

WILL LOWE
Maastricht University
KENNETH BENOIT
London School of Economics and Political Science
SLAVA MIKHAYLOV
University College London
MICHAEL LAVER
New York University

Scaling Policy Preferences from Coded Political Texts

Scholars estimating policy positions from political texts typically code words or sentences and then build left-right policy scales based on the relative frequencies of text units coded into different categories. Here we reexamine such scales and propose a theoretically and linguistically superior alternative based on the logarithm of odds-ratios. We contrast this scale with the current approach of the Comparative Manifesto Project (CMP), showing that our proposed logit scale avoids widely acknowledged flaws in previous approaches. We validate the new scale using independent expert surveys. Using existing CMP data, we show how to estimate more distinct policy dimensions, for more years, than has been possible before, and make this dataset publicly available. Finally, we draw some conclusions about the future design of coding schemes for political texts.

Almost anyone interested in party competition, whether this takes place in legislatures, the electoral arena, or government, needs sooner or later to estimate the policy positions of key political actors, whether these be individual legislators or the political parties to which they affiliate. Indeed, “how to best measure the policy preferences of individual legislators and of legislative parties” (Loewenberg 2008, 499) forms one of the central problems of legislative research. This is particularly true for scholars of comparative legislative research. While in the American settings policy preferences of legislators have been conceptualized as individual-level variables, tight party discipline in many non-American contexts makes it difficult to derive estimates of legislators’ ideal points that are distinct from aggregate policy stances of the parties to which they belong. Over the last two years this journal has devoted particular attention to the problem of

measuring the policy preferences of legislators (e.g., Alemán et al. 2009; Carroll et al. 2009; Carrubba, Gabel, and Hug 2008; Clinton and Jackman 2009; Hix and Noury 2009; Saiegh 2009; Schickler and Pearson 2009). Here we contribute to this discussion by focusing on the estimates of policy positions of parties in legislatures on different dimensions over time.

In comparative legislative research, there are many sources of data from which estimates of the policy positions of key political actors—be these legislators or legislative parties—can be derived. These include, among others: mass surveys; expert surveys; political text; roll-call votes; and bill sponsorship (see Benoit and Laver 2006, for a review). By far the most abundant source of data on policy positions, both cross-sectionally and over time, is political text. Text is a direct by-product of political activity by the political actors whose positions we wish to estimate, whether this text takes the form of speeches, debates, written submissions, written rulings, or—by far the most commonly used in the profession for estimating party policy positions—election manifestos issued by political parties. These manifestos outline policies that parties will enact once elected to legislative or executive office and serve as the empirical basis for many models of party competition in legislative and other policymaking settings.

The wide availability of these materials in electronic form has led to a large number of automated and semiautomated methods for scaling positions from political texts based on the statistical analysis of word patterns (e.g., Bara, Weale, and Biquelet 2007; Benoit and Laver 2003; Hilliard, Purpura, and Wilkerson 2007; Hopkins and King 2010; Klemmensen, Hobolt, and Hansen 2007; Laver and Garry 2000; Lowe 2008; Martin and Vanberg 2007; Monroe and Maeda 2004; Pennings and Keman 2002; Quinn et al. 2010; Slapin and Proksch 2008; Yu, Kaufmann, and Diermeier 2008). Despite this growth in automated methods, however, the most common means of analyzing political text remains manual content analysis (Krippendorff 2004; Neuendorf 2002). In a traditional manual content analysis, a predefined categorical coding scheme is applied to segments of text by trained human coders (e.g., Baumgartner, Green-Pedersen, and Jones 2008). The most comprehensive and most frequently used such dataset comes from the Comparative Manifesto Project (Budge et al. 2001; Klingemann et al. 2006, hereafter CMP) which contains the results of coding more than 3,000 election manifestos for more than 650 parties in over 50 countries. CMP data form the basis for hundreds of published studies by third-party authors and are almost always used to estimate policy positions for political parties on left-right scales.

Almost everyone using CMP data does so for the same reason: they want to estimate *positions* of parties on different common policy dimensions. Doing this typically implies assuming that a set of party positions, whether a cross-section or a time series, can be located on some (continuously defined) metric scale. Such a scale allows analysts to make statements to the effect that, for example: party *A* is “moving” towards the left; parties *A* and *B* are “closer” to each other than either is to party *C*; given parties *A*, *B*, and *C*, the “median legislator” in the set of three parties is at *X*; and so on. Spatial theories of policy preferences typically assume that party positions exist on a continuous scale, usually an interval scale, although content coding schemes such as the CMP record only absolute and relative category counts of discrete text units. To convert these observed category counts into points on a continuous policy dimension, therefore, some scaling procedure is required. The CMP data offer several general political scales based on aggregating counts of text categories. The most widely used of these is the CMP’s left-right “Rile” scale, constructed by subtracting the sum of 13 “left”-associated categories from the sum of 13 “right”-oriented categories.¹ There are many different ways to construct such scales, however, and the choice of scaling procedure involves decisions that must be defended on methodological and substantive grounds.

In this article we present a new method for scaling continuous left-right policy positions from political text coded into discrete categories and demonstrate its superiority to current approaches. Comparing our measure to previous scales, we demonstrate that our proposed scale not only better satisfies general political, linguistic, and psychological criteria, but also that it exhibits superior empirical properties when applied to the CMP data. We validate our new scale externally through comparison to independent expert surveys. Not only can our new approach be applied to improve existing policy estimates for the most commonly used CMP scales, it can also be used with existing CMP data to unlock reliable positional estimates on new policy dimensions. These new and improved scales will provide researchers in legislative studies with not only more valid measures of the policy positions than ever before, but also unlock measures for previously unused dimensions of policy that can be used to test empirical models of party competition, legislative coalitions, government formation, and executive-legislative relations. To make the scale immediately useful to applied researchers, we provide a full dataset, described in an appendix and in Tables 1 and 3, of these newly scaled policy positions with 21 new left-right scales, at least half of which

have never before been used in applied, published research. Following the method for estimating uncertainty from political text of Benoit, Laver, and Mikhaylov (2009), we also provide confidence intervals for every new estimate. Finally, by justifying and demonstrating what types of coding categories are best compared to create continuous scales, our findings provide direct lessons for the future design of improved political text coding schemes.

How Should Policy Mentions Be Counted?

The CMP's manual coding process involves several stages. In the first step, a human coder is given a political party manifesto, which he or she then divides into discrete, nonoverlapping text units known as "quasi-sentences." Quasi-sentences are textual units that express a policy proposition and may be either a complete natural sentence or part of one. Once identified, the quasi-sentence is then assigned to one of 56 mutually exclusive policy categories, distributed across seven broad policy domains such as "Political System" or "Economy." CMP data thus take the form of counts of sentences in categories, a unit of analysis that is intermediate between the more holistic analysis offered by an interpretative approach and more detailed syntactic analyses (Popping 2007; van Atteveldt, Kleinnijenhuis, and Ruigrok 2008) and purely lexical approaches (Laver, Benoit, and Garry 2003; Slapin and Proksch 2008). Category counts are then converted to percentages by dividing by the total number of sentences in the manifesto. These category percentages are then either interpreted directly as conveying information about the policy preferences of their authors or may be additively scaled to construct more general indices.

Normalizing counts this way makes sense under three conditions that we will not, for the purposes of this article, dispute: first, the sentence is the fundamental unit of policy assertion; second, different sentences assigned to the same category are exchangeable or independently distributed conditional on their policy category; third, the total number of sentences assigned to any policy category contains no information about the policy preferences that a platform expresses. The precise choice of how to construct a left-right scale from the normalized sentence counts, however, requires decisions to be made in the construction of scales. Scaling category counts, that is, choosing a procedure to transform observed category counts into estimates of unobserved policy positions, means addressing two independent questions about the content and the form of a scale. These are two funda-

mental questions to which we will return as we evaluate different methods of scaling left and right policy.

First, how should sentences be counted when constructing a scale for a particular policy domain? Should one category be considered against an absolute standard, or relative to the counts in a different category, or perhaps relative to the entire document? Second, what is the functional form of the relationship between position and counts? In particular, what is the nature of the marginal effect on sentence counts of changes in a party's position in the policy domain linked to the sentence counts? While these two key issues frame a debate that has previously occupied methodologists concerned specifically with scaling policy positions from the CMP data (e.g., Kim and Fording 2002; McDonald and Mendes 2001a), the debate applies much more generally to *any* effort to construct continuous scales from text coded into discrete categories. In what follows, we reexamine both issues from both a substantive political standpoint and also from linguistic and psychological perspectives.

