Chapter 2
Design and Analysis

Based on theoretical considerations, the previous chapter elaborated that a small number of objects are likely to be relevant for voters when they cast their ballots: parties/governments, leaders/politicians, and policies/ideologies. I also argued that the impact of considerations concerning these objects can be expected to vary across voters and contexts. The objective of this chapter is to explicate how this expectation of heterogeneity in the individual calculus of voting is conceptualised and analysed empirically. The large number of potential approaches to the study of voters’ reasoning is constrained rigorously by the availability of data. Presenting the possibilities and shortcomings of the obtainable data at the outset allows me to be more concise in the remainder of this chapter by paying less attention to unfeasible paths of analysis. For that reason, a first section of this chapter describes briefly the empirical data that are used throughout this book. A second section discusses the measurement and interpretation of the calculus of voting. This involves an account of how voters’ thoughts on parties/governments, leaders/politicians, and policies/ideologies are approximated empirically. Third, intricacies of the cross-national comparison of the calculus of voting are discussed. This entails particularly the clarification of causal relations between individual and contextual levels of analysis. Fourth, I will argue why a particular statistical tool (multilevel conditional logit modelling) is most appropriate for the analyses to be undertaken in this study. Finally, I will
propose a new solution for the ubiquitous problem of how to treat missing data.

2.1 Data: The Comparative Study of Electoral Systems (CSES)

The set of data sources that enables the research of this book is very small. The design of the study requires individual level data on the calculus of voting in many different contexts. The Comparative Study of Electoral Systems currently provides the most suitable data in this respect. The CSES is a collaborative program of cross-national research conducted by election study teams in over fifty democracies. The data of a first round of CSES were collected between 1996 and 2002 and at the time of writing this report, data of thirty-eight elections in thirty-four political systems is available. Some of these were not useful because of the lack of variables of central importance (orientations towards parties, politicians, and policies, and of course, vote choice) or because they pertain to presidential, rather than parliamentary elections. A detailed report on the selection of surveys is included in Appendix 2. Parliamentary elections from the following political systems are analysed: Australia,

---

6 Other potentially relevant studies are the European Election Study (EES), the Eurobarometer, the European Social Survey (ESS), and the World Value Surveys (WVS). Since Eurobarometer, European Social Survey, and World Value Survey do not directly focus on voting behaviour they are less suited for my purpose than CSES and EES. The decision to use CSES instead of EES is based on the much larger variety of political systems included in CSES. The EES is limited to the member states of the EU, which are somewhat similar in terms of institutional settings and social structure. The CSES, conversely, provides data from democracies around the world.

7 The CSES so far encompasses two modules: a first module has been conducted from 1996 until 2002 and a second module since 2002. In these modules, two somewhat different questionnaires have been fielded.
Canada, Britain, the Czech Republic, Denmark, Germany, Hong Kong, Hungary, Iceland, Israel, Japan, Korea, Mexico, the Netherlands, New Zealand, Norway, Peru, Poland, Portugal, Romania, Russia, Slovenia, Spain, Sweden, Switzerland, Taiwan, and the Ukraine.

The dataset consists of micro as well as macro data of analysed countries. Individual-level data derives from the post-election surveys in the countries participating in CSES. As data and codebooks of CSES are available to the scientific community without restrictions, a detailed description of the data collection in each of the countries does not have to be provided here. Such information is reported in the CSES Codebook (Sapiro & Shively 2002) that can be obtained from the WebPages of the CSES project just like the dataset itself.\(^8\)

2.2 Conceptualising the Calculus of Voting

At the outset of any analysis of heterogeneity in voters' reasoning, elements of such reasoning have to be measured empirically. Four aspects of this measurement are of particular relevance and are discussed below. First, how can choice processes that are not directly observable nevertheless be studied empirically? The second section deals with the measurement and interpretation of the analysed vote function consisting of evaluations of parties/governments, leaders/politicians, and policies/ideologies. Third, the consequences of potential bias due to omitted variables are discussed. The final section deals with the relation between indicators selected for the vote function and theoretical traditions in electoral research.

