

Michael Herrmann¹ and Holger Döring²

¹Department of Politics and Public Administration, University of Konstanz, Box 83, 78457 Konstanz, Germany.

E-mail: michael.herrmann@uni-konstanz.de

²Department of Social Sciences and SOCIUM Research Center, University of Bremen, Bremen, Germany

Abstract

We develop a new measure of party position based on a scaling of ideology tags supplied in infoboxes on political parties' Wikipedia pages. Assuming a simple model of tag assignment, we estimate the locations of parties and ideologies in a common space. We find that the recovered scale can be interpreted in familiar terms of "left versus right." Estimated party positions correlate well with ratings of parties' positions from extant large-scale expert surveys, most strongly with ratings of general left-right ideology. Party position estimates also show high stability in a test-retest scenario. Our results demonstrate that a Wikipedia-based approach yields valid and reliable left-right scores comparable to scores obtained via conventional expert coding methods. It thus provides a measure with potentially unlimited party coverage. Our measurement strategy is also applicable beyond Wikipedia.

Keywords: ideal point estimation, ideology scores, party positions, wordscores

1 Introduction

The idea that parties occupy different positions on an ideological continuum is fundamental to theories of the political process. The need for scoring parties on such a continuum is evidenced by the sheer diversity of approaches and the creativity political scientists devote to obtaining such scores, which range from expert codings of party manifestos (Volkens *et al.* 2013), to left-right placements of parties by expert judges (Kitschelt 2014) or voters (Lo, Proksch, and Gschwend 2014), to supervised scaling of word frequencies in party manifestos (Laver, Benoit, and Garry 2003), to unsupervised scaling approaches that extract positions from expert-coded manifesto content (Däubler and Benoit 2017) or raw word frequencies (Slapin and Proksch 2008) in political manifestos, and to analyses of roll call votes (Bräuninger, Müller, and Stecker 2016) and parliamentary speech (Lauderdale and Herzog 2016; Peterson and Spirling 2018). Together, these approaches enable the measurement of party positions across time, space, levels of government, and policy areas.¹

We propose a new measure of party left-right position, one that offers greater party and country coverage than any existing measure. The measure is derived from semistandardized information about party ideology available on the English Wikipedia.² In particular, we draw on ideological keywords (e.g., socialism) that Wikipedia editors use to tag parties and to link them with Wikipedia articles on those ideologies. We develop an ideal point model of how these tags are assigned and use it to scale over 2,000 parties and their associated ideologies on a latent dimension. Our scaling approach is based on the idea that co-occurrences of ideological keywords across parties

¹ Similar efforts have been made in order to obtain left-right scores for actors other than parties, for example, voters, political candidates, and industries (e.g., Bonica 2014; Temporão *et al.* 2018).

² Wikipedia is the largest knowledge database in the world. The English Wikipedia alone has over 5.8 million articles (<https://en.wikipedia.org/wiki/wikipedia>, accessed March 2019). Wikipedia is also the most used knowledge database in the world with over 500 million unique visitors each month, ranking among the top 10 most visited websites for years (<http://alexa.com/topsites>, accessed March 2019).

reveal information about the closeness of parties and ideologies. Keywords that often occur in the same parties (e.g., socialism vs. social democracy) should be closer in space than keywords that rarely occur together (e.g., socialism vs. conservatism). Likewise, parties sharing the same ideological keywords should occupy more similar positions in political space than parties sharing few keywords. To capture these dependencies, we formulate a model in which parties are more likely to get tagged with keywords that are close to them. In a second step, we extend the model to address the fact that some keywords on Wikipedia are inherently ordered.

Our analysis demonstrates that keyword summaries provided by Wikipedia editors enable valid and reliable inferences about party left–right position. Based on our model of keyword assignment, which allows for misclassification and differences in keyword informativeness, we recover a scale from Wikipedia classifications that conforms with common intuitions of left versus right. We show that estimates of party position on this scale correlate with ratings of party position from the largest available expert surveys, and most strongly with ratings of general left–right position. We further demonstrate the reliability of our estimates over repeated measurements with Wikipedia data collected months apart. Together, our results indicate that Wikipedia classifications allow for extracting left–right scores comparable to scores obtained via conventional expert coding methods.

Our findings are in line with studies showing that political information on Wikipedia is often factually correct³ (Brown 2011; Göbel and Munzert 2021; Poschmann and Goldenstein 2019), and that Wikipedia can be used for extending political science measurement to a much larger universe of cases (Munzert 2018). Our results also tie in with studies demonstrating the validity of crowd-sourcing information in political research (e.g., Sumner, Farris, and Holman 2020; Winter, Hughes, and Sanders 2020). In particular, a number of studies find that assessments by large groups of nonexpert coders (i.e., crowd-coding) of the content of party manifestos or parliamentary debates can give estimates of political position as good as assessments by expert coders (cf. Benoit *et al.* 2016; Haselmayer and Jenny 2017; Lehmann and Zobel 2018; Horn 2019).

Compared to existing approaches, a Wikipedia-based approach offers potentially unlimited coverage of parties. To get some sense of the scope afforded by our measure, note that the largest available databases in political science together include over 4,000 unique parties (Döring and Regal 2019). Extant measures of party left–right position cover only a minority of these. For example, the largest data source in terms of countries covered, the Democratic Accountability and Linkages Project (DALP), provides expert ratings of left–right position for 506 parties from 88 electoral democracies (see Table 1). The data source with the largest party coverage, the Manifesto Project, includes 1,170 parties from 60 countries. The use of Wikipedia allows us to expand that coverage considerably: based on the largest available compilation of parties, drawn from major political science datasets (Party Facts: Döring and Regal 2019), we identify about 3,900 parties (third column) that possess a Wikipedia article. Of these, about 2,100 (fourth column) contain scalable information on party ideology. With the steady expansion of Wikipedia, this number is likely to grow.

While our focus in this paper is on Wikipedia data, the measurement approach we pursue is applicable more generally. For example, selected experts or hired crowd-coders could be tasked with the indexing of parties with keywords. The keywords could be predetermined by the researcher or chosen by the coders. Aside from parties, researchers could also scale other actors, be it politicians, judges, or interest groups. Our approach should be applicable whenever

3 A number of studies examine the accuracy of Wikipedia vis à vis traditional expert sources (for a review, see Mesgari *et al.* 2014). Unlike the majority of these studies, which employ a “small-*n*, every-detail approach” (Brown 2011, 340) in which one or several field experts scrutinize the entire content of a small and often selective sample of Wikipedia articles, our analysis is more akin to Brown’s “large-*n*, specific-facts approach” (Brown 2011, 340).

Table 1. Coverage of the largest data source on parties (Party Facts), the largest data sources on party left-right positions (MAP: Manifesto Project, DALP: Democratic Accountability and Linkages Project, and CHES: Chapel Hill Expert Survey), and our measure.

Continent	Party Facts	Wikipedia	Scalable	MAP	DALP	CHES
Africa	672	601	261	6	68	0
Americas	854	696	381	90	103	0
Asia	871	732	375	216	109	10
Europe	1,916	1,712	1,079	836	214	467
Oceania	170	144	51	22	12	0
TOTAL	4,483	3,885	2,147	1,170	506	477

La République En Marche!

From Wikipedia, the free encyclopedia

La République En Marche ^[a] (frequently abbreviated **REM**, **LRM** or **LREM**, officially **LaREM**; possible translation: "The Republic on the move!"), sometimes called **En Marche** ! (French: [ɑ̃ maʁʃ]; English translation: "Forward!",^{[7][8]} "Onward!",^[9] "Working!" or "On The Move!"),^[10] is a centrist,^[11] liberal^[12] and social-liberal^[13] political party in France. It was founded on 6 April 2016 by **Emmanuel Macron**, a former Minister of Economy, Industry and Digital

Affairs, who was later elected President of France in the 2017 French presidential election with 66.1% of the second round vote. The party considers La République En Marche a progressive movement, uniting both the left and the right.^[14] The party ran candidates in the 2017 legislative elections^[15] including dissidents from the Socialist Party, The Republicans and minor parties. It won an absolute majority in

La République En Marche !

LA RÉPUBLIQUE
En Marche !

Executive Officer Vacant

General Secretary Stanislas Guerini

President in the National Assembly Gilles Le Gendre

President in the Senate François Patriat

Founder Emmanuel Macron

Founded 6 April 2016; 2 years ago

Headquarters 63, Rue Sainte-Anne
75002 Paris, France

Youth wing Les Jeunes avec Macron

Membership (2019) ▲ 411,892 claimed adherents^[1]

Ideology Liberalism^{[2][3]}
Social liberalism^[4]
Pro-Europeanism^[5]

Political position Centre^[2]

European affiliation None

Ideology

Liberalism^{[2][3]}

Social liberalism^[4]

Pro-Europeanism^[5]

Political position

Centre^[2]

Figure 1. Infobox in the Wikipedia article for La République En Marche! and associated tags.

political entities are being indexed with keywords that relate them to some unobserved latent dimension.