Previous Approaches to Scaling Policy Measures

In the discussion of scaling measures we assume that for each policy dimension there exists a “left” and a “right” direction represented by at least one CMP category.² We will denote the *number* of sentences in a manifesto assigned to the “left” and “right” categories constituting a policy issue as L and R , respectively, and the total number of sentences in all categories as N . (There is also an “other” category count O to completely partition the sentences, such that $L + R + O = N$.) For instance, for a policy dimension of more to less protectionism, L would be the number of sentences coded to “406 Protectionism: Positive,” while R would be the number of sentences coded to “407 Protectionism: Negative,” and the corresponding “PER” variables defined as $\frac{L}{N}100\%$ and $\frac{R}{N}100\%$, respectively. The *output* of any scaling procedure is an estimate of the position which we will refer to as θ , superscripting to indicate the scaling procedure and subscripting as necessary to indicate the policy dimension.

Previous Scaling Procedures

The CMP was designed to reflect “saliency theory,” a particular view of how parties compete and therefore how they express their policy preferences, asserting that “all party programmes endorse the

same position, with only minor exceptions” (Budge et al. 2001, 82). Parties are assumed to differentiate themselves by emphasising issues on which they have the best reputation with voters (Budge 1994). Because positioning is a matter of emphasis, the answer to the first general methodological question posed above must be that the frequency of quasi-sentences in one policy category should be compared to all other sentences in the manifesto. Budge (1999) suggests that a party’s position according to saliency theory, $\theta^{(s)}$, should be defined as

$$\theta^{(s)} = \frac{R - L}{N}.$$

This saliency measure is based on the difference in counts between left and right sentences counts normalized by the total number of sentences in the manifesto on any issue or on none.³ From this definition it is clear that the answer to the second general question posed above is that each count in L or R has the same marginal effect: $1/N$. The quantity $\theta^{(s)}$ is equal to zero when there are exactly the same number of left- as right-coded sentences, -1 when there is only one issue on which the party is perfectly “left,” and 1 when there is one issue and the party is perfectly “right.” In practice, however, the extreme values are never reached because party competition almost never occurs on one dimension only. For instance, the distribution of the CMP’s “Rile” left-right index, a measure that encompasses 26 different coding categories, has an empirical range of about $[-.5, .5]$.

There is a more subtle constraint on $\theta^{(s)}$ hidden in this formulation. All theories accept that if an issue becomes less important then a party will devote fewer sentences to it. That is, the relative counts $R + L$ assigned to the contrasting policy pairs R and L , for a specific policy subset of all policy dimensions in a manifesto, will shrink. But because $R + L$ is also by definition the maximum range of $R - L$, then deemphasizing an issue will push $\theta^{(s)}$ to a more centrist position by moving it closer to 0, even though the proportion of left and right sentences, the raw material for expressing a position, have not changed. For the composite “Rile” scale, this means that counts of the 30 categories *not* in the scale still affect estimated party positions. For instance, a 200-sentence manifesto with 100 right sentences and no left sentences would have a Rile score of $(50 - 0) = 50$, but the same manifesto with 50 sentences added that are neither left nor right would change its Rile score to 40 (Benoit and Laver 2007; McDonald and Mendes 2001b; Ray 2007)—suggesting that the party shifted 20%

toward the left. In the CMP, this approach is carried to an extreme by including even uncodeable content in the definition of a manifesto.⁴

Primarily in order to address this problem, Kim and Fording (2002) propose an alternative measure that restricts the difference to sentences from the constituent left and right categories (see also Laver and Garry 2000). This *relative proportional difference* estimate of position is

$$\theta^{(R)} = \frac{R - L}{R + L}.$$

The measure also ranges from -1 to 1 , but makes explicit the range constraint hidden in $\theta^{(S)}$. Dividing by $R + L$ decouples the measure from variation in the importance a party assigns to any issue area. The only remaining influence of variable issue importance is that the overall number of sentences available to *express* a position is increased or reduced. To take an extreme case, only three positions are expressible within a budget of two sentences: either both are left, both right, or one is assigned to each category, leading to estimated positions of $-1/2$, 0 , or $1/2$. Coarse sampling does not necessarily imply anything about the party's actual position on the issue but rather limits the level of nuance and specificity that it can be expressed in a manifesto and the precision that may be inferred from it by readers and researchers. According to spatial theory assumptions the party has a position on the issue dimensions, but has chosen to use its supply of sentences on other dimensions. Finally, unlike $\theta^{(S)}$ this measure will not necessarily create an apparent move to a more centrist position if the party decides to focus on other policy areas.

In terms of the two methodological questions above, $\theta^{(R)}$ compares category counts only to counts in the opposing category rather than to counts of all quasi-sentences. The marginal effect of another sentence on the left or right side of the issue is therefore $1/(R + L)$.

Although $\theta^{(R)}$ appears to fix the problem of sentences in unrelated or uncoded categories affecting position estimates, it shares the assumptions embodied in $\theta^{(S)}$ about the fixed marginal effect of another coded sentence and the existence of fixed endpoints. This has the unfortunate effect of forcing the $\theta^{(R)}$ to -1 when $R = 0$ irrespective of the value of L , or to 1 when $L = 0$ irrespective of the value of R , leading to spikes at the boundaries of the scale. That the scale has boundaries at all is a basic problem with both procedures that attempt to measure policy positions that are more naturally conceptualized in

an underlying continuum. The essential insight behind $\theta^{(R)}$ is surely correct—the position of a party on a policy dimension should depend only on L and R . The problem is that the nature of the quantity being estimated is not respected in the measure. A different answer to the second general question is needed.

A Scaling Method Based on Log Odds-Ratios

To motivate a new scaling method, consider the process of reading a party manifesto for changes in policy content, as a voter might do, for example, if trying to identify any change in some party's policy position on the European Union. If the party's previous platform contained 50 sentences in favour of increased European integration, and 20 emphasizing its disadvantages, then a new manifesto containing 50 sentences in favor and 21 against would barely register as an indicator of policy change. But if the previous platform had contained 10 and 4 sentences for and against the EU, and the new platform 10 and 5, then a policy change is more plausible. This suggests that the balance between assertions in favour of the EU and against it between platforms is usefully summarized not by the difference between sentence counts, but rather by their ratio. The effect of adding one more sentence in the first case decreases the ratio of pro to anti-EU sentences by about 5%, and in the second by 20%. By this reasoning, the marginal effect of one more sentence is decreasing in the amount that has already been said on the topic. Proportional or relative emphasis on different topics does indeed determine a reader's estimate of position, but such changes must be perceivable against the background of existing policy emphasis.

This simple linguistic intuition about reading and writing manifestos can be supported by evidence from psychology. The decreasing marginal effect of an extra unit is a general property of many perceptual quantities such as temperature, heat, or loudness studied by psychophysicists.⁵ The Weber-Fechner law (Fechner 1965; Stevens 1957) formalises this observation: the size of the “just perceivable difference” of a subjective quantity is a constant proportion of the quantity already present.⁶ Consequently we should operate in proportions, not levels, and work with a logarithmic scale relationship between the underlying quantity and subjective estimations of it. For loudness, this relationship is the familiar decibel scale, which relates perceived loudness as the log of the physical power of the sound.⁷ Following this logic it should also be possible to consider the “just perceivable policy

difference,” the proportional change necessary to infer a difference in position on an issue between two party platforms.

The Logit Scale of Position

Our logic suggests that from the point of view of a party manifesto writer wanting to communicate a position effectively, it is important to manipulate not so much the *absolute quantity* of sentences allocated ($R + L$), but rather their relative *balance*, or R/L . Increasing $R + L$ allows a wider range of expressible policy positions, but manipulating R/L expresses the position itself. Furthermore, because we are primarily interested in inferring positions, we view it as most natural to consider proportional changes on a *symmetrical* left-right scale. One natural measure for this purpose is the empirical logit:

$$\theta^{(L)} = \log \frac{R + .5}{L + .5} \quad (1)$$

$$= \log(R + .5) - \log(L + .5). \quad (2)$$

Like $\theta^{(R)}$, $\theta^{(L)}$ is conditional because it only considers sentences that are assigned to left or right. Unlike $\theta^{(S)}$ and $\theta^{(R)}$, however, the logit scale $\theta^{(L)}$ has no predefined end points: given enough sentences, it is possible to generate positions of any level of extremity.⁸ In this respect, $\theta^{(L)}$ better reflects spatial politics assumptions about the possible range of ideal points. However, although any real valued policy position can be represented, expressing extreme positions requires exponentially more sentences in L or R to move the policy position the same distance left or right as can be seen by considering its alternative formulation (2) as a difference measure.⁹

We should note that although $\theta^{(L)}$ is defined as a (logged) ratio, it offers *interval* not ratio level measurement. In particular, $\theta^{(L)} = 0$ should not automatically be identified as a substantively centrist policy position. In the absence of an external anchor, e.g., to policy outcomes, a centrist position would be some function of the mean or median position on an issue of the parties contesting the election. How this position will be expressed in R , L terms will depend on historically contingent country-level factors.¹⁰

Using the logit function to transform count data represents a novel approach to scaling left-right policy positions, but logit transformations are found in many inferential models used to estimate latent party positions. Log odds-ratios form the basis of the most

commonly used statistical models of bounded count data (Agresti 1996; Fleiss, Levin, and Paik 2003), item response and unfolding (Elff 2008), and have been studied directly by Monroe, Quinn, and Colaresi (2008).¹¹ Nevertheless, $\theta^{(L)}$ is explicitly *not* itself a model of the structure of policy positions but rather a way to measure them that is compatible with several theories of spatial politics. We do not pursue such models here because we are unwilling either to introduce the independence of irrelevant alternatives (IIA) constraint on policy dimensions that would be imposed by logit models or to estimate explicitly the distribution of party positions on multiple dimensions as required by probit models. Consequently we also take no position on important substantive issues such as the underlying dimensionality of the policy space and the correlational structure connecting issue dimensions (Elff 2008; Gabel and Huber 2000) or the dynamics of party positions over time. Our more modest goal here is to improve the future use of the hugely popular CMP dataset, after demonstrating a better way to scale policy positions than the CMP's existing, flawed approach. Furthermore, our confrontational pairing method provides scales for more policy dimensions than ever before used from the CMP dataset. Whether these new positions are comparable over time, or accurately reflect the underlying dimensions of politics, are separate questions that are broader than we can feasibly address here.