*Measuring Voters’ Decision-Making Processes*

This book aims to investigate heterogeneity in the calculus of voting. This implies that I am not primarily interested in the

\(^8\) For detailed information see [http://www.umich.edu/~cses/](http://www.umich.edu/~cses/).
choice that people make in the voting booth, but rather in the process that gives rise to this outcome. As only the outcome of this decision-making process is observable, the question arises how to analyse the process itself. Two solutions to this problem are usually suggested: first, analysing respondents’ self-report on the considerations for their vote choice, and second, to estimate the decision process from observable information about the outcome and about putative causes.

The usefulness of the first approach is limited. The calculus of voting is likely to be more complex and will possibly incorporate a larger set of considerations than are likely to be captured by one open-ended question that asks respondents about the reasons of their vote choice. Moreover, it is unlikely that all aspects of the choice process are fully conscious. For respondents to accurately report their calculus of voting would require an enormous detachment and capacity to enquire their own behaviour. Rahn et al. (1994: 532) show on the basis of US data that “voters’ reports of the reasons for their preferences were principally rationalisations.”

Kaplan (1964) makes in this respect the distinction of ‘act meaning’ and ‘action meaning’. The former is the respondent’s account of his behaviour and motivations. The latter is the analyst’s account thereof. Apart from the difficulty of knowing one’s own motivations, there are the following two problems. First, ‘act meaning’ is cast in idiosyncratic terms that are not necessarily comparable between respondents. In search of ‘action meaning’, the analyst cannot guarantee that the same variables are used to analyse all cases, thus ensuring inter-subject comparability. Second, there are interview problems at hand. Neither interviewer nor respondent are trained to really probe motivations and also the interview setting is not conducive for doing that.

Due to these limitations of self-reports on voters’ decision-making processes, I will follow the path of statistical inference to retrieve characteristics of voters’ reasoning from
observable information on respondents. Linking a set of considerations with vote choice allows the estimation of their effects and therefore to reconstruct the decision process post hoc. Multivariate patterns of covariance form the basis of this ‘reverse engineering’ that results in a description of the decision-process.

Measuring Voters’ Views on Parties, Politicians, and Policies

Since it is aim of this study to investigate differences in the calculus of voting across individuals and political systems, this calculus has to be measured in a standardised form that applies to all units of analysis (individuals and political systems). In other words, vote choice has to be regressed on an identical set of explanatory variables that are potentially relevant to citizens in different political systems. Modelling the calculus of voting as a mixture of orientations towards parties/government, leaders/politicians, and policies/ideologies as suggested in Chapter 1 seems an approach that is practical for country comparative analysis. In spite of the large number of potentially relevant empirical evaluations, images, affects, etc with respect to parties/governments, leaders/politicians, and policies/ideologies, the data available from the CSES project have only few indicators that are comparable across countries. The following three indicators measure voters’ views on political objects.

---

9 To what extent these variables are actually relevant will be assessed empirically in Chapter 3 by means of country-specific analyses of the selected vote function.
10 Even if data on voters’ orientations towards specific issues (e.g., euthanasia) would be at the disposal of a researcher for many countries, the question is to what extent such data is truly comparable. For instance, due to the specific context of elections, the issue of euthanasia was highly salient in the Dutch parliamentary election of 1998 but irrelevant for many other parliamentary elections analysed. It is questionable to what
• *Party leaning.* The first object of voters’ reasoning—parties/government—is measured by the variable of party leaning, known to most scholars as the party identification question of the Michigan model of vote choice (Campbell et al. 1960).

• *Leader sympathy.* A sympathy rating (like-dislike) of party leaders measures voters’ evaluations on the second political object—politicians/leaders—on an eleven-point scale.

