Online Appendix

Party Positions from Wikipedia Classifications of Party Ideology

Michael Herrmann

University of Konstanz

Holger Döring

University of Bremen

A Connections to other scaling models

Our model of ideology tag assignment is related to extant models for word counts in party manifestos through the following generalized linear model

$$\text{link}(E(y_{ij})) = a_j + b_j x_i + c_j x_i^2, \tag{4}$$

where x is an unobserved latent variable, a, b, c are unknown parameters, and y is a matrix of either

1. word (j) presences and absences in party infoboxes (i), or
2. word (j) counts in party manifestos (i).

Case 1 yields our model of tag assignment: Since, for binary data, $E(y_{ij}) = \Pr(y_{ij} = 1)$, our model as stated in Eq. (3) is equivalent to Eq. (4) with a logit link function and the constraint $c_j < 0$.

Case 2 yields the model in Lowe (2008) and the Wordfish approach: Applying a log link function and constraining $c_j < 0$, Eq. (4) becomes Eq. (12) in Lowe (2008, 366). As shown there, the Wordfish model results from replacing $c_j < 0$ with $c_j = 0$.

B Identification

The identifying restrictions stated in section 4.1 imply the following offsetting transformations on the other parameters. Let the unidentified parameters be denoted with a star and the identified parameters without. Consider first Eq. (1). For identification to leave outcome probabilities unchanged, we require that

$$\alpha_j - \beta_j (o_j - x_i)^2 = \alpha_j^* - \beta_j^* (o_j^* - x_i^*)^2.$$

Substituting the identifying restrictions $x_i = (x_i^* - \bar{x}^*)/\mathrm{SD}(x^*)$ and $o_i = (o_i^* - \bar{x}^*)/\mathrm{SD}(x^*)$, the left-hand side of the above relation becomes

$$\alpha_j - \beta_j (o_j^* - x_i^*)^2/\mathrm{SD}(x^*)^2,$$

from which we deduce that $\beta_j = \beta_j^* \mathrm{SD}(x^*)^2$ and $\alpha_j = \alpha_j^*$ are the offsetting transformations required for identification.

Consider now Eq. (2). For outcome probabilities to remain unchanged, we require that

$$\tau_k - \gamma x_i = \tau_k^* - \gamma^* x_i^*,$$

which, after imposing the identifying restriction $x_i = (x_i^* - \bar{x}^*)/\mathrm{SD}(x^*)$, implies the offsetting transformations $\gamma = \gamma^* \mathrm{SD}(x^*)$ and $\tau_k = \tau_k^* + \gamma^* \bar{x}^*$.

C JAGS code

The input data for Model 1 are `y`, a binary matrix with parties *i* in rows and ideology tags *j* in columns, `D`, a vector containing the dimensions of `y`, and `con`, the column number of the tag conservatism.

```
model {

  # likelihood
  for (i in 1:D[1]) {
    for (j in 1:D[2]) {
      logit(pi[i, j]) <- delta[j] + lambda[j] * xstar[i] - beta[j] * xstar[i] ^ 2
      y[i, j] ~ dbern(pi[i, j])
    }
  }

  # priors
  for (i in 1:D[1]) {
    xstar[i] ~ dnorm(0, 1)
  }
  for (j in 1:D[2]) {
    delta[j] ~ dnorm(0, 0.2)
    lambda[j] ~ dnorm(0, 0.2)
    beta[j] ~ dlnorm(0, 0.5)
  }

  # transformation and identification of posterior draws
  xbar <- mean(xstar)
  sdx <- sd(xstar)
  ocon <- lambda[con] / (2 * beta[con])
  polarity <- ifelse(xbar < ocon, 1, -1)
  for (i in 1:D[1]) {
    x[i] <- polarity * (xstar[i] - xbar) / sdx
  }
  for (j in 1:D[2]) {
    a[j] <- delta[j] + beta[j] * (lambda[j] / (2 * beta[j])) ^ 2
    b[j] <- beta[j] * sdx ^ 2
    o[j] <- polarity * ((lambda[j] / (2 * beta[j])) - xbar) / sdx
  }
}
```