Instead, we focus on the scaling procedure that connects basic data of the CMP—counts of sentences in categories—to the policy positions that form the substantive quantities of interest. Even without making model assumptions, we can show that $\theta^{(L)}$ is a far better predictor of party policy positions than previous measures. We do make one concession towards model structure by adding 0.5 to all counts, a standard statistical practice for the analysis of contingency tables (Agresti 1996) that can also be motivated as a measure to reduce bias when estimating category proportions (Brown, Cai, and Das-Gupta 2001; Firth 1993). This smooths $\theta^{(L)}$ slightly towards 0 and makes position estimates created from very small counts more stable, while barely affecting those derived from more reasonable numbers of sentences.

A Log Scale of Policy Importance

In addition to having different *positions* on each of a given set of policy dimensions, political actors may also differ in terms of the relative *importance* they attach to these dimensions. As Laver and Hunt (1992) demonstrated, some issues are simply more important to

some parties than to others, quite independent of their party positions on these dimensions, a distinction long-recognized by other scholars (e.g., Grofman 2004; Riker 1996). We thus expect “green” parties to treat the environmental dimension as the most important policy domain, and indeed this is part of our implicit definition of the set of green parties. Likewise, we expect far-right parties to treat immigration and social values as the most important dimensions. Both liberal and far-right parties might consider social values to be very important, yet take very different positions on this dimension. Scholars concerned with the policies of political actors are typically concerned with both position and importance. Empirical methods often draw this distinction very explicitly, as with the expert surveys of Laver and Hunt (1992), Benoit and Laver (2006), and Hooghe et al. (2008).

Notwithstanding this very clear analytical distinction between the importance, or salience, of a policy dimension and party positions on that dimension, the widely used policy *scales* (as opposed to raw *data*) generated by the CMP are fundamentally grounded in the CMP’s “saliency theory” of party competition (*MPP*, 76). This explicitly conflates party positions on policy dimensions and the relative salience of these dimensions. The core idea of saliency theory is that, in a given setting, parties will endorse only single sides of each issue, such as reducing crime, providing for the national defense, or protecting the environment. Parties differentiate themselves by emphasizing the issues on which their stances are most credible (*MPP*, 7). Consequently, the “taking up of positions is done through emphasizing the importance of certain policy areas compared to others” (Budge 1994, 455).¹² Operationally, “saliency” theory suggests that the relative mention of different policy areas in manifestos provides a direct measure of their importance to the party. Despite this prediction that issues are overwhelmingly one-sided, however, the CMP’s coding scheme makes numerous practical concessions to the fact that many issues are clearly two-sided, such as positions on free trade, on the level of government regulation, or on attitudes toward European integration. The existence of paired categories in the CMP scheme covering opposite sides of the same issue complicates the straightforward assessment of policy salience based on counting relative mentions of a single policy category. Our solution is simple: to group mentions of an issue, whether positive or negative, and to consider their sum as a direct indicator of policy importance. Our scale also follows the psychological and linguistic rationale for logarithmic emphasis as explained previously, however.

Our suggested measure of policy importance is

$$\theta^{(l)} = \log \frac{R + L + 1}{N},$$

with a value of 1.0 added to the numerator for consistency with the 0.5 for R and L in the position formulation. This measure follows directly from the relative emphasis logic of saliency theory and also conforms to the linguistic model we have already outlined by increasing logarithmically in extremity with additional mentions.

Estimating Scale Uncertainty

It has become widely accepted that text-based measures of policy quantities should come with associated estimates of uncertainty, rather than simply being presented as if they contained no stochastic element or measurement error (see Benoit, Laver, and Mikhaylov 2009). For this reason we also provide a means of computing standard errors and confidence intervals associated with our new scales of position and importance.

If a parametric measure of uncertainty is required, we suggest a simple Bayesian approach: a standard Beta prior over the proportions of L and R sentences with parameters $a_R = a_L = a$ implies a posterior distribution over position that is well approximated as

$$\begin{aligned} \theta^{(L)} | R, L &\sim \text{Normal}(\mu, \sigma^2) \\ \mu &= \log \frac{(R + a)}{(L + a)} \\ \sigma^2 &= (R + a)^{-1} + (L + a)^{-1} \end{aligned}$$

when $R + L \geq 10$. Setting $a = 0.5$ corresponds to a symmetrical invariant Jeffreys prior over party position (Jeffreys 1946). This distribution above suggests the 95% credible interval

$$[\theta^{(L)} - 1.96\sigma, \theta^{(L)} + 1.96\sigma]$$

which corresponds closely to the classical confidence intervals (when they are defined) while being numerically more stable (Newcombe 2001).

Many counts of quasi-sentences representing R or L , however, may be zero or close to zero in observed data, implying nonsymmetric bounds that will affect the parametric computation of confidence inter-

vals. An alternative to the parametric estimation that we propose is to use bootstrapping methods (Efron and Tibshirani 1994) to provide nonparametric intervals by resampling R and L categories in each policy dimension. In the dataset provided with this article and in the analyses presented here, we compute nonparametric confidence intervals and standard errors for all position and importance scales represented in the article, using the approach outlined by Benoit, Laver, and Mikheylov (2009).

New Policy Scales

We have constructed a set of 13 policy scales from the CMP dataset, each representing a distinct dimension of policy on which parties may take positions. These are detailed in Table 1. For each scale, we have identified a pair of CMP categories expressing policy opposites and classified the elements of each pair as either *Right* or *Left*. The pairings in Table 1 are natural and probably closer to what was originally intended by the designers of the CMP's coding scheme, although most are seldom or never used in this way. This alternative to the saliency approach has often been termed the "confrontational" approach to policy (Budge et al. 2001; Gemenis and Dinas 2010) and involves parties declaring competing positions on the same issue. In this view of policy, what matters is not whether each party purports to emphasize the issue or downplay it, but rather what the party's specific policy stances are relative to the extreme positions on any given issue, for instance what degree of permissiveness or restrictiveness regulation it favors regarding the issues of euthanasia, homosexual marriage, and abortion (Laver 2001, 66) or whether a party favours expanding the power of European-level institutions or instead reinforcing national sovereignty. Our logit scale extends and generalizes this logic while applying the notion of *relative* difference that also scales policy extremity in a way that relates to repetition in a nonlinear fashion.

In addition to these natural opposites, there are many categories for which natural policy alternatives could have been identified when the CMP coding scheme was being designed, but which do not in fact exist in the coding scheme. We identify these categories in Table 2. With the sole exception of 408 Economic Goals, these categories all relate to matters of public policy that are inherently positional.

The rationale for the CMP's unwillingness to define polar opposites for these coding categories appears to be that one position seems likely to be almost universally unpopular. Consider corruption or the

TABLE 1
Paired Policy Dimensions and Corresponding Variable Names in the Dataset

| Policy Dimension | “Left” Position | “Right” Position | Variable Name |
|----------------------|---|--|-------------------|
| Foreign Alliances | 101 Foreign Special Relationships: Positive | 102; Foreign Special Relationships: Negative | foreignalliances |
| Militarism | 105 Military: Negative | 104 Military: Positive | militarism |
| Internationalism | 107 Internationalism: Positive | 109 Internationalism: Negative | internationalism |
| EU | 108 European Integration: Positive | 110 European Integration: Negative | logeu |
| Constitutionalism | 203 Constitutionalism: Positive | 204 Constitutionalism: Negative | constitutionalism |
| Decentralisation | 301 Decentralisation: Positive | 302 Centralisation: Positive | decentralization |
| Protectionism | 406 Protectionism: Positive | 407 Protectionism: Negative | protectionism |
| Keynesian Policy | 409 Keynesian Demand Management: Positive | 414 Economic Orthodoxy: Positive | keynesian |
| Nationalism | 602 National Way of Life: Negative | 601 National Way of Life: Positive | nationalism |
| Traditional Morality | 604 Traditional Morality: Negative | 603 Traditional Morality: Positive | morality |
| Multiculturalism | 607 Multiculturalism: Positive | 608 Multiculturalism: Negative | multiculturalism |
| Labour Policy | 701 Labour Groups: Positive | 702 Labour Groups: Negative | laborpolicy |
| Welfare State* | 504 Welfare State Expansion: Positive | 505 Welfare State Limitation: Positive | logwelfare |
| Education spending | 506 Educational Provision Expansion: Positive | 507 Education Expenditure Limitation: Positive | education |

*This differs from the CMP’s welfare scale in that the CMP’s version is not confrontational (and does not include pers504).

