

PARTY POLICY IN MODERN DEMOCRACIES

Kenneth Benoit
Trinity College, Dublin
kbenoit@tcd.ie

Michael Laver
New York University
ml127@nyu.edu

|

April 11, 2006
(For publication by Routledge)

FINAL VERSION

Chapter 4

Measuring Policy Positions

In the Part I of this book, we explored theoretical issues generated by describing political competition in terms of “policy spaces” and surveyed different means of mapping these spaces into the world of real political actors. Set in this context, the primary purpose of this book is to provide a comprehensive, accurate and reliable database of the policy positions of political parties. The previous chapter outlined a range of general issues relating to the measurement of empirical policy spaces and described the advantages of the expert survey method. In this chapter, we describe our own use of expert surveys to estimate both the locations of political parties on key dimensions of policy, and the relative importance of each dimension for each party.

DESIGNING THE EXPERT SURVEY

Our approach to measuring empirical policy spaces was to conduct a set of systematic surveys, one survey for each political system under investigation, of specialists on the politics of the country concerned. The objective was to collect the information required to locate all politically relevant political parties on a wide range of policy dimensions associated with party competition, in as many countries as possible with a tradition of free elections and competitive party systems. The surveys were conducted from 2002-2003, covered 47 countries, resulted in 1491 valid expert responses, locating 387 different political parties on scales relating to a total of 37 unique policy dimensions. This section describes key operational decisions determining the design of our surveys. In Appendix A, we discuss more detailed aspects of our survey methodology and describe diagnostic tests we performed on the dataset we generated.

Choosing countries

As with its predecessor (Laver and Hunt 1992), our study was designed to cover as many countries as possible in which competitive elections are held. We were able to include 47 democracies from four continents, far more comprehensive coverage than any other survey of its type. Our survey did not include countries from Latin America, Africa, or Asia (outside of Japan). The decision to exclude these regions was mainly a practical consequence of limited resources. Given the reality that it would not be feasible to include all of the world’s democracies in our survey—some 121 of the world’s 192 governments in 2003, according to the Freedom House survey (Freedom House (U.S) 2003)—we made the decision to cover one region, Europe, as thoroughly as possible. Our primary focus was thus on European democracies, covering every country from the Iceland to Russia, plus Turkey and Israel.

Five countries outside of Europe, broadly defined, were also surveyed: the long-standing English-speaking democracies of United States, Canada, Australia and New Zealand, as well as Japan. While almost all of the political systems we surveyed were independent states, there were some exceptions. In the United Kingdom, for instance, we surveyed Northern Ireland and Britain separately. The Serbian region of Serbia and Montenegro was also surveyed separately from its federation.

Because of the decision to cover all, and not just the major, countries of Europe, our survey mapped policy spaces that have never previously been explored. Our placements of parties in countries such Albania and the former republics of Yugoslavia, for example—especially Macedonia and Moldova—provide data on countries usually excluded from cross-national data-gathering exercises. Table A1 in Appendix A provides the full listing of the countries we surveyed.

Identifying experts

The first step in each survey was to identify the population of experts whose judgments we wished to aggregate. Since the quality of expert survey data is directly influenced by the quality of the expert panel, this selection must be performed very carefully. The typical expert in our survey was an academic specializing in political parties and electoral politics of his or her country. These people were generally familiar with surveys and most would previously have participated, conducted or employed the results from, surveys of some form or other dealing with politics in their areas of expertise. Our method for identifying experts was the following. First, we contacted the national political science association of the country, if such an organization existed, with a request for their membership lists.¹ When no political science association existed, as was the case in many countries, especially formerly post-communist states, we used a “snowball” strategy where we started with a very short list of well-known experts from the country, and asked each to name as many additional experts as possible. We then contacted each of those experts, and asking them to name in turn as many additional experts as possible. In addition, we scoured the lists from universities and non-profit organizations in search of additional experts. The vast majority of our experts were drawn from academia or research institutes and, with extremely few exceptions, excluded both journalists and political actors.

Identifying political parties

The parties we asked experts to locate included all that were politically relevant, a definition that included parties meeting any one of three criteria. First, we automatically included every existing²

¹ When such list included very large numbers of members, such as that of the American Political Science Association, we selected members according to their self-declared areas of expertise—specialists in American politics from the U.S. survey, for instance.

² The qualification “existing” is required since some parties who may have won seats or votes at the most recent election had changed or ceased to exist by the time of our survey.

national political party that won seats in the national legislature at the country's most recent election. Second, we included every existing national party that had won at least one percent of the vote nationally at the country's most recent election. Finally, we included any other parties that local experts informed us were politically relevant despite not meeting the other two criteria.³ Most of the parties included using the third criterion were found in post-communist contests where party systems were not fully consolidated or were undergoing rapid change. In such cases we relied on the advice of local experts to guide in the definition of political relevance.

The decision to represent all politically relevant parties generated a long list of parties in some countries, requiring experts to locate up to a dozen parties or more. For instance, experts were asked to locate the policy positions of 13 parties in Slovakia, Italy, and Bosnia, 12 in Israel and Poland, and 11 parties in six other countries. The list of parties in a few countries was quite small – two in the United States and three in Cyprus – but the typical number was much higher. The median number of parties, over the 47 party systems we covered, was eight. We included as many parties as we feasibly could because even small parties can be crucial for the analysis of political competition. It is quite common for parties winning tiny seat shares to play important roles in forming governments, and/or other key roles in important aspects of party competition. Parties' electoral fortunes also fluctuate, sometimes quite wildly, between elections; excluding smaller parties based on vote or seat share at one point in time could lead us to exclude parties playing significant roles in party competition over a longer period. Our answer, therefore, was to err on the side of caution and include as many parties as might conceivably be relevant.⁴

Identifying policy dimensions

Survey instruments of any sort require careful design, and expert surveys are no exception. The questions of interest must be clearly identified and item wordings must be carefully chosen to elicit valid responses on these questions. The *a priori* nature of our expert survey method for measuring policy positions means that the questions of interest concern the policy dimensions on which parties are to be located. Selecting these dimensions and defining them carefully are thus key aspects of our research design, which reflects several importance choices.

First, we made the decision to measure party policy positions in *high-dimensional spaces*. The main aim in our study was to measure party positions on policy on as many separate dimensions as

³ On the advice of German experts, for instance, we included five parties in Germany that did not meet the first two criteria, including the far-right National Democratic Party that had strong showings in Saxony and several other Lander.

