Pola Lehmann/Henrike Schultze

Handbook:

The MARPOR Party-Voter Data Set

Using election and survey data to measure congruence between citizens and parties

Version: 2012-09-30

Citation

When using data from the MARPOR Party-Voter Data Set please cite the data set as well as the respective primary sources:

MARPOR Party-Voter Data Set:

Volkens, Andrea/Lehmann, Pola/Regel, Sven/Schultze, Henrike/Werner, Annika (2011): The MARPOR Party-Voter Data Set . Manifesto Project (MRG/CMP/MARPOR), Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

Survey Data:

European Values Study 1981-2008, Longitudinal Data File. GESIS Data Archive, Cologne, Germany, ZA4804 Data File Version 1.0.0 (2011-04-30) DOI:10.4232/1.14804.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 1 FULL RELEASE [data set]. Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor]. August 4, 2003.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 2 FULL RELEASE [data set]. June 27, 2007.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 3 ADVANCE RELEASE [data set]. March 31, 2011.

The Mannheim Eurobarometer Trend File 1970-2002. Schmitt, Hermann. Mannheimer Zentrum für Europäische Sozialforschung (MZES). Scholz, Evi/Leim, Iris/Moschner, Meinhard. GESIS – Leibniz Institute for the Social Sciences (former "Zentrum für Umfragen, Methoden und Analysen" and "Zentralarchiv für Empirische Sozialforschung") (http://www.gesis.org/eurobarometer).

Volkens, Andrea/Lacewell, Onawa/Lehmann, Pola/Regel, Sven/Schultze, Henrike/Werner, Annika. 2011. The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR). Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB). (http://manifestoproject.wzb.eu/).

World Values Survey 1981-2008 Official Aggregate v.20090901, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid. (http://www.wvsevsdb.com/wvs/WVSData.jsp?Idioma=I).

When referring to this handbook please cite the following:

Lehmann, Pola/Schultze, Henrike (2012): The MARPOR Party-Voter Data Set Handbook, Version 2012-09-30. Berlin: WZB.

Introduction

Over the last funding period since October 2009, MARPOR compiled a combined data set which merges data from the Manifesto Data Collection with survey data. With this combined Party–Voter Data Set, we provide new possibilities for the study of the quality of representation. For the first time it is possible to address the question of how well voters (party supporters or unaligned citizens) are represented by parties on a broad empirical basis. This is possible as the combined data set provides information on the relationship between voters and parties on the individual level. The data is available cross-nationally and over-time. Using survey and election data, MARPOR therefore provides the relevant information to assess questions of representation and of the functioning of the delegation mechanisms from the represented voters to the representing parties in a broad comparative perspective.

State of the Representation Literature: Theory, Concepts and Measures for Studying Programmatic Representation

To link MARPOR research to the tradition of studying representation we briefly summarise the line of discussion in modern representation literature below.

Approaches to Analysing the Relationship between Voters and Parties

In the political science literature a broad body of theoretical debate and numerous studies address the issue of representation in modern democracies. In this literature '[p]olitical representation is [generally understood as] the process in which one individual or group (the representative) acts on behalf of other individuals or groups (the represented) in making or influencing authoritative decisions, policies, or laws of a polity.' (Thompson 2001, 11696). With the emergence of modern democracies representation processes need to be legitimised by popular vote. As Weßels phrases it: 'Political representation is at the heart of liberal democracies' (2007, 833). He shares this understanding with Pitkin (1967), who supports the democratic ideal of having a legislature that accurately reflects the preferences of the citizenry as a whole. In order to achieve such representation in modern democracies free, fair and competitive elections become pivotal (Pitkin 1967, Powell 2000). With this development modern

democracies establish complex institutional settings to structure the election and delegation mechanisms in order to ensure legitimacy (Dahl 1989). Accordingly, with the evolution of modern party democracies a part of representation research focuses on the relationship between the elected and the citizens. Scholars investigate how well the latter are represented by the former in the electoral delegation process. It posits citizens as principals and elected individuals or parties as agents and functions as the core mechanism in the selection of political elites and the distribution of power.

With regard to the relationship between voters and elected individual or collective representatives (e.g. parties) two bodies of empirical literature on representation can be distinguished. On the one hand there is the institutionalist perspective on electoral representation that focuses on vote–seat disproportionality to measure how well voters' party preferences are transformed in the composition of the parliament. On the other hand, there is the substantive representation literature that specifically analyses the correspondence of parties' policy preferences and behaviour with the substantive policy preferences of citizens. Voting in this understanding is the linking mechanism between the two that can assure issue congruence and high levels of representation through responsiveness and accountability.

Based on the normative claim that democratic representation should ensure that citizens' substantive preferences are represented by parties (and subsequently by parliaments and governments) scholars sought to test empirically whether such 'good representation' is achieved in everyday democratic processes. The empirical literature on substantive representation has a long tradition and covers different approaches and indicators for measuring substantive representation. Within the framework of 'substantive representation', the issue congruence between represented and representative is used as an adequate indicator for the quality of representation (Golder/Stramski 2010; Powell 2004; Weßels 1999).

State of the Empirical Representation Literature

Empirical studies on issue congruence emerged in the America of the 1960s. One of the first empirical analyses in this field addressed the 'dyadic' congruence between constituencies and elected representatives using correlation coefficients (Miller/Stokes

1963). This study is criticised with regard to two points. The first line of criticism addresses the conceptualisation of representational connections. The 'geographically' structured delegation between constituencies and representatives leads to a dyadic conception of representation disregarding the nature of collective representation of citizens through parties and party representatives as in the case in Western Europe (Barnes 1977; Dalton 1985; Weissberg 1978). Therefore, it is debatable whether the concept of dyadic representation can be used in the European context. The second line of criticism is of a methodological nature, arguing that the use of the correlation coefficient is insufficient to capture representational relationships. As it is not able to take into account the absolute distance between represented and representative the operationalisation of the quality of representation as a correlation leads to a puzzling result. Under certain circumstances, even if represented and representative are considerably far apart, an analysis might result in a high correlation that seems to indicate a high quality of representation. This is the case as long as the variance of the position of the representative is explained by the variance of the represented (Achen 1978, Dalton 1985).

These critics have inspired further contributions to the empirical literature on substantive representation. Scholars have built on Miller/Stoke's epoch-making study with theoretical-conceptual as well as methodological developments. With regard to conceptual advancements, the concepts of collective representation and responsiveness, as opposed to dyadic congruence measures, have been suggested for the analysis of the relationship between voters and parties, In this context, for example, Barnes (1977), Converse/Pierce (1986), Dalton (1985), Thomasson (1994, 1999, 2005) and Weisberg (1978) discussed the implications multi-party competition has on representation, as in the case of Western European democracies. In order to address the different nature of the representational mechanism in party democracies Dalton (1985), Thomassen (1994) and Schmitt/Thomason (1999) developed the responsible party model in order to analyse the more complex model of delegation from voters to parties to governments. Moreover, Holmberg (1999), interested in the nature of the representational relationship between voters and parties, introduced the AJUS system to analyse issue agreement (in contrast to simply analysing issue congruence) in order to detect to

which extent voter preferences are mirrored by the preferences of the representatives as can be seen in the distribution curves.

