

Strengths and weaknesses of approaches to measuring policy positions of parties

Andrea Volkens*

Social Science Research Center, Wissenschaftszentrum Berlin, Reichpietschufer 50 D-10785 Berlin, Germany

Abstract

This article examines the measurement quality of the three main approaches to estimating policy positions of parties: expert surveys, the conventional content analysis of election programs by the Manifesto Research Group/Comparative Manifestos Project, and computer-assisted content analysis of election programs. Based on a literature review in tabular form containing quotations ordered according to major measurement problems, this contribution discusses the merits and shortcomings of the three approaches. The systematic comparison shows that all three approaches have their particular strengths and weaknesses. As a rule, the strength of one approach is the weakness of the others and vice versa. Therefore, the three approaches are not opposed to one another but complementary, so that all three are necessary for future research.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Parties; Policy positions; Measurement quality; Validity; Reliability

1. Introduction

Due to the importance of political parties in representative democracies, methodological research on measuring policy positions of parties has been continually growing during the last 20 years. But parts of the existing body of literature which compares different approaches to measuring policy positions are biased. Authors tend to stress the deficiencies of other approaches in order to highlight the assets of their own. This article offers tabular overviews of the strengths and the weaknesses of three approaches – expert surveys (ES), the conventional content analytic approach of the Manifesto Research Group/Comparative Manifestos Project (MRG/CMP) resting on a classification

scheme, and computer-assisted content analysis of election programs (CACA) – with quotations systematically ordered according to previously mentioned major measurement problems.

The next section presents three tables. The first table lists the conclusions of the analysed papers concerning results of cross-validations, comprehensiveness of estimates, costs, and possibilities for further research. The second table registers multiple validity problems which were detected with regard to determining the policy dimensions and positions of parties on them, to defining the parties and time scales, to differentiating between cause, effect, text, and context, to distinguishing positions and their saliencies, and to using the MRG/CMP classification scheme. The third table specifies results of reliability tests. In addition, new evidence on MRG/CMP data validity and reliability is provided in the course of discussing the good and bad aspects of

* Fax: +49 30 25491 345.

E-mail address: volkens@wz-berlin.de

the three approaches. The final section of the paper summarizes the utility of the three approaches.

2. Tabular overviews of comparisons on measurement quality

The following three tables are based on a selection of literature that concentrates on comparing the measurement quality of at least two of the three approaches under consideration and across at least two countries. The tables present the three approaches to measuring policy positions of parties in the rows, so that they can be compared according to the referring major measurement problems given in the lines. Entries in cells are quotes or shortened versions of quotes with authors, years of publication, and numbers of pages of the most recent publications. For the sake of brevity, identical arguments in older publications are not recorded. Arguments in favour of one approach are given in standard type, critical judgements in italics. The overview starts with a table quoting the conclusions of this selection, [Tables 1–3](#) look more closely at the specific aspects of validity and reliability on which these conclusions are based.

2.1. *Conclusions of comparisons*

Up to now, almost all comparisons of approaches to measuring policy positions of parties dealt with left-right positions. Since, in contrast, half of the contributions to this special volume — those by Ray, Marks et al., Whitefield et al., and Netjes/Binnema — are devoted to estimates of positions on European integration, this review of previous publications concentrates on left-right scales. Summaries of cross-validation all emphasize that the left-right scales as estimated by the three main approaches correlate highly.

The high degree of correlation is an indication that all three approaches measure roughly the same concept. It further attests to the overall reliability of all estimates. The high levels of the correlations are all the more noteworthy since agreement on the meaning of the concept is lacking in political science. Two contributions in this special volume — those by Laver/Benoit and by Keman — shed light on differences in left-right concepts in particular.

The CACA made much progress over the last ten years, and is able to provide valid estimates for left-right and economic policy positions, but, up to now, cannot deliver valid estimates for other policy dimensions. In this respect, CACA is in need of further methodological developments.

Many authors start from the assumption that the expert judgements measure the ‘true’ policy positions and detect a number of deviant cases in the approach of the MRG/CMP because far-left parties are systematically located closer to the centre by MRG/CMP data. There is no reason to assume, though, that all coders did a good job with all parties except with the far-left. This difference rather indicates that the ideological position as judged by the experts is more extreme, but that the far-left parties seem to compete for the same segment of voters as the social-democratic parties. With the exception of [Bartolini and Mair \(1990\)](#), the possibility of differences between ideological and policy spaces ([Mair, 2001](#), pp. 23–24) and its consequences have never been under close scrutiny, although Sani and Sartori discussed a related difference already 20 years ago, namely, the difference between domains of identification and domains of competition ([Sani and Sartori, 1983](#), p. 330).

Other differences, often attributed to MRG/CMP measurement errors, may be as telling. For example, the MRG/CMP data shows the Italian system of the 1990s to be in a state of flux, with many parties trying to capture new positions. Generally speaking, the strength of this approach is its ability to record changes in competitive strategies, for which it can provide long time-series. Experts, however, judge most parties to stick to the same positions over long periods of time. As compared to quickly changing programmatic positions, expert surveys seem to provide generalisations of party positions over time. The contribution of [McDonald/Mendez/Kim](#) in this volume examines this distinction.