The primary input data for Model 2 are `y_ideo`, a binary data matrix with parties *i* in rows and ideology tags *j* in columns, and `y_lr`, a matrix with two columns. Column `y_lr[, 2]` is a stacked vector of lr-position tags (coded 1 through 7) that are observed for each party. The vector is stacked to allow for multiple observations of lr-position tags within parties. A corresponding index vector `y_lr[, 1]` identifies each party. Auxilliary inputs `D` and `con` are as stated in Model 1.

```
model {

  # likelihood
  for (i in 1:D[1]) {
    for (j in 1:D[2]) {
      logit(pi[i, j]) <- delta[j] + lambda[j] * xstar[i] - beta[j] * xstar[i] ^ 2
      y_ideo[i, j] ~ dbern(pi[i, j])
    }
  }
  for (n in 1:N) {
    mu[n] <- gamma * xstar[y_lr[n, 1]]
    y_lr[n, 2] ~ dordered.logit(mu[n], tstar[1:6])
  }

  # priors
  for (i in 1:D[1]) {
    xstar[i] ~ dnorm(0, 1)
  }
  for (j in 1:D[2]) {
    delta[j] ~ dnorm(0, 0.2)
    lambda[j] ~ dnorm(0, 0.2)
    beta[j] ~ dlnorm(0, 0.5)
  }
  for (k in 1:6) {
    tau[k] ~ dnorm(0, 0.04)
  }
  tstar[1:6] <- sort(tau)
  gamma ~ dnorm(0, 0.04)

  # transformation and identification of posterior draws
  xbar <- mean(xstar)
  sdx <- sd(xstar)
  ocon <- lambda[con] / (2 * beta[con])
  polarity <- ifelse(xbar < ocon, 1, -1)
  for (i in 1:D[1]) {
    x[i] <- polarity * (xstar[i] - xbar) / sdx
  }
  for (j in 1:D[2]) {
    a[j] <- delta[j] + beta[j] * (lambda[j] / (2 * beta[j])) ^ 2
```

```
      b[j] <- beta[j] * sdx ^ 2
      o[j] <- polarity * ((lambda[j] / (2 * beta[j])) - xbar) / sdx
    }
    for (k in 1:6) {
      t[k] <- tau[k] - gamma * xbar
    }
    g <- polarity * gamma * sdx
  }
```

D Assessing parameter convergence

We ran four chains for 20,000 iterations each, after a burn-in of 2,000 iterations. Keeping every 10th posterior draw from each chain, we obtain 8,000 posterior samples for each parameter. Given the large number of parameters, we assess potential non-convergence via scale reduction factors (see below), with additional inspection of traceplots for tag parameters (not shown).

Tables D.1 and D.2 give a summary of the scale reduction factors $\hat{R}$ for all parameters in Models 1 and 2. By inspection, there are no signs of non-convergence. Scale reduction factors are generally close to one and no value exceeds 1.1. Tables D.3 and D.4 give an indication of the efficiency of the MCMC samplers for each model. Effective samples sizes suggest that 8,000 posterior samples yield sufficient information to quantify parameter uncertainty at conventional levels, in particular for parameters x and o.

Table D.1: Scale reduction factors $\hat{R}$ for parameters of Model 1

	Min	Mean	Max	50%	75%	90%	95%	99%	Npar
α	1.00	1.00	1.01	1.00	1.01	1.01	1.01	1.01	27
β	1.00	1.01	1.02	1.00	1.01	1.01	1.02	1.02	27
o	1.00	1.00	1.01	1.00	1.00	1.01	1.01	1.01	27
x	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1,367