With regard to methodological advances, scholars have introduced distance and distribution measures that allow for an analysis of how close voters and representatives are in absolute and relative terms. Achen (1977, 1978) is the first who argued against the correlation coefficient as a single measure for representational relationships between represented and representatives. Accordingly, he proposed to complete the picture on representation by introducing several new measures. With the concepts of proximity and centrism his approach allows for an investigation of how close voters and representatives are in absolute terms. The measures are different insofar as the first is particularly sensitive to the dispersion of the voter preferences, while the latter utilises this dispersion as a control measure for reasons of comparability. In order to address representation in relative terms he introduced the concept of responsiveness, operationalised as the regression between the positions of the representative and the represented. Most of the following empirical contributions on issue congruence, such as those by Dalton (1985), Herrara/Herrara/Smith (1992), Thomasson (1994, 1999, 2005), Weßels (1995), Holmberg (1999), and Powell (2000) build on Achen's approach and use his concepts or variations of different distance and distribution measures.

To date, a broad literature exists on representation with different perspectives on voter–representative[1] relationships from a theoretical view as well as an empirical–conceptual perspective. Despite of, or more likely, due to the sheer number of studies and contributions in this field scholars as of yet have reached no agreed understanding of representation and conceptualisation of congruence. As a consequence, there is an ongoing debate in the literature on different measures to analyse issue congruence as an indicator of representation.

The latest pivotal contribution to the discussion on how to adequately operationalise congruence in the context of Western party democracies stems from Golder and Stramski (2010). They criticise that scholars as to now give little attention to the conceptualisation of congruence in the context of their research question and research

---

[1] No matter whether it is in form of delegate, party or government.

object. They argue that the conceptualisation should first take the nature of representation in the respective political system into account and should secondly depend on the research question. As a consequence, Golder and Stramski (2010) conceptualise their understanding of representation in delimitation from other approaches inspired by concepts of dyadic and collective representation in the American political science literature. As a meaningful conceptualisation of representation in Western, and especially Western European party democracies they distinguish five different congruence measures, depending on the number of represented and representatives involved in the representation relationship. They argue that, depending on the research interest at hand, each of the different measures has its specific use:

1 The one-to-one congruence describes the absolute distance between an individual voter and her/his representative as the most direct (and relatively simple) concept of representation.

2 With regard to many-to-one relationships it is possible to analyse whether the policy of one representative / representative body is congruent with citizen preferences.

   2.1   The absolute median citizen congruence measures how well the citizenry in general is represented in its most preferred policy position by the representative, but without accounting for the distribution of preferences within the citizenry.

   2.2   Since this might be problematic in some (or perhaps most) research contexts Golder and Stramski define their second measure, absolute citizen congruence as the average absolute distance between individual citizen preferences and the position of the representative.

   2.3   But as this measure is dependent on the dispersion of citizen preferences it might evaluate the representational performance of one specific representative in comparison to another too critically in cases where the citizens' preferences are spread within a broader range. Therefore, Golder and Stramski advocate the measure of relative citizen congruence for comparative analyses of representational performance. It is operationalised

as the absolute difference between citizen and representative relative to the dispersion of citizens' preferences.

3 Conceptualised as a many-to-many relationship the third measure proposed by Golder and Stramski takes into account the representational performance of the collective body of representatives and allows for the comparative analysis of the shape and the location of citizen and representative preferences.

In sum, with its most recent developments towards clarifying the concept of and agreeing on methods for the measurement of representation, the literature provides a very elaborate framework and instrument for an accurate analysis of the quality of representation in modern party democracies.

Programmatic Congruence as a Measure for the Quality of Representation

The interest in party-voter congruence in the context of the MARPOR project is based on the literature in which congruence is established as a measure for representation. The relationship between parties and citizens in modern democracies is defined by a process of interest articulation and aggregation (Webb/Farrell/Holliday 2002). Parties play, or at least are intended to play, a central role in the political process by transforming popular interests into public policies (Hurley/Hill 2003; Kitschelt 2000; Lawson/Poguntke 2004; Montero/Gunther 2002; Pierce 1999; Römmele/Farrell/Ignazi 2005). With respect to this relationship, voter[2]-party congruence is both the indicator for the quality of the representation of citizens by parties as well as the 'central mechanism through which representation [of the mass] via parties can occur' (Rohrschneider/Whitefield 2010, 2). The closer parties are to their voters, it is assumed, the more likely they transform citizens' interests into policy outcomes. Vice versa, the bigger the gap between parties and voters, the less likely citizens can affect political decisions.

Following these considerations, the distance between parties and voters could be described in rather general terms as the extent to which the former resemble the latter with respect to shared characteristics, interests or experiences. In the context of our

---

[2] or citizen or party supporter – in the following, we use the term voter, but referring to citizens in general or party supporters alike. But as the main delegation mechanism occurs via voting, party-voter congruence is the central concept of interest.

research, however, we employ a narrow understanding of representation using it in its substantive meaning of policy representation. Accordingly, by representation we understand the mechanism by which the parties as representatives advance the policy preferences that serve the interests of the represented voters. The following analysis provides an exploratory examination of the level of political representation in terms of policy congruence between voters and parties. As a measure for policy congruence this analysis uses the distance between voters' preferences and the parties' programmatic supply. Since the left–right dimension has shown to be a common reference dimension for both party competition as well as party and self-placement by voters, we aggregate the opinion agreement of voters' preferences and parties' manifestos on the 'super-issue' of left–right placements and positions (for many see Bobbio 1996; Castles/Mair 1984; Gabel/Huber 2000; Fuchs/Klingemann 1990; Huber/Inglehart 1995; Inglehart/Klingemann 1976; Müller 1994; Sani/Sartori 1983; Thomassen 2005; Warwick 2002).

The Combined Party–Voter Data Set

In the following section we provide an overview of the specific data set MARPOR has created to address the question of how well voters are substantively represented by parties. On the basis of manifesto data and survey data from four of the most relevant international surveys, this new data set is suitable for measuring programmatic congruence between parties and voters. The particular strength of the data set is that it allows scholars to address this subject on different levels of analysis on a common empirical basis.

Data Set Aim

The organisation of the data set is chosen to contribute to the most recent discussion in the empirical study of substantive representation. The representation literature differentiates three levels of analysis to address questions of representation: firstly, the individual level of voters, secondly, the aggregate meso level of single parties or governments and thirdly, the aggregate macro level of party or political systems. So far, research on representation was mostly limited to either an aggregated level when

wanting to compare data over-time and between countries or to an analysis focusing on special cases when being interested in the individual level.

In this context, Golder and Stramski (2010, 90 and 93) criticise that the majority of the current comparative studies on substantive representation relate the median or mean position of voters to the government or the median or mean position of representatives in order to measure congruence. Researchers often made this decision not for conceptual reasons, as established in the median voter/median mandate literature, but for reasons of data restrictions only. This has, as Golder and Stramski point out, important implications for the accuracy of the measure and possible conceptualisation of the research at hand. Such an aggregate measure does neither allow to account for the distribution of voters' nor for representatives' preferences. Especially in the context of Western European party democracies, however, the complexity of the delegation mechanisms requires a conceptualisation of collective representation that is able to consider adequately the distribution of preferences of both, represented and representatives.

The Party–Voter Data Set accounts for Golder and Stramski's criticism. By combining comparative survey and election data, we can expand the possibilities for analyses to the individual level in a broad comparative perspective, as well as to different levels of aggregation. As a consequence, the level of analysis can be chosen on the basis of conceptual considerations in accordance to the research question at hand. To highlight the specific use of the data set for representation research, we will give a description of the newly combined Manifesto Data Collection and survey data set in the following. We conclude with discussing methodological implications that result from using different data sources to address the distance between voters' preferences and parties' programmatic supply as a measure for congruence. These considerations might be useful information for users on how to employ the data set in their own research.