If we now look at the conclusions on future prospects, both the expert judgements and the computer-assisted content analysis of manifestos carry the day, whereas the MRG/CMP seems to be in a rather dull missionary position. Authors summarize future prospects of the three approaches in terms of comprehensiveness of estimates, costs, and possibilities for further research. Concerning comprehensiveness, all three approaches miss the 100 per cent bench. Up to now, in all of the ES data some small countries are missing due to the unavailability of experts. This kind of incomprehensiveness also leads to reliability problems if the number of experts is too low. The CACA shares all the problems of comprehensiveness which are attributed to the MRG/CMP because both approaches rely on the same documents. Firstly, not every party publishes an election program, secondly, not all policy areas are mentioned in all of the programs, and, thirdly, many programs are short. On the positive side, these

Table 1

Conclusions on measurement quality of approaches to estimating policy positions of parties

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Cross-validation	<p>Mc-M01: 97–100 high correlation between different C-M, L-H, and H-I expert scales (close to .95%)</p> <p>H-I95: 77–79 0.94 correlation between C-M and H-I</p> <p>G-H00: 98 high correlations and previous studies provide evidence for the accuracy of surveys</p> <p><i>G-H00: 98 assumption that surveys provide accurate measures; whether this assumption is true, of course, is impossible to know</i></p>	<p>B01b: 217 in all applications MRG data have produced empirically plausible and theoretically consistent results</p> <p>G-H00: 94 ‘vanilla’ method of using MRG data consistently produces the best estimates of party positions and they are quite good as compared to other accepted approaches</p> <p>Mc-M01: 103 5 deviating parties, removal yields virtual identity between L-B and H-I scales</p> <p>Mc-M01: 104–106 L-B CMP high correlation <i>with few deviant cases</i></p> <p>R99: 290 0.8 correlation for positions on European integration, 0.25 on saliency, factor loading experts 0.95, MRG data 0.93</p> <p>G-H00: 102 little evidence for country-specific error, <i>except Ireland and to a lesser extent Italy</i></p> <p>V-G-M01: 201–202 implausible placements of Italian parties</p> <p>G-H00: 94 MRG data tend to locate extreme (left) parties closer to the ideological center than do other expert survey-based approaches</p> <p>L-G00: 631 MRG identifies Labour as having conservative social values</p>	<p>L01a: 244 early results surprisingly promising</p> <p>L01a: 244 high levels of cross-validation</p> <p>G01: 191 replicates economic policy positions of British, Irish, German, and Norwegian parties as derived from expert surveys</p> <p>L-G00: 631 techniques correspond to a remarkable degree on economic issues, <i>but differences on social issues, respectable nonetheless</i> ranging from 0.67 to 0.88</p> <p><i>Ba01b:155 failure to match computerized to manual codings; automatic dictionary on word counts does not distinguish between British parties; puts Conservatives to the left</i></p>
Comprehensiveness	<p>L-H92: 34 can be comprehensive if designed properly</p> <p>C-M84: 87, M-C97: 152–153 difficulty of getting enough estimates for small countries, small parties, or when they are too difficult to locate</p> <p>H-I95: 76 target of 5 experts per country met in 35 of 42, 1 completely missing</p> <p>R99: 286 1 out of 18 countries missing</p> <p>G-H00: 94 suffers important limitations: infrequent, different formats, only over the last 15 years</p> <p>R99: 286 33 out of 258 questionnaires returned blank because experts felt unqualified</p>	<p>R99: 284 not all parties publish election programs, not all parties mention European integration in their programs</p> <p>R99: 285 on the question of internal dissent manifestos are totally mute</p>	<p>Same as MRG/CMP</p> <p>Same as MRG/CMP</p> <p>L-B-G03: 315 uncertainty of estimates increases when texts are short</p>

Table 1 (continued)

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Costs	L01a: 243 quick and convenient	L01a: 243 costly L-B-G03: 311 highly labour-intensive K-P01: 164 manual coding is time consuming, boring and potentially unreliable	L-B-G03: 311 vast volumes of text easily, cheaply, instantly available G01: 183 ever-increasing volume of politically relevant text now available in electronic form B01b: 211, 222 prospect of rapid, reliable codings with vast savings in cost L-B-G03: 312 also in need of heavy human input
Possibilities for further research	M01: 24 quick, easy, comprehensive, precise numerical answer M01: 19 insights they offer are necessarily limited and contingent M01: 24-25 experts use some implicit combination of main approaches to estimating party policy positions ...such as past coalition behaviour, party programmes and ideology, and both mass and elite perceptions, they reflect a crude synthesis	G-H00: 95 provides the only means of estimating party left-right positions over a long time period in a large number of countries B01b: 210 solid basis and reliable estimates as a general standard for validating other methods Ba01a: 221 in the future used at a much lower level for micro-analysis and checking reliability and validity Ba01b: 155 indispensable role of the MRG/CMP data in validating computerized representations G01: 183 it would take years for even a small army of researchers to code parliamentary debates L-G00: 619 laborious expert hand-coding not viable for huge numbers of actors such as all members of a legislature, internal party competition dynamics, and changes between elections	L01a: 244 exciting opportunities G01: 184, L-B-G03: 312 method of application to unknown languages L-G00: 619 will allow vast volumes of text to be coded to move beyond the unitary actor assumptions Ba01a: 221 virtually unlimited possibilities L-B-G03: 312 new technique that takes words as data without heavy human input, can calculate confidence intervals by comparing reference texts to virgin texts P-K02: 76 enables new flexible ways of coding documents L01a: 243 has a long way to go before it can convince the sceptics B01b: 222 need to test estimates to know that they are valid and useful

