, 2010; Hooghe et al., 2002; Kitschelt, 1994).

Approaches to measuring positions on the basic dimensions of the policy space vary in the extent to which they rely on prior assumptions about the meaning of left and right (e.g., Bakker, 2009; Budge and Meyer, 2013; Franzmann and Kaiser, 2006; Jahn, 2011). Some authors start from theorizing the substantive content of the policy

dimension(s). As an example, consider the very widely used fixed “Rile” (right-left) index of the Manifesto Project (hereafter MP; Budge et al., 1987, 2001; Klingemann et al., 2006; Volkens et al., 2013). In promoting this measure, Budge and Meyer (2013: 89) emphasize its connections to opposing arguments about the appropriate level of market regulation and other matters, as reflected in writings by early modern political philosophers. At the other end of the spectrum, advocates of data-driven approaches recommend inferring the left-right dimension in a purely inductive manner. Gabel and Huber (2000: 96) suggest that left-right is simply “the ‘super-issue’ that most constrains parties’ positions across a broad range of policies.” This

---

Paper submitted 2 September 2020; accepted for publication 27 May 2021

## Corresponding author:

Thomas Däubler, University College Dublin, School of Politics and International Relations, Belfield, Dublin 4, Ireland.

Email: [thomas.daubler@ucd.ie](mailto:thomas.daubler@ucd.ie)

inductive route is also taken by scaling approaches that treat natural language text as the data to be analyzed (e.g., Slapin and Proksch, 2008). As discussed in detail by Lowe (2016), most existing methods toward measuring party positions, whether from content-analytic data or word frequencies, are variations of a common model of relative emphasis.

The debate about the appropriate way of measuring left and right from hand-coded party manifestos has focused on questions of reliability and validity (see e.g. Flentje et al., 2017; Franzmann, 2015; Gabel and Huber, 2000; Gemenis, 2013; Mikhaylov et al., 2012). In this research note, we point to a largely overlooked advantage provided by applying inductive approaches to content analysis data: using minimal *a priori* assumptions, they can help us learn about the actual content of the main dimension(s) of the policy space in an empirical manner. While Gabel and Huber (2000) and Albright (2008) also apply inductive methods to the full set of MP categories, they neither report nor discuss results for the item parameters.<sup>1</sup> For this purpose, we develop tailor-made Bayesian variants of the “wordfish” model (Slapin and Proksch, 2008), and apply these to data covering the full set of policy categories.

We begin by establishing that the derived left-right measure is valid, by comparing it to independent expert survey estimates. Here, it does at least as well (and in most countries better) than the MP’s Rile index. Next, contrasting each policy category’s *inferred* contribution to left-right to the associations that the Rile index *assumes*, we underscore the insight gained by applying a different tool to the same content-analytic data input. To demonstrate that scaling approaches can be applied to other annotation or coding schemes, we present results for manifesto-based data from the Comparative Policy Agendas Project (CAP) (Baumgartner et al., 2008), which provides more fine-grained data on issue coverage than the MP. Although the documents were not originally coded with the purpose of measuring policy *positions* at all, we can use them to measure well both where Belgian and Danish parties are located in the policy space and which issues separate them. Finally, we extend the measurement approach to two dimensions, inferring an economic and a socio-cultural axis of political competition. The model allows the researcher to select categories that can be linked to both latent variables while ensuring that the two dimensions remain easy to interpret.

## Scaling the “super-dimension” of left-right policy

### *Data: Category counts from manually coded party policy statements*

Party manifestos have long formed the main source of textual data for estimating left-right ideology, both inductively and deductively (e.g. Franzmann and Kaiser, 2006; Gabel and Huber, 2000; Gemenis, 2013; Jahn, 2011; Laver et al.,

2003; Mölder, 2013; Slapin and Proksch, 2008). Another large-scale qualitative content-analytic scheme in political science is the Comparative Policy Agendas Project (CAP, see Baumgartner et al., 2008). The CAP aims to identify the topic focus of policy documents, media coverage and political events such as cabinet meetings (Baumgartner et al., 2008). While the majority of the documents coded by the CAP do not represent party policy statements, their dataset does include some manifestos.

We draw on qualitative coding from both the MP and the CAP, to show how unsupervised scaling models originally developed in the quantitative, computational tradition can be applied to two different types of manually coded content-analytic data, and thus bridge automated text analysis methods with qualitative content analysis.

### *A measurement model for unordered categorical outcomes*

We use a measurement model that allows simultaneous estimation of the left-right positions of political parties, as well as the contributions of each more specific aspect of policy—such as expansion versus limitation of the welfare state, environmental protection versus an orientation on growth, or support versus opposition to traditional morality—to this dimension. For learning about the nature of the dimension of left-right policy, the category contributions are our key focus. The scaling model is adapted from “text as data” approaches that scale word frequencies, but replaces the usual “bag of words” approach with one that treats *hand-coded sentences* as the units to be scaled.

We observe a set of category codings for the manifesto text. Following Benoit et al. (2009), we start from the notion that the party intends to communicate a certain position, called  $\theta_i$  in the manifesto  $i$ . This position is fundamentally unobservable and uncertain, but will be communicated through the text. As writing proceeds, the party makes various policy statements referring to different issues, generating observable data in the form of counts  $Y_{ij}$  of statements from document  $i$  in different policy categories  $j$ . As Lowe (2016) points out, the fundamental assumption that observed relative frequencies of text units (in our case statements in a certain category) in a document decrease with the quadratic distance between the author and that unit in the latent policy space forms the common logic of a large set of scaling approaches. If we expand this squared difference of document and category location, we obtain the product of a document and a category parameter as the main component of interest in a reduced form equation (Elff, 2013; Lowe, 2016):

$$\log(\mu_{ij}) = \alpha_i + \psi_j + \beta_j \theta_i, \quad (1)$$

This log-linear model for the expected value  $\mu_{ij}$  of the counts  $Y_{ij}$  is the “wordfish” model introduced by Slapin and Proksch (2008), which was originally motivated as a fixed

effects model. The  $\alpha_i$  captures variation in document length,  $\psi_j$  takes into account that some policy categories are generally more frequent than others, and the item parameter  $\beta_j$  quantifies the direction and the extent to which the usage of category  $j$  depends on the position  $\theta_i$ .<sup>2</sup>