TABLE 2
CMP Scales with No Natural Policy Opposites

| Policy Issue | CMP Category |
|-------------------------------|--|
| Imperialism | 103 Anti-Imperialism: Anti-Colonialism |
| Peace | 106 Peace: Positive |
| Freedom/Human Rights | 201 Freedom and Human Rights: Positive |
| Democracy | 202 Democracy: Positive |
| Efficiency | 303 Governmental and Administrative Efficiency: Positive |
| Corruption | 304 Political Corruption: Negative |
| Political Authority | 305 Political Authority: Positive |
| (General) Economic Goals | 408 Economic Goals |
| Corporatism | 405 Corporatism: Positive |
| Technology and Infrastructure | 411 Technology and Infrastructure: Positive |
| Cultural Policy | 502 Culture: Positive |
| Social Justice | 503 Social Justice: Positive |
| Law and Order | 605 Law and Order: Positive |
| Social Harmony | 606 Social Harmony: Positive |
| Agricultural Policy | 703 Farmers: Positive |
| Middle Class Policy | 704 Middle Class and Professional Groups: Positive |
| Affirmative Action | 705 Underprivileged Minority Groups: Positive |

environment: Since no party is likely to support corruption or call for trashing the ecosystem, “saliency” theory assumptions seem plausible for such policy issues. A closer look, however, reveals a more nuanced picture. On environmental policy, for instance, parties do not always produce purely one-sided statements. Many parties do in fact take progrowth stances that contain thinly veiled antienvironmental messages. For instance, the 1988 Danish Liberal Party manifesto contains this statement: “*Environmental policy should not result in Danish companies being worse off than the companies in the countries with which we compete.*”¹³ The Danish Liberal Party is clearly not proenvironment, preferring instead to let the natural environment suffer in exchange for the economic benefits that presumably come from easing environmental regulations on firms. This direct preference for industry over the environment is in fact how other schemes for measuring environmental policy have expressed the environmental policy dimension: as contrasting priorities for environmental protection (at the cost of economic growth) versus economic growth (at the cost of environmental damage; Benoit and Laver 2006; Laver and Hunt 1992). We believe that this logic of contrasting extremes applies quite generally.

Not every quantity of end-user interest from the CMP may exist in the form of text units assigned to one of two bipolar categories.

Indeed, most users are only interested in the CMP dataset for its aggregate left-right scale. Fortunately, our measure works equally well for aggregated categories of R and L when each R and L consists of more than one component category count. Furthermore, as with the “Rile” index that includes quantities such as “305 Political Authority: Positive” that have no opposite category in the CMP coding scheme, many of these measures may denote right-measured positions yet not be usable in any simple, bipolar scale. For multicategory indexes, $\theta^{(L)}$ is defined the same way after aggregating category counts into a composite L and R :

$$\theta_{\text{index}}^{(L)} = \log \frac{\sum_j R_j}{\sum_k L_k}.$$

As with simple scales involving only two categories, the zero point on this scale is not substantively privileged and should not necessarily be identified with a centrist policy position. This is particularly clear when different numbers of categories are used in the numerator and denominator.

In Table 3 we have listed a set of proposed additive indexes that are amenable to use with the logit scale, wherever possible identifying the source where this index was developed. We have also proposed several new scales of our own, such as “Free Market Economy” and “State-provided Services.” Our proposed scale of environmental protection follows the confrontational pairing logic by treating the two proenvironmental categories “Antigrowth Economy: Positive” (416) and of course “Environmental Protection: Positive” (501) together to capture antigrowth politics, ecologism as “left,” and the environmentally opposed paradigm of economic growth is represented in the CMP by the category “Productivity: Positive” (410).¹⁴ In the next section we compare our scale to previous formulations and also compare the scale estimates to independent measures of position and importance from expert surveys.

Validating the Logit Policy Scale

Comparing Scales

Before turning to validation against experts it is helpful to compare the properties of $\theta^{(S)}$, $\theta^{(R)}$, and $\theta^{(L)}$ as measurements. The problems we have identified with both $\theta^{(S)}$ and $\theta^{(R)}$ are fairly easy to

TABLE 3
Additive Scaled Dimensions and Corresponding Variable Names in the Dataset

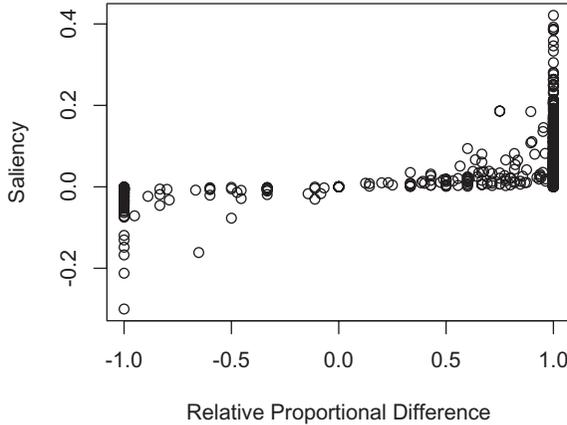
| Policy Dimension | "Left" Position | "Right" Position | Source | Variable Name |
|------------------------------|--|--|-----------------------|---------------|
| Free Market Economy | 401 Free Enterprise: Positive + | 403 Market Regulation: Positive + | (Proposed) | freemarket |
| | 402 Incentives: Positive | 412 Controlled Economy: Positive + | | |
| | | 413 Nationalisation: Positive + | | |
| | | 415 Marxist Analysis: Positive | | |
| Environmental Protection | 501 Environmental Protection: Positive + | 410 Productivity: Positive | (Proposed) | environment |
| | 416 Anti-Growth Economy: Positive | | | |
| State Involvement in Economy | 403 Market Regulation: Positive + | 401 Free Enterprise: Positive + | Benoit & Laver (2007) | stateconomy |
| | 404 Economic Planning: Positive + | 402 Incentives: Positive + | | |
| | 406 Protectionism: Positive + | 407 Protectionism: Negative + | | |
| | 412 Controlled Economy: Positive + | 414 Economic Orthodoxy: Positive + | | |
| | 413 Nationalisation: Positive + | 505 Welfare State Limitation: Positive | | |
| | 504 Welfare State Expansion: Positive + | | | |
| | 506 Education Expansion: Positive + | | | |
| State-provided Services | 701 Labour Groups: Positive | | (Proposed) | stateservices |
| | 504 Welfare State Expansion: Positive + | 505 Welfare State Limitation: Positive + | | |
| | 506 Education Expansion: Positive | 507 Education Limitation: Positive | | |

(continued on next page)

TABLE 3 (continued)

| Policy Dimension | “Left” Position | “Right” Position | Source | Variable Name |
|-----------------------------|---|---|-----------------------|---------------|
| Left-Right: CMP “Rile” | 103 Anti-Imperialism: Anti-Colonialism + 105 Military: Negative + 106 Peace: Positive + 107 Internationalism: Positive + 202 Democracy: Positive + 403 Market Regulation: Positive + 404 Economic Planning: Positive + 406 Protectionism: Positive + 412 Controlled Economy: Positive + 413 Nationalisation: Positive + 504 Welfare State Expansion: Positive + 506 Education Expansion: Positive + 701 Labour Groups: Positive | 104 Military: Positive + 201 Freedom and Human Rights: Positive + 203 Constitutionalism: Positive + 305 Political Authority: Positive + 401 Free Enterprise: Positive + 402 Incentives: Positive + 407 Protectionism: Negative + 414 Economic Orthodoxy: Positive + 505 Welfare State Limitation: Positive + 601 National Way of Life: Positive + 603 Traditional Morality: Positive + 605 Law and Order: Positive + 606 Social Harmony: Positive | Laver & Budge (1992) | logrile |
| Planned v. Market Economy | 403 Market Regulation: Positive + 404 Economic Planning: Positive + 412 Controlled Economy: Positive | 401 Free Enterprise: Positive + 414 Economic Orthodoxy: Positive | <i>MPP</i> | logplaneco |
| Social Liberal-Conservative | 103 Anti-Imperialism: Anti-Colonialism + 105 Military: Negative + 106 Peace: Positive + 107 Internationalism: Positive + 202 Democracy: Positive | 104 Military: Positive + 201 Freedom and Human Rights: Positive + 203 Constitutionalism: Positive + 305 Political Authority: Positive + 601 National Way of Life: Positive + 603 Traditional Morality: Positive + 605 Law and Order: Positive + 606 Social Harmony: Positive | Benoit & Laver (2007) | libcons |

FIGURE 1
Comparing the Relative Proportional Difference to the Saliency Scale for “National Way of Life” (categories 601 and 602)



illustrate by comparing them for a range of values across almost any scale. Notably, $\theta^{(R)}$ has a problem of reaching its limits for the extremes when $L > 0, R = 0$ or $R > 0, L = 0$. While the problem with $\theta^{(S)}$ is that it registers linear changes with each additional extreme-coded text unit in such situations, $\theta^{(R)}$ registers no changes at all. Hence, a manifesto registering five exclusively left text units would be the same as one registering 500. To demonstrate this using the CMP data, Figure 1 plots the relative proportional difference versus the “saliency” scalings of the confrontational pair for “National Way of Life: Positive/Negative” (categories 601 and 602). The vertically stacked points at the limits of the scale (at -1.0 and 1.0) show that additional mentions cause linear increases for the saliency scale, but no change for the relative proportional difference measure.