2 Party Ideology Classifications on Wikipedia

We focus on information from “party infoboxes” (see Figure 1 for an example). Infoboxes are placed at the top of a Wikipedia article and give a quick summary of facts on a topic. According to Wikipedia guidelines, they

“contain important facts and statistics of a type which are common to related articles. For instance, all animals have a scientific classification (species, family and so on), as well as a conservation status. Adding an [infobox] to articles on animals therefore makes it easier to quickly find such information and to compare it with that of other articles.”⁴

Unlike the textual format of a basic Wikipedia page, infoboxes restrict editors to submitting information for a defined set of categories; for parties, these are, for instance, the party’s name, its founding year, or the name of its leader. The categories of a party infobox are fixed and cannot

4 <https://en.wikipedia.org/wiki/Help:Infobox> (accessed March 2019).

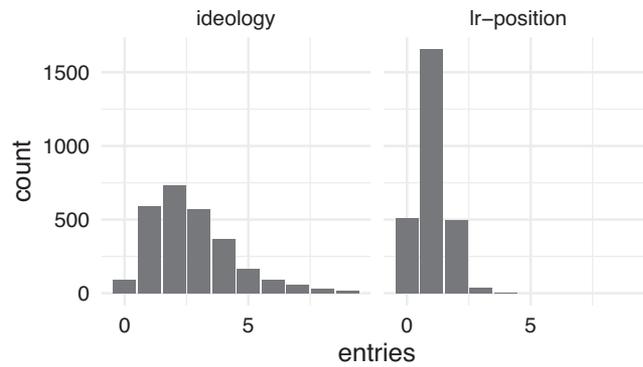


Figure 2. Number of tags per party.

be altered by editors.⁵ Their content should be “comparable,” “concise,” “relevant to the subject,” and “already cited elsewhere in the article.”⁴ While the categories of a party infobox are predefined, their use is optional. Hence, some categories, or even the entire infobox, may be missing from a party’s Wikipedia article.⁴

Of interest to us are the infobox entries (henceforth tags) provided in the categories “ideology” (henceforth ideology) and “political position” (henceforth lr-position). As can be seen from Figure 1, these categories enable Wikipedia editors to index parties with political philosophies and inclinations, and to establish links between their respective Wikipedia pages. For example, the French party La République En Marche!, formed by Emmanuel Macron in the run-up to his bid for presidency, is tagged with the ideologies liberalism, social liberalism, and pro-Europeanism, with links given for all tags. In terms of lr-position, the party is tagged as centre, including a link to an associated Wikipedia page.

There are no restrictions (that we know of) on which or how many ideology tags a party can have. Our data suggest that editors use tags sparingly, and that they often draw on existing tags rather than generating new ones. The large majority of parties in our data only receive between one and four ideology tags. Furthermore, some tags are assigned much more often than others (see Figures 2 and 3).

In the political position category, editors mostly draw upon a set of seven tags for classifying a party as far-left, left-wing, centre-left, centre, centre-right, right-wing, or far-right (see Online Appendix E for details). Compared to ideology tags, lr-position tags tend to get used somewhat less often: among all parties that exhibit a tag, four out of five have an lr-position tag, while nearly all (97%) have an ideology tag. Roughly three out of five parties receive one lr-position tag, and roughly one out of five parties receive two such tags (see Figure 2). In the latter case, the assigned tags nearly always represent adjacent positions on the political spectrum. Only very few parties receive more than two lr-position tags.

3 A Model of Tag Assignment

We assume that tag assignment is driven by similarity: a party gets tagged if its platform is perceived to be in agreement with the tagged ideology’s basic tenets. We further assume that agreement between party platforms and ideologies can be represented as distances on a latent dimension. Let $i = 1, 2, \dots, N$ be an index of parties, let $j = 1, 2, \dots, J$ be an index of ideology tags, and let y be an N by J matrix with binary entries y_{ij} indicating the occurrence of tag j in party i ’s Wikipedia article. We represent these occurrences by an ideal point model similar to the one

⁵ The party infobox template is protected, see: https://en.wikipedia.org/wiki/Template:Infobox_political_party (accessed March 2019).

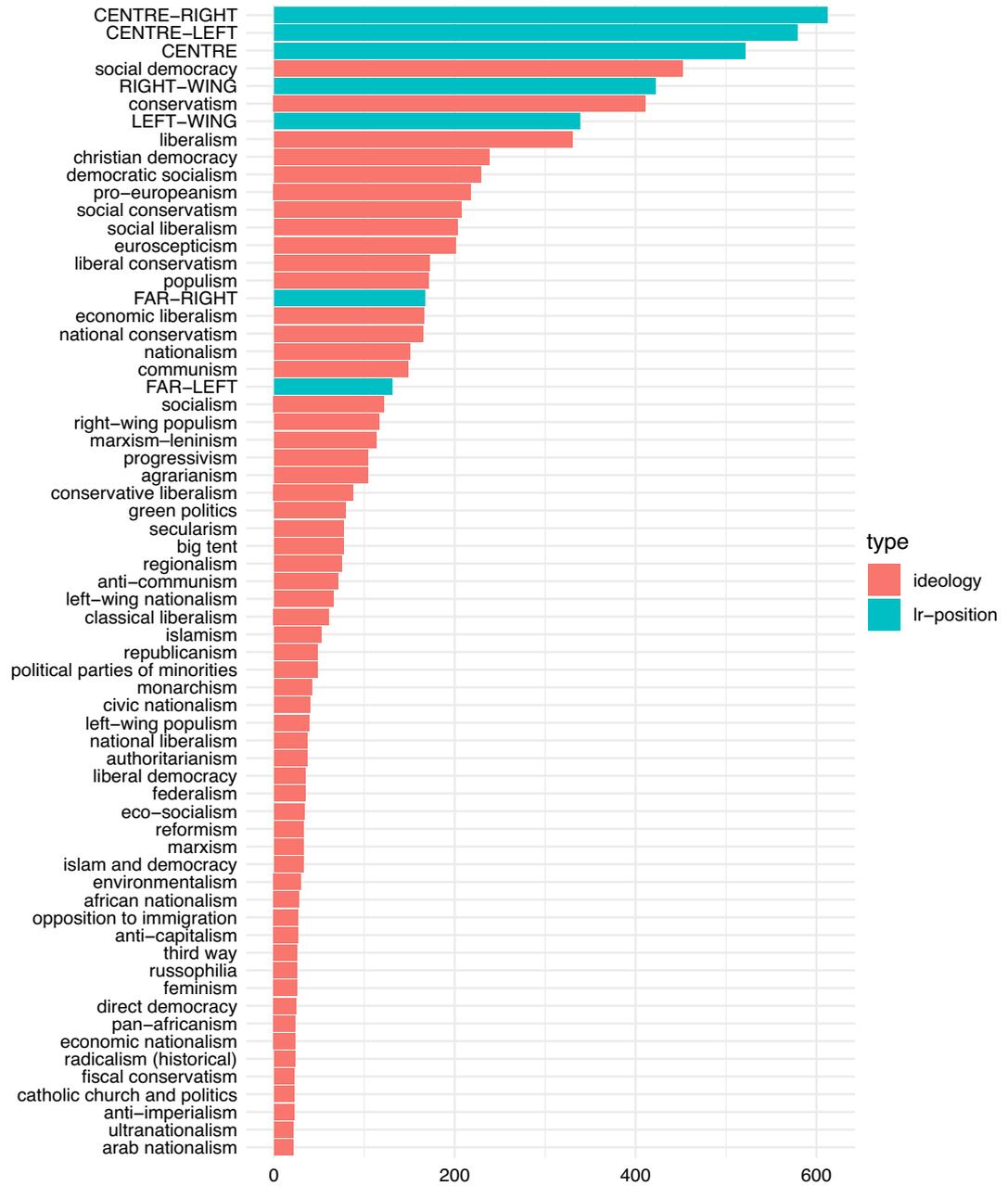


Figure 3. Frequency of usage for tags that are used at least 20 times.

suggested in Lowe (2008),

$$\Pr(y_{ij} = 1) = F(\alpha_j - \beta_j(o_j - x_i)^2), \tag{1}$$

where F is the inverse logit transformation, x_i and o_j are the positions of party i and ideology j on the latent dimension, β_j is a tag-specific discrimination parameter, and α_j is a tag-specific constant.

As can be seen, the probability of observing tag j in party i is maximized when their positions on the latent dimension coincide (i.e., when $x_i = o_j$). Thus, o_j is the modal value of the response function, and the maximum probability at o_j is $F(\alpha_j)$. The discrimination parameter β_j measures how strongly the probability of observing tag j in party i depends on their closeness on the latent

dimension, that is, how rapidly the probability of observing j decreases as we move away from α_j . A high value of β_j implies a peaked response curve in which a party's probability of exhibiting tag j decreases quickly with increasing distance from α_j ; a low value implies a wide response curve in which the probability that a party exhibits tag j does not depend much on its position on the latent dimension. Tags with low β are thus less informative about party position. If $\beta_j = 0$, tag j occurs with probability $F(\alpha_j)$, regardless of party position. Parameters β and α thus allow tags to differ in their informativeness and prevalence.