• *Left-right distance.* Voters’ considerations with respect to the third political object—policies/ideologies—are indicated by the distance between respondents’ positions and the perceived positions of parties on an eleven-point left-right scale.\(^{11}\)

The restrictive nature of these three empirical indicators is severe in the light of the peak number of potentially relevant orientations, evaluations, affects, and images that people may have with respect to these three classes of objects. However, this problem derives from the lack of comparable survey data. Considering the difficulties of international collaborative data projects, such as CSES, the restriction to a few comparable variables is a common one that cannot be solved here. Yet, I argue that the three indicators of party leaning, leader sympathy, and left-right distance are not a less-than-ideal solution. If one would be allowed to pick only one indicator for each class of object, it would be likely that it would be these three that were chosen by collaborators of the CSES project. In other words, in view of the literature, these indicators can be considered to be of extent the effect of euthanasia on vote choice lends itself to be interpreted as indicative for an issue orientation of vote choice across countries.

\(^{11}\) More detailed information on all three variables is reported in Appendix 2.
particular importance. This is basically the case because these variables are frequently reviewed as very general or even generic orientations (e.g., Jackson 1975b; Knutsen 1995b).

Voters' views on politics are not brought into being by a single election and they are not reinvented in each election all over again. Certain orientations are embedded in structures (or schemata) of long-term memory and are known to be highly stable over time (e.g., Miller et al. 1986; Lau 1989; Huckfeldt et al. 1999). Such chronically accessible points of orientation may be, for example, an ideological stand or the attachment with a party. A characteristic of these general orientations is that they moderate specific considerations that may be due to the actual situations of an election. In other words, voters evaluate, for instance, the issue position of one party in the light of their general leaning/aversion towards this party. Likewise, voters' evaluation of, for example, the integrity of one political leader will be affected by their overall orientation towards this politician. Specific considerations (view of party's issue position, integrity of the candidate) will, however, also contribute to long-term orientations (party leaning, general politician orientation) but the predominant causal direction between general and specific orientations is that the former determine the latter (e.g., Jackson 1983; Peffley & Hurwitz 1985).

Electoral research suggests that party leanings, leader evaluations, and left-right distances resemble such general orientations. The measure of the last political object—policies/ideologies—the distance between parties' and voters' positions on the left-right continuum is often described as a 'super issue'\textsuperscript{12} (e.g., Laver & Budge 1992). As Fuchs and

\textsuperscript{12} In the literature, the term 'ideology' is also frequently used to describe the left-right scale. The view of the left-right scale as ideology follows Downs' (1957) notion of ideologies as being generalisations of single issues. In this respect the left-right scale may be regarded as ideology. However, to avoid confusions with the use of the term in other instances ('ideology' as referring to '-isms' such as communism, liberalism, etc.) I
Klingemann (1990) and Knutsen (1995b) argue, left-right covers not only the traditional socio-economic cleavage but incorporates increasingly new political issue dimensions. Several scholars demonstrate by means of causal modelling that the left-right scale is antecedent to many specific policy or issue orientations (e.g., Kunz et al. 1993). Similarly, the variable of party leanings is considered to be a general measure of party orientations. The variable has been designed to specifically measure long-term views on parties (Campbell et al. 1960). Numerous studies indicate that such an interpretation holds in the sense that party leanings are antecedent to many specific evaluations of parties (e.g., Jackson 1975b; Page & Jones 1979). Finally, I assume that the same patterns hold for orientations towards politicians and that sympathy ratings of political leaders are general evaluations that affect more specific perceptions of politicians such as integrity, reliability, or competence.

In sum, the literature on electoral research reviews party leanings, like-dislike evaluations of politicians, and the left-right scale as general (or even generic) measures for orientations towards parties, politicians, and policies. Yet, one may still object that a model of vote choice that incorporates only three explanatory variables omits several factors that may be relevant for vote choice. To what extent this may be a problem for the analysis of this book is briefly discussed in the following paragraphs.

Omitted Variable Bias
The use of only a limited number of explanatory variables to model vote choice risks results to be biased owing to omitted variables. Such bias potentially occurs if (a) omitted variables are significantly affecting the dependent variable and if they are also strongly related to the independent variables included in the model. Yet, (b) only omitting causally antecedent

refrain from the use of the term ‘ideology’ and will refer to the left-right scale as general policy orientation.
explanatory variables may bias the estimation of causally decedent explanatory variables (e.g., King et al. 1994). Such omitted variable bias results in incorrect parameter estimates that limit the ability to make predictive inferences.