Several other problems with the existing scales also emerge from an inspection of Figure 1. First, because mentions of “national way of life” are relatively low in absolute frequency across manifestos, and because $(R - L)/N \leq (R + L)/N$, the low frequency of these statements relative to all other statements severely shrinks $\theta^{(S)}$ toward zero. The saliency measure is insensitive to changes for policy dimensions with low absolute frequency and misleadingly assigns a difference score close to the zero point. While we have shown this here for only the national way of life categories 601 and 602, it also applies to the CMP’s biggest scale, Rile, encompassing

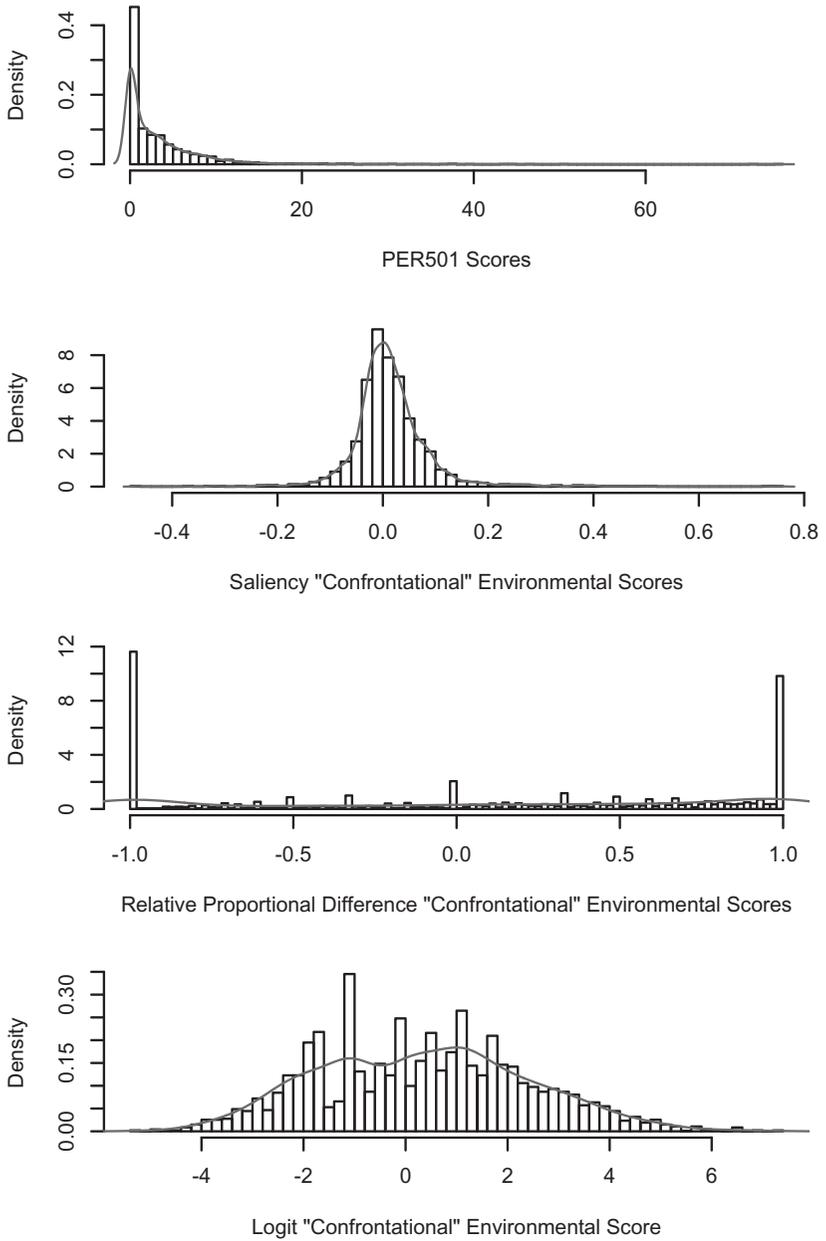
26 categories in all. While in theory this scale runs from -1.0 to 1.0 (as a proportion), in practice the range spans only from $-.5$ to $.5$ for almost every manifesto measured.

A second problem, again with $\theta^{(S)}$ can be seen at the extremes defined as $L > 0, R = 0$ for a left extreme, or $R > 0, L = 0$ for a right extreme. Extremity on the saliency measure $\theta^{(S)}$ increases at a linear rate with each additional text unit in the extreme category. Substantively, the suggestion is that the same change occurs when the extreme-only category text units increase from 5 to 10 units, as when it increases from 105 to 100 units. This assumption of linear change in position given observed text unit counts is neither sensible nor supported by perceptual theory (see our discussion of the Weber-Fechner law above).

For each of these problems we have identified in $\theta^{(S)}$, a corresponding problem can also be found in $\theta^{(R)}$. The middle-range problem of lack of sensitivity for $\theta^{(S)}$ is exactly reversed in $\theta^{(R)}$: small differences between R and L become highly influential on $\theta^{(R)}$ when these are scaled as ratios of relative content $R + L$. An extreme example makes the point: imagine a series of manifestos from a party that had no real interest whatsoever in, and effectively no position on, protectionism. Irrelevant stochastic factors in text generation, or in the coding of the text, could plausibly result in a few essentially random counts of text units into each of the protectionism categories. The effect on $\theta^{(R)}$ will be drastic in this situation, massively leveraging the error because it is only concerned with relatively proportional content.

If we compare the overall distribution of data for one of our composite scales—the environmental policy scale described above—we can see a fairly stark contrast between the spread of values for different scales that reinforces the patterns we observed in examining the National Way of Life scale. In Figure 2, we compare the distribution of scores for “PER501 Environmental Protection: Positive” to a “confrontational” scale constructed from opposing categories. Our new scale of environmental protection is based on adding the two proenvironmental categories “Antigrowth Economy: Positive” (PER416) and “Environmental Protection: Positive” (PER501) as capturing antigrowth politics and “ecologism” and contrasting this with the environmentally opposed paradigm of economic growth, represented in the CMP by the category “Productivity: Positive” (PER410). Figure 2 not only shows the better dispersion of the logit scale, but also demonstrates anew the problems we have already seen: bunching around zero of the saliency scale, as well as the

FIGURE 2
Distributions of Scales for Proposed Environmental Scales



bunching around the extremes of the relative proportional difference scale.

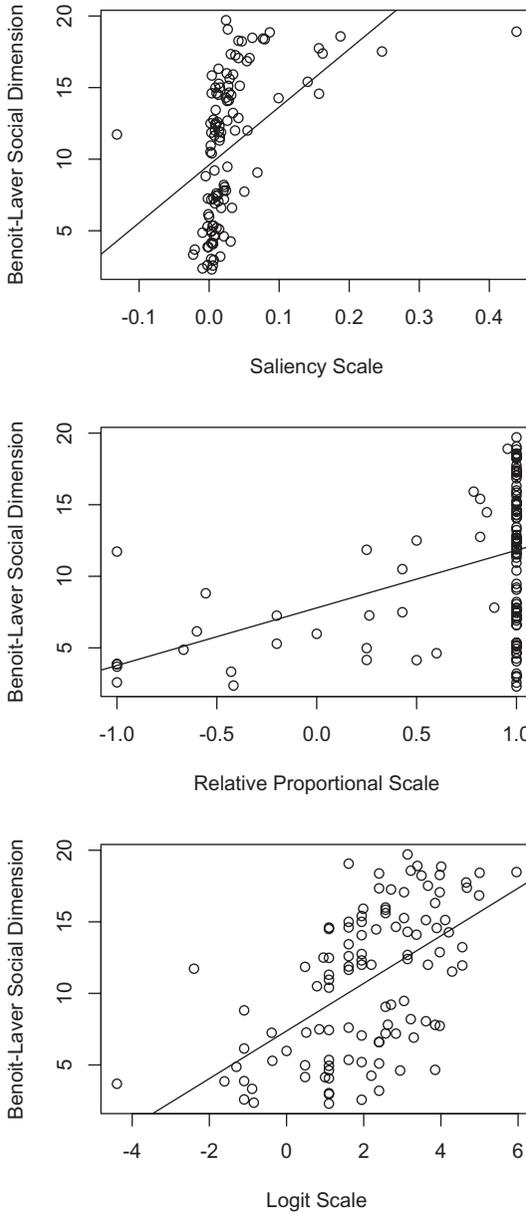
Comparisons to Expert Surveys of Policy

Up to this point we have only compared one scale with another. To judge more conclusively whether a particular scale measures what we hope it measures, we can compare the CMP-based scales to independent, external measures of party positions based on expert surveys. Expert surveys such as Benoit and Laver (2006) have been shown to provide valid and reliable measures of party policy positions, but existing measures are limited in their time frame to the two decades since 1990. Only text-based measures such as the CMP have the potential to provide valid estimates of policy positions going further back in time. Limited comparisons of expert survey estimates to CMP measures were conducted by Benoit and Laver (2007), who tested the saliency-based Rile measures against expert survey ratings of left-right from Benoit and Laver (2006) and found a high correlation and lack of bias between the two measures. Because the large number of categories tends to wash out differences in large additive indexes such as Rile, here we perform the same comparison using smaller, more policy-specific scales.