⁴ In a few cases, this strategy meant that we asked experts to place parties that were on the verge of extinction. In our survey of Canada in late 2003 and early 2004, for instance, experts located both the Canadian Alliance and the Progressive Conservative parties, despite the decision of these two parties in December 2003 to disband and integrate into a new party called the Conservative Party of Canada (not included in our survey). Likewise in France, our 2002 survey included the *Union pour la démocratie Française* (UDF), despite the contemporaneous dissolution of this party and its reformation into the UEM.

our local experts deemed politically relevant for any given country. The downside of this decision was to increase the length and complexity of our survey instrument, potentially affecting reliability. (We report checks for this potential effect in Appendix A and the results imply that the effect was present but very small.) The upside was to obtain separate locations for political parties on wide range of policy dimensions whose substantive meanings were deemed separable. This approach sharply distinguishes our study from other expert surveys of party policy, since these typically bundle policy positions into lower dimensional spaces. Castles and Mair for instance, asked experts to locate parties on a single left-right scale (Castles and Mair 1984). This approach was also followed by Inglehart and Huber (1995), although they also allowed respondents to include one other dimension of the respondent's own choice. Marks *et al.* on the other hand, asked experts to locate parties on two dimensions, one dealing with socio-economic left-right positions and the other with a contrast between "traditional/authoritarian/nationalist" and "green/alternative/liberal" values (Marks *et al.* Forthcoming). All of these studies measure policy positions in spaces determined *ex ante* to be low-dimensional. Our approach, in stark contrast, was to estimate party policy positions on the most elemental dimensions that local experts deemed validly measurable; no valid configuration of policy positions is excluded because the underlying dimensions were bundled together *by design*. This is of course a quite distinct matter from the *empirical observation* in some particular setting that some policy dimensions may tend in practice to be bundled together in practice, in the sense that party positions on these dimensions are highly correlated.

A second important feature of our expert survey is that it *deployed a core set of policy dimensions in every country surveyed*. The great advantage of this is to collect observations using directly comparable policy scales for all the countries studied. This enables the possibility of cross-national comparisons, something that within-country inductive methods such as roll-call vote analysis or factor analysis do not allow. The potential disadvantage of this approach is that, in the attempt to devise policy dimensions that can be applied to all party contents, we risk ending up with a set of scales so generic that they do not apply meaningfully to any particular national context. The policy dimension contrasting moral liberalism and conservatism, for example, referred to party positions "on matters such as abortion, homosexuality, and euthanasia." In pre-testing of the set of dimensions to be deployed in each country, many experts told us this question only partially applies to "their" country. The issue of abortion, for instance, was virtually absent from the political discourse of parties in Hungary and some other countries, although homosexuality (issues such as gay marriage) was occasionally discussed. Similarly, gay rights might no longer be a live political issue in some Scandinavian countries or the Netherlands.⁵ In some countries, such as the United States, all three

⁵ In Belgium, abortion continues to feature in policy discourse, although euthanasia does not. Belgium reported 347 legal acts of voluntary euthanasia in 2004 (*Economist*, October 15, 2005, p40) generating little or no controversy, while a single act of euthanasia in the United States—the withdrawal of feeding tubes from Terry Schiavo—split the nation in half. Switzerland is another country where euthanasia—legal since 1942—is not an issue of political controversy.

issues defining the “liberalism vs conservatism” scale were salient and divisive. The temptation in tailoring policy dimensions to each country is to include in the question wording only issues known to be salient in that context. The strong reason not to do this, however, is that it compromises scale comparability when data are used across countries or even across time in the same country. Our solution to this problem was to err on the side of comparability, maintaining a common dimension text wording to apply to all political parties in all countries for the *core* set of policy dimensions investigated. We made only very rare exceptions to this, at the strong insistence of local experts.⁶

A hard core of four substantive policy dimensions was thus deployed in every country covered in our survey; a fifth was deployed almost as widely. The core four were:

- economic policy (interpreted as the trade-off between lower taxes and higher public spending);
- social policy (interpreted as policies on matters such as abortion, gay rights, and euthanasia);
- the decentralization of decision making;
- environmental policy (interpreted as the trade-off between environmental protection and economic growth).

The fifth dimension, deployed in almost all countries, concerned state involvement in economic regulation, interpreted in non-post-communist states as “favoring high levels of market regulation versus deregulation”, and in post-communist cases as “favoring state ownership of business and industry versus private ownership”.⁷ Our implementation of this dimension thus highlights another important feature of our approach: the decision to apply a different set of policy scales to the 19 post-communist countries in our study. Although the four core dimensions were also deployed in all post-communist countries, a secondary set of dimensions was used in these that were not deployed in other countries. This is because countries engaged in transitions to democracy and free market systems from state socialism experience policy competition of a different nature (Kitschelt et. al. 1999) and can involve higher electoral volatility, party system instability, and less stability with regard to issue positioning (see Zelinski 2002). The additional dimensions applied to the post-communist cases were:

- treatment of former communists (permitting the full participation of former communist rulers in democratic politics, versus prohibiting their involvement);
- access to foreigners to purchase and own of land;

⁶ In New Zealand, for instance, issues of social values are conspicuously kept out of all political discourse, and we were told it would be both meaningless and potentially bad form to ask experts about party policy on such issues as abortion and homosexuality. Accordingly, this was the only country where this dimension was not applied.

⁷ The deregulation/privatization dimension was omitted only from the surveys in Austria, Germany, Denmark, Ireland, France, and Portugal

- freedom of the mass media to publish any material it sees fit, versus regulation of media in the public interest;
- nationalism versus cosmopolitanism (in the approach to history, culture, and national consciousness);
- promotion of religious versus secular principles in politics;
- promotion of urban versus rural interests.

A dimension relating to immigration policy was also applied in many countries. This contrasted support for the full integration of immigrants and asylum-seekers with support for returning immigrants to their countries of origin.