Legend: B Budge; Ba Bara; C-M Castles-Mair; G Garry; G-H Gabel-Huber; H-I Huber-Inglehart; K-P Kleinnijenhuis-Pennings; L Laver; L-B Laver-Budge; L-B-G Laver-Benoit-Garry; L-G Laver-Garry; L-H Laver-Hunt; M Mair; M-C Mair-Castles; Mc-M McDonald-Mendes; P-K Pennings-Keman; R Ray; V-G-M de Vries-Gianetti-Mansergh.

traits point to party competition and to intra-party conflicts. For example, the fact that Finnish parties hardly ever mention foreign policy tells us that foreign policy is an uncontested policy dimension in this particular country. Parties also seem to avoid issues on which they are internally divided. A closer inspection of ES data on internal divisions could provide systematic answers to quite a number of interesting hypotheses when compared to MRG/CMP data. On the negative side, the

length of programs restricts analytic possibilities to a limited number of policy dimensions.

Only the MRG/CMP approach is judged to be costly. However, during the last ten years, the Social Science Research Center Berlin spent no more than 6500 euro a year for coding contracts and one tenure position to compile data sets covering each election and all relevant parliamentary parties in 50 countries. To be sure, this figure underestimates the costs of

Table 2

Validity of approaches to estimating policy positions of parties

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Determining the policy dimensions	<p>H-I95: 76 open-ended responses about nature and meaning of dimensions necessary</p> <p><i>L01a:</i> 242 imposed, match to actual policy concerns unknown</p> <p><i>H-I95:</i> 76 previous research imposes particular scales on respondents rather than letting them identify the major dimensions of political conflict</p> <p><i>L01a:</i> 243, <i>B01b:</i> 211 ambiguity about what, precisely, experts are locating</p> <p><i>Mc-M01:</i> 108–109 reason to worry whether experts are able to separate views about social from economic policy position</p>	<p>G-H00: 98 tests of country and time specificity: best estimates obtained by pooling all the MRG data, estimation technique matters</p> <p>B01b: 213 data produce good representations of the policy spaces envisaged by the classical theory</p> <p><i>L01a:</i> 241 data reduction problems, reading the tealeaves in interpretations based on face validity</p> <p><i>Br01:</i> 117 ideological dimensions take on their substantive meaning within specific geographical, historical political contexts</p> <p><i>Br01:</i> 129 multidimensional scaling more suitable than spatial analysis, esp. factor analysis to study party-voter relations</p>	<p><i>L01a:</i> 242 determining the algorithms</p> <p><i>L01c:</i> 8 deterministic or probabilistic method of generating the dictionary</p> <p><i>B01b:</i> 211 need for a known standard against which counting procedures can be checked</p> <p><i>L-B-G03:</i> 315 reference text should span the dimensions</p>
Left-right dimension	<p>H-I95: 77 let respondents list the key issues that divide parties</p> <p>M01: 27 increasingly professional</p> <p><i>H-I95:</i> 90 different types of criteria used by experts</p> <p><i>B00:</i> 107–108 comparability of judgements at different time points due to varying criteria questionable</p> <p><i>H-I95:</i> 76 no effort to identify the substantive meanings of scales in prior studies</p>	<p>B01a: 54 a left-right policy space emerges from the MRG codings which fits both theoretical concerns and historical records;</p> <p><i>Mc-M01:</i> 101–103 if expert judgements accepted as ‘true’, CMP scales have too much economic and too little social content, therefore slightly off the mark, but close approximations</p> <p>H-I95: 83 issues related to economic and class conflict by far the most frequently named by experts for the meaning of left and right</p>	Same as above
Left-right positions of parties	<p><i>H-I95:</i> 75 meaning of left and right varies over time and across cultures</p> <p><i>M-C97:</i> 151 in early surveys, all social-democratic parties could end up being assigned a similar left-right position regardless of particular positions in practice</p> <p><i>H-I95:</i> 80 validity of data difficult to check, crude test</p> <p><i>M01:</i> 23 range of the space may be a function of fragmentation of the party system</p>	<p><i>P03:</i> 67–68 extremely implausible results for some parties-time points in many party systems although they produce a left-right dimension</p> <p><i>P03:</i> 83–86 scores do not measure policy positions but willingness to change positions</p> <p>G-H00: 96 question whether MRG coding scheme provides the issue-specific information necessary for generating left-right positions</p>	Same as above

Table 2 (continued)