Table D.2: Scale reduction factors $\hat{R}$ for parameters of Model 2

	Min	Mean	Max	50%	75%	90%	95%	99%	Npar
γ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1
τ	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	6
α	1.00	1.01	1.08	1.00	1.01	1.03	1.04	1.07	28
β	1.00	1.00	1.03	1.00	1.00	1.02	1.02	1.02	28
o	1.00	1.01	1.08	1.00	1.02	1.03	1.04	1.07	28
x	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	2,147

Table D.3: Effective sample sizes for parameters of Model 1

	Min	Mean	Max	1%	5%	10%	25%	50%	Npar
α	333	3,583	8,000	339	359	443	978	2,440	27
β	196	2,120	8,000	206	243	267	501	1,638	27
o	795	4,524	8,000	822	923	1,141	2,164	3,729	27
x	890	6,366	8,000	1,546	2,242	2,903	4,668	8,000	1,367

Table D.4: Effective sample sizes for parameters of Model 2

	Min	Mean	Max	1%	5%	10%	25%	50%	Npar
γ	5,321	5,321	5,321	5,321	5,321	5,321	5,321	5,321	1
τ	2,856	5,350	8,000	2,875	2,950	3,044	3,670	5,007	6
α	288	5,159	8,000	293	532	963	3,093	5,298	28
β	101	4,319	8,000	117	164	175	1,181	3,734	28
o	414	5,691	8,000	447	608	1,168	3,757	7,454	28
x	1,016	6,675	8,000	1,914	2,779	3,472	5,355	8,000	2,147

E Data collection and coding of tags

As stated in the main text (section 5), we consider only tags that include a link to an associated Wikipedia article and we code a party as tagged with the ideology or lr-position to which the link refers.

Technically, a tag offers two pieces of information: a link and an associated label that is presented in the infobox. Our coding of tags focuses on the targets of links rather than their labels. On their own, labels are merely words that editors can choose freely in order to describe a party's ideology. This lack of constraint in labels can make it difficult or even impossible for other users to verify their meaning and accuracy. Links resolve this difficulty by requiring editors to commit to already existing content on Wikipedia. Their use ensures that the veracity of tags in a party's infobox can be assessed and verified by other Wikipedia users. To reduce measurement error, we therefore use link targets instead of labels.

Links sometimes refer to Wikipedia urls that redirect to another Wikipedia article. For example, the urls for 'liberalist' and 'liberal politics' both redirect to the Wikipedia article Liberalism. In such cases, the final destination article is what we code as a party's tag.

We use Python and R scripts to extract the desired information from Wikipedia and to create a dataset for analysis. Our data collection proceeds in five steps:

1. We retrieve the entire infobox content (all categories) from parties' Wikipedia articles.

2. We create a dataset that stores the downloaded information and the raw Wikitext markup code for all infoboxes.

3. We select all entries for the infobox categories ideology and position and extract their links.

4. We download the titles of the Wikipedia articles (including redirect targets) to which the links refer.

5. We create a stacked dataset with the following variables: party, category (ideology vs. position), link, tag (i.e., article title) from which we construct all input data used in the estimation of Models 1 and 2.

To distinguish ideology tags from lr-position tags, we employ the following definition: tags referring to the Wikipedia articles Far-left politics, Left-wing politics, Centre-left politics, Centrism,

Table E.5: Tags that are observed in the infobox category on political position

	position	ideology
centre-right politics	612	0
centre-left politics	578	1
centrism	455	66
right-wing politics	422	0
left-wing politics	337	1
far-right politics	166	1
far-left politics	131	0
big tent	28	49
syncretic politics	6	5
third position	5	3
radical centrism	3	5
left of center (turkey)	1	1
pro-europeanism	1	217
third way	1	25

Centre-right politics, Right-wing politics, or Far-right politics are lr-position tags. All other tags are ideology tags. For ease of exposition, we refer to lr-position tags as far-left, left-wing, centre-left, centre, etc. (i.e., we refer to their labels), rather than using their full article titles in Figures 3 and 5.