Data Sources

MARPOR is first and foremost interested in the programmatic congruence between parties and voters. To measure this, we use information on parties' policy positions and voters' policy preferences on the 'super issue' of left–right. The first can be derived from

manifesto data while the latter information is regularly provided in survey data. With regard to policy positions derived on the basis of manifestos the new data set includes the Manifesto Data Collection in its most up-to-date version (version 2012a, see https://manifesto-project.wzb.eu/). With regard to surveys, four different studies were included[3]:

- Firstly, all data from the Eurobarometer (EB) Trend file, which covers the original 15 European member countries and Norway, spanning the years 1970-2002, with the exact time covered for each country dependent on its accession to the European Union.

- Secondly, Modules I, II and III from the Comparative Study of Electoral Systems (CSES). In these several more countries are included, such as the East European countries and some non-European democracies that are not part of the EB data.

- Thirdly, the World Value Survey (WVS) Five Wave Aggregated File 1981-2008 and the European Values Study (EVS) Longitudinal Data File 1981-2008 were included. These two data files cover the five waves of the World Values Survey and four waves of the European Values Study, and thus data from 87 countries all over the world between the years 1981 and 2008. Not all these countries are included in the Manifesto Data Collection so we cannot include them in the combined data set yet. But once new countries are added to the Manifesto Data Collection, we can make use of the additional data for these countries in the WVSs.

Measuring Parties' Programmatic Supply and Voters' Demand

In order to measure distances between parties' programmatic supply and the policy preferences of their voters we had to identify the relevant information for this. Accordingly, left-right placements of parties in the Manifesto Data Collection, as well as information on voters' left-right placements and party preferences from the surveys were selected for the linking procedure.

---

[3] For detailed citation information see page 2.

The Party Position

The Manifesto Project (MRG/CMP/MARPOR) analyses policy statements of parties in manifestos employing a content analytical method. The content of the manifestos is divided into statements identifying one political argument. To each statement coders assign one of 56 policy categories from the classification scheme the project developed to analyse programmatic supply in party competition. Based on salience theory the relative frequency of statements belonging to one category is used to identify policy positions in different policy fields. To ensure a comparative analysis the complete range of political issues is considered, rather than single issues. Therefore, parties' statements are used to identify the parties' positions on an aggregate left–right dimension. The Manifesto Project uses a theoretical scale from –100 (left) to +100 (right). Up to today the empirically observed span, however, ranges between –75 and +65.

The Voter Position

All surveys included in this study use a 0 or 1 (left) to 10 (right) scale. In CSES interviewees are asked the following: 'In politics people sometimes talk of left and right. Where would you place yourself on a scale from 0 to 10 where 0 means the left and 10 means the right?'. EVS and WVS use the same scale, asking: 'In political matters, people talk of "the left" and "the right." How would you place your views on this scale, generally speaking? [1 'Left' … 10 'Right']'. In EB the respondents are asked: 'In political matters people talk of "the left" and "the right. How would you place your views on this scale? [1. LEFT … 10. RIGHT]'.

Identifying Voters through Expressed Party Support

To connect voters to their parties, survey information needs to be linked to the Manifesto Data Collection. As the latter are data for parties in parliamentary elections we drew on the vote choice questions for parliamentary elections from the surveys. In all four surveys we identified different questions as relevant for this:

In the three CSES modules certain questions ask for the vote choice regarding the first parliamentary chamber. In all three modules respondents are asked about their vote choice in the current election, module II and III also include the vote choice question for the previous election. For all these questions CSES discerns between the possibility for the respondents to vote directly for a party list and the possibility to vote for a

candidate. In the latter case CSES uses the party of the candidate the respondent voted for. The questions we considered for the combined data set are the following. CSES1 for the current election A2030: 'In systems where respondent had option of voting directly for a party list in district-level elections, party list that respondent voted for.' and A2031: 'In systems where respondent had option of voting directly for a candidate or candidates in district-level election, party of candidate respondent votes for.' CSES2 for the current election B3006_1 and B3006_2 as well as for the last election B3018_1 and B3018_2: 'This variable[s] report[ ] the vote(s) cast by the respondent in lower house elections.' CSES3 for the current election C3023_LH_PL and C3023_LH_DC as well as C3032_LH_PL and C3032_LH_DC for the last election: 'This variable[s] report[ ] the vote(s) cast by the respondent in lower house elections.'

From the World Value Surveys we selected the questions asking for the vote intention relating to the next parliamentary election as relevant for the combined Party-Voter Data Set. In the WVS, there are two questions posed with regard to a distinct vote intention; one for the first vote preference (e179), another for the second (e180). With slight variations in phrasing between the different survey phases the question reads: 'If there were a national election tomorrow, for which party [...] would you vote?'[4] If the respondent gave a missing answer to one or both of these questions, WVS contains a further question asking for a more general party preference (e181): 'If don't know: Which party appeals to you most?'

For questions taken from the Eurobarometer surveys related to the respondents' party choice in the next election. It reads: 'If there were a "general election" tomorrow (say if contact is under 18 years: and you had a vote), which party would you support?' If the respondent did not answer this question, he/she was asked about a less distinct vote inclination: '<If DK or NA in VOTEINT> Which party would you be inclined to vote for?' In addition to these questions, we used those relating to the vote choice at the last election to identify party support. With some minor changes in the exact wording this question

---

[4] WVS 1995: If there were a [country] election tomorrow, for which party on this list would you vote? Just call out the number on this card.
WVS 2000: If there ware a national election tomorrow, for which party on this list would you vote? Just call out the number on this card. If DON'T KNOW: Which party appeals to you most?.
EVS 1999: If there was a general election tomorrow, which party would you vote for?. COUNTRY SPECIFIC LIST OF POLITICAL PARTIES.

is: 'Which party did you vote for at the last general election in (year of last general election)?'[5]

All surveys provide a list of the different parties from which the respondents could choose to indicate their vote choice. By linking the party codes given there to those given in the party list of the Manifesto Data Collection we could link survey respondents to the CMP parties. We linked the codes by identifying corresponding party names, party history and organisation as well as election results in the Manifesto Data to the frequencies in which parties where named by respondents in the surveys that were conducted in close proximity of time.

---

[5] EB11: Were you able to go and vote on the last general election on <date> or where you prevented? If voted: For which party did you vote? <France: On the first ballot?>
EB18 to EB27: <Only asked in Belgium: Which party R voted for in the last election?>
EB24 to EB29: <Only asked in Spain and Portugal: Which party R voted for on the last election?>
Since EB30: Which party did you vote for at the last general election in (year of last general election)?