Eq. (1) posits a rather restrictive data-generating process for the statement counts. Variation is due to positioning as reflected in the product term, the overall “verbosity” of the document and the baseline frequency of the category. In practice, other factors also play a role, such as “agenda” effects making some categories more prominent in some election campaigns or coder effects arising from varying interpretations of the category scheme. We can think of these as random effects disturbing the observed counts of statements within documents, leading to additional variation. To take this into account, we relax the restrictive equidispersion assumption (that the variance of the counts equals their mean) of the Poisson distribution. Since this research note is mainly interested in the categories’ links to the latent dimension, we also model the extra variation as differing across categories.<sup>3</sup>

We model this using the negative binomial distribution, with expected mean  $\mu_{ij}$ , and an additional variance parameter  $\phi_j$  for each category. The probability mass function as defined by Cameron and Trivedi (1986: 32–33) is

$$\begin{aligned} \Pr(Y_{ij} = y_{ij} | \mu_{ij}, \phi_j) \\ = \frac{\Gamma(y_{ij} + \phi_j)}{\Gamma(y_{ij} + 1)\Gamma(\phi_j)} \left( \frac{\phi_j}{\mu_{ij} + \phi_j} \right)^{\phi_j} \left( \frac{\mu_{ij}}{\mu_{ij} + \phi_j} \right)^{y_{ij}} \end{aligned}$$

The expected value of the counts is given by  $E(Y_{ij}) = \mu_{ij}$ , and the variance by  $\text{Var}(Y_{ij}) = \mu_{ij}(1 + \mu_{ij}\phi^{-1})$  (the “NB2” variance function). The parameter  $\phi_j^{-1}$  represents the extra variance in the data relative to the special case of the Poisson ( $\phi_j^{-1} = 0$ , which is equivalent to the basic “wordfish” model).

We opt for a Bayesian approach to inference, for several pragmatic reasons: the coding is easy to implement, the uncertainty in all parameters is quantified as an inherent feature of the approach, and the procedure is easily amenable to incorporating prior information. In supplementary materials, we describe in more detail how to constrain and simulate the parameters of the model and how to interpret the category parameters in the context of multiple response categories. Replication code and R functions for general use can be found on <https://github.com/kbenoit/bnbwordfish>.

## Estimating left-right as a latent variable

In this section we fit the scaling model to manual content analysis data for party manifestos from the MP, to demonstrate how individually coded sentences can be used as units to infer the one-dimensional latent variable  $\theta_i$ . Our benchmark compares the inferred positions to independent measures from expert surveys.

## Validation: Comparison to expert surveys

Before we draw any inferences about the nature of the policy space, we need to establish that the inductive scaling approach successfully recovers general left-right policy positions. We do this by validating our inferred positions of  $\theta_i$  with expert survey estimates, and then compare the fit to that of Rile, the MP’s left-right index computed directly from a fixed basket of left and right policy categories.

The MP coding scheme comprises 56 core policy categories, which are used by human coders to annotate sentence units according to which policy they pertain.<sup>4</sup> Individual policy codes are combined into the left-right policy index known as “Rile,” by far the most widely used empirical measure of left-right party positioning in political science. Rile is based on 13 categories on each side of the spectrum, which are designated as having content intrinsically associated with either right or left ideology; the remaining 30 policy categories are treated as unrelated to left-right. Exemplifying the *a priori* approach, Budge and Meyer (2013: 89) argue that the construction of this fixed left-right scale reflects opposing arguments by highly influential early modern theorists, including Marx and Engels on the left and Disraeli and Spencer on the right.

Fitting the main model (Eq. 1) to the 56 MP policy counts (with data pooled across countries), we infer the policy positions  $\theta_i$  for each party  $i$  on a single dimension of policy. We work with two separate sets of manifestos, those from Western democracies in the entire post-1972 period and a post-1990 set of documents also including post-communist countries.<sup>5</sup> We find a very good correspondence with external measures of policy in the form of both face validity and correlations with completely independent expert survey estimates of left-right positions.<sup>6</sup> Across countries, we find a correlation of  $r = .81$  for manifestos from Western democracies in the post-1972 period ( $N = 546$ ), and a similar value  $r = .76$  for the post-1990 set of documents from a broader range of countries (total  $N = 759$ ). In this regard, the inferred latent dimension also outperforms the Rile measure, which reaches values of  $r = .67$  and  $r = .64$  only. Our measure using inductive scaling of *all* policy categories thus corresponds well to what experts judged to be the left-right dimension, without having to use any *ex ante* expert judgment as to which policy categories should be associated with this dimension.