Our above definition is consistent with the modeling assumption that lr-position tags encode successive intervals on an underlying continuum (see section 3). It is also consistent with where we find tags in infoboxes. Table E.5 lists all the tags that we observe in the infobox category on political position. It shows their respective frequency in that category, as well as their frequency in the ideology category. We find that lr-position tags, as defined above, are used almost exclusively in the position category of an infobox (with the exception of centrism). Conversely, ideology tags, as defined above, are used almost exclusively in the ideology category, as the table lists only seven such tags.

An alternative way of defining ideology and lr-position tags would be solely with reference to their infobox category. As Table E.5 shows, such a definition would lead to inconsistencies. Tags used across both categories would have to be treated in the estimation as ideology as well as lr-position tags, depending on their category. The tag centrism in particular is used in both the ideology and position category in a total of 36 parties. Furthermore, tags in the lower half of Table E.5 do not refer to successive intervals on the underlying dimension, which makes them incompatible with our assumptions about the ordering inherent in lr-position tags (see section 3).

Treating these tags as ideology tags allows us to estimate their locations on the underlying scale freely.

F Point estimates for ideologies

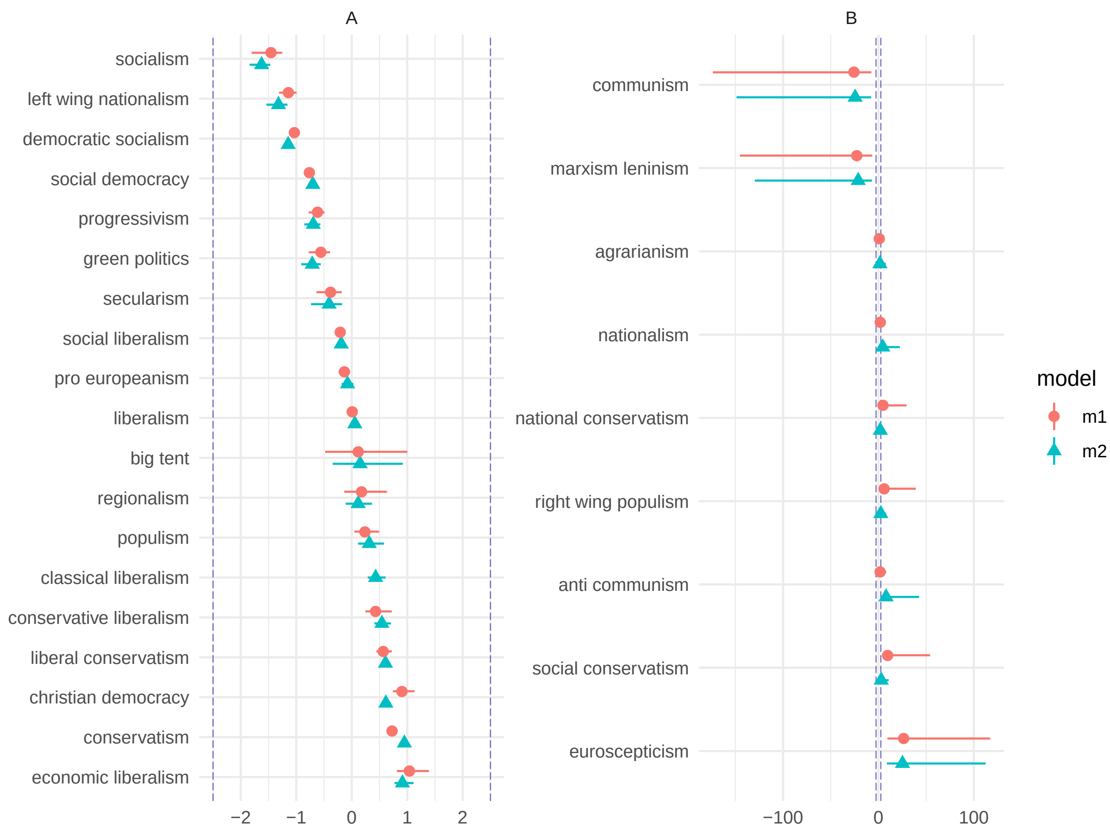

Figure F.1: Point estimates and 95% credibility intervals for the locations (i.e., optima) of ideology and position tags. Panel A shows tags whose 95% credibility intervals lie within the range of party positions (-2.5 to 2.5), as indicated by the dashed lines; panel B shows tags whose estimated 95% credibility intervals lie beyond the range of party positions.