Table 1: Overview on Voter Position Question

	Variable Name	A3031	B3045	C3013
CSES	Voter Position Question	In politics people sometimes talk of left and right. Where would you place your-self on a scale from 0 to 10 where 0 means the left and 10 means the right?	In politics people sometimes talk of left and right. Where would you place your-self on a scale from 0 to 10 where 0 means the left and 10 means the right?	Where would you place yourself on this scale? (scale: 0=left; 10=right)
	Missing Codes	98: Don't know 99: Missing	97: Refused 98: Don't know 99: Missing	95: Volunteered: haven't heard of left-right 97: Volunteered: refused 98: Volunteered: don't know where to place 99: Missing

	Variable Name	lrs		
EB	Voter Position Question	In political matters people talk of "the left" and "the right". How would you place your views on this scale?		
	Missing Codes	96 / .b: Refused 98 / .a: DK/NA 99: INAP		

	Variable Name	e033		
WVS/EVS	Party Mapping Question	In political matters, people talk of "the left" and "the right". How would you place your views on this scale, generally speaking?		
	Missing Codes	–5: Missing/Unknown –4: Not asked in survey –3: Not applicable –2: No answer –1: Don't know		

Table 2: Overview of Party Mapping Question

	Variable Name	A2030	A2031	B3006_1 B3018_1	B3006_2 B3018_2	C3023_LH_PL C3032_LH_PL	C3023_LH_DC C3032_LH_DC
CSES	Party Mapping Question	Party list voted for – district If applicable and respondent cast a ballot: In systems where respondent had option of voting directly for a party list in district-level elections, party list that respondent voted for.	Party of candidate voted for – district If applicable and respondent cast a ballot: In systems where respondent had option of voting directly for a candidate or candidates in district-level election, party of candidate respondent votes for.	Current election / Previous election: Vote Choice – Lower House 1 If applicable and respondent cast a ballot: This variable reports the vote(s) cast by the respondent in lower house elections. See Election Study Notes for more information.	Current election/ Previous election: Vote choice – Lower House 2 If applicable and respondent cast a ballot: This variable reports the vote(s) cast by the respondent in lower house elections. See Election Study Notes for more information.	Current election/ Previous election: Vote choice lower house – Party list If applicable and respondent cast a ballot in the Lower House legislative election: These variable reports the respondent's vote choice for party list in Lower House elections.	Current election/ Previous election: Vote choice lower house – District candidate If applicable and respondent cast a ballot in the Lower House legislative election: These variable reports the respondent's vote choice for district candidate in Lower House elections.
	Missing Codes	70 (with exception of the Netherlands) and 96: other 98: Respondent cast invalid ballot/did not vote 99: Applicable but not ascertained – don't know, refused 00: Not applicable to this electoral system		89: independent candidate 90: other 94: Inconsistent response: R reported casting a ballot but R did not vote 96: Respondent cast invalid ballot 97: Refused 98: Don't know 99: Missing		89: independent candidate 90: other 92: Respondent cast invalid ballot 93: Respondent cast blank ballot 97: Volunteered: refused 98: Volunteered: don't know 99: Missing	

Table 2 continued:

Variable Name	voteint	inclvote	lastvote
EB	Party Mapping Question	If there were a 'General Election' tomorrow (say if contact is under 18 years: and you had a vote), which party would you support?	‹If dk or na in voteint› Which party would you be inclined to vote for?	Which party did you vote for at the last general election in (year of last general election)?
	0: NA 990: Other party 995: Empty ballot/no party preference 996: Did not vote/would not vote/was not able to vote 997: Refused 998: DK/(DK+NA – if NA not coded separately 999: INAP		

Table 2 continued:

	Variable Name	e179	e180	e181
WVS/EVS	Party Mapping Question	If there were a national election tomorrow, for which party on this list would you vote? First choice	If there were a national election tomorrow, for which party on this list would you vote? Second choice	If don't know: Which party appeals to you most?
		−5: Missing -4: Not asked in survey -3: Not applicable -2: No answer -1: Don't know 1: No right to vote 2: I would not vote 3: I would cast a blank ballot/White vote 4: None 5: Other 6: All political parties		

Data Set Description

In order to analyse the quality of programmatic congruence comparatively, across countries and over time, we created this new data set. It comprises the most relevant information from election and survey data on party-voter connections and placements on the left-right dimension.

Data Set structure
The basis of the data set is the Manifesto Data Collection. The final Party-Voter Data Set hence includes entries for the countries, elections, and parties with their left-right placement that are also included in the Manifesto Data Collection. To this election data we matched survey data on voters' expressed vote choices and policy positions at the individual level of single survey respondents. The matching of the surveys to the Manifesto Data Collection was undertaken for respondents' expressed party vote choice relating to the current, next and/or last election. Survey data were linked to the election date regarding the timing of the expressed vote choice. For this we used the date of field work in relation to the election date. In cases were no survey data was available to be linked to a specific election or party contained in the Manifesto Data collection these (election and party) observations on the side of the CMP-Data were dropped from the Party-Voter Data Set.

Whenever the matching of expressed vote choice to an election in the Manifesto Data Collection was possible the Party-Voter Data Set provides one row of observations for each survey respondent. Since respondents were able to express vote choices several times (e.g. their vote choice at the previous and at the next election) the observation is included as often as there is a valid vote choice answer. Due to the matching procedure, the data is organised in a one-to-one relationship between parties and voters at an election date, provided the survey respondents had expressed a distinct party preference with regard to this very election. With this it is possible to analyse how well parties represent their supporters in election manifestos. This seems to be an accurate congruence measure as voters are assumed to be satisfied with the quality of representation if there is an institution that represents congruent positions which they can vote for. However, depending on the research question at hand, it can also be

interesting to include the unaligned voters, who in the surveys provided no vote choice but for example a self-positioning on the left-right scale. These cases are only linked using the date of field work and the direction of the question asked (vote choice in next, current or last election) in relation to the election date as described below.

<u>Variable Description</u>

In the following we give a brief description of the variables in the data set. Variable names are given in the brackets:

For each observation the information given from the Manifesto Data Collection is the country name (country), election date (edate) and the election year (marpor_year), the respondent's party vote intention on the basis of the CMP party codes (party) and the rile score of the named party (rile).

For the survey data we added the following information. The ID of the used data set (dataset_id), the country in which the survey was conducted as named in the survey data set (survey_country), the year in which the survey took place (survey_year), and two variables giving the exact data of the beginning and end of the survey fieldwork in the specific country (fieldwork_in and fieldwork_out). In cases where no exact date was available we interpolated the date based on the vague information given in the survey documentation (e.g. when the month and year was reported, we took the first and the last day of this month, when only the year was reported, we took June, 1st and 30th of that year as proxies. Interpolating proxies was necessary for the technical linking procedure, so that the linking can account for the connection between the election date and the timing of the vote choice expression. From fieldwork_in and fieldwork_out we deducted a third variable which gives the difference in days between the fieldwork and the election (fieldwork_diff). A specific variable (current_vote_type) denotes the survey variables we used to capture the respondents' expressed party preference with regard to the current, next or last election, respectively.

In addition to this technical information, the only variable included from the surveys by now is the self-placement of respondents on the left-right scale. Based on this information we created a new variable that reports the left-right self-placement on a common scale (selfrile). Recoding was necessary as the corresponding variables in the

surveys are coded in different scales. First, the original values from the different surveys have been recoded according to Giebler, Haus and Weßels (2009, p. 248) to a 1–9 scale. Furthermore we harmonised the missing codes. As with the party preference, a technical variable denotes the survey variables we used to identify respondents' self-placement (selfrile_survey_type) so users can trace the sources and original value of the self-placements.

Further technical variables provide more detailed information on the match of the survey data sets with the Manifesto Data Collection. First, a variable gives a general overview from which data set the information for an observation is taken (observation_origin). A more fine-grained variable indicates when matching based on expressed party preference was possible, when this was not and if so why. This was the case when a respondent did not name a party to which he could be matched or if this party is not included in the Manifesto Data Collection (missing_type). A variable is included that unifies the missing codes for repondents' party vote choice as reported in the surveys (survey_partypref_missing). This variable allows to distinguish more specifically on the individual level between voters, nonvoters and respondents for whom due to technical reasons no party preference is reported in the Party-Voter Data Set. Two variables indicate whether or not in the surveys linked to the election data the questions about party preference in different vote choice questions and/or self-placement of the respondents on the left-right scales were posed at all (party_na and selfrile_na). To allow users to distinguish between respondents from the survey data sets as objects of the analysis with several possible observations for expressed vote choices in the Party-Voter Data Set we included the unique identifier variables for survey respondents from the original data sources (A1005, B1005, C1005, id, s007).