Not only does our measure correlate highly with independent positional estimates from expert surveys, but also it corresponds to these more closely than does the Rile index. Additional analyses by country, based on the post-1990 set of manifestos from countries with at least four matched parties, are shown in Figure 1. When the Rile score better matched variation in the expert scores, we plotted it as a green triangle, otherwise, we plotted it as a red square. Except for four countries of Eastern Europe (Hungary, Croatia, Bulgaria and Romania), where the fit

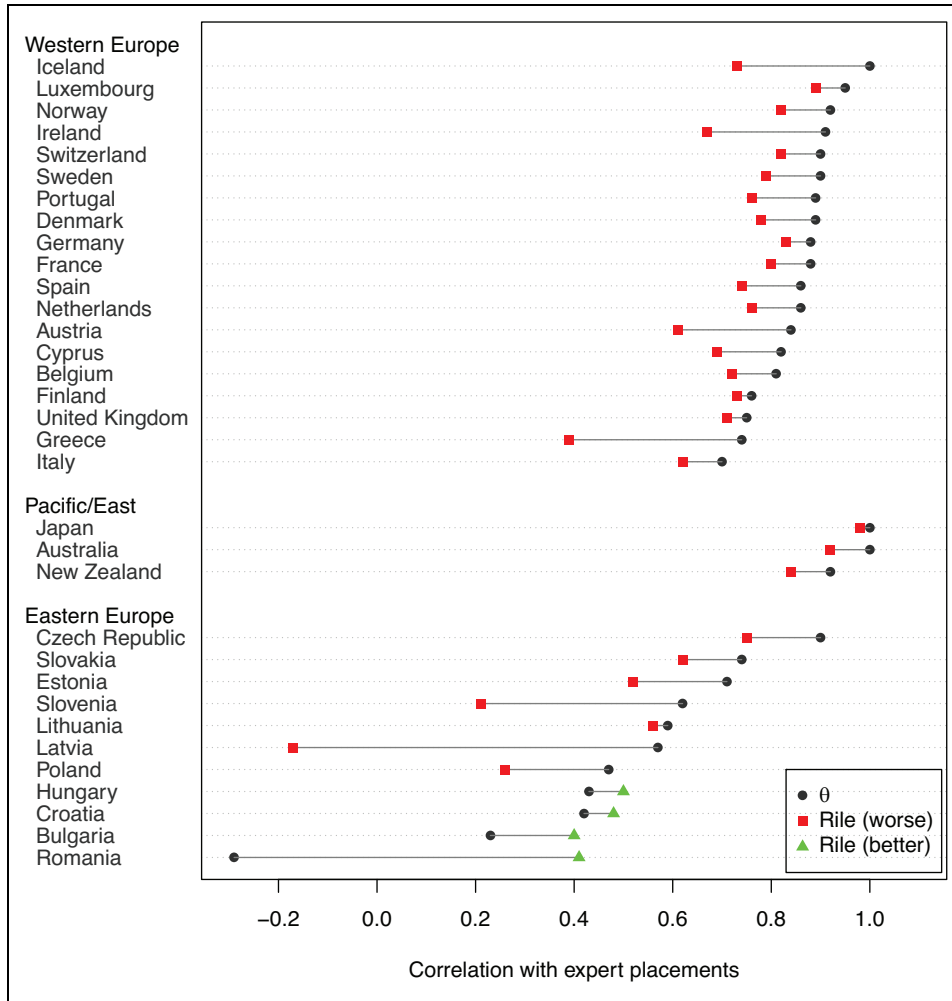


Figure 1. Comparison of scaled results versus Rile correlations with expert survey estimates of left-right policy, by country.

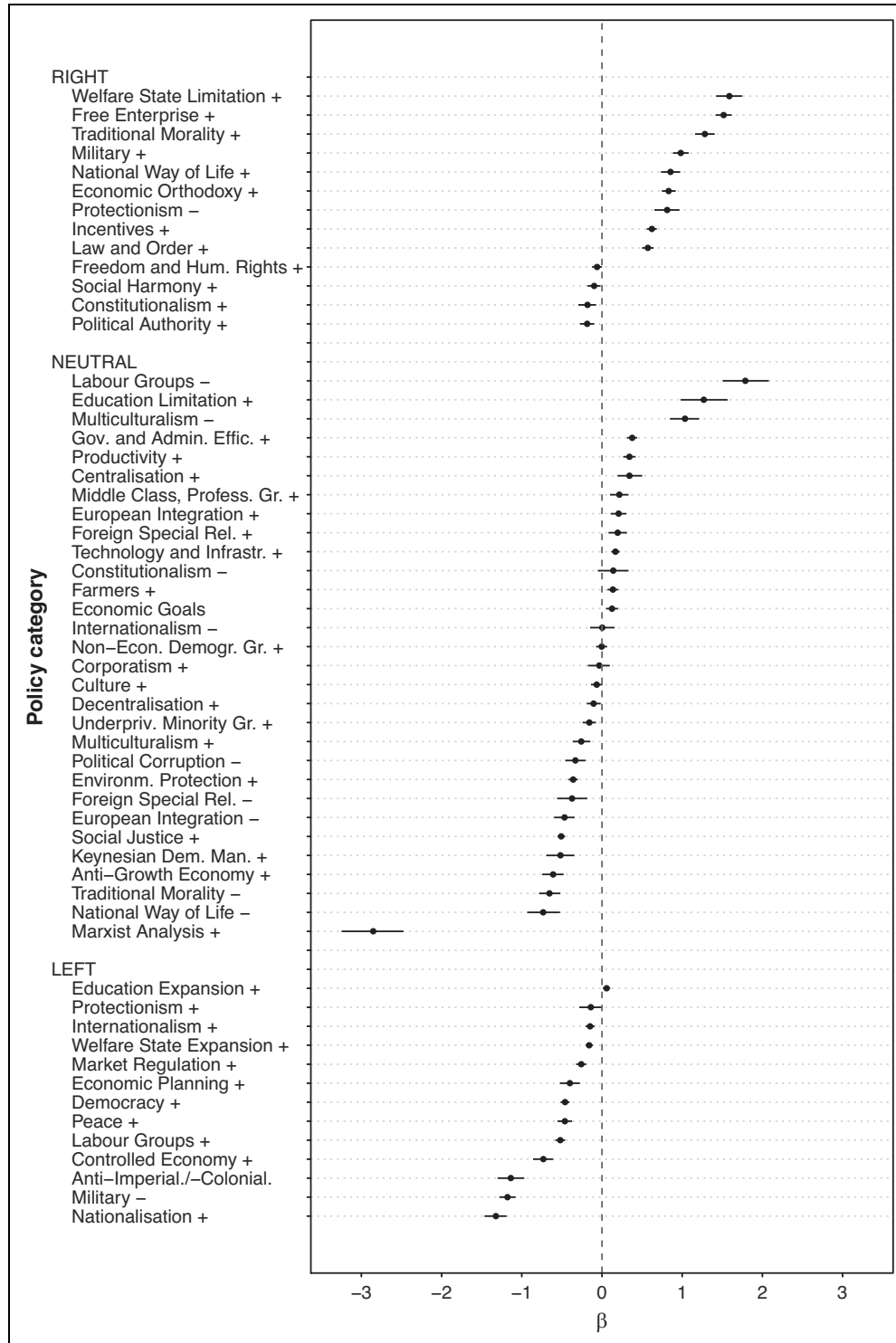
was among the poorest of all countries for both scales, our inferred position provided better correspondence with expert survey results than the Rile measure. For several countries, including Ireland, Greece and Latvia, the improvement in the correlation is above 0.2.