G Comparison to CA estimates

Figure G.2 compares our Bayesian estimates to estimates obtained through correspondence analysis of the data matrix. For parties from left to center both approaches yield very similar results, and CA apparently gets better at distinguishing center-right parties from more right-wing parties, once lr-position tags are included (notice the longer linear slope in the right-hand panel). Yet, CA does not satisfactorily distinguish right-wing and far-right parties, 'squeezing' them together relative to left-of-center parties. Figure G.3 shows that CA nevertheless yields useful estimates of party positions that successfully distinguish left from right. The overall correlation between CA estimates and expert ratings is not much worse than for our Bayesian estimates, especially after including lr-position tags. Yet, the correlation does not hold up for both sides of the political spectrum. In particular for right-of-center parties with an expert score of 6 or higher, the correlation disappears.

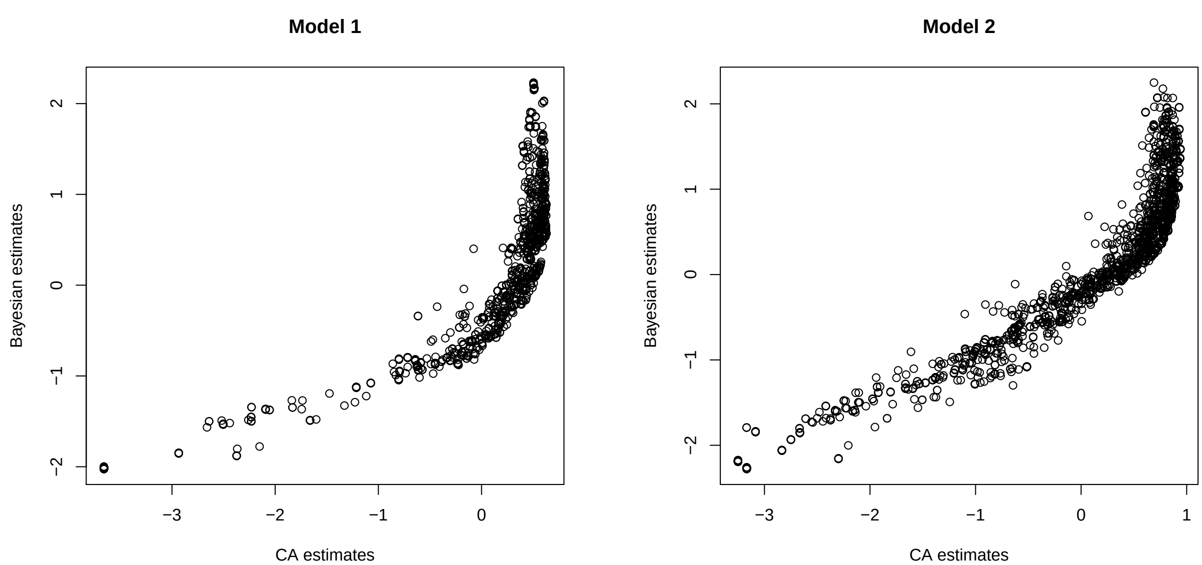

Figure G.2: Comparison of party position estimates to their starting values, i.e., estimates obtained via correspondence analysis

