OXFORD

Mapping Policy Preferences

Estimates for Parties, Electors, and Governments 1945–1998

6

Checking the Party Policy Estimates: Convergent Validity

Michael D. McDonald and Silvia M. Mendes

INTRODUCTION

In a sense most of this book has been concerned with the validity of the Manifesto data and the estimates derived from them. Chapter 1 has matched the major Left-Right indicator with the historical record (predictive validity); Chapter 2 has reported their use in research (hypothesis validity); Chapters 3 and 4 have gone into the plausibility of the underlying Saliency Theory and the match between the coding distributions and its expectations (face validity). What this chapter deals with in particular is construct validity—whether the codings correlate with other measures of the same conceptual construct (Weber 1990: 18).

The construct in question is the Left-Right scale which has been central to most chapters of the book and used in most applied research. It also, as noted in Chapter 1, bases itself on all the variables of the coding frame either directly or indirectly. As argued above, there are merits in a holistic approach to checking such a large and open ended data-set. We want to know if it generally gets things right when used in a comparative over time context rather than evaluating the details of particular coding decisions.

There is also a practical reason for concentrating on the Left-Right scale rather than on the other policy estimates given below or on particular categories of the coding frame. It is the only measure based on the data where there are enough alternative indicators to support an investigation of construct validity. While the same approach could be applied in principle to the other policy scales, the empirical basis is lacking. They must therefore rely for the time being on the other forms of validation discussed above—though, like the rest of the data, their standing will be enhanced if the general Left-Right measure comes successfully through the checks imposed here.

We do include a number of other Left-Right scales derived from Manifesto data in the comparisons. These are derived from suggestions made by other investigators as to how the coding categories might be combined in this context. As they draw on a somewhat different range of categories and combinations to the MRG scale, their collective performance can be taken as a more generalized check on the validity of the entire data-set, rather than of one particular measure derived from it.

One difficulty with comparing measures however is that not all of them have equal standing. The 'expert judgements' of parties' Left-Right positions (Castles and Mair 1984; Laver and Hunt 1992; Huber and Inglehart 1995) against which we validate the Manifesto-based scales have been severely criticised for mixing up declared party policy with their actual behaviour (fatal if one wants to explain behaviour by party positioning) and being imprecise about the time period and criteria involved in the judgements, among other things (Budge 2000). The other Left-Right scales derived from the Manifesto data, which we also use in comparisons, are generally exploratory attempts at measuring party positions on inductive factor-analyses of the data as it then was (Budge and Robertson 1987; Bartolini and Mair 1990). Given the problems involved with the other measures, the 'saliency' measure employed here could well be accepted as a superior standard by which the others should be judged, rather than using *them* to judge *it*, which is what we propose to do here.

The general status of the measures is not a question to argue about at this point. In order to set up the test we will simply assume that expert judgements directly reflect the 'real' positions of the political parties and see how far the saliency-based scale measures up to them. To put it in context we will also include the other left-right measures based on the Manifesto data, regardless of their general standing. Shaping our research design in this way makes our test of convergent validation a harder one for the main Left-Right scale to pass. If it does, it should gain greater credibility than if it had been accorded a more privileged status.

RELATING THE EXPERT PARTY POLICY SCALES TO EACH OTHER

Following this reasoning we are going to accept the Left-Right party positioning on expert scales as the meaning of Left-Right that we intend to measure. That simplifies the question; so that we ask: Can the CMP data be used to measure the Left-Right positions of parties as those Left-Right positions are understood by experts? Of course, taking the expert scales as the standard for evaluation without knowing much about their reliability and validity runs the risk of inferring that any mismatch between the expert and the CMP scales results from problems with the CMP data. It might well be the case that different expert scales measure Left-Right positions differently, contain a good deal of random noise, or both. That would force the unrealistic requirement that the CMP scales match moving targets. Therefore, we begin by exploring the reliability and validity of three expert scales.