The above ideal point model is closely connected to extant approaches to scaling word frequencies in political manifestos.⁶ Lowe (2008, 365f) discusses how the Wordscores and Wordfish approaches can be subsumed under an ideal point model with the same parametric structure as Equation (1). Our application of the model is different in that we only observe binary word presences and absences.⁷ The main substantive difference to the commonly used Wordfish approach is that Equation (1) allows words to have distinct locations. As explained in Lowe (2008), Wordfish implements a reduced version of the ideal point model that allows word usage only to rise or fall along the latent dimension. By contrast, we assume that keywords (i.e., ideologies) occupy positions on the underlying dimension such that their occurrence may rise and fall, that is, parties may be too far to the right as well as too far to the left for being tagged with a particular ideology. Word frequency scaling approaches typically also include constant terms for the manifestos to correct for the obvious fact that longer manifestos contain more words (Slapin and Proksch 2008). We do not apply such a correction as parties do not vary much in their propensity of being tagged: inspection of the marginal distribution shows that the number of tags per party varies little (see Figure 2) and is small compared to the number of available tags (see Figure 3 for the most prominent ones).

As explained in the previous section, Wikipedia offers two kinds of tags: ideology and lr-position. Unlike ideology tags, lr-position tags directly encode regions (i.e., far-left, left-wing, centre-left, etc.) on the underlying scale. We shall treat these tags as coarse, graded indicators of party position in one of seven consecutive intervals along the latent dimension and assume that a party gets tagged if its position is perceived to fall within the interval implied by the tag. To achieve this, we represent the assignment of lr-position tags by an ordered logit model with unknown x (Treier and Jackman 2008; Caughey and Warshaw 2015). Numbering tags from left to right, and letting $z_i = k$ denote the presence of tag $k = 1, 2, \dots, 7$ in party i 's Wikipedia article,

$$\Pr(z_i = k) = F(\tau_k - \gamma x_i) - F(\tau_{k-1} - \gamma x_i), \quad (2)$$

where γ is a discrimination parameter and τ are tag-specific cut points with $\tau_0 = -\infty$, $\tau_7 = \infty$, and $\tau_{k-1} < \tau_k$, for all k .

Our ordered logit model in Equation (2) is essentially a one-item version of the graded response model from item response theory (Samejima 1969). In this model, the probabilities of tag assignment are single-peaked, except for the left- and rightmost tags. The model parameterizes these outcome probabilities via $\Pr(z_i \leq k) = F(\tau_k - \gamma x_i)$, the cumulative probability of observing tag k or lower (i.e., the probability that party i is tagged as k or further to the left). These cumulative probabilities define a set of binary logit models, each with its own intercept τ_k and a common slope parameter γ . The slope parameter γ measures how rapidly the probability of tag assignment changes in response to party position on the latent dimension. Values of γ further from zero imply steeper cumulative response curves and more peaked outcome response curves.

- 6 Similar approaches can also be found in the literatures on item response theory (i.e., unidimensional unfolding; see, e.g., Andrich 1988) and community ecology (i.e., Gaussian ordination; see, e.g., ter Braak 1985; ter Braak and Šmilauer 2015).
7 In Online Appendix A, we elaborate on this connection by showing that Equation (1) and the model studied in Lowe (2008) are special cases of the same generalized linear model.

The intercepts τ_k define a set of adjacent intervals on the latent dimension, corresponding to the outcome values. The boundaries of these intervals can be interpreted as the points at which the cumulative probabilities are tied, and they are given by τ_k/γ , for $k = 1, 2, \dots, 6$. For example, a party located at $x = \tau_2/\gamma$ has $F(\tau_2 - \gamma x) = F(0) = 0.5$, and thus a 50 : 50 chance of being tagged as left-wing or something further to the left, as opposed to being tagged as centre-left or something further to the right. By construction, the interval boundaries are defined with respect to the cumulative responses, not the observed responses. However, the boundaries are also related to the observed responses in that the probability of outcome k peaks in the middle of the interval associated with it.

As explained in the previous section, for some parties, we observe more than one lr-position tag. In these instances, we treat each outcome value as an independent realization of z , conditional on x , and model their joint probability. Formally, this means that instead of z_i in Equation (2), we model the outcome variable z_{il} , where $l = 1, 2, \dots, 7$ indexes a party's first, second, and so on observed lr-position tag.

In what follows, we pursue two approaches to recovering party positions from Wikipedia articles:

1. Estimating Equation (1) with data on ideology tag assignment.
2. Jointly estimating Equations (1) and (2) with data on ideology and lr-position tag assignment.

The first approach recovers the positions of parties and tags solely from their co-occurrences without any prior assumptions about their locations on the latent dimension. This approach has applicability beyond Wikipedia. It can be used whenever political entities are indexed with a set of keywords, following a logic of “pick any(-thing that applies)” (Levine 1979).

When applied to Wikipedia tags, a downside of the approach is that it discards the additional information contained in lr-position tags. Since we know the ordering of these tags, their relative locations on the latent dimension do not need to be estimated. This allows us to constrain the estimation problem and recover the locations of parties and tags with greater precision. The fact that we know the ordering of lr-position tags also helps with the identification of the latent dimension. To see why, note that a scaling of ideology tags alone can reveal such a dimension only if neighboring tags occur jointly in some parties. The inclusion of lr-position tags makes this precondition unnecessary: even if ideology tags occurred in completely separate (i.e., nonoverlapping) groups of parties, their ordering could still be inferred from the lr-position tags with which they co-occur.⁸ The key assumption behind combining ideology and lr-position tags in this way is that both sets of tags can be placed on the same latent dimension. As a simple check of this assumption, we compare the placement of ideology tags under both estimation approaches. Strong differences in results would be an indication that the two sets of tags might not form a common scale.

4 Estimation

We employ Bayesian Markov Chain Monte Carlo (MCMC) methods to obtain a posterior distribution for all model parameters (Albert and Chib 1993). MCMC simulations are performed with JAGS, Version 4.3.0 (Plummer 2017). JAGS code used to estimate both models is provided in Online Appendix C.⁹

⁸ We thank an anonymous reviewer for pointing out this particular fact to us.

⁹ See Herrmann and Döring (2021) for code and data to reproduce all results.

4.1 Identification and parameterization

Like all latent variable models, Equations (1) and (2) are not identified without some restrictions on the parameters. Three constraints are necessary to identify all parameters in a unidimensional model (Rivers 2003). To resolve invariance to addition, we center the underlying scale at 0; to resolve invariance to multiplication, we standardize the scale to units of $SD(x)$; and to resolve invariance to reflection, we impose the constraint $\bar{x} < o_m$, where m is the index value for the tag conservatism. The identifying restrictions are imposed on each posterior draw, with offsetting transformations on the other parameters to keep outcome probabilities unchanged (see Online Appendix B for further details).

To facilitate convergence to the target distribution, we estimate a reparameterized version of Equation (1). As shown in Lowe (2008, 366), Equation (1) can be equivalently stated as

$$\Pr(y_{ij} = 1) = F(\delta_j + \lambda_j x_i - \beta_j x_i^2), \quad (3)$$

where $\delta_j = \alpha_j - \beta_j o_j^2$ and $\lambda_j = 2\beta_j o_j$, and where the negative sign on β follows from the assumption of concavity (i.e., the response function must be single peaked, not single dipped). Equation (3) has the same number of parameters as Equation (1) but is quadratic only in x , while Equation (1) is quadratic in x and o . In practice, we find that parameterizing the model as in Equation (3) yields faster convergence to the target distribution both in terms of iterations and runtime.¹⁰ We therefore use Equation (3) as our estimation equation. To obtain estimates of the parameters in Equation (1), we transform posterior draws of λ , δ , and β into posterior draws of α , β , and o , using the reparameterization relations stated above, and subsequently apply the identifying restrictions.

4.2 Priors

We assign standard normal priors to x (Albert and Johnson 1999), and normal priors with mean zero and variance 5 to δ and λ . To enforce the concavity constraint in Equation (3), we assign log-normal priors with log-mean zero and log-variance parameter 2 to β , which implies a prior variance on β of about 47. To see what these priors mean, consider the prior variation of o relative to x (cf. Clinton and Jackman 2009, 601–602): Monte Carlo simulation shows that the prior variances on λ and β imply a prior 95% credibility interval for o of about $[-9.6, 9.6]$. Party positions are thus *a priori* interior to ideologies with the prior variance of ideologies being large relative to that of party positions. Since the locations of ideologies and parties are only identified relative to each other, because the scale of the latent dimension is unknown, the wide range of ideologies relative to party positions suggests that the priors are permissive enough to allow the data to inform the estimation result. Using wider priors on ideologies leaves results unchanged (but slows down convergence). Likewise, using somewhat tighter priors also yields similar results.