**The policy content of the left-right dimension**

We now use the model to discover which policy emphases contribute to an inductively scaled left-right policy dimension and how. We can also test the respective assumptions made by existing indices of left-right policy, such as the MP’s Rile index. In fitting the manifesto data to the single dimension that appeared meaningfully to differentiate parties, Laver and Budge (1992: 25) emphasize that this process was “based solely on the intrinsic plausibility and coherence of the sets of issues that define the underlying policy dimension.” Presumably, this is why categories such as “Political authority: Positive (305)” are considered

“right-wing”: because in the sample examined, this was the pattern of their association. Using our method that automatically includes and infers the weighted contribution of each policy input to the left-right dimension, we can compare our scaling results to those of the original MP’s categorization.

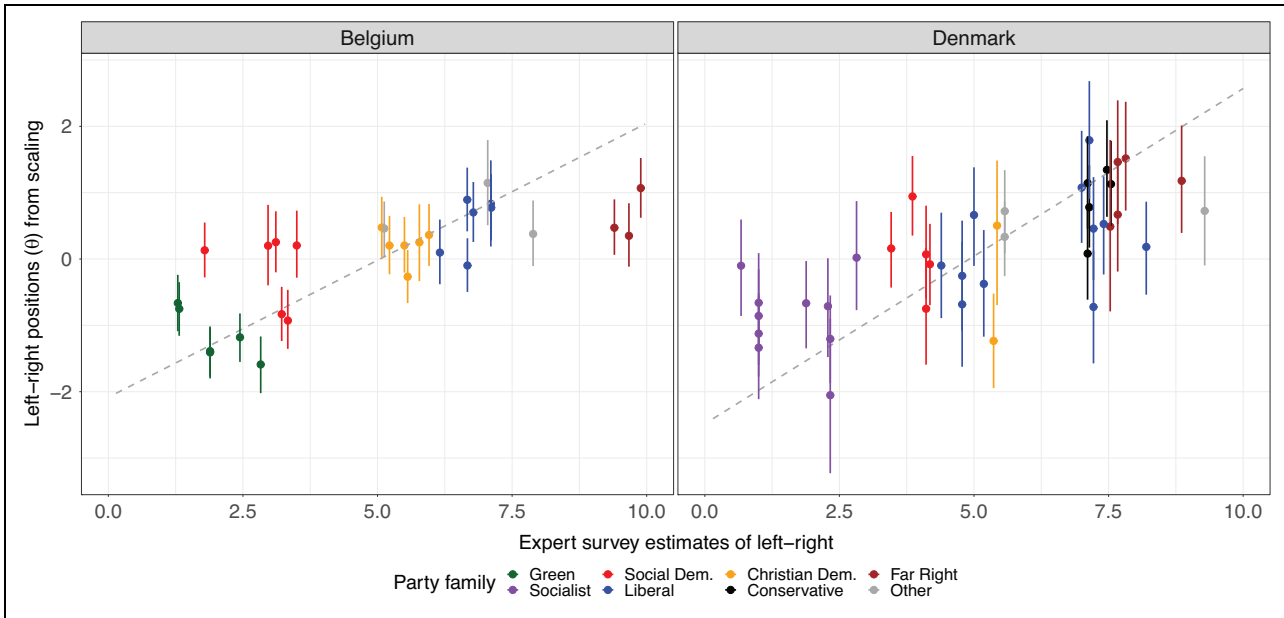
Applying the one-dimensional scaling model to the coded policy category counts, pooled across countries, we observe category positional parameters that largely confirm the Rile’s left and right policy category choices. Figure 2 plots the category discrimination parameters  $\beta_j$  for each policy category, fit to the post-1972 set of manifestos. As can be seen by the positioning of the item parameters relative to the dividing line (which represents the mean  $\beta$  across items), most of the Rile left categories are indeed associated with left positions, and most of the right categories with right positions. However, some are far less informative than others, and some are not estimated as corresponding to left or right as intended by the Rile index. The latter include Political Authority: Positive (305) and



**Figure 2.** Item discrimination parameters ( $\beta_j$ ) by Rile category, for the post-1972 set of documents (posterior means and 95% BCIs).

Constitutionalism: Positive (203), which are meant to be right-leaning items in the “Rile” scheme, and Education Expansion: Positive (506), which is supposed to be a leftist category.

Among the 30 policy categories the MP excludes from the Rile index, furthermore, Figure 2 shows several to be very strongly associated with left-right policy positions: Marxist Analysis: Positive (415) and National Way of Life:



**Figure 3.** Left-right positions for Belgian (1991–2007) and Danish (1994–2011) parties scaled from the Comparative Agendas Project dataset (posterior means with 95% BCIs) compared to expert survey placements. Dashed gray line represents linear fit. Party family information is from MP Data, with minor changes.