Since 1980 Frank Castles and Peter Mair (1984), Michael Laver and W. Ben Hunt (1992), and John Huber and Ronald Inglehart (1995) have produced expert scales of party positions. They provide a common coverage of 84 parties in 16 Western countries operating as democracies throughout the post-war period. The

Castles-Mair and Huber-Inglehart scales expressly focus on the Left-Right location of parties. Laver and Hunt asked their experts to place the parties along eight dimensions. They suppose that their public ownership dimension is the most indicative of the usual conception of Left-Right (Laver and Hunt 1992: 122). Relating their public ownership, social issues, and taxes-versus-spending dimensions to the other scales through regression analysis confirms this idea, though all three dimensions have some relationship to the overall Left-Right indices casting some doubt incidentally on whether experts can distinguish between the separate policy dimensions well enough to place parties quite independently on them.

To bring the Laver-Hunt measures into line with the Left-Right content of Castles-Mair and Huber-Inglehart, we calculated a weighted average of the public ownership/social value/tax-service scales in order to generate a general Left-Right scale. The public ownership and tax/spending scales have a weight of 1.5 and the social scale has a weight of 1.0. That gives three times as much weight to economic as to social issues, a fact that accords with the coefficient weights found in the two regression equations.

Correlations between the three expert indices are:

C–M/H–I	0.93
C–M/L–H	0.94
H–I/L–H	0.94

To assess the reliability and stability of these measures we can apply the Heise (1969) measurement model, which distinguishes between the two. This model assumes a Markovian process, in which a party's change from today's position to tomorrow's will be unaffected by its position yesterday. As each of the expert scales can be associated with a different time period prior to its date of publication, we can check change over time (begging the question of whether or not all the experts focused on the whole of the post-war period).

The Heise measurement model produces stability and reliability estimates of:

Reliability	=	0.947
Stability, early 80s to late 80s'	=	0.988
Stability, later 80s to early 90s'	=	0.994
Stability, early 80s, to early 90s	=	0.982

The measurements are highly reliable; 94.7 per cent of the variation is systematic and just over 5 per cent is random. Once the modest unreliability is taken into account, the positioning of the parties is almost perfectly stable. This can be taken as both good news and bad news. To the good, there is very little randomness in the experts' placements of the parties. The bad news is that a problem would arise if one were to attempt to use the expert scales as a basis for analyzing party strategic movements. There are essentially no observable movements other than those due to a small amount of measurement error. Across a decade's time, the experts saw the parties in essentially the same relative

locations. Perhaps the parties never moved, or perhaps the experts are reporting an over-time general statement about party locations in the post-war period.

THE MANIFESTO LEFT-RIGHT SCALES

Various scales have emerged from factor-analytic and theoretically based explorations of the Manifesto data conducted at various time points and with different concerns. For reasons explained in previous chapters the theoretically driven but inductively informed measure developed in Laver and Budge (eds.) (1992) has developed as the leading one. The coding categories included in it and the other Left-Right scales are shown in Table 6.1.

What is interesting is that each of the scales grouped rather different combinations of coding categories, drawing in almost the whole range of Manifesto variables into one or other of them. This makes an assessment of their overall validity in some sense a check on the validity of the Manifesto data-set as a whole, even if our main interest focuses on the first scale.

The Manifesto Research Group (MRG) appear to have created a global form of a Left-Right scale, inasmuch as they include coding categories from all seven policy domains of the data-set. Budge and Robertson, whose analysis actually identifies a left-wing isolationism position and a capitalist traditionalist position (1987: 404), are nearly as catholic. The Left versus Capitalist positions cover all but the welfare domain. Bartolini and Mair's (1990) Left-Right measure decidedly focuses on economic matters. Coding categories from the economic domain alone are included in their Left-Right scale. Finally, Laver and Garry propose that Left-Right party positions are best described in terms of concerns for state intervention versus capitalist economics plus negative mentions of welfare. They alternatively treat concerns over social and cultural values as a separate liberal versus conservative aspect of Left-Right politics.