With the technical information we provide in the data set, users can adapt it and specify the cases in accordance to their particular research focus. The variables specifying the type of missing observations can be used to exclude certain observations from the analysis. On the level of the single observation they allow to distinguish within the group of observations without a distinct vote choice for the respondent between voters and non-voters. Furthermore, one can distinguish cases where technical reasons are the cause for the missing observation such as the lack of correspondence in the CMP

and surveys' party lists or that the question about left–right self-placement was not posed. Last but not least, researchers can limit the entries in the data set to one observation for each individual using the unique respondent identifier. Furthermore, technical variables can be used to link further information from the source data sets, in cases where these are of interest depending on the respective research programme. The unique survey respondent identifiers allow users to add other variables from the respective survey data. Since the original variable names, format and values are adopted exactly as they are in the surveys, including further information from them should be possible via a merge-command. Adding additional variables from the CMP/MARPOR data set is possible using country, edate and party as merging variables.

Table 3: Overview over Variables in the Party Voter Data Set (PVds)

Variable Name	Description	Range	Codebook	Missing Codes
country	Reports the CMP country code.	11 to 171	See Manifesto Project Codebook available via: https://manifesto-project.wzb.eu	none
edate	Indicates the date of the election.	31mar1968 to 27sep2009	See Manifesto Project Codebook available via: https://manifesto-project.wzb.eu	none
marpor_year	Shows the year of the election.	1968to 2009		none
party	Denotes the MARPOR party code of the party which the respondent identified as being close to or having voted for.	11110 to 171601	See Manifesto Project Party List, available via: https://manifesto-project.wzb.eu	–99
rile	Indicates the rile score of the specific party in the respective election.	–100 to +100	See Manifesto Project Codebook available via: https://manifesto-project.wzb.eu	
dataset_id	Details in which survey the respective respondent participated.	11 to 31	11: CSES 1 12: CSES 2 13: CSES 3 21: EB 31: WVS/EVS	none
survey_country	Reports the survey country code.		Three-digit letter code (abbreviation of country name) for CSES; two-digit numeric code for EB; and three-digit numeric code for WVS/EB	.
survey_year	Identifies the year from which the survey originates.	1970 to 2009		none

Variable Name	Description	Range	Codebook	Missing Codes
fieldwork_in	Indicates the date on which the survey's fieldwork began.	01feb1970 to 28sep2009	Taken either from survey data set or documentation, in cases, where no fieldwork-dates were reported, the dates are interpolated.	none
fieldwork_out	Reports the date on which the survey's fieldwork ceased.	30mar1970 to 28oct2009	Taken either from survey data set or documentation, in cases, where no fieldwork-dates were reported, the dates are interpolated	none
fieldwork_diff	Shows the difference in days between the fieldwork and the election date.[6]	-792 to 792	(For limitation of the time span for linking survey and election data see the following pages.)	none

---

[6] In case no exact fieldwork was given, we used June, 30th of the survey year as an estimate.

Variable Name	Description	Range	Codebook	Missing Codes
current_vote_type	Denotes the survey variable that was used to identify the party alignment of the respondent.		A2030 A2031 B3006_1 B3006_2 B3018_1 B3018_2 C3023_LH_PL C3032_LH_PL C3023_LH_DC C3032_LH_DC voteint inclvote lastvote e179 e180 e181	.
selfrile_survey_type	Reports the survey variable that provided the information on respondent's self-placement on left-right scale.		A3031 B3045 C3013 e033 lrs	.
selfrile	Identifies the left-right position of the survey respondent	1 to 9	left-right self-placement of voters, original values from the different surveys recoded according to Giebler, Haus and Weßels (2009, p. 248) to a common 1-9 scale	-99 -88 -77 -66 -55

Variable Name	Description	Range	Codebook	Missing Codes
observation_origin	Differentiates whether observation in PVds originates from CMP or survey only or is a match of both data sources.	1 to 3	provenance information where a specific observation comes from 1: cmp only, 2: survey only or 3: both, cmp and survey data.	
missing_type	Reports how linking of cmp party and survey party preference relates.	3 to 100	3: cmp party missing (respondent in survey indicated party preference that has no equivalent in CMP party list) 4: mapping variable is missing value (.) (respondent (intentionally) did not indicate any party preference or answer was not reported, inconsistent/invalid/not applicable) 100: MATCH! cmp and survey entry successful	.
survey_partypref_ missing	Reports when survey respondent did not indicate a clear party preference when asked about vote choice, (differentiates between other preference than options listed in party list or missing code).	–99 to –11	unified missing code for voters' party preference reporting original missing type for party preference from surveys when respondents were asked about vote choice: –99: Missing / no answer –88: Don't know –77: Refused to answer –66: Respondent cast invalid ballot –55: Respondent did not vote [–54: No right to vote] –44: Inconsistent response	.

Variable Name	Description	Range	Codebook	Missing Codes
			–33: Other party / all / none of the parties	
			–22: Not applicable	
			–11: Unspecified missing code in survey	
party_na	Reports whether question about party preference when asked about vote choice was asked in survey or not.		1: question was not asked in survey.	.
selfrile_na	Reports whether question about respondents' self-placement on the left-right scale was asked in survey or not.		1: question was not asked in survey.	.
A1005	Unique identifier variables for respondents / observations from surveys		unique respondent identifier for survey cses1	
B1005	s. above		unique respondent identifier for survey cses2	
C1005	s. above		unique respondent identifier for survey cses3	
id	s. above		unique respondent identifier for survey eb	
s007	s. above		unique respondent identifier for survey wvs/evs	

Methodological Challenges of Linking Survey to Election Data

Linking survey to election data in order to account for the distribution of voters' party preferences on different levels of analysis, MARPOR faced three major conceptual and methodological challenges.

Levels of Aggregation and Measuring Congruence

Data linkage is achieved by using the expressed vote preference as the identifier for the individuals' party preference. In order to measure congruence between voters and parties we organised the data in a first step as a one-to-one relationship between survey respondents and parties. The data set, organised in this fashion, provides information about single parties' policy positions and individual voters' policy preferences on the 'super issue' of left-right. In a second step and in the course of further analyses, this allows for aggregation on the level of parties, party systems or government systems. Using information on survey respondents' expressed party preference, if available, makes it first and foremost possible to focus the analysis on how well parties represent their supporters. However, for the purpose of studying further aspects of representation and in particular the quality of representation on the aggregate level, the survey respondents, for whom no vote choice but only a self-positioning on the left-right scale is reported, are also included in the data set. They can (provided that the question about vote choice is posed in the respective survey wave or module – in a respective country and a certain point in time) be treated as unaligned citizens, who are only linked using the date of field work in relation to the election date as described below. In studies on party-voter congruence they can then, for example, be linked to the median party position which at least allows to analyse how well these 'non-party supporters' are represented by the party system in general. As a result, the data set organisation allows for the operationalisation of all congruence measures that are commonly discussed in the empirical representation literature.