To illustrate this point consider the following example: the model on the calculus of voting in this book does not account for economic issues, which many applications have shown to be relevant for voting (e.g., Fiorina 1981; Jung 1982; Lewis-Beck 1988; Paldam 1991). It is likely that voters’ evaluations of parties in terms of competence in this field are positively related to party leanings. Respondents may, for example, ascribe higher economic abilities to the party they feel close to and therefore see performance with their partisan-glasses, or, they may prefer certain parties because of perceived economic performances. Hence, one has to conclude that missing out variables on economic evaluations fulfils the first condition of omitted variable bias. One could easily think of other examples of omitted variables that are correlated with the dependent variable of vote choice and the independent variables of party leaning, leader evaluation, and left-right distance.

The causal relationship in the example presented (the perceived economic performance of a party and the leaning towards a party) is, however, highly relevant for the question of whether omitting the effect of voters’ perception of parties’ economic performances on vote choice biases the estimated effect of party leanings. This is only the case if the former can be considered antecedent to the latter (condition b). In other words, only if voters feel close to a party because of perceived economic performances (and that they do not view economic performance with their partisan-glasses), the estimated effect of party leanings on vote choice may be biased.

As alluded to in the previous subsection, I consider explanatory variables of the estimated vote function (party leanings, leader evaluations, and left-right distances) to be
very general (or even generic) orientations towards parties, politicians, and policies. The prevailing wisdom in the literature on (political) attitudes and opinions is that general evaluations, for the most part, determine specific evaluations. Hence, omitting variables on specific evaluations does not (or not strongly) bias the estimated effect of general evaluations of parties, politicians, and policies. In the example provided in the previous paragraphs, omitting a variable on the perceived economic performance of parties does in all likelihood not bias the estimated effect of the general party evaluation (party leaning) in models reported in this book.\(^{13}\)

All things considered, parameter estimates of the vote function of party leanings, leader evaluations, and left-right distances are possibly somewhat overestimated due to the

\(^{13}\) If one thinks that certain antecedent variables of party leanings, leader evaluations, and left-right distances are omitted in the analysis, one can make an informed guess about the direction and the size of possible omitted variable bias (King et al. 1994). Positive correlations of omitted variables with dependent and independent variables result in overestimating the effects of variables that have been included (party leaning, leader evaluation, and left-right distance). At the same time, a quite different problem will unavoidably result in estimates of effects that are too low in comparison to their true value. The methods of analysis that will be used throughout this book are not capable of taking the consequences of measurement error into account when estimating effect parameters. Like all variants of the regression model, they assume that the dependent variable is measured without error. Violations of this assumption result in estimates of effects that are less efficient and more uncertain. Moreover, even though measurement error in the independent variables does not lead to bias, it increases the standard error of estimates in ways that are not calculated by regression algorithms themselves. For all practical purposes, regression models assume all variables to be measured without error, an absolutely implausible assumption. Measurement error, however, results in attenuation of observed associations between variables, hence to underestimation of effects. How the effects of these different problems —omitted variables and measurement error— add up is impossible to say in general terms. They do tend to eliminate each other. For the rest, this book is as much plagued by these problems as eventually all of the relevant literature, which also relies heavily on regression approaches.
biasing effect of all kinds of factors that impinge on vote choice but for which I cannot control because they are not available in the data. However, while only three considerations of vote choice are included in the model, the analysis presented in Chapter 3 will show that about half of the variance of vote choice is explained by these three variables alone. This is a reasonably high variance reduction for a model of vote choice and leaves not much room for omitted variables. The vote function that I estimate cannot claim comprehensiveness. Voters are likely to apply all kinds of different considerations in addition to the three that will be estimated here. Some of these will be of a political nature, some will be apolitical in character. But this lack of comprehensiveness does not necessarily invalidate the efforts in this study, namely to explore heterogeneity in the importance of these considerations included across individuals and political contexts.