The scales we check out in this chapter take the combinations of variables proposed by the various authors in Table 6.1, but score them in the subtractive way originally suggested by Laver and Budge (1992: 25–30). They identified 26 coding categories that go into their measurement definition of Left-Right. They add 13 Left items and subtract from this quantity the sum of 13 Right items. A party that makes 200 total statements with 100 (or 50 per cent) of them about Left items and 40 (or 20 per cent) about the Right items receives a score of +30 (i.e. 50 minus 20).

This difference or subtractive measure is consistent with saliency theory. Of all the statements the party made, on balance, 30 more units were devoted to Left matters than to Right matters. Imagine that at the next election this party says exactly the same things it had said last time but adds 200 new statements about an issue that is not of concern to the Left-Right scale (e.g. favourable statements about protecting the environment). Now the party is making 400 total statements, and relative to that total they are making only half as many Left statements (25 per cent) and half as many Right statements (10 per cent) as they did for the first election. The party's Left-Right position is recorded as moving from +30 to +15. That is, the party is scored as considerably less left-leaning at the second election compared to the first. It has moved toward the centre by virtue of devoting attention to policy matters that are not within the categories relevant to the Left-Right scale.

This is clearly not the only way categories could be aggregated. For one thing they could be weighted by their factor loadings rather than treated equivalently, as the original investigators in fact scored them. In this analysis however we treat them all in a subtractive form, with equal weighting of categories, which makes them more equivalent for later comparison.

THE MANIFESTO SCALES: VALIDITY

In order to evaluate and reliability of the scales constructed from the CMP data, we use the five Left-Right scales described in Table 6.1. We have created scores for each of the five scales by, first, summing the Right items and the Left items and then calculating a subtractive measure (Left-Right). Because the expert scales show no sign of change in the party positions, as if the experts have summarized the typical positions of the parties, we use each party's 1972–92 period average over the Manifesto scales for testing their scale validities. Requiring a party to have a Manifesto throughout this 20-year period reduces the number of parties we analyse from 84 to 66. Once we know something about validity, we turn back to the Heise reliability and stability test and apply it to the five scales under consideration.

Figure 6.1 illustrates the factor loadings (principal axis with varimax rotation) of all five Manifesto scales. The analysis includes the three expert Left-Right scales and, in order to help define the factor space, the Laver and Hunt (1992) public ownership and social ratings.

For ease of exposition at this point, the two factors can be discussed as 'pure' indicators of an economic dimension (the horizontal axis) and a social dimension (the vertical axis). In that view, the expert scales appear to be a mix of those two dimensions, with slightly more weight attributable to the economic as compared to the social. Relative to the expert scales, the CMP scales are more economic-laden. Indeed, they are even closer to the horizontal axis than is the Laver-Hunt public ownership expert scale.

Given that we are accepting the expert scales as the meaning of Left-Right, the Manifesto scales are slightly off the mark on the validity question. To determine by how much they miss the mark of the expert scales, we follow Guildford and Hoepfner's (1969) advice and rotate the horizontal dimension so that it goes directly through the centroid formed by the three expert scales. The rotation substitutes the criterion of defining a dimension by how well it hits a theoretical mark as opposed to how well it accounts for particular types of

Table 6.1. CMP Coding categories included in Left-Right scales across four studies