Data Linkage and Dates of Data Generation

Linking survey to election data is challenging with respect to the fact that, in order to capture the concept of congruence correctly, the information on policy preferences required to analyse congruence have to be generated close in time. But, with the

exception of data stemming from CSES, party positions and voters' preferences and attitudes are collected at different points in time. The data collection on voters' preferences of all other surveys is not conducted in close proximity to elections but at some point during the course of the legislative term. To deal with this problem, Powell (2000) suggests that information from surveys can be linked to election data adequately when the fieldwork of the surveys is conducted within a two year time span around the election. We adopted his approach for linking survey to election data but in addition considered the chronological relation between expressed party preferences and the programmatic supply. Using the date of fieldwork, we linked the survey data to the election date within the two year time frame (giving a two month tolerance). We linked to the following election for party preferences expressed when respondents were asked for future vote choices. For party preferences expressed related to the last election, the survey data is linked to this election.

Data Linkage and the Left-Right Dimension as Common Reference
Despite the fact that we assume the left-right dimension to be a common reference frame for both, party competition as well as party positioning and self-placements by voters, we face the challenge that by combining manifesto and survey data we have to consider two points.

Firstly, it may be true that MARPOR adopts the underlying assumption that survey respondents have an understanding of the left-right dimension, an assumption generally supported by the literature. But, as outlined by Best/McDonald/Budge (2012), we assume in addition that voters use the national context of political issues and party competition as a reference for their self-placement. In contrast to this, the Manifesto Project scale is designed purposely to cover a theoretically derived absolute left-right dimension to allow for cross-country and over-time comparison of the competition between the relevant parties.

Secondly, as the Manifesto Data Collection does not cover all parties, but only those represented in parliament, we cannot expect the empirically observed party policy space to comprise the total policy space in a country. Other and more extreme issues and positions, not addressed by parties that are represented in parliament, might be of relevance. First, one must expect that some of the parties that are not included in the

Manifesto Data Collection, as they never entered parliament, are still so visible in a country that they influence the perceived policy space. Additionally, there are always some voters who will see themselves positioned outside the spectrum of offers parties make to them. In these cases where the relevant parties only occupy a very small part of the left–right span that defines the policy space in a country, it seems illogical to expect voters to restrict their self–placement to this predefined limited policy space.

In sum, we assume that surveys use a different range and segment of the left–right dimension than manifesto data. To overcome the differences in survey and manifesto based left–right–scaling, the data has to be rescaled to a unified one. Unfortunately, it is impossible to grasp a 'true' policy space that accurately reflects the diversity and relative distances of both parties' and voters' positions in every country with the existing survey and election data. To remedy this and to allow for cross-country comparisons, the only possibility is to employ the closest possible simulation. In general, we would like to advise researchers to use the information from the election data to identify cross-country and over-time variance and to draw upon information from the survey data to estimate the polarisation of the policy space for rescaling. Relating these data can be achieved by using bridging observations (Gschwend/Lo/Proksch 2012; Quinn et al. 1999; Quinn 2004). By way of an exemple, we already applied one possibility of rescaling in this tradition. Through linear regression and using the party positions as they are perceived by survey respondents as bridging observation we rescaled the survey respondent self–positioning to the RILE–CMP scale (Lehmann/Schultze, 2013). Other rescaling techniques (with and without the use of bridging observations and on the basis of different mathematical functions for the transformation) are possible as well. The specific procedure of how to rescale the data, however, has to depend on the respective research question. Therefore, we would like to encourage the users of the data set to establish and test their own rescaling procedures.

Data Coverage and Update Plans

On our website, we provide an overview over the data coverage of the Party–Voter Data Set in its current state. Please consult the documentation of metadata, which gives

information on the countries and election years, parties and party-voter combinations for which we have manifesto and survey data in the Party-Voter Data Set, under https://manifesto-project.wzb.eu/.

In its current version (end of first funding period by 30/09/2012), the MARPOR Party-Voter Data Set already provides some longer time series for Western Europe. For Eastern Europe and non-European countries this is as of yet not the case. But we aim to close the gap with updating and expanding the Party-Voter data set in the next funding period.

As a start, we will concentrate on consolidating the data set based on updated versions of data sets that are already included in the Party-Voter Data Set. For example, further WVS data for non-European countries and some minor cases still missing in the CSES data linkage can be included with an update of the Manifesto Data Collection. Therefore, we will concentrate on such updates for the combined data set. Thereafter, we will focus on updates of the already included survey data sets, which will be launched in several modules and waves in the future. This will be the case for the upcoming CSES modules, but also for further updates of the WVS Aggregated File. In line with regular updates of the Manifesto Data Collection this will already lead to a considerable expansion of the data set.

Further expansion plans project an inclusion of selected new survey data sets. Additionally, available data for Central and Eastern European countries can be included in the combined data set by using the Central and Eastern Eurobarometer Trend-file which adds data from these countries for the years 1990-1997. By the time the Manifesto Project is expanding its scope to Latin America and acquiring respective manifesto, we will also include the Latinobarometros in the combined data set. The inclusion of further survey data sets will be discussed in more detail in the next project phase starting in October 2012.

Bibliography

Achen, Christopher H. (1977), "Measuring Representation: Perils of the Correlation Coefficient". in: *American Journal of Political Science* 21(4): 805-815.

———. (1978), "Measuring Representation". in: *American Journal of Political Science* 22(3): 475-510.

Barnes, Samuel H. (1977), *Representation in Italy: Institutionalized Tradition and Electoral Choice.* Chicago: University of Chicago Press.

Best, Robin E. / Budge, Ian / McDonald, Michael D. (2012), "Representation as a median mandate: Taking cross-national differences seriously". in: *European Journal of Political Research* 51(1): 1-23.

Bobbio, Noberto (1996), "Left and Right: The Significance of a Political Distinction", Cambridge: Polity Press.

Budge, Ian/Klingemann, Hans-Dieter/Volkens, Andrea/Bara, Judith/Tanenbaum, Eric with Fording, Richard C./Hearl, Derek J./Kim, Hee Min/McDonald, Michael/Mendez, Silvia. (2001), *Mapping Policy Preferences. Estimates for Parties, Electors, and Governments 1945–1998.* Oxford: Oxford University Press.

Castles, Francis C./Mair, Peter. (1984), "Left-Right Political Scales: Some Expert Judgements." in: *European Journal of Political Research* 12: 83–88.

Converse, Philip E., and Roy Pierce. (1986), *Political representation in France.* Cambridge, Mass. [u.a.]: Belknap Pr. of Harvard University Press.

Dahl, Robert. 1989. *Democracy and its Crisis.* New Haven: Yale University Press.

Dalton, Russell J. (1985), "Political Parties and Political Representation". in: *Comparative Political Studies* 18(3): 267 -299.

Fuchs, Dieter/Klingemann, Hans-Dieter. (1990), "The Left-Right Schema." in: Jennings, M. Kent/van Deth, Jan W./Barnes, Samuel H. (eds.). *Continuities in Political Action. A Longitudinal Study of Political Orientations in Three Western Democracie.*, Berlin: de Gruyter: 203–234.

Gabel, Matthew J./Huber, John D. (2000), "Putting Parties in their Place: Inferring Party Left-Right Ideological Positions from Party Manifestos Data." in: *American Journal of Political Science* 44: 94–103.

Gschwend, Thomas/Lo, James/Proksch, Sven Oliver (2012), "Europe's Common Ideological Space." http://www.piredeu.eu/Database/Conf_Papers/V2_1-ESSpaper1.pdf.