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Judged-stated versus 'real' policy position	<i>M01: 25 based on past coalition behaviour, party programmes and ideology, and both mass and elite perceptions, provides crude synthesis and short-cut</i>	<i>L-G00: 620 whether manifestos reflect 'real' positions is a fruitless discussion since this is an elusive, even metaphysical, notion</i>	Same as MRG/CMP
	<i>L-S90: 245 most likely to be conditioned by historical experiences of coalitions</i>	<i>L-H92: 31 parties can be held accountable for their published policies</i>	Same as MRG/CMP
	<i>B00: 111 snapshot lacks clear definition</i>	<i>L-G00: 620 difficult for party members to resile not sure what this word is from its policies; party leaders can be charged with failure to implement</i>	Same as MRG/CMP
		<i>Mc-M01: 90 explicit or implicit criticism that programs tell us about what parties have said but not about policy positions as such</i>	Same as MRG/CMP
		<i>T-L73: 215–216 differences important for coalition formation may not be found in manifestos, addressed to the electorate</i> <i>B-B01b: 65–69, L-B92 plausible results for coalition formation</i>	Same as MRG/CMP
Definition of party	<i>B01a: 50 ambiguity about the precise identity of the actor</i>	<i>B01a: 51, B01b: 211 only collective policy statement that parties as such ever make, combined views of the party as an organisation, endorsement by a representative gathering of the party</i>	Same as MRG/CMP
	<i>B00: 105–107 leaders, activists, or voters</i>	<i>B01b: 211 party as a collective actor as opposed to policies advocated by internal factions or individuals inside the party</i>	Same as MRG/CMP
Time scale	<i>L01a: 243, B00: 104, B01b: 211 ambiguity about time point</i>	<i>B01a: 50, B01b: 211 statements made at a particular point in time by a specific person or organisation</i>	Same as MRG/CMP
	<i>L01a: 243 cannot sensibly locate very large number of actors at different time points in the same space</i>	<i>B00: 110–111 tied to a particular date whose referents are clear</i>	Same as MRG/CMP
	<i>M01: 19 extension of snap-shots to other time periods</i>	<i>Mc-M01: 112 probably the only data set for observing dynamics and party movements</i>	Same as MRG/CMP
	<i>B00: 110–111 descriptive and static rather than analytic and dynamic</i>	<i>B01b: 215 traces movements of parties over a series of elections</i>	Same as MRG/CMP
	<i>Mc-M01: 100 no observable movement over time</i>	<i>Mc04: 16 whereas experts judge positions to be static, they are in fact volatile as is shown by CMP data, volatility produces accurate representation and is reflective of the faction that has gained temporary control</i>	Same as MRG/CMP
	<i>R99: 286 experts were asked to evaluate European integration positions at four time points, slight decrease in standard deviations over time</i>	<i>L-G00: 620 issued by each party at each election of the post-war period, provides historical evidence of movements, unparalleled way to retrieve data on party policy in the past</i>	Same as MRG/CMP

(continued on next page)

Table 2 (continued)

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
	<p><i>Mc-M01: 97–100, H-I95: 77–79</i> <i>high correlation between judgements 10–15 years apart</i></p> <p><i>Mc-M01: 100 little dynamic variation</i> <i>M-C97: 154 desirability of regular updates</i></p>		<p><i>L-B-G03: 314 assumption that manifesto at election t is valid reference for t + 1 unlikely to be 100% correct</i></p>
Cause and effect	<p>M01: 24 experts are well-read, well-informed, experienced and for the most part intelligent observers</p> <p><i>L01a: 239 entanglement of independent and depended variables</i> <i>B01a: 50 estimates policy positions at least in part from the actual behaviour they are designed to explain</i> <i>B00: 109–110 intentions and preferences or overt behaviour</i> <i>B01b: 211 confuse preferences with the actual behaviour</i> <i>B01a: 50 ambiguity about criteria used to locate actors</i> <i>See above</i></p>	<p><i>B01b: 211 studying texts has the advantage that they are statements of policy made at a particular place and time by a specific person or organisation</i> <i>L-G00: 622 precise time point so that time line between cause and effect can be established</i></p>	<p>Same as MRG/CMP</p> <p>Same as MRG/CMP</p>
Context		<p><i>G01: 184–185 coders take context of the text into account, potential of bias for a source known to the coder</i></p> <p><i>L01a: 240 preconceptions about ideological standpoints</i></p> <p><i>L-G00: 624, 625, 626 uses subjective judgement of a human coder, difficult to confine coder to context of text as opposed to other knowledge</i> <i>R99: 284 evaluation of coders may be a somewhat subjective and contextual exercise, subjective reading of texts</i> <i>L-B-G03: 311 texts are treated as discourses to be understood and interpreted</i></p>	<p><i>G01: 184 problem of ambiguous meaning of words may not in practice be as great as it seems, words in practice used by particular actors</i> <i>G01: 185, L-G00: 625 coding text out of context not susceptible to any biases</i> <i>G01: 184 no account of the external political and social context</i> <i>R01: 152 problems of homography, context, contingency, and order</i> <i>L-G00: 626 changing political meanings of words over time and space; good case for reconstructing the dictionary as often as we have independent data sources against which to cross-validate it</i></p>
Unit of analysis	No problem at all	<p><i>P-L01: 133 pledges or all text to be analysed</i> <i>L-G00: 620 621 MRG on a sentence by sentence base, Harmel-Janda scales based on the entire manifesto</i> <i>L-G00: 624 definition of 'quasi-sentence' might be source of unreliability</i></p>	<p><i>L01c: 8 word or more complex unit</i> <i>B01b: 211 results and maps differ considerably due to different bases of word-counts</i> <i>R01: 152 ambiguity of individual words or whole sentence, painstaking work to devise rules</i></p>

Table 2 (continued)