	CMP Var No.	Domain	MRG	Budge & Robertson	Bartolini & Mair	Laver & Garry	
						Economic	Social
	102	Foreign		Foreign special relations, con			
	103	Foreign	Decolonization	Decolonization			
	105	Foreign	Military, con				
	106	Foreign	Peace				
	107	Foreign	Internationalism, pro				
	110	Foreign		European Community, con			
	202	Freedom	Democracy				
	204	Freedom					Constitutionalism, con
	304	Govt					Government corruption
	403	Economic	Regulation of capitalism	Regulation of capitalism	Regulation of capitalism	Regulation of capitalism	
Left	404	Economic	Economic planning		Economic planning	Economic planning	
Items	406	Economic	Protectionism, pro		Protectionism, pro	Protectionism, pro	
	408				Economic goals		
	409	Economic			Keynesian economics		
	410	Economic			Productivity		
	411	Economic			Tech & infrastructure		
	412	Economic	Controlled economy		Controlled Economy	Controlled Economy	
	413	Economic	Nationalization	Nationalization	Nationalization	Nationalization	
	504	Welfare	Welfare, pro				
	506	Welfare	Education, pro				
	602	Fabric		Nat'l way of life, con			Nat'l way of life, con
	604	Fabric		Traditional morality, con			Traditional morality, con
	701	Groups	Labour groups, pro	Labour groups, pro			
	104	Foreign	Military, pro				
	108	Foreign		European Community, pro			
	203	Freedom	Constitutionalism, pro				Constitutionalism, pro
	201	Freedom	Freedom & human rights	Freedom & human rights			
	301	Govt		Decentralization			

Table 6.1. *(continued)*

	CMP Var No.	Domain	MRG	Budge & Robertson	Bartolini & Mair	Laver & Garry	
						Economic	Social
Right Items	303	Govt		Government efficiency			
	304	Govt		Government corruption			
	305	Govt	Government authority	Government authority			Government authority
	401	Economic	Free enterprise	Free enterprise	Free enterprise	Free enterprise	
	402	Economic	Incentives	Incentives	Incentives	Incentives	
	407	Economic	Protectionism, con		Protectionism, con	Protectionism, con	
	414	Economic	Economic orthodoxy	Economic orthodoxy	Economic orthodoxy	Economic orthodoxy	
	505	Welfare	Welfare, con			Welfare, con	
	601	Fabric	Nat'l way of life, pro				Nat'l way of life, pro
	603	Fabric	Traditional morality, pro	Traditional morality, pro			Traditional morality, pro
	605	Fabric	Law & order				Law & order
	606	Fabric	Social harmony	Social harmony			Social harmony
	607	Fabric		Multiculturalism, pro			

variance (where the varimax criterion is maximizing the squared factor loadings).

The expert scales are quite well defined by the Left-Right axis. They load highly on it and near zero on an orthogonal axis. This corroborates the inferences that they are highly reliable (communalties>.9) and valid (nearly pure measures) of what we take to be Left-Right party positioning. The loadings of the CMP scales show them to be measuring something similar but slightly angular. In that sense, none of them are precisely valid measures of Left-Right. They have, it would appear, too much economic content and/or too little social content to match what the experts have in mind for Left-Right.

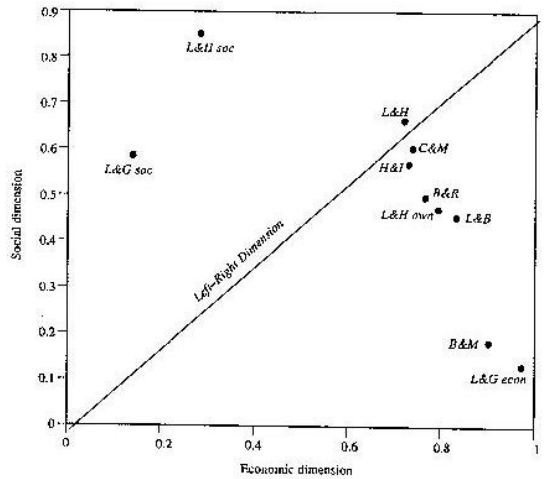
This conclusion should not be overdrawn. The Manifesto measures, especially those of Budge-Robertson and the MRG scale, are close approximations to the Left-Right party positions given by the expert scales. The Bartolini-Mair and Laver-Garry Manifesto scales were designed to measure principally the economic or the social positioning of parties, and they do.