For γ in the ordered logit model, we use a normal prior with mean zero and variance 25. For the category cut points, we choose normal priors with mean zero and variance 25, subject to the order constraint $\tau_1 < \tau_2 < \dots < \tau_6$.

4.3 Starting Values

We employ correspondence analysis (CA) to generate starting values for party positions. CA is a deterministic dimension-reduction technique, similar to principal components analysis (Greenacre 2010). Our use of it is motivated by a result in ter Braak (1985) proving that first-dimension coordinates from a CA of the binary data matrix yield approximate maximum likelihood

¹⁰ Following Bafumi *et al.* (2005, 176), we use potential scale reduction factors to compare convergence under both parameterizations for a given number of iterations.

(ML) estimates of x and o if the data-generating process adheres to Equation (1).¹¹ This property of CA has made it particularly popular in scaling applications involving sparse data matrices (i.e., matrices with many more zeroes than ones), which are common in other fields (see, e.g., Smith and Neiman 2007; ter Braak and Šmilauer 2015) and which we also encounter in our application. ML estimators tend to be numerically unstable in these situations, while CA is guaranteed to yield a solution regardless of how large or sparse the data matrix is (ter Braak and Šmilauer 2015). Moreover, CA tends to approximate ML more closely when the data are sparse (ter Braak 1985, 863).¹² In our MCMC estimation approach, numerical instability is not an issue; however, convergence to the target distribution can be slow with sparse data. To facilitate convergence, we follow the proposal in ter Braak (1985) and use CA estimates of x and o as starting values. Online Appendix G compares our initial CA estimates to the final Bayesian estimates.

To generate starting values for the remaining parameters, we first estimate logistic regression models—one for each ideology—of the form given in Equation (1) using CA estimates of o and x as inputs. This gives us estimates of α and β . We then transform those estimates, using the reparameterization relations given in Section 4.1, to obtain starting values for δ and λ . To obtain starting values for γ and τ , we estimate the ordered logistic regression model in Equation (2) using CA estimates of x as inputs.

4.4 Convergence

We set up four parallel chains and run each model for 22,000 iterations, discarding the first 2,000 draws as burn-in. We thin the result, keeping every 10th draw, to obtain 8,000 samples from the posterior distribution. Assessment of posterior draws via traceplots and potential scale reduction factors suggest convergence of all parameters to their target distribution (see Online Appendix D for details).

5 Selection of Parties and Tags

Collecting information from Wikipedia requires a list of parties that defines our target population. To maximize coverage, we draw on the largest list of parties that is currently available in political science: the Party Facts database (version 2019a; Döring and Regel 2019; Bederke, Döring, and Regel 2019).

Party Facts is a collaborative project that aims to solve the problem of delineating the universe of political parties. It offers an authoritative reference list of relevant parties in the world, based on the parties included in the CLEA, ParlGov, Manifesto Project, PolCon, and a number of smaller datasets. Party Facts includes nearly all parties that won at least 5% seat share in a national election, as well as parties with at least 1% vote share for some countries. Of all political science datasets, it currently provides the largest coverage of political parties worldwide, with over 4,500 parties and URLs to the Wikipedia pages of about 3,900 of these (see Table 1).

We collect all ideology and lr-position tags for parties that have a Wikipedia URL in Party Facts and an infobox on Wikipedia. Data collection took place on April 27, 2019. We consider only tags that are linked to an associated Wikipedia article and code a party as tagged with the ideology or lr-position to which the link refers (see Online Appendix E for further details).

11 The result of ter Braak (1985) also applies if the observed data are frequency counts rather than presences or absences. Lowe (2008) introduces CA to the literature on political text scaling by showing its similarity to the Wordscores method, as well as discussing its properties as an estimator of word and document scores.

12 That said, CA estimates can never be equal to ML estimates and they show some well-known biases. In one dimension, the CA solution compresses the ends of the scale, pulling estimates of x and o that lie at the boundaries toward the middle of the scale (ter Braak 1985, 864). This defect does not take away the usefulness of CA in identifying the underlying dimension, but it highlights the added value from estimating the full parametric model in Equation (1) (on this point, see Lowe 2008, 369).

To get some insight into the variety of tags and their usage, [Figures 2 and 3](#) give a breakdown of the raw numbers. The first thing to note is that tags are used sparingly. As [Figure 2](#) shows, Wikipedia editors typically use one to four tags to describe party ideology; only a small fraction of parties show more than seven ideology tags. For lr-position tags, the modal frequency is one, but a considerable number of parties also receive two such tags. A closer inspection of these latter cases reveals that editors almost always assign adjacent lr-position tags to parties (e.g., centre-right and right-wing).

In addition to being used sparingly, tags are also used unequally: the 20 most often-used ideology tags, which make up less than 5% of all observed tags, account for over 50% of total tag usage. Among this core of widely applied tags, we find well-known, major ideologies such as liberalism, conservatism, and social democracy (see [Figure 3](#)).

To reduce the potential for bias and inaccuracy in our data, we focus on tags that are widely used on Wikipedia, as such tags should get more exposure and, as a result, be better known and more easily scrutinized than tags referring to rare or obscure ideologies. To achieve this, we implement a threshold of 50-tag occurrences. Tags observed in fewer parties are omitted from the analysis. This restriction allows us to rule out narrow ideologies that depend on national context (e.g., Basque nationalism), as well as peculiar ones (e.g., eurocommunism) about which only a small number of editors probably know enough to be able to apply them adequately, and spot and correct mistakes.¹³ For parties, we choose an inclusive threshold of at least two tags. In sum, this gives us scalable information for more than 2,100 parties (see [Table 1](#)).

6 The Resulting Scale

We begin by inspecting the scale that we obtain for its face validity. [Figures 4 and 5](#) summarize the estimation result via the estimated response curves and party positions for Models 1 and 2. Each response curve indicates the probability that a particular tag is assigned to a party with a given position on the underlying dimension. For example, a party at position 0 in [Figure 5](#) has an estimated probability of about 60% of being tagged with liberalism, about 30% probability of being tagged with social liberalism, about 20% probability of being tagged with pro Europeanism, and so on.¹⁴ The use of politically informative keywords thus lends a substantive interpretation to areas of the scale based on the ideologies that are particularly prominent there. In turn, this is reflected in party positions.

[Figures 4 and 5](#) generally support the idea of a global left–right dimension underlying Wikipedia classifications of party ideology. Ideologies that are traditionally associated with the labels “left” and “right” (e.g., communism, socialism, social democracy, liberalism, conservatism, and national conservatism) line up in the familiar order. These ideologies’ response curves are also steep in the sense that their probability of being assigned to a party rises and falls quickly along the underlying spectrum. The ideologies thus help discriminate parties on the underlying dimension. By contrast, ideologies such as agrarianism or regionalism have flat response curves and thus contribute less information on where to place parties on the underlying dimension. The fact that major ideologies line up in the familiar order suggests that party position estimates can be interpreted in terms of left versus right. In other words, the scale has face validity.

We emphasize that we obtain substantively the same result, regardless of whether we include lr-position tags or not (compare [Figures 4 and 5](#)). Ideology tags alone are thus sufficient to obtain a meaningful left–right scale. However, the fact that the assignment of ideology tags follows a familiar left–right reasoning suggests that Wikipedia’s lr-position tags can be used to tighten our

- ¹³ We employ the 50-tag threshold to be conservative. Our results remain substantively and quantitatively the same if we include tags that are used fewer than 50 times.
- ¹⁴ Point estimates for the locations of all ideologies (i.e., the locations where the probability curves peak) along with credible intervals are given in Online Appendix F.