Giebler, Heiko/Haus, Elmar/Weßels, Bernhard (2010), "2009 European Election Candidate Study – Codebook (Advanced release, V2)", WZB: Berlin, (www.piredeu.eu); GESIS: Mannheim (ZA5048, see http://info1.gesis.org/dbksearch19/download.asp?id=37351)

Golder, Matt/Stramski, Jacek. (2010), "Ideological Congruence and Electoral Institutions." in: *American Journal of Political Science* 54(1): 90–106.

Herrera, Cheryl Lyn/Herrera, Richard/Smith, Eric R. A. N. (1992),"Public Opinion and Congressional Representation." in: *Public Opinion Quarterly* 56(2): 185 -205.

Holmberg, Sören. (1999), "Collective Policy Congruence Compared". In: Miller, Warren E./Pierce, Roy/Thomassen, Jacques/Herrera, Richard/Holmberg, Sören/Esaiasson, Peter/Weßels, Bernhard (eds.). *Policy representation in Western democracies*, Oxford: Oxford University Press: 87–109.

Huber, John D./Inglehart, Ronald. (1995), "Expert Interpretations of Party Space and Party Location in 42 Societies." in: *Party Politics* 1: 73–111.

Hurley, Patricia A./Hill, Kim Quaile. (2003), "Beyond the Demand-Input Model: A Theory of Representational Linkages." in: *Journal of Politics* 65: 304–326.

Inglehart, Ronald/Klingemann, Hans-Dieter. (1976), "Party Identification, Ideological Preference and the Left-Right Dimension among Western Mass Publics." in: Budge, Ian/Crewe, Ivor/Farlie, Dennis (eds.). *Party Identification and Beyond*. New York: Wiley: 243–273.

Kitschelt, Herbert. (2000), "Linkages between Citizens and Politicians in Democratic Polities." in: *Comparative Political Studies* 33: 845–879.

Klingemann, Hans-Dieter/Volkens, Andrea/Bara, Judith/Budge, Ian/McDonald, Michael. (2006), *Mapping Policy Preferences II. Estimates for Parties, Electors, and Governments in Eastern Europe, the European Union and the OECD, 1990–2003*. Oxford: Oxford University Press.

Lawson, Kay/Poguntke, Thomas (eds.). (2004), *How Political Parties Respond. Interest Aggregation Revisited*. London: Routledge.

Lehmann, Pola/Schultze, Henrike (2013), "Evaluating Congruence Measures Compara tively", in: Volkens, Andrea, et al.: *Mapping Policy Preferences 3: Measurement Solutions for Manifesto Analyses: 54 Democracies, 1945–2012*, Oxford: Oxford University Press.

Miller, Waren E./Stokes, Donald E. (1963), "Constituency Influence in Congress." In: *The American Political Science Review* 57(1): 45-56.

Montero, José Ramón/Gunther, Richard. (2002), "Introduction: Reviewing and Assessing Parties." in: Gunther, Richard/Montero, José Ramón/Linz, Juan J. (eds.). *Political Parties. Old Concepts and New Challenges.* Oxford: Oxford University Press: 1–38.

Müller, Wolfgang C. (1994), "Political Traditions and the Role of the State." in: *West European Politics* 17: 32–51.

Offe, Claus. (2003), "Democratic Institutions and Moral Resources". in: Offe, Claus (ed.). *Herausforderungen der Demokratie.* Wiesbaden: VS Verlag für Sozialwissenschaften.

Pierce, Roy. (1999), "Mass-Elite Linkages and the Responsible Party Model of Representation." in: Miller, Warren E./Pierce, Roy/Thomassen, Jacques J. A./Herrera, Richard/Holmberg, Sören/Esaiasson, Peter/Weßels, Bernhard (eds.). *Policy Representation in Western Democracies.* Oxford: Oxford University Press: 9–32.

Pitkin, Hanna F. (1967), *The Concept of Representation.* Berkeley: University of California Press.

Powell, G. Bingham. (2000), *Elections as Instruments of democracy.* New Haven/London: Yale University Press.

———. (2004), "Political Representation in Comparative Politics." in: *Annual Review of Political Science* 7: 273-96.

Quinn, Kevin M./Martin, Andrew D./Whitford, Andrew B. (1999), "Voter Choice in Multi-Party Democracies: A Test of Competing Theories and Models" in: *American Journal of Political Science* 43(4): 1231-1247.

Quinn, K. (2004), "Bayesian factor analysis for mixed ordinal continuous response". in: *Political Analysis* 12(4): 338–353.

Römmele, Andrea/Farrell, David M./Ignazi, Piero,eds. (2005), *Political Parties and Political Systems. The Concept of Linkage Revisited.* Westport/London: Praeger.

Rohrschneider, Robert/Whitefield, Stephen. (2010), "Representation in Europe: How well do Parties Represent Unaligned Voters?". Paper presented at the APSA's 106th Annual Meeting at September 2-5, 2010 in Washington, D.C.

Sani, Giacomo/Sartori, Giovanni. (1983), "Polarization, Fragmentation, and Competition in Western Democracies." in: Daalder, Hans/Mair, Peter (eds.). *Western European Party Systems. Continuity and Change.* London/Beverly Hills: Sage: 307–340.

Schmitt, Hermann/Thomassen, Jacques A. (1999), *Political representation and legitimacy in the European Union.* 1999th edition. Oxford: Oxford University Press.

Thomassen, Jacques A. (1994), "Empirical Research into Political Representation: Failing Democracy or Failing Models". in: Jennings, M. Kent/Mann, Thomas E. (eds.). *Elections at home and abroad : essays in honor of Warren E. Miller.* Ann Arbor: University of Michigan Press.

————. (1999), "Political Communication between Political Elites and Mass Public. The Role of Belief Systems." in: Miller, Warren E./Pierce, Roy/Thomassen, Jacques J.

A./Herrera, Richard/Holmberg, Sören/Esaiasson, Peter/Weßels, Bernhard (eds.). *Policy Representation in Western Democracies.* Oxford: Oxford University Press: 33–58.

———. (ed.). (2005), *The European Voter. A Comparative Study of Modern Democracies.* Oxford: Oxford University Press.

Thompson, Richard F. (2001), "Political Representation." in: Smelser, Neil J./ Baltes, Paul B. (eds.): International Encyclopedia of the Social and Behavioral Sciences. Amsterdam: Elsevier: 11696-11698.

Warwick, Paul V. (2002),"Towards a Common Dimensionality in West European Policy Spaces." in: *Party Politics* 8: 101–122.

Webb, Paul/Farrell, David M./Holliday, Ian. (2002), Introduction: Political parties in advanced industrial democracies". in: Webb, Paul/Farrell, David M./Holliday, Ian (eds.). *Political parties in advanced industrial democracies*, Oxford et al.: Oxford University Press: 1-15.

Weissberg, Robert. (1978), "Collective versus dyadic representation in Congress." in: *American Political Science Review* 72: 535-47.

Weßels, Bernhard. (1995), "Support for Integration: Elite or Mass-Driven?" in: Niedermayer, Oskar/Sinnott, Richard (eds.). *Public opinion and internationalized governance*, Oxford et al.: Oxford University Press: 137-162.

———. (1999,) "System Characteristics Matter. Empirical Evidence from Ten Representation Studies." in: Miller, Warren E./Pierce, Roy/Thomassen, Jaques J. A./Herrera, Richard/Holmberg, Sören/Esaiasson, Peter/Weßels, Bernhard (eds.). *Policy Representation in Western Democracies.* Oxford: Oxford University Press: 136–160.