We want to press on, however, and see whether the Manifesto data can be used to produce an even more valid Left-Right measure, with validity still defined in terms of the expert scales. The prime candidates for adjustment are the Budge-Robertson and MRG scales. The intuitively appealing way to move their outcomes closer to the expert scales is to add social items, remove economic items, or both. Those sorts of additions and deletions should move the locations up along the social dimension and therefore closer to the positions of the expert scales. We made several adjustments on this sort, and they sometimes moved the factor analysis position closer to the Left-Right dimension that passes through the expert scales' centroid. However, none of those efforts resulted in the Manifesto scale being as close to the Left-Right dimension as any one of the three expert scales themselves.

A residual analysis revealed that for both scales, and for virtually all the adjustments we made to them, five parties were being scored consistently differently by the experts and CMP scores. The Italian Communists (PCI) and Danish Centre Democrats (CD) were scored as considerably farther Left by the experts than by the CMP scores. The experts placed the Italian PCI as nearly as far Left as the French PCF; the CMP scales have the PCI decidedly more toward the centre. The Danish CD are seen by the experts as a centrist party, similar to the Norwegian SP and the Swedish CP, but the CMP scales score them as right of centre, as far right as, say, the Danish KF and FRP. The conservative Italian MSI are about as Right-leaning as a party gets, according to the experts. The CMP scales place the MSI to the right of centre but not at the extreme right. Finally, the Finnish KESK and Norwegian Høyre are each scored as right of centre by the experts, whereas the CMP scales have both parties as Centre-Left. (c.f. Gabel and Huber (2000) for similar discrepancies).

After removing the five parties that mismatch on the expert versus Manifesto scales, factor analysis of the subtractive CMP scales reported in Figure 6.1

Factor loadings of expert scales and CMP subtractive scales

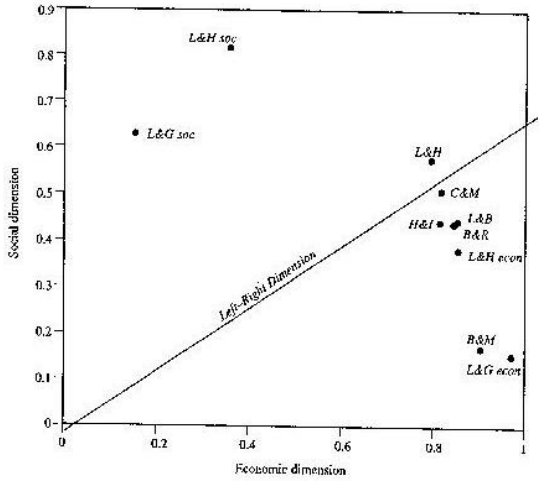


CMP Subtractive scales numerical value of factor loadings

Scale	Varimax rotation		Rotation to expert scales centroid	
Castles-Mair	.731	.610	.952	-.013
Laver-Hunt	.713	.670	.977	.044
Huber-Inglehart	.723	.577	.924	-.033
Budge-Robertson	.759	.505	.905	-.111
MRG	.825	.462	.927	-.186
Bartolini-Mair	.895	.192	.804	-.437
Laver-Garry econ	.965	.141	.824	-.521
Laver-Garry social	.133	.592	.486	.363
Laver-Hunt own	.787	.479	.909	-.149
Laver-Hunt social	.275	.858	.767	.472

Figure 6.1. Factor analysis results for expert and manifesto scales

Factor loading plot



Numerical values of factor loadings

Scale	Varimax rotation		Rotation to expert scales centroid	
Castles-Mair	.809	.515	.959	.002
Laver-Hunt	.787	.581	.976	.070
Huber-Inglehart	.836	.444	.944	-.072
Budge-Robertson	.806	.448	.921	-.053
MRG	.845	.451	.955	-.071
Bartolini-Mair	.896	.176	.851	-.331
Laver-Garry econ	.963	.160	.899	-.380
Laver-Garry social	.145	.633	.461	.457
Laver-Hunt own	.846	.387	.922	-.126
Laver-Hunt social	.349	.822	.735	.508

Note
The five excluded parties are: Danish CD, Finnish KESK, Italian PCI, Italian MSI and Norwegian Høyre.