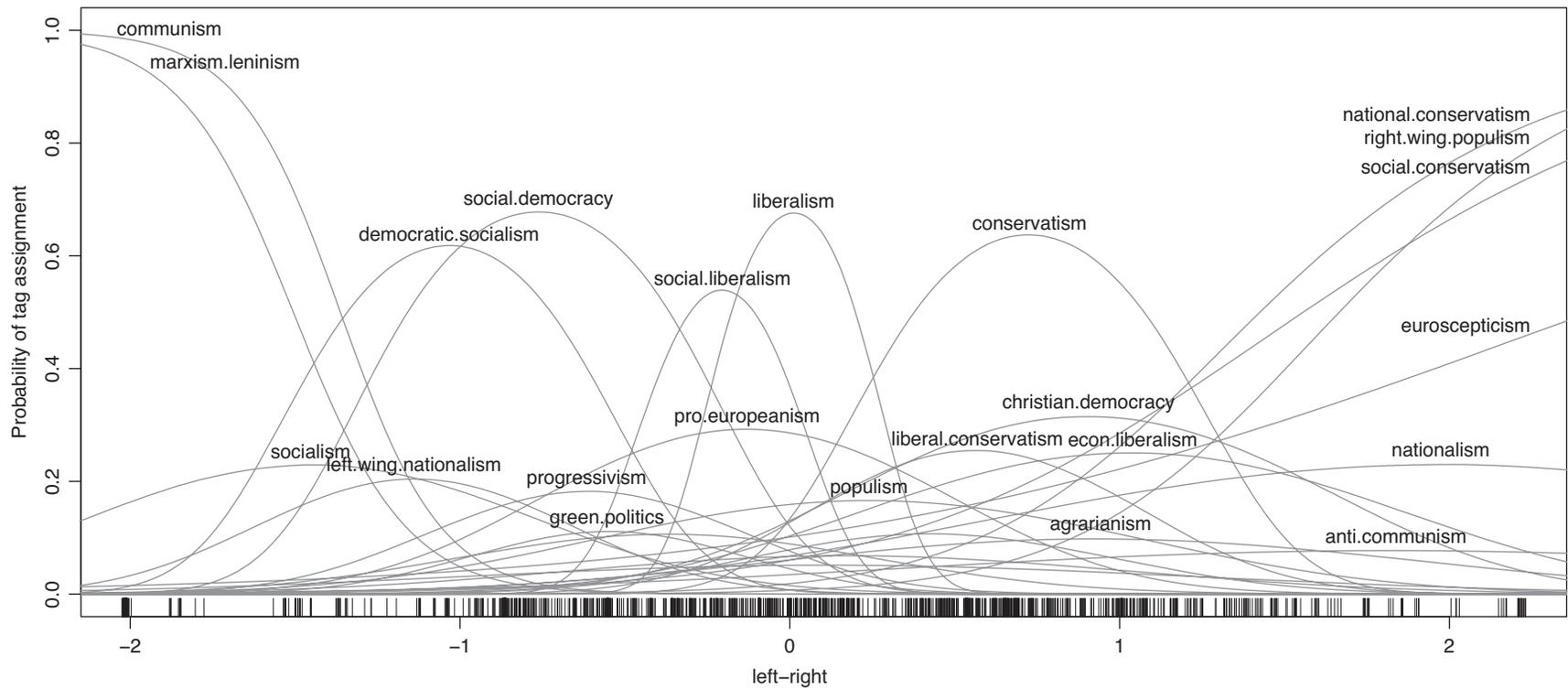


Figure 4. Response curves and estimated party positions (indicated by tick marks) from a scaling of ideology tags only; 1,367 parties and 27 tags (some response curves are unlabeled to avoid clutter).

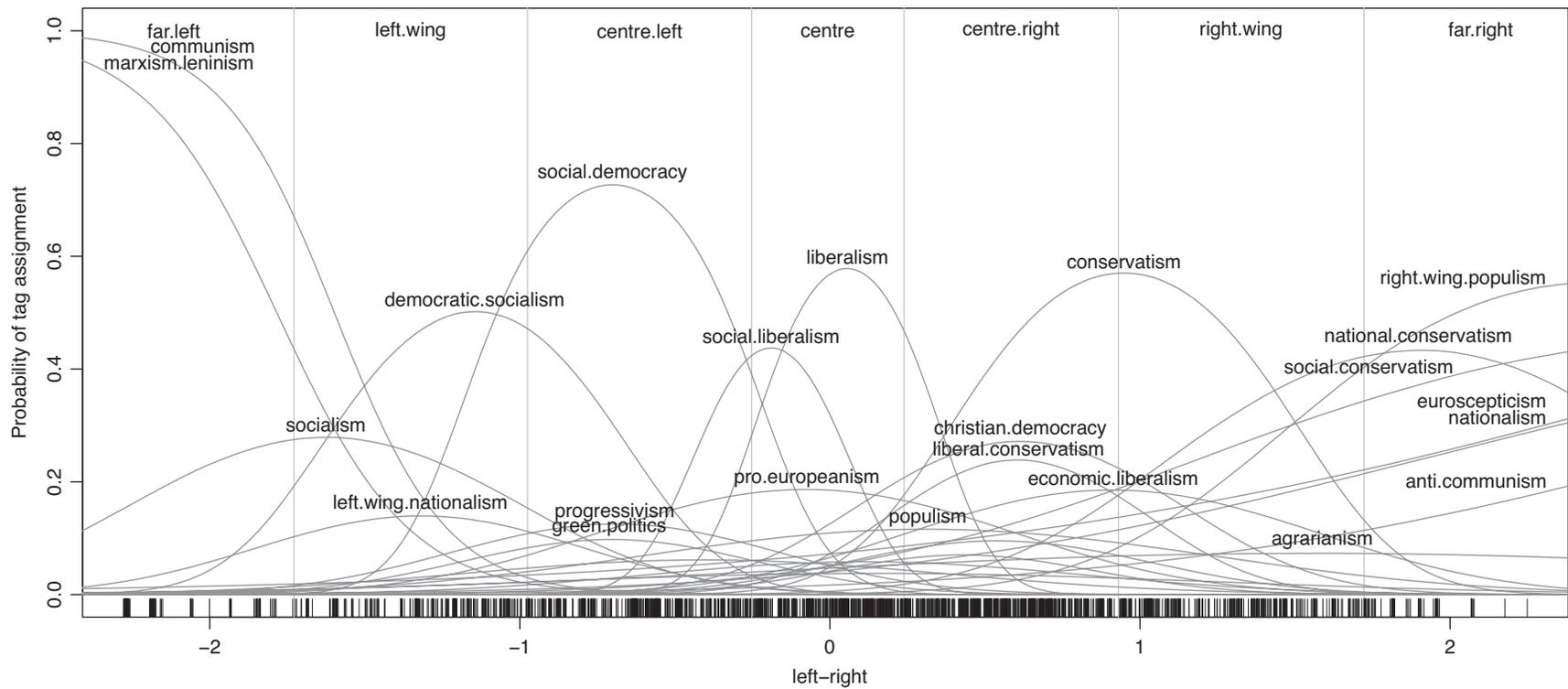


Figure 5. Response curves for ideology tags, estimated intervals for lr-position tags, and estimated party positions (indicated by tick marks) from a joint scaling of ideology and lr-position tags; 2,147 parties and 35 tags (some response curves are unlabeled to avoid clutter).

inferences about the underlying scale (besides allowing us to include even more parties in the estimation) by construing them as ordered indicators of adjacent intervals on the underlying dimension—the assumption that gives rise to Model 2.

As [Figure 5](#) shows, the inclusion of lr-position tags leaves the ordering of ideologies largely unchanged but helps distinguish some ideologies more clearly from one another. For example, social democracy and democratic socialism are now separated more clearly, with the former falling firmly into the centre-left bracket and the latter into the left-wing bracket. Likewise, Christian democracy now clearly peaks left of conservatism and within the centre-right bracket. Lastly, the inclusion of lr-position tags reduces the probability of tag assignment for most ideology tags, especially for some on the far-right end (i.e., right-wing populism, national conservatism, etc.). This is due to the higher frequency with which lr-position tags are observed compared to most ideology tags, as well as their patterns of co-occurrence. In particular, the rightmost ideology tags often occur in conjunction with the far-right tag as well as the right-wing tag. This leads to more equal probabilities of far-right and right-wing parties being tagged with one of these ideologies.

7 Validation with Expert Surveys

Having established our estimates' face validity, we now consider their convergent and discriminant validity (cf. [Adcock and Collier 2001](#)). Accordingly, a valid measure should correlate highly with other measures of the same construct, while showing lower correlations with measures of related but different constructs.

To test this, we draw on expert ratings of party left–right position as well as party position on other dimensions from the two largest existing expert surveys, the DALP ([Kitschelt 2014](#)) and the Chapel Hill Expert Survey (CHES) ([Polk et al. 2017](#)). The DALP covers 506 parties from 88 electoral democracies worldwide. Ratings were collected between 2008 and 2009 and may not be entirely accurate with respect to parties that have recently altered their position significantly. Nevertheless, no other expert survey provides broader country coverage. From the CHES, we use the most recent set of ratings, which were collected in 2014 and cover 268 parties from 31 European democracies.

Regarding convergent validity, [Figure 6](#) shows that our estimates correlate well with expert ratings of party left–right position. Estimates from Model 2 generally fit expert ratings more closely than those of Model 1, suggesting that the inclusion of lr-position tags adds useful information above and beyond ideology tags. [Table 2](#) further shows how the correlation between our scores and expert ratings varies over countries. Since expert surveys ask for placements of parties only within a given polity (i.e., each expert rates parties from his or her country), pooling of ratings across countries may introduce measurement error due to differential item functioning ([Hare et al. 2014](#); [Struthers, Hare, and Bakker 2019](#)).¹⁵ Comparing estimates on a country-by-country basis, the median correlation is above 0.9 and the 25th percentile is above 0.75 in all four comparisons. Thus, for three-quarters of the countries covered by either the DALP or the CHES, our estimates of party positions accord well—and for half the countries covered, they accord very well—with the judgment of country experts.