In the European cases we covered—the vast bulk of countries in the survey—we also measured party policy in relation to the European Union. Numerous expert surveys have measured party policy towards European integration, including surveys by Marks et al. (forthcoming), Marks, Wilson, and Ray (2002), Ray (1999), (Meyer et al. 2002), and Rohrschneider and Whitefield (2005). None, however, covered all of the 39 European countries included in our survey, nor measured at the same time the non-EU dimensions of policy that we include. Following extensive discussions with specialists on the impact of the EU on European national politics, and in the context of our participation in a quite separate research project on precisely this matter⁸, our questions on European integration were designed to measure party policy towards Europe on two critical dimensions:

- scope of EU policy authority: whether the domain within which the EU can authoritatively make policy decisions should be expanded or restricted; and
- approach to EU governance: whether EU institutions should provide direct links to citizens through representative institutions such as the European Parliament, or should be controlled instead by national governments.

We also included a measure of a party policy on the issue of expanding the role of the EU in collective security, peacekeeping, and other military affairs. For countries not yet in the European Union, including the 10 states that joined in May 2004⁹, we asked only a single question relating to support or opposition for joining the EU.

In all cases, dimensions chosen for a specific country were pre-screened by at least two country experts who provided feedback on fit of dimensions to the particular country context. In rare cases this resulted in the deletion of dimensions, but more often it resulted in the addition of “local” dimensions important only in that country. In Northern Ireland, Britain, and Ireland for example,

⁸ This was the EU-funded “Domestic Structures and European Integration” (DoSEI) project, directed by Tomas Koenig.

⁹ This also included countries that had been invited to join but had declined – Norway and Switzerland.

parties were also located on a dimension regarding policy towards Northern Ireland. Similar country-specific questions were added in Canada (regarding Quebec sovereignty), Palestinian statehood (Israel), and national identity (Japan). Other dimensions applied in only a few cases included civil liberties, relations with Russia, defense policy, and health care.

Two scales on which we asked experts to locate parties were quite distinct from the others. The first was defined as an explicit left-right dimension. Following the questions asking experts to place the parties on specific policy scales, we asked experts to locate each party's position on a general left-right scale. Explicitly leaving the precise interpretation of left and right to the country specialists being surveyed, we asked them to "locate each party on a general left-right dimension, taking all aspects of party policy into account." Including this scale allowed us to draw important conclusions about the meaning of "left" and "right" in specific national contexts, discussed at some length later in this book.

The second distinctive scale we included asked experts to locate each party in terms of how close it was to their own personal policy preferences, taking all aspects of party policy into account. The purpose of this question was, following Laver and Hunt, to test for respondent bias by checking whether expert placements of parties on substantive dimensions were correlated with their personal sympathy for a party's policies (Laver and Hunt 1992). Appendix A provides the results of a huge battery of tests we ran to test for evidence of respondent bias, something also discussed later in this chapter. The conclusion of these tests, however, is that there was no systematic bias in expert placements introduced by their personal political views—a conclusion replicating similar tests conducted by Laver and Hunt (Laver and Hunt 1992).

Each scale deployed in each survey was given a precise title, and was anchored at each end with two precise substantive definitions of the scale endpoints. To take one example, the main economic policy dimension we deployed was defined as (1) "Promotes raising taxes to increase public services" and (20) "Promotes cutting public services to cut taxes." Precise English language wordings for all scale definitions and endpoints are given in Appendix A.

Questionnaire format

Following a cover letter explaining the survey and its purpose, a sample sheet listed the parties to be located on each policy dimension, and provided an example of how the parties should be located on the scale dimensions provided in the questionnaire. For the same reasons as Laver and Hunt, and in order to maintain comparability with subsequent surveys using the same format, we used a scale running from 1 to 20, with the lower position indicating the typically "left" position and the higher value the traditionally "right" position (Laver and Hunt 1992). The use of the twenty-point scale permits experts to distinguish the positions of many more parties than would a scale with fewer scale

positions, something that becomes essential when the number of parties approaches two digits – as it did in many of the countries we covered.

Respondents were presented with one policy dimension at a time, and asked to indicate where on the scale each party was located. In addition to the party's policy *position*, respondents were also asked to assess the *importance* of that policy dimension to each party, also on a 20-point scale. Finally, as we have already indicated, respondents were asked to place each party on an overall left-right scale as well as to indicate how close they felt were each party's policy positions to their own views.

In a departure from the Laver and Hunt (1992) survey design, we did not ask respondents to consider separately the positions of party leaders and party voters. Not only were there few discernable differences between the two in the Laver-Hunt data, but only leader positions have tended to be used by researchers employing the Laver-Hunt data.¹⁰ Nor did our survey ask respondents to provide information about internal party politics or activists, data which were also never used in subsequent research using the Laver-Hunt data. Given the onerous task of completing surveys asking the location typically eight parties on typically ten policy dimensions, with two scales (position and importance) for each dimension – plus the general left-right and sympathy dimensions – we felt that sacrificing other little-used information was justified by the need for quality measures on the key data, which have to do with policy locations and importance.

In an effort to increase both the quality and quantity of expert respondents, we translated the survey questionnaires into native languages in every case where local experts advised doing so. This led us to translate the questionnaires and cover letters into a total of 23 different non-English languages (see Table A1 of Appendix A and Table 4.1).

Table 4.1. Questionnaire formats, country frequency

Questionnaire Format	Language		Total
	Non-Native English	Native (incl. English)	
Paper	1	18	19
Web	13	17	30
Total	14	35	*49

Note: *Belgium and Germany counted twice since two different languages used for surveys in each country.

Each translation was performed by a native speaker of the target language, and checked by at least one other native speaker who had the original English-language document at his or her disposal for comparison. Non-native English experts were thus typically not forced to respond to an English-only

¹⁰ The only exception of which we are aware is a recent article by one of the authors (Laver 2005).

questionnaire; indeed, in cases where the questionnaire was translated into a native language, many of the experts we surveyed did not speak English. Non-English native languages were used to reach 35 expert target groups in all. English was used as a substitute for native languages mainly in the Scandinavian countries, Flanders, the Netherlands, and Denmark, for convenience, and in Malta and Cyprus, and Turkey for practical reasons. In Romania, local experts advised us to deploy the survey in English, although we later translated the survey into Romanian for deployment in Moldova.