———. (2007), "Political Representation and Democracy". in: Russel J. Dalton/Hans-Dieter Klingemann (eds.). *The Oxford Handbook of Political Behaviour.* Oxford: Oxford University Press.

Data Sources

European Values Study 1981-2008, Longitudinal Data File. GESIS Data Archive, Cologne, Germany, ZA4804 Data File Version 1.0.0 (2011-04-30) DOI:10.4232/1.14804.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 1 FULL RELEASE [dataset]. Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor]. August 4, 2003.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 2 FULL RELEASE [dataset]. June 27, 2007.

The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 3 ADVANCE RELEASE [dataset]. 2011.

The Mannheim Eurobarometer Trend File 1970–2002. Schmitt, Hermann. Mannheimer Zentrum für Europäische Sozialforschung (MZES). Scholz, Evi/Leim, Iris/Moschner, Meinhard. GESIS – Leibniz Institute for the Social Sciences (former "Zentrum für Umfragen, Methoden und Analysen" and "Zentralarchiv für Empirische Sozialforschung") (http://www.gesis.org/eurobarometer).

Volkens, Andrea/Lacewell, Onawa/Lehmann, Pola/Regel, Sven/Schultze, Henrike/Werner, Annika. 2011. *The Manifesto Data Collection. Manifesto Project (MRG/CMP/MARPOR).* Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB). (http://manifestoproject.wzb.eu/).

Volkens, Andrea/Lehmann, Pola/Regel, Sven/Schultze, Henrike/Werner, Annika. 2011. The MARPOR Party–Voter Data Set . Manifesto Project (MRG/CMP/MARPOR), Berlin: Wissenschaftszentrum Berlin für Sozialforschung (WZB).

World Values Survey 1981–2008 Official Aggregate v.20090901, 2009. World Values Survey Association (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid. (http://www.wvsevsdb.com/wvs/WVSData.jsp?Idioma=I).

Appendices

Country List

The latest version of the country list is documented and regularly updated in a separate data set that is accessible via https://manifest-project.wzb.eu/datasets/pvds.

Party List

The latest version of the party list is documented and regularly updated in a separate data set that is accessible via https://manifest-project.wzb.eu/datasets/pvds.

Election Dates and Survey Fieldwork Dates

The latest version of the election – fieldwork dates list is documented and regularly updated in a separate data set that is accessible via https://manifest-project.wzb.eu/datasets/pvds.

Summary of Meta Data

The documentation of the latest version of the MARPOR Party-Voter Data Set, providing an overview on the data coverage in terms of the number of countries, elections, parties and individual respondents included as well as a summary on missing data for party preferences, left-right positions in contrast to successfully linked party–voter combinations on the level of the individual voter is accessible via https://manifest-project.wzb.eu/datasets/pvds.

Recoding of Missing Codes for Survey Respondents Party Vote Choice

Variable: Respondents' expressed party preference when asked about vote choice – indistinct preference

Missing codes	New code	CSES1	CSES2	CSES3	EB	WVS/EVS
Missing / no answer	–99		99: Missing	99: Missing	0: NA	–5: Missing / –2: No answer
Don't know	–88		98: Don't know	98: Volunteered: don't know	998: DK/(DK+NA – if NA not coded separately	–1: Don't know
Refused to answer	–77	99: Applicable but not ascertained – don't know, refused	97: Refused	97: Volunteered: refused	997: Refused	
Respondent cast invalid ballot	–66	98: Respondent cast invalid ballot/did not vote	96: Respondent cast invalid ballot	92: Respondent cast invalid ballot 93: Respondent cast blank ballot	995: Empty ballot/no party preference	3: I would cast a blank ballot/White vote (7: null vote)
Respondent did not vote	–55				996: Did not vote/would not vote (was not able to vote)	2: I would not vote
No right to vote	–54					1: No right to vote
Inconsistent response	–44		94: Inconsistent response: R reported casting a ballot but R did not vote			
Other party / all / none of the parties	–33	90: other 70: other (except for the Netherlands)	90: other 89: independent	90: other 89: independent	990: Other party	4: None 5: Other (6: All political parties)
Not applicable	–22	00: Not applicable to this electoral system			999: INAP	–3: Not applicable
Unspecified missing code in survey	–11	.	.	.	.	

Recoding of Survey Respondents Left–Right Self–Placements

PVds – New coding: Selfrile [0–9]		CSES1: Old coding: A3031 [0–10]	CSES2: Old coding: B3045 [0–10]	CSES3: Old coding: C3013 [0–10]	EB: Old coding: lrs [1–10]	WVS/EVS: Old coding: e033 [1–10]
1	left	0	0	0	1	1
2		1	1	1	2	2
3		2; 3	2; 3	2; 3	3	3
4		4	4	4	4	4
5		5; 6	5; 6	5; 6	5; 6	5; 6
6		7	7	7	7	7
7		8	8	8	8	8
8		9	9	9	9	9
9	right	10	10	10	10	10
–99	No answer/missing	99	99	99	99	–3; –5 (;–4)
–88	Don't know	98	98	98	98 .a	–1
–77	refused				96 .b	–2
–66	No understanding of left–right		96	95		
–55	unspecified missing-type in survey	.	.	.	.	.
–44	Not asked in survey					–4

Weights

We provide the data from the surveys in its original, unweighted version. For your research, please regard the weigthening information provided by the survey sources.

1) The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 1 FULL RELEASE [dataset]. Ann Arbor, MI: University of Michigan, Center for Political Studies [producer and distributor]. August 4, 2003.

   Link: ftp://ftp.nes.isr.umich.edu/ftp/cses/studies/module1/data/cm1_cod2.txt

   • weights described in detail in variables A1010 to A1014

2) The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 2 FULL RELEASE [dataset]. June 27, 2007.

   Link: ftp://ftp.nes.isr.umich.edu/ftp/cses/studies/module2/data/cses2_codebook_part2_variables.txt

   • weights described in detail in variables B1010 to B1014

3) The Comparative Study of Electoral Systems (www.cses.org). CSES MODULE 3 FIRST ADVANCE RELEASE [dataset]. 2010.

   Link: ftp://ftp.nes.isr.umich.edu/ftp/cses/studies/module3/data/cses3_codebook_part2_variables.txt

   • weights described in detail in variables C1010 to C1014

4) The Mannheim Eurobarometer Trend File 1970–2002. Schmitt, Hermann. Mannheimer Zentrum für Europäische Sozialforschung (MZES). Scholz, Evi/Leim, Iris/Moschner, Meinhard. GESIS – Leibniz Institute for the Social Sciences (former "Zentrum für Umfragen, Methoden und Analysen" and "Zentralarchiv für Empirische Sozialforschung") (http://www.gesis.org/eurobarometer).

   Link to PDF Codebook: http://info1.gesis.org/dbksearch13/download.asp?id=16275

   • weights described on page 59

5) World Values Survey 1981–2008 Official Aggregate v.20090901, 2009. *World Values Survey Association* (www.worldvaluessurvey.org). Aggregate File Producer: ASEP/JDS, Madrid. (http://www.wvsevsdb.com/wvs/WVSData.jsp?Idioma=I).

   Link: http://zacat.gesis.org/webview/index.jsp?object=http://zacat.gesis.org/obj/fCatalog/Catalog5

   • see register on the left

   • weight to be found under:

   EVS 1981–2008 Longitudinal Data File

   => Variable Description => [ZA4804] Weight=> Weight OR Weight [with split ups]