We have compared the CMP-based indexes to the Benoit and Laver (2006) expert survey estimates of party position on the issue of social liberalism, one of two fundamental axes of political competition (the other being economic left-right) on which they place parties in every country. Some variant of this noneconomic dimension has been identified as a distinct, basic axis of political competition in numerous studies (e.g., Inglehart 1984; Marks, Wilson, and Ray 2002). Figure 3 plots the Benoit and Laver social liberalism dimension scores against each of the three scales based on counts of “604/603 Traditional Morality: Negative/Positive.”

The patterns from the plots are consistent with the interscale comparisons examined earlier. The saliency scale is highly bunched around zero, suppressing variation even when huge differences are identified by the Benoit and Laver scores. The relative proportional scale in the middle panel shows spotty variation in the middle ranges, with a very high proportion of values at the right side where Benoit and Laver indicate a complete range of differences but the relative proportional scale has reached its maximum value. Finally, the logit scale looks approximately linear, has no bunching at the extremes, or dispersed points in the middle. Its scale is centred to

FIGURE 3
Comparison of CMP Scales of Traditional Morality with Benoit and Laver (2006) Social Liberalism Dimension



the right of zero, reflecting the higher proportion of text units of “Traditional Morality: Positive” (and many exclusively so), but this does not perturb the scale’s linear relationship with the expert survey scores. Residual analysis suggests that the relationship between expert survey scores and $\theta^{(L)}$ are both linear and homoskedastic.¹⁵

Finally, we perform the comparison with one of our simple additive scales that is not strictly constructed from a bipolar pair. This is the dimension of environmental protection, where the “left” side has two components. This is also an interesting category for comparison since, as previously discussed, the CMP’s saliency approach identifies only one possible side to this issue. In Figure 4 we compare the CMP’s default single-category scale of 501, to the relative absolute proportional difference scale constructed as per Table 3, and to the corresponding logit scale. The two versions of the saliency scale (per501 and the difference scale we have proposed) are not particularly poor, although 501 is clearly bounded to be positive, with several values where the zero boundary suggests anti-environmental policies when in fact the party said nothing at all about the environment. The saliency difference scale (middle plot) is left-skewed with some extreme proenvironmental values skewing the pattern. The new scale (bottom plot), however, shows a much better behaved linear relationship with the Benoit and Laver scores, without the skew from the saliency or relative proportional difference plots. The logit scale suggests the more sensible conclusion that a manifesto containing 40% proenvironmental sentences is not 10 times more proenvironment than a manifesto with just 5%, but rather only about 2.3 times more. And by comparison to the one-sided policy issue approach suggested by pure saliency theory, the comparison in Figure 4 also reinforces our earlier argument that even seemingly one-sided issues perform better when recast in terms of confrontationally opposite categories.

As a final form of external validation, we also compare our suggested importance scale to the separately measured policy importance estimates from Benoit and Laver (2006). These are plotted, for four directly comparable measures, in Figure 5. The positive linear relationship for these scales suggests that the proposed measure of importance based on total mentions on either side of an issue do indeed form a valid indicator of the political importance of this issue to a party. Our proposed importance measure provides a scale for importance that is more valid linguistically, based on logarithmically increasing extremity. It also unlocks a general measure of importance

FIGURE 4
Comparison of CMP Scale of Environment with Benoit and Laver
(2006) Environmental Dimension

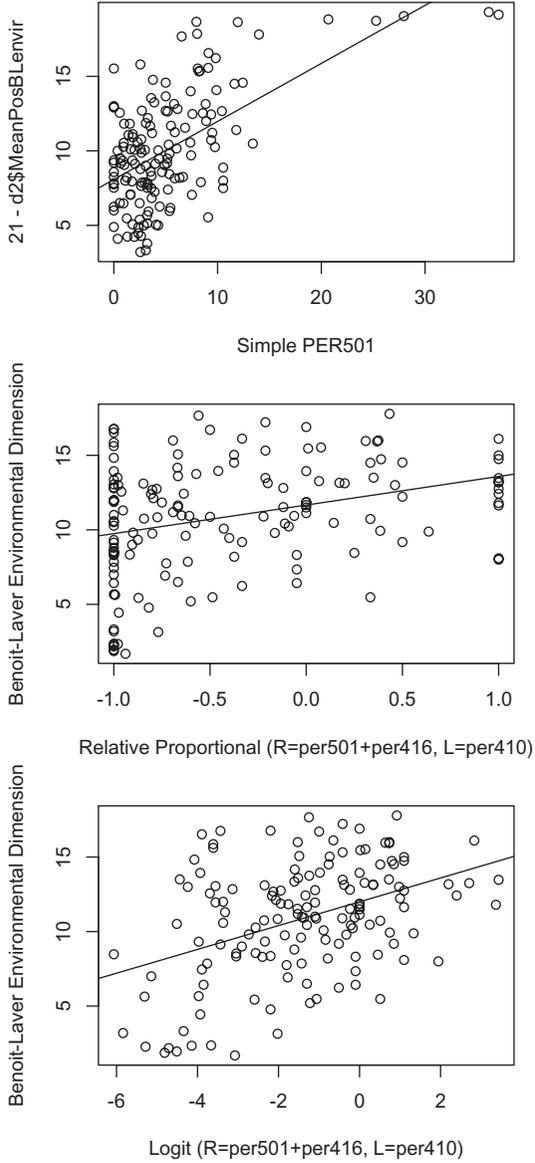
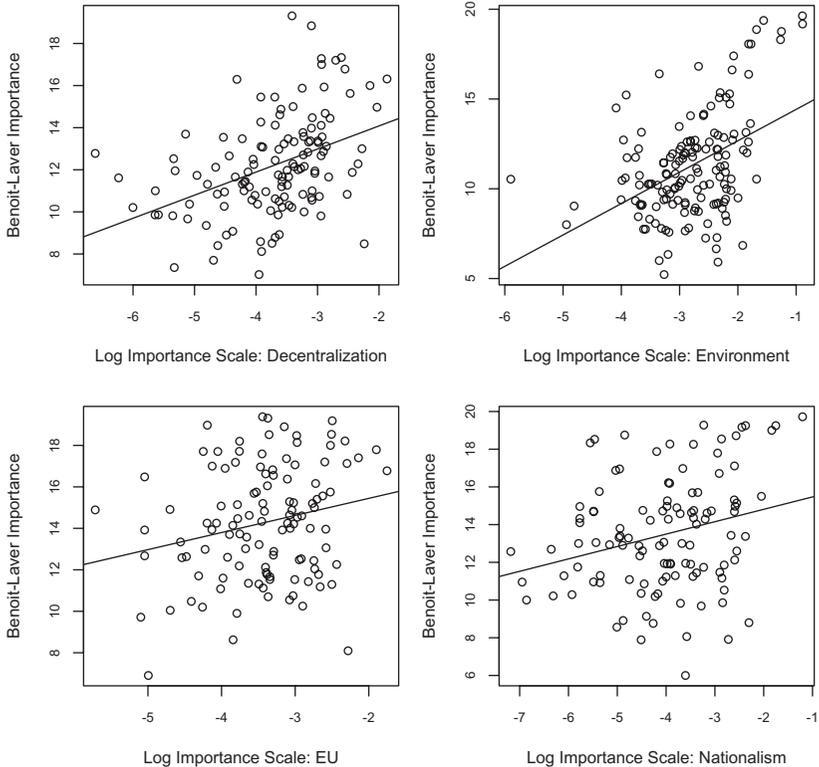


FIGURE 5
 Comparison of Proposed Log Importance Scale
 with Benoit and Laver (2006) Importance Measures
 (only for cases where total *R* and *L* mentions are non-zero)



from the CMP data that has never before been made systematically available.