Parenthesis: Parties, Politicians, and Policies and the Michigan Model of Vote Choice

The Michigan model of vote choice noticeably influences measures of the calculus of voting. This is foremost true for the variable of party leanings (Campbell et al. 1960). The choice of variables does, however, not imply that I adopt a Michigan-view of voting in this book. The selection of this vote function is rather driven by the fact that variables meet with the required criterion of analysis, as outlined before. Party leanings, like-dislike evaluations of politicians, and left-right distances and are variables that most generally measure voters’ views on parties, politicians, and policies. This theoretically non-aligned analysis of the selected vote function has implications for the interpretation of these variables and for the supposed relationship between them.14

14 This theoretically non-aligned view of variables included in the selected vote function is supported by literature on these measures. As indicated in Chapter 1, several scholars express doubts about a narrow interpretation
First, I adopt a broad interpretation of measures for voters’ views on parties, politicians, and policies and refer to them as general evaluations of or orientations towards these objects. The general character of these measures, as alluded to above, supports such an interpretation. Therefore, party leanings are not interpreted as lifelong psychological identifications but as general party evaluations. Likewise, like-dislike evaluations of party leaders are interpreted as generalised evaluations of politicians and left-right distances are seen as general policy orientation.

Second, in line with the theoretically non-aligned view on measures of general party, politician, and policy orientations I do not suppose any relationship between these variables as frequently suggested in the literature on the Michigan model of vote choice. The initial proposition was that (long-term) party identifications influence (short-term) evaluations of candidates and issues (Campbell et al. 1960).

---

of, for instance, the measure of party identifications as suggested by Campbell and his colleagues (1960) (e.g., Budge et al. 1976). This is particularly true for applications of the measure outside the USA (Thomassen 1976; Gluchowski 1978; van der Eijk & Niemöller 1983).

For that reason I use the term ‘party leaning’ instead of ‘party identification’ throughout the book. To avoid becoming repetitive I also refer to party leanings as party adherence, general party evaluation, party orientation, and party attachment.

One may object that the wording of survey questions asks respondents to evaluate these political objects on basis of different standards. Such wording mentions ‘being close’ in case of party leanings, ‘like-dislike’ in case of leader evaluations, and ‘positions’ in case of left-right distances. These differences are, however, not systematically discriminated throughout this book. I will refer to connotations in wording at points where they may arguably have consequences and may conflict with interpretations of explanatory variables as general evaluations of parties, politicians, and policies.

Again, research on the Michigan model of vote choice suggests that such propositions may not hold empirically. In opposition to this traditional view, non-recursive interpretations of the relationship between these different elements of the Michigan model have been proposed (Jackson 1975a; Fiorina 1981).
Due to uncertainties about the relationship between political objects in voters’ reasoning, party, leader, and policy evaluations are not modelled or interpreted as being antecedents or consequences of each other.

2.3 The Comparative Analysis

Since it is a central expectation of the study that the context of elections moderates the individual calculus of voting, a comparative perspective of individual voting behaviour is indispensable. Comparative analysis is understood in this book as a multilevel data problem in the same vein as propagated by Przeworski and Teune (1970). Variation in the dependent variable (individuals’ calculus of voting) is assumed to be a function of lower-level (individual) as well as higher-level (contextual) determinants. The two problems to be solved in such a perspective are: (a) identifying causal relationships between different levels of analysis and (b) distinguishing effects between these levels.

Relationships between Contextual and Individual Levels

The first problem related to the multilevel nature of the comparative analysis is to specify relationships between different levels. Does context directly affect individual vote choice, making it an intrinsic part of the reasoning of voters? Or, does context moderate the individual reasoning of the decision but not directly affect vote choice? To illustrate the problem of causal relationships, consider Figure 2.1.

---

18 In Stein Rokkan’s (1963) terms, this can be labelled as analysis of macro (characteristics of elections), micro (characteristics of voters) explanations respectively of a micro-micro relationship (evaluations of choice options affecting vote decisions).