The survey questionnaires were deployed in two different formats, with the format selected for a particular national context depending on advice received from local experts. One method involved sending printed paper surveys through the regular post, along with pre-addressed (but not pre-stamped) envelopes for returning the surveys. Paper surveys received were then checked and keyed by hand into our database of responses. Of the 49 national contexts where the surveys were deployed¹¹, 19 were deployed using the traditional paper-and-post method. The other method of deployment was via e-mail, using individually addressed e-mails containing the cover letter as well as a customized World-Wide Web link leading to our survey page. The web page was designed to look and function as closely to the paper survey as possible. This had the tremendous advantage of combining speed and economy of deployment, with elimination of labor costs and error from data entry since each completed survey was directly recorded in our database. Feedback from respondents on the web-based system, furthermore, was in general very positive.¹² The payoff of the translations and custom-fitted deployment methods can be seen in Table A.1; these are high relative response rates compared with previous political expert surveys. In the typical country context, 28 percent of the experts surveyed returned validly completed forms, an excellent response rate given the lengthy and demanding nature of the survey questionnaire.

DEALING WITH BIAS AND RANDOM ERROR

Describing a population as opposed to inferring population characteristics from a sample

Analyzing the results of expert surveys involves complex methodological issues of sampling and statistical inference. Expert surveys differ in an important way from opinion surveys. Their objective is to uncover an assumed underlying “truth”, the spatial location of a party’s policy position, rather than to infer attitudes of a population from information about a sample of this, as is the case with traditional survey methodology. If we were certain every expert would provide a perfectly accurate location of each party policy position, then we would need to ask only one expert in each country. In

¹¹ Belgium and Switzerland are counted twice since we deployed surveys in two different languages in these countries, depending on the respondent’s region

¹² Entirely open-source software was used for hosting the survey web pages and database, namely MySQL as a database server, Apache for serving web pages, and Tomcat for providing the server-side JSP web scripting functionality. These were hosted on a fairly typical desktop Dell Pentium 4 computer running Linux, connected to the Internet from Trinity College and left running continuously for nearly 18 months. Total infrastructure cost (excluding the cost of the Internet connection): about \$1,200.

practice of course, expert judgments about the location of the same party on the same policy dimension will vary for all sorts of reasons; this is why it is important to survey not only a number of experts, but also experts holding a range of different perspectives. Our confidence in our estimates as an accurate description of a true party position increases according to the classical rules of sampling, with confidence intervals shrinking as we increase the number of expert respondents.

Nonetheless we must also remember that the set of respondents to an expert survey is typically assumed to be the *population* of qualified experts in the field under investigation – for instance, all experts from some valid register of country specialists. Of course, the notion of the “population of experts” in any context is almost metaphysical. At the very least, however, the results of the expert survey represent a summary of the judgments of the members of the population surveyed. Because of this key difference between expert surveys and traditional opinion surveys, a statistical summary of expert survey scores is not an *estimate* but a *description* of population characteristics. This has implications for how we use expert surveys to construct valid measures of party positions and deal with issues of both bias and random error.

Evaluating bias among experts

It is self-evident that our expert respondents are not random samples from a population, at least not a population whose characteristics we are attempting to estimate. Experts, by their very definition as experts, are highly *unrepresentative* of the population of citizens, voters, or politicians in any given political setting. As we have just noted, however, we do not want to use our sample of experts to characterize a population, but rather to characterize unobservable yet meaningful quantities by systematically consulting the knowledge of experts. The classic problem of sample bias is not a concern, even if the experts we consult hold strong political preferences, as long as these preferences do not interfere with their expert knowledge or the ways in which they deploy that knowledge to complete our survey. Our concern with bias thus has to do with whether the information supplied by respondents is linked systematically to some respondent characteristic—for example political preferences—with the consequence that inferences from the set of respondents are biased with respect to this characteristic. Our chief concern is not so much to obtain an *ideologically representative sample* of experts, but to ensure there is no systematic relationship between the experts’ own ideologies and their judgments about party policy positions. For example, if we survey party positions in a country whose experts are predominantly left-wing socialists, and if left-wing socialists tend to rank left-wing parties as more centrist and/or right-of-center parties as more right-wing, than they “really” are, then our inferences from the expert survey will be biased. In order to be able to test for this type of respondent bias in our study, we included a “sympathy scale” that asked to experts to

place all parties on a scale indicating their own closeness to each party's policies.¹³ All other things being equal—in other words, for the same country, party, dimension, and scale—an expert's placement of a party should be unrelated to his or her closeness to that party. If experts' political sympathies influence their placements of parties, on the other hand, then for any given placement we will observe systematic differences between experts' placements, according to their political closeness to the party concerned.

In the United States, to take a striking example, despite a broadly even split of partisan affiliation in the voting population, the experts we consulted for the most part responded that they felt much closer to the Democratic than to the Republican Party. This can be seen in the striking bimodal distribution of party sympathy scores shown in Figure 4.1. This gives us interesting political information about US political scientists, of course, but such information is not at all our concern in this study. As long as our US experts' feelings about the two parties are unrelated to their placements of these parties on our policy scales, then our resulting estimates will be unbiased. If, however, experts sympathetic to the Democrats tend to place the Democrats towards the center and the Republicans towards the extreme right, while Republican-leaning experts do the exact opposite, then this would be evidence of respondent bias, since political sympathy would be correlated with party placement. The ideologically skewed nature of our population of experts would then be a problem.

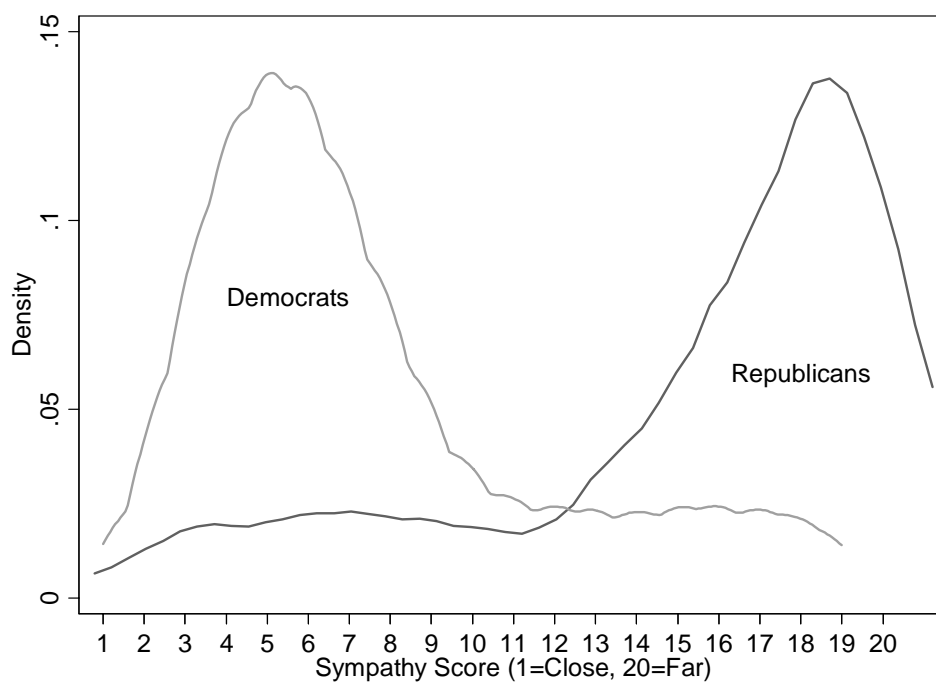


Figure 4.1: Distribution of respondent sympathy scores in the United States, $N=164$

¹³ Happily for our purposes, nearly all respondents (87.8%) also completed the sympathy dimension.