Negative (602) on the left, for instance, and Labor Groups: Negative (702) and Multiculturalism: Negative (608) on the right. So there are numerous categories not used to infer left-right context that could have contributed productively to the measurement of party positions along this single dimension. With our unsupervised approach, by contrast, we can include all available policy information, measuring each category’s relative contribution to determining the latent position  $\theta_i$  without requiring a (potentially arbitrary or controversial) list of “in” and “out” categories that determine the content of left-right.<sup>7</sup>

### Extending policy discovery to other hand-coding schemes

Using our unsupervised approach, which imposes no *ex ante* assumptions about policy content, we can also discover the nature of the left-right policy dimension using manual annotations from content analysis schemes never explicitly designed for the purpose of measuring policy positions. To demonstrate, we rely on a set of Belgian and Danish manifestos whose sentences were hand-coded by the CAP into 1 of 213 categories (Bevan, 2019).<sup>8</sup> Unlike the MP coding scheme, the CAP categories span very detailed topics including issues of political organization (e.g. “Branch relations,” code 2011), economic matters (e.g. “Unemployment rate,” code 103), social questions (e.g. “Immigration,” code 900). Additional various categories (to name a few) encompass “Sports Regulation” (code 1803), “Illegal Drugs” (code 1203) and

“Diplomats” (code 1929). These were never designed to measure positions, but rather to capture relative issue attention (Walgrave et al., 2006: 1025).

Constructing a fixed index to measure left-right positions would be very difficult from this data, since we have few prior theoretical expectations as to which of the numerous policy content categories actually convey information in terms of left-right positioning. By contrast, the inductive scaling approach allows us to measure this latent dimension and infer the contribution of each item to the scale. To do so, we apply the negative binomial scaling model to manifestos from the post-1990 period ( $N = 56$  Belgian documents from 1991 to 2007 and  $N = 56$  Danish documents from 1994 to 2011).<sup>9</sup>

Here, we focus on the validation of the results for the set of cases that could be matched with temporally close expert survey estimates, as above for the MP Data. Figure 3 plots the results, indicating high face validity. Green, Socialist and Social Democratic parties tend to be placed on the left, and the Far Right parties also appear on the right as expected. Overall, we obtain quite high correlations with the expert scores ( $r = .74$  in the Belgian and  $r = .72$  in the Danish case, with the gray dashed line indicating the OLS regression line).

Using a scaling approach, we can also infer the ideological association of specific policy categories, even when they were never explicitly designed for producing an ideological scale. Inspecting the distributions of the  $\beta_j$  parameters more closely reveals that approximately 33% of the categories discriminate on the left-right dimension in

the case of Belgium (judged on the basis of whether 95% of the posterior distribution of a  $\beta_j$  are to the left or right of zero), but only 12% do so in the model for the Danish data.

In Belgium, the three most leftist items are “Hazardous waste,” “Energy research and development,” and “Nuclear,” while the three most rightist categories are “Political campaigns,” “Juvenile crime,” and “Handicap discrimination.” The Danish results suggest the following as the items with the strongest link to the main axis of competition: “Air pollution,” “Alternative and renewable energy,” and “Family issues” on the left; “Criminal and civil code,” “Law enforcement agencies,” and “Law and crime general” on the right. Hence, to some extent in both countries questions of environment on the one side and those of law and order on the other are the ones that clearly separate parties. Not only does the scaling approach combine these appropriately to measure ideology, but also it allows direct inference about each category’s relative contribution to left- or right-associated ideology.

### Scaling multiple dimensions of left-right policy

For many applications, researchers are interested in measuring policy positions in more than one dimension. For instance, many party systems can be adequately characterized by competition along two main dimensions, an economic and a socio-cultural one (Bornschieer, 2010; Hooghe et al., 2002; Kitschelt, 1994). One approach for obtaining policy positions in multiple dimensions is to assign categories *ex ante* to different specific policy fields/dimensions, and conduct one-dimensional scaling for each of those (Elff, 2013; Meyer and Wagner, 2019), which could also be done with the approach we introduced above. This is equivalent to determining *a priori* which categories are connected to which dimensions.

Proceeding this way is not recommended, however, when we have reasons to believe that there are categories which can be linked to more than one of the latent dimensions. This is especially relevant if the second dimension is understood as a broader socio-cultural axis that goes beyond moral questions (like abortion and homosexuality) and also encompasses, for example, postmaterialism or issues like nationalism and immigration (Bornschieer, 2010; Hooghe et al., 2002; Kitschelt, 1994). In this case, we can say *a priori* what this axis should definitely exclude (categories that can uncontroversially be considered economic), but some statements, like those on the scope of environmental protection, can both reflect economic or socio-cultural positions. Hence, deriving two broader dimensions, one economic and one socio-cultural, is an excellent application for the two-dimensional scaling approach we suggest.

Extending the model to two dimensions  $d$ , we now model the expected counts as

$$\log(\mu_{ij}) = \alpha_i + \psi_j + \beta_{1j}\theta_{1i} + \beta_{2j}\theta_{2i} \quad (2)$$

in which we infer two positions  $\theta_{di}$  and two sets of additive discrimination parameters  $\beta_{dj}$  (for two dimensions  $d \in \{1, 2\}$ , compare Jackman, 2001). The challenge, as with any multidimensional measurement model, is to impose appropriate constraints in order to reach statistical identification of all the parameters (Jackman, 2001; Rivers, 2003). For the two-dimensional model, we set  $\phi_j^{-1} = 0$ , i.e. choosing a Poisson likelihood *a priori*. We provide full details on the constraints and parameterization in supplementary materials.