19 A third problem of the comparative approach is how to estimate individual and contextual effects most accurately. A description of multilevel techniques and their application in the analysis of this book is given in Section 2.4.
Figure 2.1 Individual and Contextual Determinants of the Calculus of Voting

Figure 2.1 distinguishes contextual characteristics of elections (e.g., institutional settings, economic context, etc.), individual characteristics of voters (e.g., age, education, etc.), evaluations of choice options (voters' orientations towards parties, politicians, and policies), and individual vote choice. A box in Figure 2.1 combines evaluations of choice options and vote choice to what I call the individual calculus of voting, or, the reasoning of vote choice. The arrows in Figure 2.1 represent direct and indirect effects on vote choice. Not all of these effects are studied in this book. Dotted arrows stand for possible determinants of vote choice that are ignored and solid arrows denote causal relationships that are analysed. To spell out what these paths represent and what
this study aims to explain, I will briefly discuss and illustrate ideas behind the arrows in Figure 2.1.

Each of the direct effects on vote choice signifies possible considerations of the individual vote decision. As discussed in Section 2.1 on the conceptualisation of the calculus of voting, I analyse voters’ evaluations of parties, politicians, and policies as determinants of vote choice. In other words, this study estimates the direct effect of evaluations of choice options on vote choice (solid line) omitting direct effects of characteristics of elections and voters (dotted lines).

Figure 2.1 includes four indirect effects: two mediated effects of contextual and individual characteristics on vote choice via the evaluation of choice options and two moderating effects of contextual and individual characteristics. The latter two may be described as indirect effects of vote choice or as direct effects on the individual calculus of voting. The research question of this book focuses on these moderating effects of contextual and individual

---

20 Issue voting models, for example, illustrate the (solid) line between evaluations of choice options and vote choice. These studies show that there is a direct link between orientations to choice options (parties’ policy stands) and vote choice (e.g., Enelow & Hinich 1984; Rabinowitz & MacDonald 1989).

21 The direct arrow from contextual characteristics to vote choice finds empirical support in studies on tactical voting. Voters take the context of elections in the form of constraints of institutional settings and the competitiveness of elections into account when casting ballots (e.g., Franklin et al. 1994; Blais & Nadeau 1996; Alvarez & Nagler 2000; Blais et al. 2001). Explanations of vote choice that are based on socio demographic variables, such as models of class voting, are examples for the effect of individual characteristics on vote choice (e.g., Lazarsfeld et al. 1944; Nieuweeerta & Ultee 1999).

22 The fact that I omit the direct effect of individual and contextual characteristics on vote choice entails that I also do not control for the possibility that some of the effects of individual and contextual characteristics is mediated by evaluations of choice options (dotted lines).
characteristics on vote choice (solid lines). Chapter 4 focuses on the moderating effects of individual characteristics and Chapter 5 on the moderating effects of contextual characteristics for the effect of evaluations of parties, politicians, and policies on vote choice. These chapters investigate, for instance, to what extent politically sophisticated voters weigh political objects differently than less informed voters when casting ballots, or whether certain institutional settings affect the importance of parties, politicians, or policies for actual choice.

---

23 There are four more classes of moderating effects possible in Figure 2.1. Individual traits may moderate (a) direct individual and (b) contextual effects and contextual traits may also moderate (c) direct individual and (d) contextual effects on vote choice. Due to the fact that I omit direct effects (dotted lines) of individual and contextual characteristics on vote choice, I also do not analyse to what extent these considerations are moderated. To illustrate what these indirect effects stand for, consider the direct effect class as individual trait may have on of vote choice. This effect may be moderated across individuals and contexts. The effect of class is (a) possibly larger for voters who are very conscious of their class membership and (b) for countries that know a politicised socio-economic cleavage. As direct contextual effect consider the extent to which voters rely on information about the competitiveness of elections (tactical voting). This direct effect on vote choice may vary across groups of voters that (c) differ in political sophistication and may vary between (d) political systems of proportional representation and first-past-the-post systems. One may argue that not all voters are necessarily aware of the closeness of a race (individual moderating effect) and in the second case one may argue that the question which party or candidate receives most votes (race between the two largest options) does not have the same importance for systems of proportional representation than for first-past-the-post systems (contextual moderating effect).