Our method for assessing the level of bias among sets of experts is described in detail in Appendix A, but in a nutshell involves searching for systematic relationships between the political sympathies of experts and their placements of parties on the general left-right dimension. From the expert placements of hundreds of parties that we analyzed in this way, only seven appeared to involve systematic bias that could be statistically distinguished from zero. These are listed in Table A3 of Appendix A and all are extreme right parties. In each case, the observed bias is that less sympathetic experts tend to place these parties as more extreme than do more sympathetic experts. These parties were a tiny minority of the universe of parties we cover, however. Figure A6, in Appendix A, graphs all differences for all parties by country type, and clearly shows that very few biases emerge that were statistically different from zero. Thus, returning to the US experts whose political views are shown in Figure 4.1, there was absolutely no evidence that their highly skewed party sympathy scores were affecting their left-right placements of the US parties in any systematic way.

The results of our systematic attempts to check for the effects of bias in our populations of expert respondents were therefore very reassuring. We did not find systematic evidence of bias in expert placements for the overwhelming majority of parties, and are left with a small number of far-right parties that do seem to have been rated as more extreme by unsympathetic experts. We are fortunate that this is a rare occurrence in our surveys, since such cases create a dilemma that is difficult to resolve. On one hand, academic experts, who are predominantly unsympathetic to extreme-right parties, seem to be rating these parties as more right-wing than would a hypothetical “unbiased” expert. On the other hand, such parties seem to invite biased perceptions by expert observers, so that the notion of an unbiased expert observer is indeed largely hypothetical. If we are honest, we must admit that our expert survey technique fails in such cases. This is not because of an inherent flaw in the technique. It is because, in these rare cases (which our technique is however able to identify), we are using an expert survey to measure something that may well not exist, which is an unbiased perception of a particular type of party. In this sense our diagnostics suggest that any assessments of the positions of these extreme right parties should be treated with skepticism, which is an important finding in its own right.

Characterizing measurement error

A particular feature of our approach to estimating party policy positions is that it measures the policy positions of all politically significant parties on all potentially relevant policy dimensions. The advantage of this approach is that it leaves open the empirical question of which parties find which dimensions important. The potential disadvantage is that it may increase the *error* in our measurement if we are asking experts to locate smaller political parties, about which they know relatively little, on policy dimensions of lesser importance. This will be reflected in the relative *variation* in the placements, by particular sets of country experts, of the same party on the same policy dimension.

There are two ways to interpret this variation, only one of which relates to measurement error. One interpretation is that the policy positions of a particular party may be *fundamentally vague and ill-defined*. In this case we expect experts to disagree on the position of a party, and the resulting variation in expert judgments of the party policy position has substantive meaning. The other is that variation in expert judgments of the same party policy position is a manifestation of *imprecision in our measurement instrument*. It is difficult to separate these two potential effects but, if we can be certain of anything, we can be certain that our measurement technique generates measurement error, as does any other measurement technique, and that this error may be both random and systematic. We dealt with the main potential source of systematic error when investigating the potential for respondent bias in the previous section. Random measurement error might arise for many reasons; here we list some of the main potential sources in relation to our expert survey.

- Different experts may have *different conceptualizations of the “target”*, having in mind different time periods or different characterizations of the target party when evaluating the same party policy position.
- Different experts may have in mind *different substantive interpretations of the “same” scale*, notwithstanding the explicit end points we supplied.
- Different experts may *use the same scales in different ways*. Some experts may be psychologically disinclined to use the full range of the scales, and tend to bunch all parties towards the center; other experts may do the opposite.
- Experts may disagree when placing a *commonly understood party on a commonly understood scale*. This may happen when experts simply disagree over the position of a party, having processed all available information and come up with a different answer, perhaps because they have different access to information or different internal ways of processing that information. The source of disagreement comes close to the notion of fundamental uncertainty in any given party’s policy positions.

The important thing about *random* error is that it does not have any *consistent* effects across the policy placements we are attempting to estimate. Instead, it pushes observed scores up or down randomly, in ways that are not predictable or correlated with other characteristics of the quantities being measured. In this sense, having more measurement error gives us an unbiased estimate of the main quantities we are interested in, though this estimate is more uncertain. We will be able to observe the scale of this error in the variation of different experts’ placements of the same policy positions. The potential for random measurement error can be reduced, furthermore, by increasing the size of the set of expert respondents. This increases our confidence that a statistical summary of the expert positions (for instance, the sample mean) represents a party’s true policy position.

We systematically investigated the impact of some potential sources of measurement error in our expert survey and report the results of these diagnostic tests in Appendix A. For example, we investigated the extent to which the level of variation in expert judgments of a given party was affected by the following:

- the number of judgments required of each expert in a given country (the number of parties times the number of policy dimensions), in the expectation that more onerous surveys might generate more variation in expert responses, which could be interpreted as more error;
- the size of the party, in the expectation that smaller parties might generate more variation in expert responses – perhaps because their positions are more fundamentally uncertain, or because lower levels of information about them generate more random error in expert responses
- the type of party system, in particular making the distinction between post-communist party systems and others – in the expectation that there would be more fundamental uncertainty about party positions in relatively new party systems, and thus greater variation in the expert judgments.