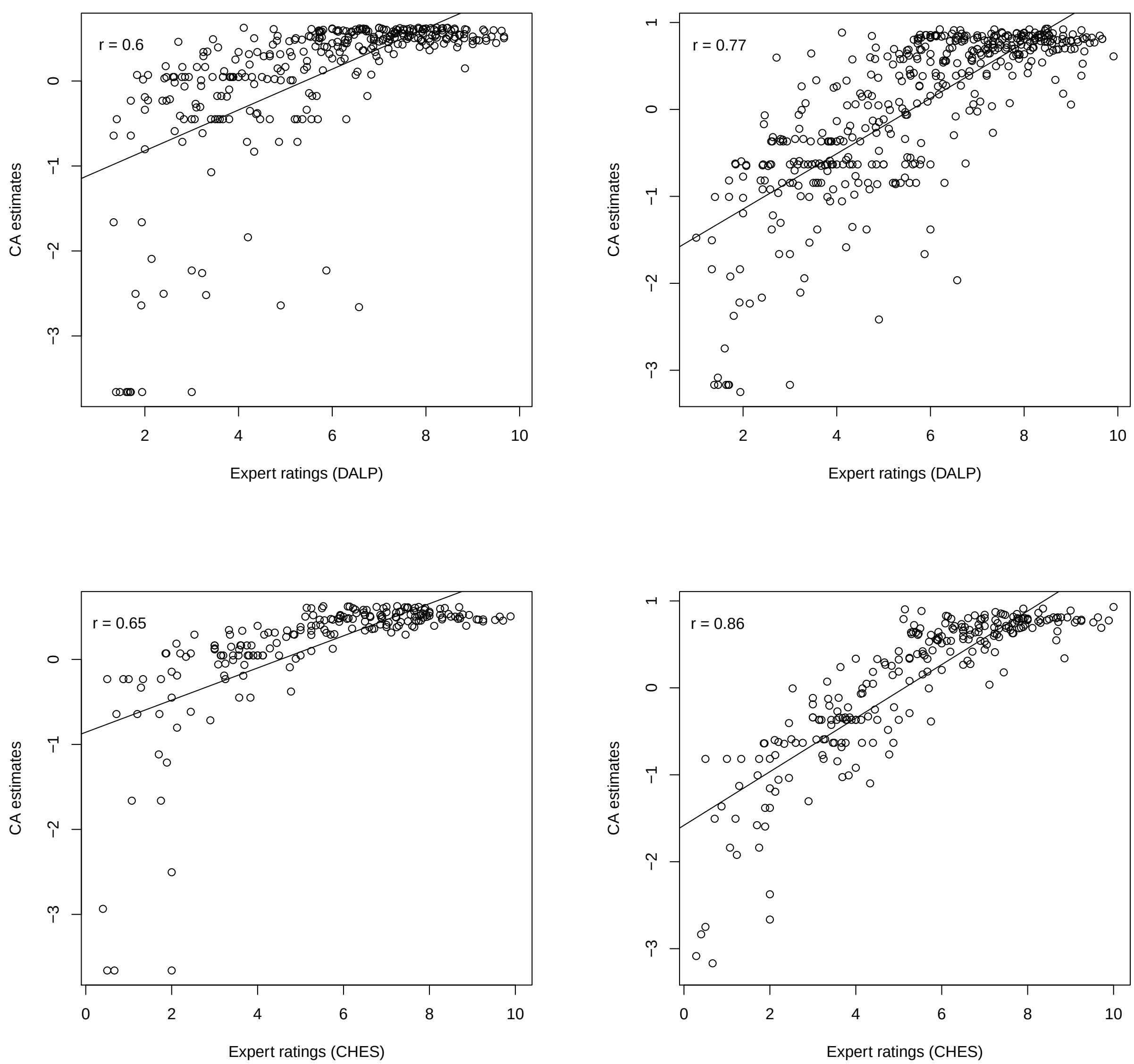

Figure G.3: Comparison of CA estimates to expert ratings of party positions: left-hand panels show results based on ideology tags only (i.e., the data used to estimate Model 1); right-hand panels show results based on ideology and lr-position tags (i.e., the data used to estimate Model 2)

H Inspection of changes in tags

In the main text (section 7) we conclude that edits to tags are mostly refinements rather than sweeping revisions of a party's ideology classification. Here we examine whether the larger deviations that we saw in the test-retest might reflect the opposite, namely controversy and divergent views among editors over a party's ideology. We consider the 13 parties mentioned in section 7 whose position changed by more than 0.5 units on the latent dimension (i.e., a move from one left-right bracket to the next).