Figure 4 plots the discrimination parameters in the two dimensions, with labels attached to items whose discrimination parameter is greater than 0.5 in absolute terms (for full results see Figure A3 in the Online Supplementary Materials). The IRT approach allows to have both categories that are linked to just one dimension and categories that can be associated with either dimension. This provides flexibility while keeping the inferred dimensions easy to interpret and in correspondence with real-world political discourse, a feature that many other scaling approaches lack. The graph shows that there are several items that are connected with both the economic and the socio-cultural dimension, and often in opposing directions. For example, negative statements about European integration may reflect economically left or socio-culturally right positions. Conversely, positive remarks about European integration can result from views that are economically rightist, or socio-culturally liberal. These results inform debates on the nature of support for the European Union (e.g. Braun et al., 2019; Hooghe et al., 2002) and could be investigated in more detail in future research. Similarly, it is interesting to see that items representing critical stances toward globalization or cosmopolitanism (Multiculturalism negative, National way of life positive, Internationalism negative) correspond to pronounced rightist positions on the second dimension, but also to leftist views about economic policy. Finally, the modeling also reveals that political parties’ calls for decentralization may originate in either rightist economic positions or strongly liberal views on the socio-cultural dimension. Overall, we find a negative correlation between the discrimination parameters in the two dimensions ( $r = -.42$ ), which is more pronounced for items included in the Rile index ( $r = -.72$  for right-side and  $r = -.64$  for left-side of these categories). This suggests that a two-dimensional model of the policy space indeed gives more analytical leverage, and the proposed scaling method provides a flexible tool to derive such a space or test theoretical propositions about it.



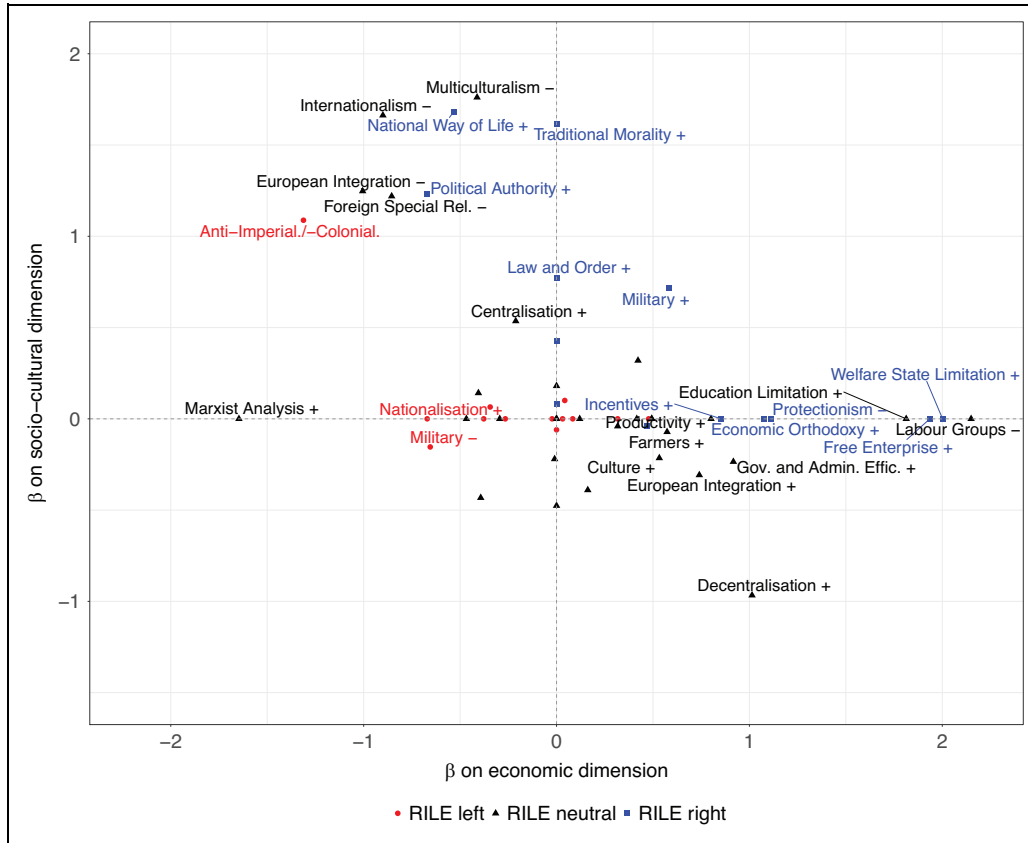


Figure 4. Discrimination parameters in a two-dimensional model.

## Conclusion

We use a statistical model for scaling category counts from hand-coded content analysis of political texts and demonstrate how it helps to infer the content of the left-right dimension in a way that avoids debates *ex ante* about the appropriate components of this dimension. We show that this method can produce valid left-right positions, based on inferring the contribution of each component policy dimension to the left-right super-dimension. This approach works with the MP codes that were designed for this purpose, but its flexibility also allows the use of annotations from content analysis schemes designed for other purposes, such as the CAP.

Using a parametric scaling model permits direct inference about the strength and direction of each content category's contribution to left-right ideology, while directly quantifying the underlying uncertainty. With some minimal constraints, the model can also be extended to two dimensions of left-right policy, mapping the policy space along a primarily material/economic as well as a socio-cultural dimension.

Our application of the measurement model also underscores the great value of manually coding manifesto texts using expert human judgment. Scaling content-analytic

data combines the best of two worlds: context-sensitive classification of specific policy statements originating from expert human judgment, with a valid and flexible procedure that infers low-dimensional policy measurements from the full set of coded statements.