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Distinction between position and salience	<p><i>M01: 20 accuracy and consistency of estimates as a function of saliency of particular policy dimension (observer uncertainty)</i></p> <p><i>B00: 108–109 unrealistic contrasts between scalar opposites, hypothetical policy contrasts which parties try to avoid in their own declarations</i></p>	<p><i>J-H-E-G95: 178 taps relative intensities with which issues are addressed</i></p> <p><i>B01a: 52 key difference between parties is varying extent to which they mention one-position issues</i></p> <p><i>B87: 24 parties gloss over areas which might favour their rivals while emphasizing those on which they feel they have an advantage</i></p> <p><i>B-B01a: 12 a misapplied or wrong-headed theory is unlikely to support relevant codes and procedures, ability to ‘work’ is an implicit check on the theory</i></p> <p><i>L01b: 66 position and salience are two distinct components of the policy of a given actor</i></p> <p><i>K-P01: 180 conclusion that political positions have both saliency and direction</i></p> <p><i>S-P04: 76 position and saliency should be measured independently to test theories</i></p>	<p><i>L-G00: 625 intervention due to ambiguous words and relying on text units results in hybrid techniques</i></p> <p><i>B01b: 211 need for a known standard against which counting procedures can be checked</i></p> <p><i>B-B01a: 11 MRG/CMP estimates and computerized ones are both based on saliency assumptions</i></p> <p><i>B01b: 213 counting the frequency of words by computer is in fact to measure relative emphasis on one policy position</i></p> <p><i>B01a: 52 saliency relates to word counts</i></p>
Classification scheme	<p><i>Same as computer-assisted content analysis: selecting policy dimensions and definitions for polar opposites involves crucial decisions by the researcher, equivalent to choice of coding scheme</i></p>	<p><i>A01: 77 summary of criticism of MRG/CMP classification scheme: robustness, too broad, too narrow, country-specific subcategories, time-specificity, unbalanced with respect to number of categories in policy areas</i></p> <p><i>L-H92: 31 coding scheme conditions the precise empirical description of the manifesto</i></p> <p><i>L-G00: 623 need for a more fine-grained coding scheme, policy concerns coded in a tripolar way</i></p> <p><i>P-K02: 55 overlapping or missing categories</i></p>	<p><i>L-G00: 619 retrospective estimation of past party systems requires extraction method from texts</i></p> <p><i>L-B-G03: 314 selecting reference texts involves crucial decisions by the researcher, equivalent to choice of coding scheme</i></p>

A Agasoster; B Budge; B-B Budge-Bara; Br van der Brug; G Garry; G-H Gabel-Huber; H-I Huber-Inglehart; J-H-E-G Janda-Harmel-Edens-Goff; K-P Kleinnijenhuis-Pennings; L Laver; L-B Laver-Budge; L-B-G Laver-Benoit-Garry; L-G Laver-Garry; L-H Laver-Hunt; L-S Laver-Schofield; M Mair; Mc McDonald; Mc-M McDonald-Mendes; M-C Mair-Castles; P Pelizzo; P-K Pennings-Keman; P-L Petry-Landry; T-L Taylor-Laver; R Ray; S-P Shikano-Pappi.

conventional content analyses for a number of reasons. First of all, developing classification schemes on which conventional content analyses rest is time consuming, and the bulk of resources invested by the MRG/CMP in the early 1980s does not show up in this figure. Second, the MRG/CMP is confined

to relevant parties in parliament. Third, although many programs have been checked and recoded, MRG/CMP estimates are usually based on the codings of one expert only. Fourth, for many countries, MRG/CMP hired PhD students who used the data for their own theses, so that small money sufficed as added

Table 3

Reliability of approaches to estimating policy positions of parties

	Expert surveys	MRG/CMP analysis of manifestos	Computer-assisted analyses
Results of reliability tests	<p>Mc-C97: 153 questionnaire should include pre-defined sets of parties, 1–10 point scales, publication of questionnaire</p> <p>Mc-M01: 97–100, H-I95: 78–79 highly reliable, high correlation between different C-M, L-H, and H-I expert scales (close to 0.95%)</p> <p>L-H92: 37, H-I95: 76 at least 5 experts per country to get reliable rating</p> <p>H-I95: 79–80 standard deviation for newly established democracies same as for developed democracies, <i>but low in non-democratic countries</i></p> <p>R99: 288 <i>careful analysis of responses revealed a few anomalous respondents, 7 and 5 ‘suspect’ out of 135 evaluations, 2% of estimates substantially affected by suspects</i></p> <p>Mc-C97: 156 <i>multitude of H-I95 party scores are based on a handful of respondents</i></p> <p>C-M84: 76 <i>high range of scores for some parties</i></p> <p>Mc-C97: 152–153 C-M84 <i>too few experts, double figure response missing, range of scores sometimes quite extensive, not all relevant parties scored by all respondents</i></p>	<p>Mc-M01: 104–106 as highly reliable as the expert scales</p> <p>H01: 124 potential for methodological error between MRG and CMP does not occur</p> <p>V01b: 109 complete recoding of German programs rendered only marginal deviations</p> <p>B01b: 212 series of internal checks and controls along with instantaneous electronic contact between coder and supervisor</p> <p>L-B-G03: 311 text analyses can be replicated, modified, improved</p> <p>L01a: 243 <i>potentially unreliable</i></p> <p>G01: 185 <i>often leads to high levels of disagreement between coders and between the same coder at two time points</i></p> <p>L-B-G03: 329 <i>identifies the meaning of the text, understands meaning in context which can be easily done by human coders, but is apt to be unreliable</i></p>	<p>G01: 185 far more reliable than hand coding</p> <p>B-B01a: 10 given procedure applied to the same text will always produce the same results</p> <p>L-G00: 625 <i>reliability issues are raised when phrases or word strings from different authors are related to some underlying dimension due to stylistic idiosyncrasies</i></p> <p>Same as MRG/CMP</p>