Detailed results are reported in Appendix A. Party size made almost no difference; smaller parties did not appear to have more uncertainty associated with their estimated positions than larger parties. The complexity of the survey (and thus of the party system under investigation) did have a very small, but nonetheless statistically significant, effect on the variation in expert judgments. This may be because the expert survey instrument becomes less reliable as the “payload” of the survey increases, or because there is more fundamental uncertainty about party policy positions in more complex party systems. By far the most striking finding, however, was that there was a significantly higher level of variation of expert judgments in the post-communist party systems. It seems plausible to attribute this to a higher level of fundamental uncertainty among experts about party policy positions in post-communist countries than to an artifact of our measurement instrument. For whatever reason, however, it shows that we are somewhat less certain about our estimates of party policy positions in the post-communist party systems.

Above all, of course, and here we apologize if many of those we are preaching to have already been converted, the most important thing to bear in mind is that there is indeed error in our data as in any other data, and that users must take this into account when drawing inferences about party policy positions. When we report an estimate of the position of a given party on a given dimension, therefore, the standard error of this estimate is every bit as important as the estimate itself – and indeed is what gives to estimated position substantive meaning. In the end, what we are able to do is to establish statistically a *range* for each party’s position on each dimension investigated –

allowing users to determine, for example, whether two party positions are, in a statistical sense, “the same” or significantly “different”.

ASSESSING THE VALIDITY OF OUR RESULTS

It is very difficult, indeed in a strict epistemological sense it may be impossible, to demonstrate definitively that a given measure of some fundamentally unobservable concept is more valid than some alternative measure. When different measures of the same unobservable property differ, it can be very difficult to know which to choose as the “best”. Nonetheless, it is also important to assess the validity of any new measure by comparing them with existing measures of the same thing, if these existing measures are accepted as in some sense valid.

As a check on the content validity of our estimates, therefore, it is helpful to compare at least some aspects our measures with measures generated by a well-known, widely-used, published measurement instrument relying on a completely different method. It is impossible to compare our estimates on all policy dimensions, since no other study of which we are aware has estimated party positions on all of these policy dimensions. However, there is a range of different quite independent estimates of the positions of parties on the general left-right dimension. A widely used collection of estimates can be found in the published left-right policy scores derived from the content analysis of party manifestoes by Comparative Manifesto Project (CMP), published in the CD-ROM issued with their most recent book, *Mapping Policy Preferences* (Budge et al. 2001). Because both our expert surveys and the CMP scales directly measure the positions of parties on a general left-right scale, comparing these estimates offers a unique opportunity to make direct comparisons between the results generated by two quite different methods. To the extent that two quite independent methods agree, we are reassured about their validity. To the extent to which they do not agree, we may examine substantive disagreements and attempt to draw inferences from these. Our approach to this comparison is straightforward; we directly compare the left-right scores and rankings produced by the two methods.

Cross-national comparisons between manifesto and expert survey estimates

We are in a position to compare our expert survey estimates, with the CMP manifesto-based estimates, of parties’ left-right positions for the set countries for which both studies estimate left-right positions (listed in Table 4.3). For each country, we use CMP left-right estimates for most recent election year covered in the CMP dataset (also listed in Table 4.3). To compare estimated left-right positions parties in the pooled cross-national dataset thereby created, we plot left-right scores from both measures, for every national party for which both CMP and expert survey estimates are available. Figure 4.2 shows this plot, using country-party abbreviations to identify each point. It also

plots a fitted linear regression line along with a 95 percent confidence interval.¹⁴ (The “cross-validation hypothesis” is in effect that there is a perfect linear relationship between the two measures.) The large cross-hairs split the plot into four quadrants, intersecting at the mid-point of each scale—0 for CMP and 10.5 for the expert survey estimates.

The first impression from inspecting Figure 4.2 is reassuring: there is a relatively good linear fit¹⁵ between the two measures and there is a nearly perfect intersection of the regression line through the (10.5, 0) midpoint of the two scales. In this sense, the two scales seem to be measuring the same thing, albeit noisily. If one set of measures were systematically different from the other, we would not see this empirical origin correspond so neatly with the theoretical midpoint of the scales. As a check on face validity, we find this pattern to provide a good indication that, compared to an established measure, our expert survey data are not measuring something fundamentally different from pre-existing, independent measurement instruments.

The second clear impression from inspecting Figure 4.2 is that, despite the good linear fit, there is a lot of apparently random noise. The expert survey estimates of parties’ left-right positions explain only 40 percent of the variance in the CMP estimates of ostensibly the same thing. Furthermore, there is no obvious pattern to this noise that emerges from the overall scatter-plot. To examine this matter further, therefore, we select individual cases of high disagreement and try to ascertain why our expert survey estimates and the CMP scores diverge. To identify these cases, we focus on the outliers in the off-diagonal quadrants: the parties classified as “left” by our expert survey and “right” by the CMP manifesto analysis, as well as the inverse of this. Table 4.1 identifies all of these parties, listing our expert survey estimates and the CMP’s manifesto-based scores (along with the standardized residual).

Drawing on our general knowledge about the parties listed in this table, there are clearly some that have been located incorrectly on at least one of the left-right scales. We have not identified any particular pattern that explains these differences, although several of the outliers in Table 4.1 seem to be parties for which immigration, nationalism, or the environment are important issues. Because none of these issues are components of the CMP left-right scale used in *Mapping Policy Preferences*, it is possible that CMP scores for parties emphasizing these issues could differ from the corresponding expert survey estimates of their left-right positions. For instance, it might explain why the CMP ranked as centrist the Austrian Greens and Belgian *Ecolo* parties, whereas our expert survey scores placed these parties on the left. Likewise, the CMP ranked as left-of-center several nationalist parties such as the Belgian VB and the New Zealand First Party, which are scored as right-wing by our experts. In nearly all cases such as this where the two independent measures disagree, and while

¹⁴ CMP scale positions are taken for the most recent election in the CMP dataset published with MPP, and thus are somewhat earlier in time than the estimates from the 2003 expert survey (see Table 4.3 for a list of CMP manifesto dates we used).

¹⁵ The OLS regression produces the following results: $N=114$, $R^2=.40$, Root MSE=18.763, Expert survey coefficient (SE) 3.19 (.372), Constant -33.06 (4.378).

readers are entitled to draw their own conclusions, we feel quite confident that our expert estimates come closer to what most observers would regard as the “correct” policy positions than the CMP’s manifesto-based estimates. To put it crudely, when measures disagree, we are happier with a measure that puts the Italian Communist Party and the Austrian Greens on the left (as does our measure) rather than on the right (as does the CMP). Similarly, we are happier with a measure that puts Spain’s *Partido Popular* or Greece New Democracy on the right (as does our measure) rather than on the left (as does the CMP). We can find no off-diagonal case in which our expert survey measure appears to give the “wrong” answer, in this sense.