Regarding discriminant validity, we compare our estimates to expert ratings of party position on economic policy, redistribution, GAL–TAN, and identity politics. As [Table 3](#) (first column) shows, experts' perceptions of party left–right position are substantially correlated with their perceptions of party stances on all other measured dimensions. This suggests that each of the more specific dimensions reflects some aspect of left versus right. Reassuringly, our estimates also correlate with every other dimension measured in expert surveys. However, the correlations remain weaker

15 For the CHES, [Bakker et al. \(2014\)](#) show that differential item functioning in experts' ratings of parties' economic left–right position is small.

Table 2. Country-wise correlations with expert ratings (means and percentiles).

		Mean	10%	25%	50%	75%	90%
DALP	Model 1	0.72	0.26	0.77	0.91	0.97	1.00
	Model 2	0.74	0.42	0.76	0.90	0.97	0.99
CHES	Model 1	0.83	0.66	0.79	0.90	0.96	0.98
	Model 2	0.90	0.75	0.90	0.94	0.97	0.99

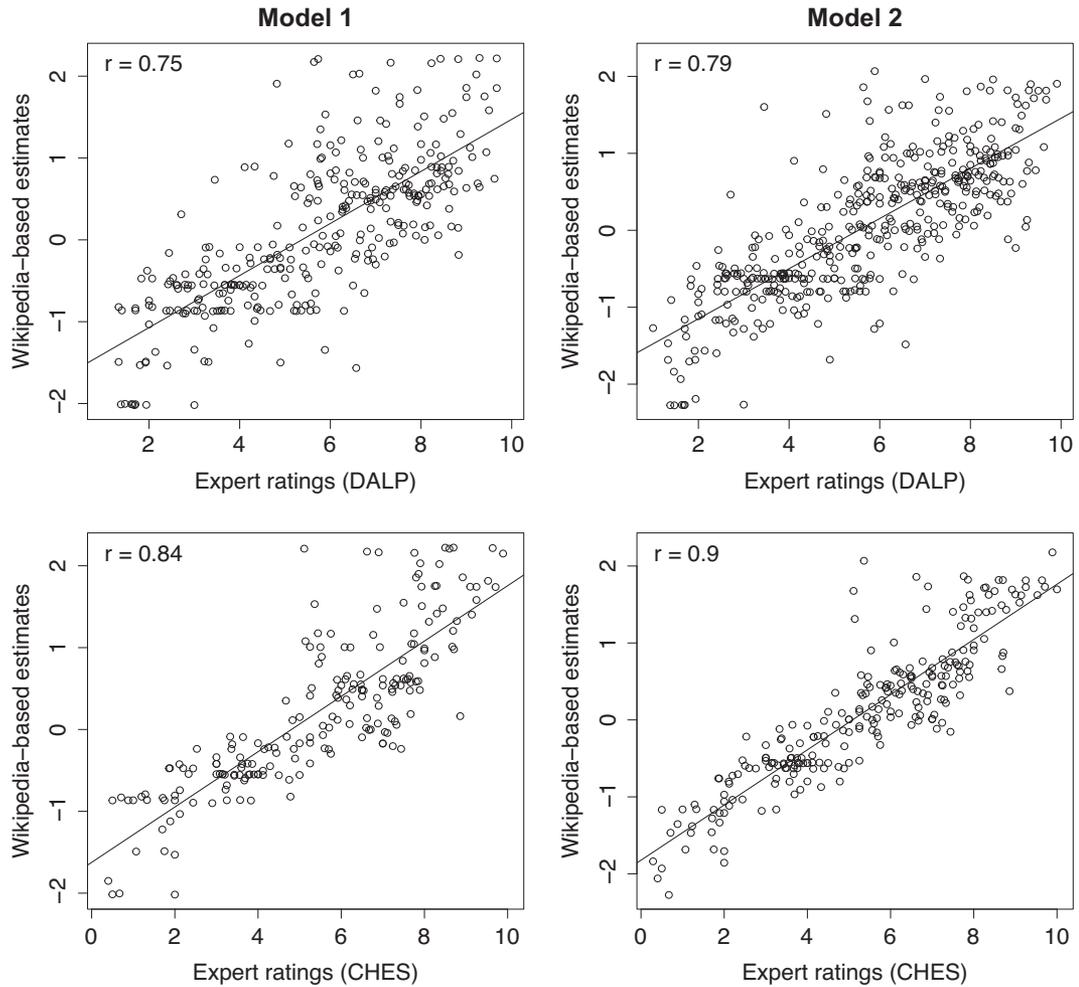


Figure 6. Comparison of party position estimates to expert ratings from the Democratic Accountability and Linkages Project (top row) and from the 2014 Chapel Hill Expert Survey (bottom row). Model 1: $N = 321$ and $N = 203$; Model 2: $N = 435$ and $N = 247$.

than those with left-right, suggesting that our estimates most likely represent general left-right position, as opposed to more specific policy stances.

Closer inspection of Table 3 reveals further nuances. While experts' judgments of party left-right position relate most strongly to their perceptions of parties' stances on economic or redistributive policy, Wikipedia-based estimates are more strongly associated with expert ratings of parties' positions on GAL-TAN or identity politics than on economic policy. With the inclusion of Ir-position tags, correlations increase somewhat between our estimates and ratings of parties' economic and redistributive stances (compare Model 1 and Model 2), but, overall, our scores remain slightly more associated with social and identity politics. Although small, these differences suggests that

Table 3. Correlations with expert ratings on other dimensions.

		Expert L-R	Model 1	Model 2
DALP	Left-right	1.00	0.75	0.79
	Economy	0.72	0.47	0.53
	Redistribution	0.67	0.44	0.49
	GAL-TAN	0.54	0.53	0.54
	Identity	0.65	0.65	0.65
CHES	Left-right	1.00	0.84	0.90
	Economy	0.81	0.53	0.63
	Spend-Tax	0.77	0.51	0.61
	GAL-TAN	0.70	0.77	0.75

Wikipedia editors' understanding of left versus right is driven more by value considerations than by economic considerations, whereas experts tend to view left versus right somewhat more in terms of economic policy and redistribution than social policy.

8 Reliability

As content on Wikipedia is open to instant updating and revision by anyone, the raw data are constantly evolving. Any day, new tags might get added to a party's infobox, others deleted. This raises the question of how robust our estimation is to Wikipedia's openness: Can we rely on estimates of party positions from a single point in time, when the raw data are subject to constant revision (i.e., additions, refinements, mistakes, and corrections of tags)?

To gauge this uncertainty, we estimate party positions using ideology and lr-position tags observed at different points in time. Specifically, we compare our current scores to scores obtained in the same way using data collected some 15 weeks earlier (on January 11, 2019). The assumption is that during this short period of time, the underlying party positions remain the same. Changes in tags will be due to editors adding new classifications or modifying what they see as incorrect classifications, either through deletion or replacement. Given Wikipedia's 500 million monthly visitors, there should be ample opportunity for observing such changes in the given time period. By comparing the positions of parties that experience a change in their tags, we gain some insight on the sensitivity of our measure to ongoing editing activity. A reliable measure should yield similar position estimates despite overt variation in tags.

Across the 15-week period considered, we find that 111 and 279 of the parties included in the estimation of Models 1 and 2 see a change in their tags. Comparing those parties' estimated scores, we find a close correspondence between them (Figure 7). The average absolute differences between scores are only 0.20 and 0.17, respectively. If we take a difference of 0.5 as the minimum for a substantial change—a difference of 0.5 roughly corresponds to a move from one left-right bracket to the next (see Figure 5)—we find that 10 and 13 parties, respectively, experience such a change in their positions (see Online Appendix H for a detailed inspection of these cases). The strong correlation between both sets of scores indicates a high test-retest reliability.

The fact that changes in tags lead to small variation in party positions suggests that the observed edits are mostly refinements rather than sweeping revisions of parties' infobox classifications.¹⁶ Perhaps this result should not be too surprising. After all, if there was little stability in parties' infobox tags, we would not expect a scaling of tags collected at an arbitrary point in

¹⁶ For example, adding the tag liberal conservatism to a party tagged with liberalism and conservatism would be a refinement that does not alter the party's estimated position to any significant degree.