B Budge; B-B Budge-Bara; C-M Castles-Mair; G Garry; H Hearl; H-I Huber-Inglehart; L Laver; L-B-G Laver-Benoit-Garry; L-G Laver-Garry; L-H Laver-Hunt; M-C Mair-Castles; Mc-M McDonald-Mendes; R Ray; V Volkens.

value. Fifth, numbers of relevant parties are growing in many party systems. Last but not least, the length of programs is growing over time. The latter two facts predict rising future costs for updating the MRG/CMP time series.¹ However, the MRG/CMP approach can be applied without much prior knowledge on content analytic procedures, whereas users of CACA have to acquire a lot of learning before any analysis can even be started. The application of CACA is also in need of time-consuming preparations if no machine-readable versions of programs are available.

¹ The latest publication covering 25 countries and the time period between 1945 and 1998 is given on a CD-ROM appended to Budge et al. (2001). The next data publication covering 50 countries for the time period between 1990 and 2003 will be provided in the same format in Klingemann et al. (2006, in print).

Concerning possibilities for further research, CACA is certainly the first choice whenever further dependent or independent variables can be operationalised on the basis of large amounts of text. There are, however, many non-textual variables and on many of these experts can make informed judgements. As opposed to both CACA and ES, the future prospect of the MRG/CMP data is supposed to be marginal, since its function is mainly seen in assisting the CACA to get valid estimates. There are, though, at least two circumstances in which the MRG/CMP data is superior: firstly, it provides time series data that can be compared to other time series data. And secondly, the classification scheme has been successfully applied to local (Agasoster, 2001) as well as European election programs (Wüst and Volkens, 2003), so that multiple levels of party competition can be compared. Thus, all three

approaches are needed because the choice of one over the others depends on the specific research question.

2.2. *Validity of approaches*

The validity of estimates considers whether the used approach measures the intended concept. To achieve a valid determination of the policy dimensions in general or of the left-right dimension and the ‘true’ policy positions of parties in particular presents challenges to all three approaches, since this is the goal of the whole procedure.

Although several methods have been suggested to improve their validity, the systematic comparison clearly shows that none of the approaches can claim absolute truth in determining the relevant policy dimensions and the ‘true’ positions of parties on them.

The major weaknesses of ES concern the validity with respect to the definition of party, time scale, and cause and effect. Expert surveys are criticised because of their ambiguity about the precise understanding of parties as collective actors, of the time period for which estimates are valid, and of the distinction between programmatic promises and resulting actions. In these instances, MRG/CMP and CACA are in an advantageous position because they rely on election programs. Election programs are official documents agreed upon by the party as a whole and published at a precise point in time, so that problems with respect to definition of the party, time scales, and cause and effect are of minor importance. Reliance on official publications of parties constitutes the most positive traits of MRG/CMP and CACA because parties in office can be held accountable for pledges mentioned in their election programs. Parties out of office often fight over the content because they deliberate the major outlines of policies that are binding to factions of the party.

Problems of context are closely related to problems of cause and effect, but take on a different meaning when MRG/CMP and CACA are compared. On this account, the old discussion of qualitative as opposed to quantitative analyses crops up. It is impossible to repeat all the arguments exchanged over several decades in one short paper. A few examples may suffice to make the point. The MRG/CMP method with which country and time specific issues are accorded to more general, abstract policy positions, which travel across cultures and times, is highly contextual. In this respect, it is a qualitative approach. Although coders are advised to avoid inferences on what statements may mean in reality, they indeed need a lot of contextual knowledge on

policies, politics, and polities beyond the respective text to decipher its meaning. Many election programs are much more specific than most people think, so that abbreviations for all kinds of projects in all kinds of policy areas appear in the text. Coders need to know about the status quo to decide whether a class frequency of 20 pupils means extending or limiting education; they need prior knowledge on what cold climate allowance, post-code treatment, or deep tax is. All conventional content analytic approaches resting on classification schemes, in fact, rely heavily on coders’ knowledge and abilities to abstract analysis. As opposed to ES, all questionable MRG/CMP results can be checked by looking at the text base.

However, relying on a text base does not only present challenges to conventional content analysis. Although choosing a text unit is a non-issue for ES, it is an impediment to conventional as well as computer-assisted analyses because the resulting estimates will always differ, more or less, according to the chosen unit. MRG/CMP and CACA, though, have more in common than just the text base. There are different computer-assisted methods, and not all are context-free word counts. The more human intervention a computer assisted method uses for determining the policy dimensions, the closer it comes to conventional content analytic approaches with their strengths and weaknesses. That computer-generated estimates based on mere word counts are not as unequivocal as one might think is shown by Budge/Pennings in this volume.