Within-country comparisons between manifesto and expert survey estimates

As well as comparing raw left-right scores for each party, we can also compare left-right scores for parties within each country. Our method for this is also deliberately simple: within each country, we compare the orderings from left to right of every party for which there are both CMP and expert survey scores. This method avoids the potential problem that left-right scores in one country are not directly comparable to left-right scores in another. By using only ordinal information, this method avoids problems of scale comparability. Table 4.2 shows the left-right orderings produced by the CMP and our expert survey scores and highlights the differences between these, and reports two measures of ordinal association summarizing the agreement between measures. Kendall’s tau-a measures the relative likelihood of the ranks agreeing, rather than disagreeing, and ranges from -1.0 to 1.0 . In Ireland, for example, where the only difference between measures is the relative placement of Fine Gael and Fianna Fáil, the two party orderings are 80% more likely to agree than to disagree. Spearman’s ρ is an ordinal measure of correlation analogous to Pearson’s r but applied to ordinal data. The countries for which the two measures were in greatest disagreement were Greece and Australia, where disagreement was 33% and 17% more likely than agreement respectively. In nine of the 23 countries compared, the expert survey and CMP rankings agreed perfectly.

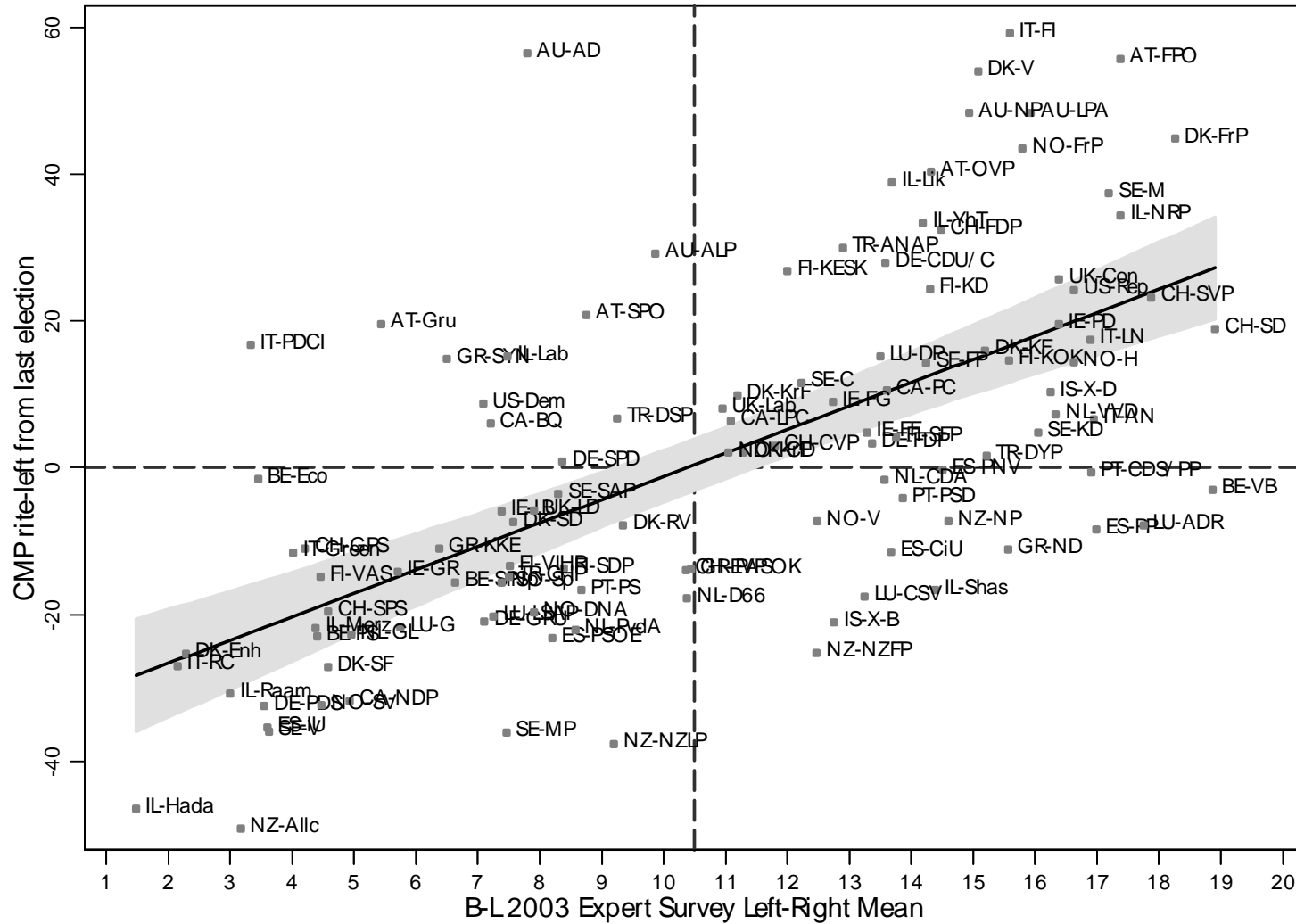


Figure 4.2 Across-Country Scatterplot of CMP Left-Right Scores versus Expert Survey Left-Right Estimates

Country	Party	Party Name	Left-Right Score		Stand- ardized Residual
			Expert Survey	CMP	
Expert Survey Says Left, CMP says Right					
AU	AD	Australian Democrats	7.8	56.6	3.47
IT	PDCI	Partito dei Comunisti Italiani	3.3	16.8	2.13
AT	Gru	The Greens	5.4	19.7	1.90
AU	ALP	Australian Labor Party	9.9	29.3	1.65
GR	SYN	Synaspismos	6.5	14.8	1.46
AT	SPO	Austrian Social Democratic Party	8.8	20.8	1.39
IL	Lab	Labor	7.5	15.2	1.31
US	Dem	Democratic Party	7.1	8.8	1.03
CA	BQ	Bloc Québécois	7.2	6.1	0.87
TR	DSP	Demokratik Sol Partisi	9.2	6.7	0.55
Expert Survey Says Right, CMP says Left					
NL	CDA	Christen Democratisch Appe`l	13.6	-1.6	-0.64
NO	V	Venstre	12.5	-7.3	-0.75
NZ	NP	New Zealand National Party	14.6	-7.2	-1.11
PT	CDS/PP	People's Party	16.9	-0.7	-1.16
ES	CiU	Convergència i Unió de Catalunya	13.7	-11.5	-1.18
LU	CSV	Christian Social People's Party	13.2	-17.6	-1.44
GR	ND	Nea Dimokratia	15.6	-11.2	-1.49
IS	X-B	Framsóknarflokkurinn	12.8	-21.1	-1.54
IL	Shas	Shas	14.4	-16.7	-1.58
ES	PP	Partido Popular	17.0	-8.4	-1.60
BE	VB	Flemish Block	18.9	-3.0	-1.63
LU	ADR	Action Comity for Democracy and Pensions Justice	17.8	-7.9	-1.70
NZ	NZFP	New Zealand First Party	12.5	-25.2	-1.71