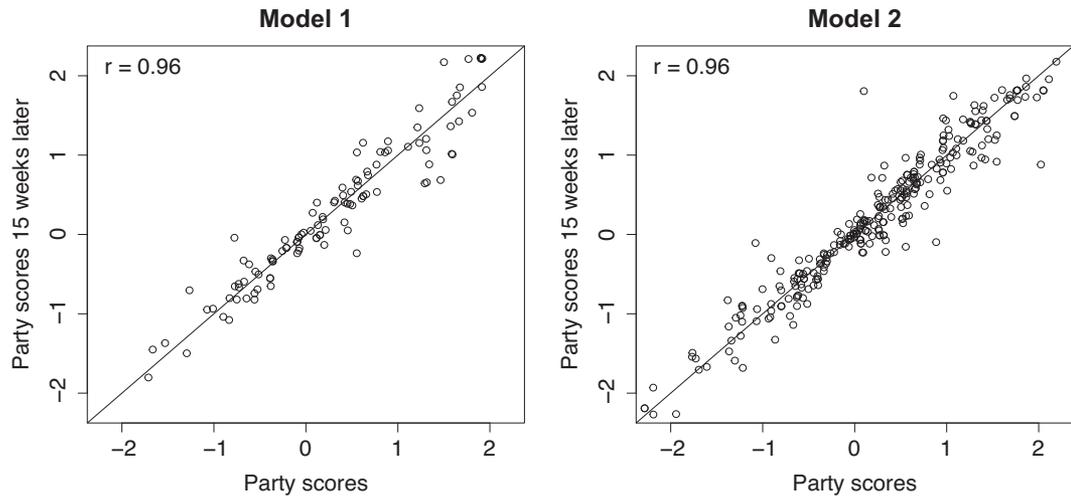


Figure 7. Reliability of party position estimates. The x-axis shows party position estimates obtained from Wikipedia classifications collected 15 weeks prior to those on which our current estimates are based, which are shown on the y-axis. A 90° line is superimposed. Model 1: $N = 111$; Model 2: $N = 279$.

time to yield a strong and coherent left–right dimension. Nonetheless, the result is reassuring. Monitoring the evolution of party scores in the long run will, of course, be necessary for maintaining confidence in the measure.

9 Generalizing Beyond Expert Surveys

Parties differ greatly in how much attention they receive by the public, with some parties being much better known than others. This could affect how much attention and scrutiny they receive on Wikipedia. Prominent parties are likely to attract greater interest from Wikipedia editors, and with more resources going into the creation and monitoring of those parties' articles, their tags are likely to be more accurate than those of less well-known parties.¹⁷ Our above comparisons with expert ratings might thus represent a best-case scenario, as expert surveys tend to include the most important parties in a polity. For parties not included in expert surveys, the accuracy of our scores might be lower.

To assess this possibility, we formulate two hypotheses. Hypothesis one is that greater scrutiny of a party's Wikipedia article improves the accuracy of its estimated left–right score. Hypothesis two is that articles of parties included in expert surveys are scrutinized more than those of other parties. If both hypotheses hold true, we can expect our scores to be less accurate for parties not included in expert surveys, and more so the stronger the observed relationships are. We use two indicators for an article's level of scrutiny (see Online Appendix I for details): its number of editors; and the (average) experience of its 50 most active editors, where experience is defined as an editor's total number of live edits on Wikipedia. For comparison, we also include two indicators of party prominence in the analysis: article pageviews and vote share as recorded in Party Facts.

To assess the first hypothesis, we study the closeness of fit between our scores and the expert ratings shown in Figure 6. If more editors and greater editor experience improve the accuracy of information on a party's Wikipedia page, one would expect the party's estimated position to correspond more closely to the experts' views and hence be closer to the regression line than that of other parties. We employ a linear regression model with variance heterogeneity (Verbyla 1993; King 1998) to examine this conjecture (see Online Appendix J for a description of the model). This approach allows us to model the residual variation around the regression line as a function

¹⁷ Another possibility for tags of less well-known parties to be less accurate is that less information is available about their ideology. This is a generic problem that arises with any measure of party position, including expert surveys.

Table 4. Predictors of party position estimates (DALP: $N = 430$; CHES: $N = 244$).

		DALP		CHES	
		Coef	SE	Coef	SE
Mean	Constant	-1.815	0.076	-1.843	0.065
	Left-right expert	0.330	0.013	0.360	0.011
Variance	Constant	0.063	0.313	-0.053	0.372
	Vote share	0.165	0.167	-0.419	0.383
	Pageviews (log)	-0.001	0.031	0.021	0.040
	Number of editors (log)	-0.023	0.035	-0.041	0.038
	Editor experience (log)	-0.084	0.031	-0.107	0.034

Table 5. Comparing parties included in expert surveys ($N = 610$) to other parties on Wikipedia ($N = 3,275$).

	In survey	Mean	SD	Min	25%	75%	Max
Vote share	Yes	0.21	0.19	0.00	0.06	0.31	1.00
Vote share	No	0.17	0.22	0.00	0.03	0.21	1.00
Pageviews (log)	Yes	7.03	2.20	0.00	5.84	8.46	12.64
Pageviews (log)	No	5.49	1.83	0.00	4.30	6.69	11.91
Editors (log)	Yes	4.27	2.05	0.00	3.50	5.63	8.51
Editors (log)	No	3.09	1.59	0.00	2.20	4.14	7.98
Editor experience (log)	Yes	7.25	1.55	3.50	6.21	7.97	12.43
Editor experience (log)	No	8.10	1.63	3.09	6.92	9.16	12.63

of additional covariates, thus providing a natural way to study correlates of inaccuracy in party scores.

We find that more editors and greater editor experience are associated with a tighter fit between our scores and both the DALP and the CHES ratings (see Table 4). By contrast, pageviews and vote share are not consistently associated with greater accuracy of party scores (the correlation between pageviews and the number of editors is 0.79). Similar results hold if we test each predictor in isolation. In both models, editor experience shows the strongest effect: the standard error of the regression decreases by 8% (i.e., by a factor of $\exp(-0.084)$) and by 10%, respectively, for a one-unit increase in (the log of) editor experience. Given the observed ranges and variances on all the variables (see Table 5), editor experience is thus associated with the biggest reduction in regression error.

To assess the second hypothesis, Table 5 compares parties covered either by the CHES or the DALP surveys to all other parties that have a Wikipedia page. Consistent with the fact that expert surveys deliberately select the most important parties in a polity, we find that the parties included in expert surveys tend to have a higher vote share and receive more pageviews, on average, than other parties on Wikipedia. Consistent with the notion of greater interest, these parties also attract more editors, on average, than other parties on Wikipedia. For editor experience, we find the opposite pattern with parties not included in expert surveys having more seasoned article editors, on average. Apart from these differences, we find that the two samples are surprisingly similar in terms of variation and spread. In fact, the samples overlap to a large extent on each of the indicators.

The results suggests that both groups of parties have something speaking for them. Parties included in expert surveys tend to attract more editors than other parties on Wikipedia. Yet, their editors tend to be less experienced. Since editor experience contributes to the accuracy of tag

assignment, Wikipedia-based scores need not be less accurate for parties not covered by expert surveys. For example, we would expect in-survey parties' scores to be 5% more accurate, on average, based on their higher number of editors, but we would expect out-of-survey parties' scores to be 9% more accurate, on average, based on their editors' greater experience.¹⁸ The observed correlation between a party's number of editors and the average experience of its editors is about -0.66 , so a party with fewer editors tends to have more experienced editors and vice versa. This as well as the large overlap between the two samples on each of the relevant indicators gives us reason to expect a Wikipedia-based measure not to perform worse (than expert surveys would) in estimating positions for parties for which we currently lack such estimates.

10 Conclusion and Outlook

Party positions are central to explanations of politics. In this paper, we entertained the hypothesis that semistandardized classifications of party ideology provided on Wikipedia carry valid and reliable information about party position. We stipulated an ideal point model that allows us to exploit variation in ideological keywords across parties and obtained scores of party position on a latent dimension whose substantive interpretation fits common intuitions of left versus right. In line with our hypothesis, we found party scores to correlate well with independent expert judgments of party left–right position and we found them to be reliable in a test–retest scenario. This demonstrates that our approach yields a novel measure of party left–right position.

As a proof of concept, our results hold great promise for future research. Party ideology and the notion of ideological differences between parties are key causal factors in explanations of the political process. They feature prominently in theories of political change and stability, legislative decision making, coalition politics, public policy, economic growth, inequality and redistribution, as well as in accounts of democratization, state building, and violent conflict. Conversely, parties' ideological positions are the object of study in many research areas, from electoral competition to the quality of representation to political fragmentation and polarization. With its potentially unlimited coverage, our measure of party position opens up new possibilities for researchers to study these and other topics on a much larger universe of cases. This includes many parties for which we currently lack ideological placements, as well as new countries not covered by extant data sources.

Our measure is not limited to parties that currently exist. Unlike many expert surveys, our scaling includes parties of the past, provided they have a Wikipedia article, and provided Wikipedia editors assign to these parties some of the same tags that they assign to other parties. In addition, our Wikipedia-based approach may provide estimates for parties founded just recently, which expert surveys may not (yet) cover. One database that does include past parties and also regularly updates its sample is the Manifesto Project. Compared to manifesto-based measures of party position, our measure does not depend on the existence of an electoral manifesto. This allows for the inclusion of parties that do not produce such documents, including nascent parties, nondemocratic parties, or parties in regimes with restricted electoral competition.

Acknowledgments

The authors would like to thank panelists at the 2019 Annual MPSA Conference, and participants in research colloquia at the Universities of Bremen, Konstanz, and Mannheim for their feedback, especially Christopher Hare, Christopher Wratil, Samuel Baltz, Markus Tepe, Dominic Nyhuis, Sascha Göbel, Thomas Malang, Benjamin Guinaudeau, Paul Bederke, Susumu Shikano, Konstantin Käppner, Christian Breunig, Thomas Gschwend, and Hans-Dieter Klingemann.