A specific problem for conventional content analytic approaches lies in the design of classification schemes, and it is this particular aspect of MRG/CMP on which doubts are often raised. Most of these doubts, however, are due to misreadings of the saliency approach, which is not non-positional but for the most part one-positional. Categories reported as missing are specific thematic concerns such as unemployment, not policy positions such as fighting unemployment by means of free market mechanisms or state intervention. Some argue that one should set up a classification scheme to test assumptions of saliency theory, but that was, in fact, done while the scheme was being developed when a number of counter-positions were introduced.

A basic check of the validity of a classification scheme is the proportion of text units left as uncoded. A classification scheme is not or no longer valid if this proportion is high or if it increases considerably over time. The percentage of uncoded (quasi)sentences, however, has been continually dropping from about ten per cent in the 1940s to the 1970s (mainly due to the exceptionally high percentages of uncoded sentences for

pre-1970 Danish and Italian programs) down to about two per cent in the most recent programs.

2.3. Reliability of approaches

As far as CACA relies on word counts, reliability, defined as the measurement problem of getting the same results whenever the method is applied, does not pose any problem at all. But both ES and MRG/CMP are confronted with it. As is emphasized in the literature, the ES became much more professional over time. Much depends on the number of experts available to judge the party positions. The range of their responses can be used as a measure for reliability and deviant judgements can be deleted.

Whereas reliability of ES can be easily checked and improved, it is a major obstacle to conventional content analysis. Therefore, many different kinds of reliability tests have been suggested, all of which have to be adjusted to particular projects. The MRG/CMP has been subjected to quite a number of them, with encouraging results. Among others, the CMP uses a specific kind of accuracy test to train coders (Volkens, 2001a,b). Contrary to studies where a text of average difficulty is chosen to test trained coders, the CMP has chosen the most difficult text it could find to trigger mistakes. The training phase starts after tests are filled in, based on evaluations of mistakes occurring in the tests. Results of these tests show no more and no less than the ability of coders to get a difficult text right after having studied the coding instructions given in the handbook.

First evaluations of test results led to revisions of the handbook. Additional rules were introduced to prevent regularly occurring mistakes. Meanwhile, 23 coders have completed the test after studying the second version of the handbook (Appendix II in Budge et al., 2001). Their average correlation to a master copy was 0.83. Of these 23 coders, 14 coders got their first contract and filled in the reliability test for the first time with an average correlation to the master copy of 0.82. Compared to the tests based on studying the former version of the handbook with an average of 0.72, the specification of additional rules seems to pay off.

In addition, nine coders got a second coding contract and filled in the reliability test a second time, so that results of their first test can be compared to the second results. Again, this is not a usual intra-coder reliability test because several years lie between the tests, whereas it is just weeks in other projects. In the first round, the average correlation of these nine coders to the master copy was 0.70, in the second round it was 0.85. All

coders did a better job the second time. Although even these improved figures certainly leave much to be desired, one should be aware of the fact that there is no way of getting 100 per cent identical results with conventional content analytic approaches. The merits of taking contexts into account are always offset by some deficiencies in reliability.

3. Conclusions

The overview of the existing literature on measurement quality of three approaches to estimating parties' policy positions — expert surveys, the conventional content analytic MRG/CMP approach, and computer-assisted content analysis of manifestos — shows that all three approaches have merits as well as deficiencies. Each approach buys the strength in some aspects of measurement quality at the cost of measurement quality in other aspects. When the three approaches are compared, it is evident that the strength of one approach is the weakness of the others, and vice versa. Therefore, the three approaches are not opposed to one another but complementary. In such situations, method mixing is the first choice whenever possible. In addition, the suitability of the approaches for employment in future research largely depends on features of further dependent or independent variables. Therefore, all three approaches are needed because the choice of one over the others hinges on the specific research question.

Up to now, divergences in estimates were usually discarded as measurement errors. The few examples mentioned in this paper show, though, that divergences are not necessarily due to a flaw in one of the estimates but can gain substantive meaning. What up to now was considered as crucial failures may in the future provide answers to questions of party democracy.

References

- Agasoster, B., 2001. A framework for analysing local party policy emphases in Scotland. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 76–89.
- Bara, J., 2001a. Tracking estimates of public opinion and party policy intentions in Britain and the USA. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 217–236.
- Bara, J., 2001b. Using manifesto estimates to validate computerized analyses. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M. (Eds.), *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York, pp. 143–156.