The framework introduced is flexible enough to permit more sophisticated parameterizations. Exploiting the ability to add hierarchically dependent parameters in a Bayesian model, for instance, item parameters can be allowed to vary across space and time. Not only should this help in producing better measures of policy positions—and their uncertainty—but also it will make it possible to get a better understanding of variation in low-dimensional policy spaces across countries. The Bayesian scaling approach also has the potential to build richer models of the underlying political processes. This could not only further improve the model fit in specific contexts, but also allow for testing substantive explanations of parties' policy shifts directly within the measurement model.

## Acknowledgments

We thank Andrea Ceron, Michael Fop, Jouni Kuha, Ben Lauderdale, Will Lowe, and conference/seminar participants around the globe for comments and suggestions, and Korinna Veller for research assistance.

## Author's Note

Kenneth Benoit is also affiliated with Australian National University, Australia.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the European Research Council grant ERC-2011-StG 283794-QUANTESS.

## Supplemental material

Supplemental material for this article is available online.

## Notes

1. By contrast, other previous measurement models applied to the MP data come into play only *after* some form of pre-processing of the categories. Bakker (2009) pre-selects manifesto categories deemed left or right before running structural equation and Bayesian scaling models. The factor analysis approach suggested by König et al. (2013) requires metric input, so these authors initially group the sentence counts into 16 dimensions. Elff (2013) and Meyer and Wagner (2019) do not seek to measure left-right but policy area-specific positions, so they separate the raw data into policy fields *ex ante*.
2. As we discuss in more detail in the Online Supplementary Materials, this approach can also be seen as a (multinomial) generalization of a prominent item response theory (IRT) model from the psychometrics literature, the “nominal response model” (Bock, 1972). Suppose we have a multiple-choice exam question with unordered answer categories. The aim is to estimate to which extent the answer categories are informative about students’ latent abilities. In the text coding case, rather than dealing with an exam consisting of several questions with different answer options, we have multiple statements but always the same set of possible categories, those of the annotation scheme.
3. See also Slapin and Proksch (2008) and Lo et al. (2016), who extend the basic “wordfish” model in a text-as-data context, introducing an overdispersion parameter that refers to the document level.
4. It also includes an additional 51 extended sub-categories added to cover policy in Central and East European countries, which we have aggregated into the main categories according to the guidelines in the MP codebook (Volkens et al. 2018b: 9).
5. The oil crisis was a major turning point in post-war politics (see Jahn, 2011: 53), and the broader second sample should have an overlapping time period for all countries. Starting from the 2018b edition of the data (Volkens et al., 2018a) we use manifestos from countries with a Polity-IV rating of at least

seven, or a Freedom House rating of at least nine (when no Polity-IV rating was available). We dropped cases based on estimates and those with missing document length information.

6. Expert data are from Benoit and Laver (2006), Hooghe et al. (2010), Bakker et al. (2015), and Polk et al. (2017). Supported by Döring and Regel (2019), we matched the expert placement to the temporally closest manifesto, if a document is available within 3 years before or after the survey date.
7. Figure A2 in the Online Supplementary Materials shows that the use of the negative binomial distribution is justified. There are several categories whose variance exceeds the mean by a factor of five or more. The degree of overdispersion is strongest for categories that do not appear at all in a large share of the documents.
8. The Belgian data were originally collected by Stefaan Walgrave and his collaborators (Jeroen Joly, Anne Hardy, Brandon Zicha, Julie Sevenans, and Tobias Van Assche). Funding came from the European Science Foundation (grant number: 07-ECRP-008), from the Flemish National Science Foundation (grant number: G.0117.11N) and from the Belgian Federal Science Policy (grant number: IUAP P7/46). The data in the Danish Policy Agenda Project have been collected by Christoffer Green-Pedersen and Peter B. Mortensen with support from the Danish Social Science Research Council and the Research Foundation at Aarhus University.
9. The model is run for a warmup period of 5000 iterations, followed by another 5000 for inference, with a thinning factor of 5.

## References

- Albright JJ (2008) *Bayesian Estimates of Party Left-Right Scores*. Unpublished paper, Bloomington: Indiana University.
- Bakker R (2009) Re-measuring left–right: a comparison of SEM and Bayesian measurement models for extracting left–right party placements. *Electoral Studies* 28(3): 413–421.
- Bakker R, de Vries C, Edwards EE, et al. (2015) Measuring party positions in Europe: the Chapel Hill expert survey trend file, 1999–2010. *Party Politics* 21(1): 143–152.
- Baumgartner FR, Green-Pedersen C and Jones BD (2008) *Comparative Studies of Policy Agendas*. Milton Park: Routledge.
- Benoit K and Laver M (2006) *Party Policy in Modern Democracies*. London: Routledge.
- Benoit K, Laver M and Mikhaylov S (2009) Treating words as data with error: uncertainty in text statements of policy positions. *American Journal of Political Science* 53(2): 495–513.
- Bevan S (2019) Gone fishing. In: Baumgartner FR, Breunig C and Grossman E (eds) *Comparative Policy Agendas: Theory, Tools, Data*. Oxford: Oxford University Press, pp. 17–34.