Recommendations

By focusing attention on producing better measures of party policy positions over time, as well as introducing new measures of party policy, our study should contribute to new developments in the field of legislative studies, especially the study of legislatures in multiparty settings. To “understand what a legislature does (and why it

does it) we need to know the policy preferences of its members” (Loewenberg 2009, 415). This need for data becomes all the more interesting at the party level in contexts with multiparty governments, coalitions, and high party discipline.

Our conclusions can be summarized as follows. First, our analysis of the use of the logit scale to estimate left-right positions from counts of textual categories, as well as our demonstration through direct comparison to other scales as well as to independent external data, suggests that the logit scale is superior and should be used in place of the “saliency” and “relative proportional difference” approaches used previously. We recommend using the logit scale for all policy categories and have provided a set of 21 such scales (in Tables 1 and 3) that can be constructed directly from the existing CMP dataset. In addition, we have calculated uncertainty estimates for all quantities using the simulation method proposed by Benoit, Laver, and Mikheylov (2009). This dataset is available immediately and offers a superior alternative to the estimates supplied by the CMP, estimates that we have shown are based on the inferior saliency-based scales, and with few exceptions not constructed in the confrontational pairing approach we recommend here.

Second, we have proposed a new and separate measure of policy importance that is consistent with our logit scale of position and demonstrated that this proposed scale correlates well with independent, external measures of policy importance from expert surveys. These importance estimates are also provided with accompanying uncertainty measures.

Finally, we have shown that the assumption that individual parties take only one side, and indeed that all parties take the same side, of an issue, is demonstrably false, even given the CMP’s own dataset. For our purposes, this implies a critique of the basic CMP coding scheme, since the existing scheme consists of a mixture of confrontational and saliency-based categories. Our analysis suggests that any revision of the coding scheme would complete the step toward a fully confrontational coding scheme, consisting only of opposing, *pro* and *contra* categories. It would also be possible to go one step further and include a neutral category for each confrontational policy scale, which could be ignored when computing $\theta^{(L)}$ but counted when considering $\theta^{(I)}$. This would address the concerns of McDonald and Mendes (2001b) about the nonreflection of neutral stances in the positional scales, as well as better reflecting overall policy importance based on counting text units.

Will Lowe <w.lowe@maastrichtuniversity.nl> is Assistant Professor in Research Methods, Department of Political Science, Maastrich University, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Kenneth Benoit <kbenoit@lse.ac.uk> is Professor and Director of the Methodology Institute, London School of Economics, Columbia House, Houghton Street, London WC2A 2AE, United Kingdom. Slava Mikhaylov <v.mikhaylov@ucl.ad.uk> is a Lecturer in Political Science, University College London, The Rubin Building, 29/30 Tavistock Square, London, WC1H 9QU, United Kingdom. Michael Laver <michael.laver@nyu.edu> is Professor of Politics, New York University, 19 W. 4th Street, New York, NY 10012.

APPENDIX: DATASET DESCRIPTION

The data described in this paper are available for download from <http://www.kenbenoit.net/cmp/scales/>. The dataset contains all of the variables described in Tables 1 and 3, with suffixes indicating whether the variable refers to position, the 95% confidence interval on position, the estimate of importance, and the standard errors of position and importance. For the variable protectionism (for example), the five associated variables are:

| | |
|---------------------|--|
| protectionism | position estimate |
| protectionism_lo95 | 95% lower bound for position estimate |
| protectionism_hi95 | 95% upper bound for position estimate |
| protectionism_SE | standard error for position estimate |
| protectionism_imp | importance estimate |
| protectionism_impSE | standard error for importance estimate |

NOTES

This paper was originally prepared for presentation at the 2008 ECPR General Conference, Potsdam. We thank Thomas Däubler and Jonathan Slapin for comments on multiple drafts of this manuscript. This research was supported in part by the Irish Research Council for Humanities and the Social Sciences.

1. Details may be found in Table 3. We return to this scale later in the text.

2. For the initial development we treat each policy area as defined by one “left” and one “right” CMP category. In fact neutral categories are also possible, and in some cases it is helpful to aggregate more than one CMP category to generate a substantively appropriate left of right count.

3. More precisely, the CMP's saliency-based scale multiplies θ by 100 to allow interpretation as a percentage.

4. The percentage of uncodeable content in the average manifesto in the CMP combined dataset is 6.8%, making the inclusion of uncoded content a real worry for many texts.

5. More complex nonlinear marginal effects that decrease the effect of early as well as later mentions have long been suggested (e.g., Jakobovits and Lambert 1963; Jakobovits and Hogenraad 1967) and are intuitively reasonable for manifesto data, perhaps as part of an interaction with issue salience. Investigating this relationship is further work.

6. Later research (Stevens 1957) has established a range of power law relationships between physical and subjective magnitudes in different modalities, not all of which exhibit decreasing marginal effects. Nevertheless we work here with the logarithmic relationship because of its simplicity, its linguistic motivation as sketched above, and most importantly, its excellent fit to policy positions generated independently by experts, considered below.

7. The scale is given units by reference to a barely perceptible reference sound.

8. In practice, however, the logit scales applied to the CMP data ranges from approximately -7 to $+7$, since few R or L categories (or indeed, N) tend to exceed $\log(1000) = 6.9$.

9. In applications we follow standard statistical practice in the analysis of contingency tables and add 0.5 to L and R , (Agresti 1996). This can be motivated as a means to reduce estimation bias (Firth 1993) or to provide better behaved interval measures of uncertainty (Brown, Cai, and DasGupta 2001).

10. Statistical latent variable modelers (e.g., Clinton, Jackman, and Rivers 2004; Slapin and Proksch 2008) make the same observation by noting that the zero point, direction, and units of the measurements are model identification constraints not substantive assertions about position.

11. In the framework of parametric models, $\theta^{(L)}$ could be seen as subpart of a multinomial logistic regression model of the category counts $[R, L, N - (R + L)]$ in party platform, where $N - (R + L)$ is the number of sentences assigned to other categories or left uncoded. Using L as a base category, $\theta^{(L)}$ will, as N increases, approximate the first linear predictor in such a model.

12. In the saliency theory approach, policy dimensions are assumed to consist of issue areas or clusters of issues (Robertson 1976, 61).

13. In Danish: "Miljøpolitikken [environmental policy] måske stille danske virksomheder dårligere, end virksomhederne i de lande vi konkurrerer med" (Venstre Manifesto 1988). We thank Jacob Rathlev for this suggestion and Martin Hansen for drawing attention to this example and for help with the translation.

14. This reflects the definition of one of the core four dimensions in the expert survey in Benoit and Laver (2006, 129).

15. Similar patterns are easily observed for numerous other scales, such as Multiculturalism: Positive/Negative against the Benoit and Laver scores for nationalism and immigration.

REFERENCES

- Agresti, A. 1996. *An Introduction to Categorical Data Analysis*. New York: Wiley.
- Alemán, E., E. Calvo, M.P. Jones, and N. Kaplan. 2009. "Comparing Cosponsorship and Roll-Call Ideal Points." *Legislative Studies Quarterly* 34 (1): 87–116.
- Bara, J., A. Weale, and A. Biquelet. 2007. "Analysing Parliamentary Debate with Computer Assistance." *Swiss Political Science Review* 13 (4): 577–605.
- Baumgartner, F.R., C. Green-Pedersen, and B.D. Jones. 2008. *Comparative Studies of Policy Agendas*. London: Routledge.
- Benoit, K., and M. Laver. 2003. "Estimating Irish Party Positions using Computer Wordscoring: The 2002 Elections." *Irish Political Studies* 17 (2): 97–107.
- Benoit, Kenneth, and Michael Laver. 2006. *Party Policy in Modern Democracies*. London: Routledge.
- Benoit, Kenneth, and Michael Laver. 2007. "Estimating Party Policy Positions: Comparing Expert Surveys and Hand Coded Content Analysis." *Electoral Studies* 26 (1): 90–107.
- Benoit, Kenneth, Michael Laver, and Slava Mikhaylov. 2009. "Treating Words as Data with Error: Uncertainty in Text Statements of Policy Positions." *American Journal of Political Science* 53 (April): 495–513.
- Brown, L.D., T.T. Cai, and A. DasGupta. 2001. "Interval Estimation for a Binomial Proportion (with discussion)." *Statistical Science* 16: 101–33.
- Budge, Ian. 1994. "A New Spatial Theory of Party Competition: Uncertainty, Ideology and Policy Equilibria Viewed Comparatively and Temporally." *British Journal of Political Science* 24 (4): 443–67.
- Budge, Ian. 1999. *Estimating Party Policy Preferences: From Ad Hoc Measures to Theoretically Validated Standards*. Number 139 in "Essex Papers in Politics and Government." University of Essex and Department of Government.
- Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum. 2001. *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford: Oxford University Press.
- Carroll, Royce, Jeffrey B. Lewis, James Lo, Keith T. Poole, and Howard Rosenthal. 2009. "Comparing NOMINATE and IDEAL: Points of Difference and Monte Carlo Tests." *Legislative Studies Quarterly* 34 (4): 555–91.
- Carrubba, C., M. Gabel, and S. Hug. 2008. "Legislative Voting Behavior, Seen and Unseen: A Theory of Roll-Call Vote Selection." *Legislative Studies Quarterly* 33 (4): 543–72.
- Clinton, J., S. Jackman, and D. Rivers. 2004. "The Statistical Analysis of Roll Call Voting: A Unified Approach." *American Journal of Political Science* 98 (2): 355–70.
- Clinton, J.D., and S. Jackman. 2009. "To Simulate or NOMINATE?" *Legislative Studies Quarterly* 34 (4): 593–621.
- Efron, Bradley, and Robert Tibshirani. 1994. *An Introduction to the Bootstrap*. New York: Chapman and Hall/CRC Hall.
- Elff, M. 2008. "A Spatial Model of Electoral Platforms." POLMETH Working Paper.