Table 4.1: Largest divergences between expert survey and CMP estimated left-right positions

Country	CMP Election Year	Kendall's tau-a	Spear- man's rho	Party Ordering from Left to Right	
				Expert Survey	CMP
GR	1995	-0.33	-0.60	KKE SYN PASOK ND	PASOK ND KKE SYN
AU	1995	-0.17	-0.32	AD ALP NP LPA	ALP LPA NP AD
BE	1998	0.00	-0.20	Eco PS SPSp VB	PS SPSp VB Eco
IT	1995	0.33	0.49	RC PDCI Green FI LN AN	RC Green AN PDCI LN FI
TR	1998	0.33	0.40	CHP DSP ANAP DYP	CHP DYP DSP ANAP
CH	1997	0.52	0.75	GPS SPS EVP CVP FDP SVP SD	SPS EVP GPS CVP SD SVP FDP
FI	1996	0.52	0.71	VAS VIHRS DP KESK SFP KD KOK Hada Raam Merz Lab Lik YhT Shas	VAS SDP VIHRS SFP KOK KD KESK Hada Raam Merz Shas Lab YhT NRP
IL	1997	0.64	0.76	NRP	Lik
NO	1998	0.71	0.89	SV Sp DNA KrF V FrP H	SV DNA Sp V KrF H FrP
SE	1995	0.71	0.86	V MP SAP C FP KD M	MP V SAP KD C FP M
DK	1998	0.72	0.90	Enh SF SD RV KrF CD V KF FrP	SF Enh RV SD CD KrF KF FrP V
ES	1998	0.80	0.90	IU PSOE CiU PNV PP	IU PSOE CiU PP PNV
IE	1996	0.80	0.90	GR LB FG FF PD	GR LB FF FG PD
LU	1996	0.80	0.90	G LSAP CSV DP ADR	G LSAP CSV ADR DP
AT	1998	1	1	Gru SPO OVP FPO	Gru SPO OVP FPO
CA	1995	1	1	NDP BQ LPC PC	NDP BQ LPC PC
DE	1995	1	1	PDS GRÜ SPD FDP CDU/C	PDS GRÜ SPD FDP CDU/C
IS	1996	--	1	X-B X-D	X-B X-D
NL	1994	1	1	GL PvdA D66 CDA VVD	GL PvdA D66 CDA VVD
NZ	1997	1	1	Allc NZLP NZFP NP	Allc NZLP NZFP NP
PT	1996	1	1	PS PSD CDS/PP	PS PSD CDS/PP
UK	1995	1	1	LD Lab Con	LD Lab Con
US	1997	--	1	Dem Rep	Dem Rep

Table 4.2: Within-country comparisons of CMP and expert survey left-right party orderings

Differences between CMP and expert survey placements

There are two main ways to explain differences our expert survey estimates of left-right positions and the CMP's manifesto-based estimated of ostensibly the same thing. The first is that, between the time of the election for which the CMP scored the party's manifesto and the time of our expert survey, some party changed their positions on the left-right scale on radical ways. We might make this case for the US Democratic Party, for instance: perhaps it was right-of-center in 1997 but moved to the left-of-center by 2003. We might find such an argument much more difficult to sustain for parties such as the Italian Communists (PDCI) or the Belgian Flemish Bloc (VB). This would amount to arguing that the Italian Communists were indeed on the right when assessed by the CMP, but moved dramatically to the left by the time of our survey; and that the Belgian Flemish Block, widely perceived as among the group of far-right parties in Europe, was in fact on the left when measured by the CMP, switching dramatically to the right at the time of our expert survey.

The second possibility is that one of the two measures is wrong for these particular cases. For any variable with a degree of measurement error, it would be entirely consistent to observe overall patterns that appear unbiased—that is, right on average as indicated by Figure 4.2—but mistaken in particular cases such as those listed in Table 4.1. We contend that whichever measure appears to be more “correct” substantively in the off-diagonal cases (Table 4.1) is the measure containing less error. Our firm view is that, substantively, when our expert survey and the CMP estimated diverge, the expert survey estimates are to more plausible substantively.

REFERENCES

- Budge, Ian, Hans-Dieter Klingemann, Andrea Volkens, Judith Bara, Eric Tannenbaum, Richard Fording, Derek Hearl, Hee Min Kim, Michael McDonald, and Silvia Mendes. 2001. *Mapping Policy Preferences: Parties, Electors and Governments: 1945-1998: Estimates for Parties, Electors and Governments 1945-1998*. Oxford: Oxford University Press.
- Castles, Francis M., and P. Mair. 1984. Left-Right Political Scales: Some 'Expert' Judgements. *European Journal of Political Research* 12:73-88.
- Freedom House (U.S). 2003. *Freedom in the world*. New York, Lanham, Md., Oxford: Freedom House Rowman & Littlefield.
- Laver, Michael, and W. Ben Hunt. 1992. *Policy and party competition*. New York: Routledge.
- Marks, Gary, Lisbet Hooghe, Moira Nelson, and Erica Edwards. Forthcoming. Party Competition and European Integration in East and West: Different Structure, Same Causality. *Comparative Political Studies*.
- Meyer, Mary A., Kenneth B. Butterfield, William S. Murray, Ronald E. Smith, and Jane M. Booker. 2002. Guidelines For Eliciting Expert Judgment As Probabilities or Fuzzy Logic. In *Fuzzy logic and probability applications: bridging the gap*, edited by T. J. Ross, J. M. Booker and W. J. Parkinson. Philadelphia: Society for Industrial and Applied Mathematics and the American Statistical Association.