- Bock RD (1972) Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika* 37(1): 29–51.
- Bornschieer S (2010) The new cultural divide and the two-dimensional political space in Western Europe. *West European Politics* 33(3): 419–444.
- Braun D, Popa SA and Schmitt H (2019) Responding to the crisis: Eurosceptic parties of the left and right and their changing position towards the European Union. *European Journal of Political Research* 58(3): 797–819.
- Budge I and Meyer T (2013) Understanding and validating the leftright scale (Rile). In: Volkens A, Bara J, Budge I, McDonald MD and Klingemann HD (eds) *Mapping Policy Preferences from Texts: Statistical Solutions for Manifesto Analysts*. Oxford: Oxford University Press, pp. 85–106.
- Budge I, Klingemann HD, Volkens A, et al. (2001) *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.
- Budge I, Robertson D and Hearl D (1987) *Ideology, Strategy and Party Change: Spatial Analyses of Post-War Election Programmes in 19 Democracies*. Cambridge: Cambridge University Press.
- Cameron AC and Trivedi PK (1986) Econometric models based on count data. Comparisons and applications of some estimators and tests. *Journal of Applied Econometrics* 1(1): 29–53.
- Döring H and Regel S (2019) Party Facts: A database of political parties worldwide. *Party Politics* 25(2): 97–109.
- Downs A (1957) An economic theory of political action in a democracy. *Journal of Political Economy* 65(2): 135–150.
- Elff M (2013) A dynamic state-space model of coded political texts. *Political Analysis* 21(2): 217–232.
- Flentje JE, König T and Marbach M (2017) Assessing the validity of the manifesto common space scores. *Electoral Studies* 47: 25–35.
- Franzmann S and Kaiser A (2006) Locating political parties in policy space: A reanalysis of party manifesto data. *Party Politics* 12(2): 163–188.
- Franzmann ST (2015) Towards a real comparison of left-right indices. A comment on Jahn. *Party Politics* 21(5): 821–828.
- Gabel MJ and Huber JD (2000) Putting parties in their place: Inferring party left-right ideological positions from party manifesto data. *American Journal of Political Science* 44(1): 94–103.
- Gemenis K (2013) What to do (and not to do) with the comparative manifestos project data. *Political Studies* 61(S1): 23–43.
- Hooghe L, Bakker R, Brigevech A, et al. (2010) Reliability and validity of the 2002 and 2006 Chapel Hill expert surveys on party positioning. *European Journal of Political Research* 49(5): 687–703.
- Hooghe L, Marks G and Wilson CJ (2002) Does left/right structure party positions on European integration? *Comparative Political Studies* 35(8): 965–989.
- Jackman S (2001) Multidimensional analysis of roll call data via Bayesian simulation: identification, estimation, inference and model checking. *Political Analysis* 9(3): 227–241.
- Jahn D (2011) Conceptualizing left and right in comparative politics towards a deductive approach. *Party Politics* 17(6): 745–765.
- Kitschelt H (1994) *The Transformation of European Social Democracy*. Cambridge: Cambridge University Press.
- Klingemann HD, Volkens A, Bara J, et al. (2006) Mapping policy preferences II. In: *Estimates for Parties, Electors and Governments in Central and Eastern Europe, European Union and OECD 1990–2003*. Oxford: Oxford University Press.
- König T, Marbach M and Osnabrügge M (2013) Estimating party positions across countries and time—a dynamic latent variable model for manifesto data. *Political Analysis* 21(4): 468–491.
- Laver M and Budge I (1992) Measuring policy distances and modeling coalition formation. In: Laver MJ (ed.), *Party Policy and Government Coalitions*. New York, NY: St. Martin's Press, pp. 15–40.
- Laver M, Benoit K and Garry J (2003) Estimating the policy positions of political actors using words as data. *American Political Science Review* 97(2): 311–331.
- Lo J, Proksch SO and Slapin JB (2016) Ideological clarity in multiparty competition: a new measure and test using election manifestos. *British Journal of Political Science* 46(3): 591–610.
- Lowe W (2016) Scaling things we can count. Unpublished paper, Princeton University. Available at: <https://dl.conjugateprior.org/preprints/all-on-the-line.pdf> (accessed on 3 July 2021).
- Meyer TM and Wagner M (2019) It sounds like they are moving: understanding and modeling emphasis-based policy change. *Political Science Research and Methods* 7(4): 757–774.
- Mikhaylov S, Laver M and Benoit K (2012) Coder reliability and misclassification in the human coding of party manifestos. *Political Analysis* 20(1): 78–91.
- Mölder M (2013) The validity of the rile left-right index as a measure of party policy. *Party Politics* 22(1): 37–48.
- Polk J, Rovny J, Bakker R, et al. (2017) Explaining the salience of anti-elitism and reducing political corruption for political parties in Europe with the 2014 Chapel Hill expert survey data. *Research & Politics* 4(1): 1–9.
- Rivers D (2003) *Identification of Multidimensional Spatial Voting Models*. Unpublished manuscript. Stanford, CA: Stanford University.
- Slapin JB and Proksch SO (2008) A scaling model for estimating time-series party positions from texts. *American Journal of Political Science* 52(3): 705–722.
- Volkens A, Bara J, Budge I, et al. (2013) *Mapping Policy Preferences from Texts: Statistical Solutions for Manifesto Analysts*. Oxford: Oxford University Press.

- Volkens A, Krause W, Lehmann P, et al. (2018a) The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Version 2018b. DOI: 10.25522/manifesto.mps.2018b.
- Volkens A, Krause W, Lehmann P, et al. (2018b) The Manifesto Project dataset - Codebook. Manifesto Project (MRG/CMP/MARPOR). Version 2018b. [https://manifesto-project.wzb.eu/download/data/2018b/codebooks/codebook\\_MPDataset\\_MPDS2018b.pdf](https://manifesto-project.wzb.eu/download/data/2018b/codebooks/codebook_MPDataset_MPDS2018b.pdf) (accessed 3 July 2021).
- Walgrave S, Varone F and Dumont P (2006) Policy with or without parties? A comparative analysis of policy priorities and policy change in Belgium, 1991 to 2000. *Journal of European Public Policy* 13(7): 1021–1038.

### Author biographies

**Thomas Däubler** is Assistant Professor and Ad Astra Fellow at University College Dublin. His main research interests are electoral systems, legislative politics, and political parties.

**Kenneth Benoit** is Professor of Computational Social Science at the Department of Methodology, London School of Economics and Political Science, and Professor, Part-Time at the School of Politics and International Relations, Australian National University. His research focuses on automated, quantitative methods of processing large amounts of textual and other forms of big data—mainly political texts and social media—and the methodology of text mining.