- Fechner, G.T. 1965. Elemente der Psychophysik, Sections VII and XVI. In *A Source Book in the History of Psychology*, ed. R.J. Herrnstein and E.G. Boring. Harvard University Press, 66–75.
- Firth, D. 1993. “Bias Reduction of Maximum Likelihood Estimates.” *Biometrika* 80 (1): 27–38.
- Fleiss, Joseph L., B. Levin, and M.C. Paik. 2003. *Statistical Methods for Rates and Proportions*. 3d ed. New York: John Wiley.
- Gabel, Matthew, and John Huber. 2000. “Putting Parties in Their Place: Inferring Party Left-Right Ideological Positions from Party Manifesto Data.” *American Journal of Political Science* 44: 94–103.
- Gemenis, K., and E. Dinas. 2010. “Confrontation Still? Examining Parties’ Policy Positions in Greece.” *Comparative European Politics* 8: 179–201.
- Grofman, Bernard. 2004. “Downs and Two-Party Convergence.” *Annual Review of Political Science* 7: 25–46.
- Hilliard, D., S.J. Purpura, and S. Wilkerson. 2007. “Computer Assisted Topic Classification for Mixed Methods Social Science Research.” *Journal of Information Technology and Politics* 4 (4): 31–46.
- Hix, S., and A. Noury. 2009. “After Enlargement: Voting Patterns in the Sixth European Parliament.” *Legislative Studies Quarterly* 34 (2): 159–74.
- Hooghe, Liesbet, Ryan Bakker, Anna Brigeveich, Catherine de Vries, Erica Edwards, Gary Marks, Jan Rovny, and Marco Steenbergen. 2008. “Reliability and Validity of Measuring Party Positions: The Chapel Hill Expert Surveys of 2002 and 2006.” University of North Carolina, Chapel Hill manuscript.
- Hopkins, D., and G. King. 2010. “A Method of Automated Nonparametric Content Analysis for Social Science.” *American Journal of Political Science* 54 (1): 229–47.
- Inglehart, Ronald. 1984. The Changing Structure of Political Cleavages in Western Society. In *Electoral Change, Realignment and Dealignment in Advanced Industrial Democracies*, ed. Russell Dalton et al. Princeton: Princeton University Press, 25–69.
- Jakobovits, Leon A., and R. Hogenraad. 1967. “Some Suggestive Evidence on the Operation of Semantic Generation and Satiation in Group Discussions.” *Psychological Reports* 20: 1247–50.
- Jakobovits, Leon A., and W.E. Lambert. 1963. The Effects of Repetition in Communication on Meanings and Attitudes. In *Television and Human Behavior*, ed. L. Arons and M.A. May. New York: Appleton-Century-Crofts, 167–76.
- Jeffreys, H. 1946. “An Invariant Form for the Prior Probability in Estimation Problems.” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences* 186 (1007): 453–61.
- Kim, Heemin, and Richard C. Fording. 2002. “Government Partisanship in Western Democracies, 1945–1998.” *European Journal of Political Research* 41 (2): 187–206.
- Klemmensen, R., S.B. Hobolt, and M.E. Hansen. 2007. “Estimating Policy Positions using Political Texts: An Evaluation of the Wordscores Approach.” *Electoral Studies* 26 (4): 746–55.

- Klingemann, Hans-Dieter, Andrea Volkens, Judith Bara, Ian Budge, and Michael McDonald. 2006. *Mapping Policy Preferences II: Estimates for Parties, Electors, and Governments in Eastern Europe, European Union and OECD 1990–2003*. Oxford: Oxford University Press.
- Krippendorff, Klaus. 2004. *Content Analysis: An Introduction to Its Methodology*. 2d ed. Thousand Oaks, CA: Sage.
- Laver, M., and J. Garry. 2000. “Estimating Policy Positions from Political Texts.” *American Journal of Political Science* 44 (3): 619–34.
- Laver, Michael. 2001. *Estimating the Policy Position of Political Actors*. New York: Routledge.
- Laver, Michael, and Ben W. Hunt. 1992. *Policy and Party Competition*. New York: Routledge.
- Laver, Michael, and Ian Budge. 1992. *Party Policy and Government Coalitions*. New York: St. Martin’s Press.
- Laver, Michael, Kenneth Benoit, and John Garry. 2003. “Estimating the Policy Positions of Political Actors using Words as Data.” *American Political Science Review* 97 (2): 311–31.
- Loewenberg, G. 2008. “Introduction.” *Legislative Studies Quarterly* 33 (4): 499–500.
- Loewenberg, G. 2009. “Introduction.” *Legislative Studies Quarterly* 34 (4): 451–53.
- Lowe, W. 2008. “Understanding Wordscores.” *Political Analysis* 16 (4): 356–71.
- Marks, Gary, Carole Wilson, and Leonard Ray. 2002. “National Political Parties and European Integration.” *American Journal of Political Science* 46 (3): 585–94.
- Martin, L.W., and G. Vanberg. 2007. “A Robust Transformation Procedure for Interpreting Political Text.” *Political Analysis* 16 (1): 93–100.
- McDonald, Michael, and Silvia Mendes. 2001a. “Checking the Party Policy Estimates: Convergent Validity.” In *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*, ed. Ian Budge, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, and Eric Tanenbaum. Oxford University Press, 127–41.
- McDonald, Michael, and Silvia Mendes. 2001b. “The Policy Space of Party Manifestos.” In *Estimating the Policy Position of Political Actors*, ed. Michael Laver. London: Routledge, 90–114.
- Monroe, B., and K. Maeda. 2004. “Talk’s Cheap: Text-Based Estimation of Rhetorical Ideal-Points.” POLMETH Working Paper.
- Monroe, B.L., K.M. Quinn, and M.P. Colaresi. 2008. “Fightin’ Words: Lexical Feature Selection and Evaluation for Identifying the Content of Political Conflict.” *Political Analysis* 16 (4): 372–403.
- Neuendorf, K.A. 2002. *The Content Analysis Guidebook*. Thousand Oaks CA: Sage.
- Newcombe, R.G. 2001. “Logit Confidence Intervals and the Inverse Sinh Transformation.” *American Statistician* 55: 200–02.
- Pennings, P., and H. Keman. 2002. “Towards a New Methodology of Estimating Party Policy Positions.” *Quality and Quantity* 36 (1): 55–79.
- Popping, Roel. 2007. “Text Analysis for Knowledge Graphs.” *Quality and Quantity* 41: 691–709.

- Quinn, K.M., B. Monroe, M. Colaresi, M. Crespin, and D. Radev. 2010. "How to Analyze Political Attention with Minimal Assumptions and Costs." *American Journal of Political Science* 54 (1): 209–28.
- Ray, Leonard. 2007. "Validity of Measured Party Positions on European Integration: Assumptions, Approaches, and a Comparison of Alternative Measures." *Electoral Studies* 26 (1): 11–22.
- Riker, William H. 1996. *The Strategy of Rhetoric: Campaigning for the American Constitution*. New Haven, CT: Yale University Press.
- Robertson, David. 1976. *A Theory of Party Competition*. London and New York: Wiley.
- Saiegh, S. 2009. "Recovering a Basic Space from Elite Surveys: Evidence from Latin America." *Legislative Studies Quarterly* 34 (1): 117–45.
- Schickler, E., and K. Pearson. 2009. "Agenda Control, Majority Party Power, and the House Committee on Rules, 1937–52." *Legislative Studies Quarterly* 34 (4): 455–91.
- 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): 705–22.
- Stevens, S.S. 1957. "On the Psychophysical Law." *Psychological Review* 64 (3): 153–81.
- van Atteveldt, W., J. Kleinnijenhuis, and N. Ruijgrok. 2008. "Parsing, Semantic Networks, and Political Authority Using Syntactic Analysis to Extract Semantic Relations from Dutch Newspaper Articles." *Political Analysis* 16 (4): 428–46.
- Yu, B., S. Kaufmann, and D. Diermeier. 2008. "Classifying Party Affiliation from Political Speech." *Journal of Information Technology and Politics* 5 (1): 33–48.