First, we consider the identity of the tags that changed. If there is controversy over tags, we might expect most disagreement over the assignment of extreme tags such as far right, far left, right-wing populism, or socialism. Among the 13 parties with the strongest change in position, the tags that are being added or removed most often are centre right (6 times), right wing (5), Christian democracy (3), liberalism (3), and pro-Europeanism (2). All other tags only change in one party (see below).

Second, we examined for each party which of its tags did and did not change. To assess whether editors diverge in their views on a party's position, we compare the ideological range of the party's tags at both points in time. The ideological range is the difference between a party's rightmost and leftmost tags. We take an increase in the ideological range over time as evidence for potential divergence in editors' views. To obtain the ideological range, we treat the point at which a tag's response curve peaks as its location (see Figure 5 or Figure F.1). For lr-position tags, this point is the midpoint of each left-right interval. For tags whose estimated location falls outside the range of party positions, we substitute the values –2.5 and 2.5. Our assessment of each party's tags is given below. For reference, we also give each party's estimated positions and expert rating if available. Overall, we find more indication of convergence than divergence of tags over time.

Convergence of tags (ideological range decreases):

- The Slovenian National Party exhibits the largest change in estimated position (from 0.10 to 1.81). It loses the tag left wing, while gaining Euroscepticism, right-wing populism, social conservatism, and right wing. Its other tag is far right. The range of tags decreases. The party's DALP expert survey score is 7.9 on a scale from 1 to 10.
- The Alliance for Change (of Guyana) (positions: –0.91 and –0.30) loses the tag socialism. Its other tags are nationalism and progressivism. The range of tags decreases.

- The Association for the Rebirth of Madagascar (positions: –1.08 and –0.11) gains the tag liberalism. Its other tags are nationalism and socialism. The range of tags decreases.

- The German Social Union (of the former German Democratic Republic), a historical party (positions: 0.96 and 1.46), loses the tags Christian democracy and centre right. Its other tags are national conservatism, social conservatism, and right wing. The range of tags decreases.

- The New Liberal Alliance (of Denmark) (positions: 0.18 and 0.72) loses the tag liberalism, while gaining right wing. Its other tags are Euroscepticism and centre right. The range of tags decreases. The party's DALP expert survey score is 7.8 on a scale from 1 to 10; its CHES score is 7.9 on a scale from 0 to 10.

- The People's Front for Democracy and Justice (of Eritrea) (positions: –1.38 and –0.83) loses the tag Marxism-Leninism. Its other tags are left wing nationalism, socialism, secularism and big tent. The range of tags decreases.

- The Popular Action (of Peru), a historical party (positions: 0.34 and –0.22), loses the tags centre right and nationalism, while gaining the tags big tent, liberalism, progressivism, and social democracy. Its other tag is centre. The range of tags decreases.

- The Revolutionary Nationalist Movement (of Bolivia) (positions: 0.32 and 0.87) loses the tag left-wing nationalism. Its other tags are centre right and social conservatism. The range of tags decreases. The party's DALP expert survey score is 6.5 on a scale from 1 to 10.

- The Rastakhiz Party (of Iran), a historical party (positions: 0.89 and –0.10), loses the tags anti-communism and right wing. Its other tags are populism and secularism. The range of tags decreases.

- Volya (of Bulgaria) (positions: 1.07 and 1.75) loses the tag centre right, while gaining the tags right wing and far right. Its other tags are Euroscepticism and right-wing populism. The range of tags decreases.