- Bartolini, S., Mair, P., 1990. Policy competition, spatial distance and electoral instability. *West European Politics* 13, 1–16.
- van der Brug, W., 2001. Analysing party dynamics by taking partially overlapping snapshots. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 115–132.
- Budge, I., 1987. The internal analysis of election programmes. In: Budge, I., Robertson, D., Hearl, D. (Eds.), *Ideology, Strategy and Party Change. Spatial Analysis of Post-War Election Programmes in 19 Democracies*. Cambridge University Press, Cambridge, pp. 15–38.
- Budge, I., 2000. Expert judgements of party policy positions: use and limitations in political research. *European Journal of Political Research* 37, 103–113.
- Budge, I., 2001a. Validating the Manifesto Research Group approach: theoretical assumptions and empirical confirmations. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 50–65.
- Budge, I., 2001b. Validating party policy placements. *British Journal of Political Science* 31, 210–223.
- Budge, I., Bara, J., 2001a. Introduction: content analysis and political texts. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M. (Eds.), *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York, pp. 1–18.
- Budge, I., Bara, J., 2001b. Manifesto-based research: a critical review. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M. (Eds.), *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York, pp. 50–73.
- Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M., 2001. *Mapping Policy Preferences: estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York.
- Castles, F.G., Mair, P., 1984. Left-right political scales: some 'expert' judgements. *European Journal of Political Research* 12, 73–88.
- Gabel, M.J., Huber, J.D., 2000. Putting parties in their place: inferring party left-right ideological positions from party manifesto data. *American Journal of Political Science* 44, 94–103.
- Garry, J., 2001. The computer codings of political texts: results from Britain, Germany, Ireland and Norway. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 183–192.
- Hearl, D.J., 2001. Checking the party policy estimates: reliability. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M. (Eds.), *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York, pp. 111–125.
- Huber, J., Inglehart, R., 1995. Expert interpretations of party space and party locations in 42 societies. *Party Politics* 1, 73–111.
- Janda, K., Harmel, R., Edens, C., Goff, P., 1995. Changes in party identity: evidence from party manifestos. *Party Politics* 1, 171–196.
- Kleinnijenhuis, J., Pennings, P., 2001. Measurement of party positions on the basis of party programmes, media coverage and voter perceptions. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 162–182.
- Klingemann, H.-D., Volkens, A., Bara, J., Budge, I., McDonald, M., 2006. *Mapping Policy Preferences II: Comparing 24 OECD and 24 CEE Countries, 1990–2003*. Oxford University Press, Oxford, in print.
- Laver, M., 2001a. How should we estimate the policy positions of political actors? In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 239–244.
- Laver, M., 2001b. Position and salience in the policies of political actors. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 66–75.
- Laver, M., 2001c. Why should we estimate the policy positions of political actors? In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 3–9.
- Laver, M., Benoit, K., Garry, J., 2003. Extracting policy positions from political texts using words as data. *American Journal of Political Science* 97, 311–331.
- Laver, M., Budge, I., 1992. *Party Policy and Government Coalitions*. St. Martin's Press, New York.
- Laver, M., Garry, J., 2000. Estimating policy positions from political texts. *American Journal of Political Science* 44, 619–634.
- Laver, M., Hunt, W.B., 1992. *Policy and Party Competition*. Routledge, New York and London.
- Laver, M., Schofield, N., 1990. *Multiparty Government: The Politics of Coalition in Europe*. Oxford University Press, Oxford.
- Mair, P., 2001. Searching for the positions of political actors: a review of approaches and a critical evaluation of expert surveys. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 10–30.
- Mair, P., Castles, F.G., 1997. Revisiting expert judgements. *European Journal of Political Research* 31, 150–157.
- McDonald, M., 2004. Parties in democracy, democracy in parties: lessons I learned from Ian Budge and the CMP data. Paper presented for the Conference and Festschrift in Honour of Ian Budge, University of Essex, Colchester, UK (May 26–28).
- McDonald, M., Mendes, S., 2001. The policy space of party manifestos. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 90–114.
- Pelizzo, R., 2003. Party positions or party direction? An analysis of party manifesto data. *West European Politics* 26, 67–89.
- Pennings, P., Keman, H., 2002. Towards a new methodology of estimating party policy positions. *Quality and Quantity* 36, 55–79.
- Petry, F., Landry, R., 2001. Estimating interparty policy distances from election programmes in Quebec, 1970–89. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 133–146.
- Ray, L., 1999. Measuring party positions on European integration: results from an expert survey. *European Journal of Political Research* 36, 283–306.
- Ray, L., 2001. A Natural Sentence Approach to the Computer Coding of Party Manifestos. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 149–161.
- Sani, G., Sartori, G., 1983. Polarization, fragmentation and competition in western democracies. In: Daalder, H., Mair, P. (Eds.), *Western European Party Systems. Continuity and Change*. Sage, London and Beverly Hills, pp. 307–340.
- Shikano, S., Pappi, F.U., 2004. Ideologische Signale in den Wahlprogrammen der deutschen Bundestagsparteien 1980 bis 2002, Working Paper No. 76, Mannheimer Zentrum für Europäische Sozialforschung, Universität Mannheim.

- Taylor, M., Laver, M., 1973. Government coalitions in Western Europe. *European Journal of Political Research* 1, 205–248.
- Volkens, A., 2001a. Manifesto research since 1979: from reliability to validity. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 33–49.
- Volkens, A., 2001b. Quantifying the election programmes: coding procedures and controls. In: Budge, I., Klingemann, H.-D., Volkens, A., Bara, J., Tanenbaum, E., with Fording, R.C., Hearl, D.J., Kim, H.M., McDonald, M., Mendes, S.M. (Eds.), *Mapping Policy Preferences: Estimates for Parties, Electors, and Governments 1945–1998*. Oxford University Press, New York, pp. 93–109.
- de Vries, M., Giannetti, D., Mansergh, L., 2001. Estimating policy positions from the computer coding of political texts: results from Italy, the Netherlands and Ireland. In: Laver, M. (Ed.), *Estimating the Policy Positions of Political Actors*. Routledge, London and New York, pp. 193–216.
- Wüst, A.M., Volkens, A., 2003. *Euromanifesto Coding Instructions*, Working Paper No. 64, Mannheimer Zentrum für Europäische Sozialforschung, Universität Mannheim.