Figure 6.2. Two-dimensional factor analysis results for expert and manifesto scales (excluding five parties)

produces the results reported in Figure 6.2. Without those five parties in the analysis, the MRG and Budge-Robertson CMP scales appear very near to the Left-Right dimension. The MRG position is virtually identical to that of the Huber-Inglehart expert scale, and the Budge-Robertson position is quite near to them.

Several conclusions are warranted. By this test the Manifesto data can be used as valid measures of party policy positions. Left-Right scales, designed to combine economic and social policy statements (e.g. Budge-Robertson and MRG), are close approximations to Left-Right party positions produced by experts. Except for five parties, these two CMP scales are a near equivalent of the Huber-Inglehart expert scale. Arguably, the mismatches on five parties could be held against a validity claim for the CMP scales, but just as arguably the expert placements of those five parties could be responses to their reputation rather than to their actual policy advocacy. This, we think, is one of those issues for which it is proper to call for further investigation.

THE MANIFESTO SCALES: RELIABILITY

The validity of the Manifesto scales have been evaluated based on the average of the parties' policy positions across a 21-year period. That leaves open the question of whether the CMP data can be used for single time points. Application to shorter time frames would make it possible for researchers to use the Manifesto to analyze movements as parties adopt various mixes of strategic and sincere positions from one election to the next. This is not possible with the expert scales, for we have already seen that experts place the parties in almost completely stable locations. If, however, the Manifesto scales cannot be applied to single elections, then they can be deemed no more nor less useful than the expert surveys. Either the expert or the Manifesto scales can be used to locate the parties, but neither data source would allow investigation of the dynamics of party movements. In other words, we need to know two additional facts about the Manifesto scales with respect to single elections: (1) Are they reliable measures of the party positions? (2) If they are reliable, do the parties move around or stay at fixed positions in the policy space?

We evaluate reliability through the same Heise measurement model we earlier applied to the expert scales. The time points are, as nearly as practicable, the most recent election prior to each expert survey—Castles-Mair prior to 1983, Laver-Hunt prior to 1990; and Huber-Inglehart prior to 1993. In countries that held no election between 1990 and 1993, the time-3 point is the last election in the CMP94 data set (Volkens 1994), and the time-2 election is the one preceding that. With these data, we generated the correlation matrix reported in Table 6.2. In turn, from the threefold sets of correlations for each scale we

Table 6.2. Correlations, reliability and stability of Left-Right party positions for each of five manifesto scales, based on a three-wave panel for 66 parties

Scale	Correlations				Reliability	Stability		
		Time 1	Time 2	Time 3		T1–T2	T2–T3	T1–T3
Budge and Robertson	Time 1	1.00			.889	.480	.831	.399
	Time 2	.427	1.00					
	Time 3	.355	.739	1.00				
MRG	Time 1	1.00			.942	.786	.845	.664
	Time 2	.741	1.00					
	Time 3	.626	.796	1.00				
Bartolini and Mair	Time 1	1.00			.654	.938	1.016	.953
	Time 2	.614	1.00					
	Time 3	.624	.665	1.00				
Laver and Garry Economic	Time 1	1.00			.874	.891	.920	.820
	Time 2	.778	1.00					
	Time 3	.716	.804	1.00				
Laver & Garry Social	Time 1	1.00			.069	–	–	–
	Time 2	.039	1.00					
	Time 3	.202	.350	1.00				

NOTE. The time points are defined by a nation's prior election closest to 1983, 1990 and 1995. In countries that held no election between 1990 and 1993, time 3 point is the last election in the CMP94 data set (Volkens, 1994), and the time 2 election is the one preceding that. The reliability and stability estimates are calculated by the measurement model formulae developed by Heise (1969).

estimate the reliability of the measure and the stability of the positions using the Heise model.