Divergence of tags (ideological range increases):

- The League of Polish Families exhibits the second largest change in estimated position (2.02 to 0.88). It gains the tags Christian democracy, conservatism, pro Europeanism, and centre right. Its other tags are Euroscepticism, social conservatism, and far right. The range of tags increases. The party's DALP expert survey score is 9.3 on a scale from 1 to 10. Inspection of the party's Wikipedia article shows that its tags are divided into "current" and "historical". According to the article, the party moved from a far right toward a more centrist position in 2010 (after the DALP survey was administered). Our approach currently ignores this additional information, lumping all tags together. The "change" in estimated position is thus due to measurement error rather than disagreement between editors.

- Change 90 (of Peru), a historical party (positions: 1.54 and 0.92), loses the tag right wing, while gaining the tags centre right and Christian democracy. Its other tags are right-wing populism and social conservatism. The range of tags increases.

Change of editor perception:

- The Liberal People's Party (of Norway) (positions: 0.56 and –0.16) loses the tags conservative liberalism, economic liberalism, and centre right, while gaining the tags pro-Europeanism, social liberalism, and centre. No tags remain unchanged. In both cases the range of tags is small. The changes are not indicative of starkly opposing views of the party's position (i.e., its position shifts from about center right to the center of the scale).

I Collection of metadata on Wikipedia articles

The metadata on Wikipedia articles that we present in the main text (section 9) was retrieved from Wikipedia using the XTools interface. R scripts were used to collect the data. We considered all parties for which the Party Facts database records a Wikipedia url. For each of these parties' articles, we requested

1. its number of pageviews,
2. its number of editors from the earliest available date until the day of tag data collection, excluding bots and spiders, and
3. the names of its top 50 editors from the earliest available date until the day of tag data collection, excluding bots and spiders.

For each of the ($>$ 100k) editors identified in step 3, we then retrieved the total number of live edits made on Wikipedia up until the date of tag data collection. To compute our measures, we take the log of each count, adding 1 to pageviews to avoid $\log(0)$. For our measure of editor experience, we compute the average of the log number of live edits over each party's top 50 editors.

J Modeling the accuracy of Wikipedia-based estimates

As stated in the main text (section 9), we assess the accuracy of our scores by modeling how closely they approximate the expert ratings. For simplicity, we assume that error in expert ratings is negligible so that all measurement error rests with our scores. If both measures tap into the same basic dimension, they should linearly map into each other. Deviations from the linear mapping then represent measurement error in our scores. We formalize these assumptions with the following model

$$\hat{x}_i \sim \text{Normal}(\mu_i, \sigma_i), \tag{5}$$

$$\mu_i = \nu e_i, \tag{6}$$

$$\sigma_i = \exp(\eta z_i), \tag{7}$$

where $\hat{x}_i$ is a party's Wikipedia-based position estimate (by construction, these estimates are normally distributed), νe_i is a scalar product that linearly maps a party's expert rating into its Wikipedia-based position estimate with ν being a two-element parameter vector and e_i a vector containing a party's expert rating (taken either from the DALP or the CHES survey) and a leading 1; likewise ηz_i is a scalar product of further parameters η and additional covariates z_i including a leading 1.

The proposed model is a linear regression with variance heterogeneity (King 1998, Verbyla 1993). Of particular interest to us is Eq. (7), which conditions the standard error of the regression on additional covariates. This feature allows us to study to what extent measurement accuracy in our scores varies with attributes of the parties' Wikipedia articles such as their number of editors or editor experience. We obtain estimates of ν and η via maximum likelihood.

K Software statement

Python 3.7

- `https://github.com/siznax/wptools`
- `https://github.com/5j9/wikitextparser`
- `https://pandas.pydata.org/`

R 3.5

- `crch`
- `curl`
- `data.table`
- `ggrepel`
- `glue`
- `jagsUI`
- `jsonlite`
- `MASS`
- `plyr`
- `rio`
- `stargazer`
- `tidyverse`
- `xtable`

JAGS 4.3.0

- `http://mcmc-jags.sourceforge.net/`