¹⁸ This calculation uses the coefficient estimates from the CHES model in [Table 4](#) and the differences in sample means from [Table 5](#).

Data Availability Statement

All replication materials are available on Dataverse (Herrmann and Döring 2021).

Supplementary Material

For supplementary material accompanying this paper, please visit <https://doi.org/10.1017/pan.2021.28>.

References

- Adcock, R., and D. Collier. 2001. "Measurement Validity: A Shared Standard for Qualitative and Quantitative Research." *American Political Science Review* 95(3):529–546.
- Albert, J. H., and S. Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88(422):669–679.
- Albert, J. H., and V. E. Johnson. 1999. *Ordinal Data Modeling*. New York: Springer.
- Andrich, D. 1988. "The Application of an Unfolding Model of the PIRT Type to the Measurement of Attitude." *Applied Psychological Measurement* 12(1):33–51.
- Bafumi, J., A. Gelman, D. K. Park, and N. Kaplan. 2005. "Practical Issues in Implementing and Understanding Bayesian Ideal Point Estimation." *Political Analysis* 13:171–187.
- Bakker, R., S. Jolly, J. Polk, and K. Poole. 2014. "The European Common Space: Extending the Use of Anchoring Vignettes." *Journal of Politics* 76(4):1089–1101.
- Bederke, P., H. Döring, and S. Regel. 2019. "Party Facts—Version 2019a." <https://doi.org/10.7910/DVN/FTQAYT>, Harvard Dataverse, V1, UNF:6:tC+wuU8rgtphyPEAUNxkIQ== [fileUNF].
- Benoit, K., D. Conway, B. E. Lauderdale, M. Laver, and S. Mikhaylov. 2016. "Crowd-Sourced Text Analysis: Reproducible and Agile Production of Political Data." *American Political Science Review* 110(2):278–295.
- Bonica, A. 2014. "Mapping the Ideological Marketplace." *American Journal of Political Science* 58(2): 367–386.
- Bräuninger, T., J. Müller, and C. Stecker. 2016. "Modeling Preferences Using Roll Call Votes in Parliamentary Systems." *Political Analysis* 24(2):189–210.
- Brown, A. R. 2011. "Wikipedia as a Data Source for Political Scientists: Accuracy and Completeness of Coverage." *PS: Political Science and Politics* 44(2):339–343.
- Caughey, D., and C. Warshaw. 2015. "Dynamic Estimation of Latent Opinion Using a Hierarchical Group-Level IRT Model." *Political Analysis* 23(2):197–211.
- Clinton, J. D., and S. Jackman. 2009. "To Simulate or NOMINATE?" *Legislative Studies Quarterly* 34(4):593–621.
- Däubler, T., and K. Benoit. 2017. "Estimating Better Left-Right Positions Through Statistical Scaling of Manual Content Analysis." *Working Paper*, University of Mannheim.
- Döring, H., and S. Regel. 2019. "Party Facts: A Database of Political Parties Worldwide." *Party Politics* 25(2):97–109.
- Göbel, S., and S. Munzert. 2021. "The Comparative Legislators Database." *British Journal of Political Science*. doi: [10.1017/S0007123420000897](https://doi.org/10.1017/S0007123420000897).
- Greenacre, M. J. 2010. "Correspondence Analysis." *Wiley Interdisciplinary Reviews: Computational Statistics* 2(5):613–619.
- Hare, C., D. A. Armstrong, II, R. Bakker, R. Carroll, and K. T. Poole. 2014. "Using Bayesian Aldrich–McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59(3):759–774.
- Haselmayer, M., and M. Jenny. 2017. "Sentiment Analysis of Political Communication: Combining a Dictionary Approach with Crowdcoding." *Quality & Quantity* 51(6):2623–2646.
- Herrmann, M., and H. Döring. 2021. "Replication Data for: Party Positions from Wikipedia Classifications of Party Ideology." <https://doi.org/10.7910/DVN/1JHZIU>, Harvard Dataverse, V1.
- Horn, A. 2019. "Can the Online Crowd Match Real Expert Judgments? How Task Complexity and Coder Location Affect the Validity of Crowd-Coded Data." *European Journal of Political Research* 58(1):236–247.
- King, G. 1998. *Unifying Political Methodology: The Likelihood Theory of Statistical Inference*. Ann Arbor, MI: Michigan University Press.
- Kitschelt, H. 2014. *Dataset of the Democratic Accountability and Linkages Project (DALP)*. Durham, NC: Duke University.
- Lauderdale, B. E., and A. Herzog. 2016. "Measuring Political Positions from Legislative Speech." *Political Analysis* 24:374–394.
- Laver, M., K. Benoit, and J. Garry. 2003. "Extracting Policy Positions from Political Texts Using Words as Data." *American Political Science Review* 97(2):311–331.
- Lehmann, P., and M. Zobel. 2018. "Positions and Saliency of Immigration in Party Manifestos: A Novel Dataset Using Crowd Coding." *European Journal of Political Research* 57(4):1056–1083.

- Levine, J. H. 1979. "Joint-Space Analysis of "Pick-Any" Data: Analysis of Choices from an Unconstrained Set of Alternatives." *Psychometrika* 44(1):85–92.
- Lo, J., S. O. Proksch, and T. Gschwend. 2014. "A Common Left–Right Scale for Voters and Parties in Europe." *Political Analysis* 22:205–223.
- Lowe, W. 2008. "Understanding Wordscores." *Political Analysis* 16:356–371.
- Mesgari, M., C. Okoli, M. Mehdi, F. A. Nielsen, and A. Lanamäki. 2014. "'The Sum of all Human Knowledge': A Systematic Review of Scholarly Research on the Content of Wikipedia." *Journal of the Association for Information Science and Technology* 66(2):219–245.
- Munzert, S.. 2018. "Measuring the Importance of Political Elites." SocArXiv. doi:10.31235/osf.io/t8gs5.
- Peterson, A., and A. Spirling. 2018. "Classification Accuracy as a Substantive Quantity of Interest: Measuring Polarization in Westminster Systems." *Political Analysis* 26:120–128.
- Plummer, M.. 2017. *JAGS: A Program for Analysis of Bayesian Graphical Models Using Gibbs Sampling*. Version 4.3.0.
- Polk, J., 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 and Politics*. doi: 10.1177/2053168016686915.
- Poschmann, P., and J. Goldenstein. 2019. "Disambiguating and Specifying Social Actors in Big Data: Using Wikipedia as a Data Source for Demographic Information." *Sociological Methods and Research*. doi: 10.1177/0049124119882481.
- Rivers, D. 2003. "Identification of Multidimensional Spatial Voting Models." Typescript, Stanford University.
- Samejima, F. 1969. *Estimation of Latent Ability Using a Response Pattern of Graded Scores*. Psychometrika Monograph Supplements, vol. 17. Richmond, VA: Psychometric Society.
- Slapin, J. B., and S.-O. Proksch. 2008. "A Scaling Model for Estimating Time-Series Party Positions from Texts." *American Journal of Political Science* 52(3):705–722.
- Smith, K. Y., and F. D. Neiman. 2007. "Frequency Seriation, Correspondence Analysis, and Woodland Period Ceramic Assemblage Variation in the Deep South." *Southeastern Archeology* 26(1):47–72.
- Struthers, C. L., C. Hare, and R. Bakker. 2019. "Bridging the Pond: Measuring Policy Positions in the United States and Europe." *Political Science Research and Methods*. doi: 10.1017/psrm.2019.22.
- Sumner, J. L., E. M. Farris, and M. R. Holman. 2020. "Crowdsourcing Reliable Local Data." *Political Analysis* 28(2):244–262.
- Temporão, M., C. V. Kerckhove, C. van der Linden, and J. M. Hendrickx. 2018. "Ideological Scaling of Social Media Users: A Dynamic Lexicon Approach." *Political Analysis* 26:457–473.
- ter Braak, C. J. F. 1985. "Correspondence Analysis of Incidence and Abundance Data: Properties in Terms of a Unimodal Response Model." *Biometrics* 41(4):859–873.
- ter Braak, C. J. F., and P. Šmilauer. 2015. "Topics in Constrained and Unconstrained Ordination." *Plant Ecology* 216:683–696.
- Treier, S., and S. Jackman. 2008. "Democracy as a Latent Variable." *American Journal of Political Science* 52(1):201–217.
- Verbyla, A. P. 1993. "Modelling Variance Heterogeneity: Residual Maximum Likelihood and Diagnostics." *Journal of the Royal Statistical Society. Series B (Methodological)* 55(2):493–508.
- Volkens, A., J. Bara, I. Budge, M. McDonald, and H.-D. Klingemann. 2013. *Mapping Policy Preferences from Texts III: Statistical Solutions for Manifesto Analysts*. Oxford: Oxford University Press.
- Winter, N. J. G., A. G. Hughes, and L. M. Sanders. 2020. "Online Coders, Open Codebooks: New Opportunities for Content Analysis of Political Communication." *Political Science Research and Methods* 8(4):731–746.