The MRG scale—that is the one used elsewhere in this book—is just as highly reliable as the expert scales. The Budge-Robertson scale and the Laver-Garry economic scale are also reasonably reliable. The Bartolini-Mair scale falls below most conventions for acceptable reliability, with less than two-thirds of its variation being systematic. The Laver-Garry social policy scale, with a reliability of only .069, is mostly noise.

There are likely to be several reasons for these varying reliabilities. First, scales formed with a large number of items, such as Budge-Robertson and MRG, tend to produce higher reliabilities compared to those with fewer items. This is true in conventional testing, and it appears to be true for the Manifesto data. Second, it may well be that the Bartolini-Mair scale really requires that different items have different weights. Third, reliability could be sensitive to the inclusion of certain items. In particular, the welfare items, excluded from the Bartolini-Mair scale but included in Budge-Robertson, MRG and Laver-Garry economic scales, are likely to add systematic variation to the party positions.

The stability estimates are as heartening for analysts as the reliability estimates are for those who have used or want to use the Manifesto scales. As measured by the Budge-Robertson and MRG scales, the parties are not completely stable. Parties do offer different positions from one time to another. This, as we have been suggesting, opens the door to the possibility of using the CMP scales to analyse party movements, particularly strategic movements. We know from other analyses (Budge, Robertson and Hearl (eds.) 1987; Budge and Laver (eds.) 1992; Klingemann, Hofferbert and Budge 1994) that over the long run parties do not stray too much from their usual ideological, Left-Right location. It is rare to see one party 'leapfrog' another. We see from Table 6.2, however, that the party positions are changeable in the short run. Together these sets of findings mean that the parties do not wander so far from their ideological base as to alienate their core constituents, but they do take up different positions at different times.

CONCLUSION

The main message to draw from the analyses presented here is that the Manifesto data can be and have been used to provide valid and reliable measurements of party policy positions. Accepting the problematic claim of expert assessments of party policy positions as the standard for what it means for a party to be on the Left, on the Right, or in the Centre, we have shown that the Budge and Robertson as well as the MRG measure of Left-Right party locations are quite similar to what the experts say. And, once we take account of the expert versus CMP differences with respect to five parties—the Danish CD, Finnish KESK, Italian PCI and MSI, and Norwegian Høyre—these two CMP

measures place the parties in essentially the same way on the Left-Right positions. On the question of reliability, we demonstrate that the three expert surveys are highly reliable measurements of party Left-Right positions, with reliabilities close to 95 per cent. We also show that the MRG Left-Right 'Saliency' scale is just as reliable. With respect to validity and reliability there is little that distinguishes the results of expert surveys from, at least, that scaling of parties. The evidence here tells us that to the extent one has confidence in the party positioning from expert surveys there is every reason to have just as much confidence in party positioning based on the Manifesto data.

Validity and reliability are necessary measurement conditions. Our analysis strongly suggests that expert surveys are deficient and the Manifesto measures are useful on two important and additional concerns involved in measuring party policy positions. The experts place the parties in such stable locations that there is little hope of using expert surveys to investigate party policy dynamics. The CMP is probably the only viable data source for observing such dynamics and analyzing the party movements predicted by various theoretical propositions.

Furthermore, there is also a strong indication that expert surveys may produce suspect results about where the parties stand on different dimensions of politics and policy. The Laver and Hunt (1992) survey was expressly and carefully designed to elicit expert responses to party policy positions along several seemingly distinct policy dimensions. Our results indicate some reasons to doubt whether expert respondents actually can and do make clear and sharp distinctions between these.

All-in-all, the Manifesto data and scales offer analysts everything that can be offered by expert surveys. But that is not all. On the matter of party policy dynamics the CMP data are all we have, and they are good—valid and reliable.