24 Figure 2.1 also includes a reciprocal effect between individual and context, which is not analysed in this book. The example of the previous footnote on the association between class voting, class-consciousness, and a politicised class-cleavage may illustrate this point as well. It may be that the extent to which individuals have a distinct class-consciousness affects the likelihood that class cleavage is politicised in a country. It may, however, also be that individuals in a political system with a strong class cleavage have a higher likelihood of developing a class-consciousness.
The choice to analyse these three arrows does not mean that other causal relationships do not exist or that they are not a fruitful field of research. It only refers to the constraints of available data and the reasoned simplification of a complex subject such as the calculus of voting. The primary focus of this research is to identify determinants (individual and contextual) of the calculus of voting (the effect of party, politician, and policy evaluations on vote choice) and not to provide a comprehensive explanation of voting behaviour.

*Individual, Compositional, and Contextual Effects*

Having discussed the first problem of the comparative approach, identifying causal relationships between different levels of analysis, I now turn to the second problem, attributing variation in the dependent variable to independents at different levels of analysis. Analysing the calculus of voting across individuals and across contexts inherently refers to two sources of variation and two levels of analysis. The first source of variation is individual and relates to variables such as education, age, attitudes, beliefs, etc. The second source is contextual and includes, for example, institutional characteristics and characteristics of the party system. Contextual characteristics only vary at the contextual level, thus they are constant within contexts. Individual characteristics, however, do not only vary on the individual level but also in terms of their distribution on the contextual level as illustrated in Figure 2.2.
**Figure 2.2** Different Effects in the Comparative Analysis

### Level of Analysis

<table>
<thead>
<tr>
<th>Basis of Variation</th>
<th>Individual</th>
<th>Contextual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Individual</td>
<td>(Relative-) Individual Variation</td>
<td>Compositional Variation</td>
</tr>
<tr>
<td>Contextual</td>
<td>Constant</td>
<td>Contextual Variation</td>
</tr>
</tbody>
</table>

Consistent with Marsh’s (2002) terminology, I will use the term ‘compositional variation’ to refer to individual specific variance across higher-level contextual units, such as differences in the distribution of age or education across countries.\(^{25}\) The term (relative-) individual variation describes individual-level factors adjusted for their contextual variation. Iversen (1991) calls this contextual analysis with relative effects: what matters for variation of the dependent variable is not necessarily the absolute value of an individual-level, independent variable but the value *relative* to the distribution of the individual-level variable in the higher-level unit. The

\(^{25}\) In the literature the terms ‘pseudo-contextual variation’ or ‘distributional variation’ are used equivalently (Johnson et al. 2001). Lazarsfeld and Menzel (1969: 427) refer in this respect to the “analytical properties of a collective (precinct) composed of individuals.” Contextual properties may not always relate to the mean of an individual-level variable in an analytical unit, but also to the standard deviation, a correlation, etc. (Lazarsfeld & Menzel 1969).
distribution of individual level variables across contextual units is usually approximated by the mean of the variable in each context (Iversen 1991; Snijders & Bosker 1999; Zorn 2001). (Relative-) individual effects are therefore explanatory variables centred by the group mean within contextual units.

**Problems of Compositional Variation**

A general problem of analysis across countries is the comparability of measurements of individual-level variables. This concerns demographic variables like education as well as opinion questions such as responses on the left-right scale. Such problems occur in the form of calibration problems as well as diminished reliability and validity of compositional variation. Such incomparabilities across countries may originate from differences in the applicability of answer categories, differences in data collection, and differences in the meaning of individual-level measures. These three points are briefly outlined in the following paragraphs.

Answer categories of survey questions are not always applicable to all countries. This results in calibration problems in comparative survey research. Whereas the comparability of answer categories on ‘age’ or ‘sex’ of respondents is widely regarded as unproblematic, the equivalence of educational degrees or the frequency of church attendance is definitely more problematic, because response categories to such survey questions may differ across countries. As a case in point, consider the comparison of educational degrees between the USA and Germany. The US educational system knows two stages (high school degree, university degree) whereas the German systems know three stages (Hauptschulabschluss/ Mittlere

---

26 In fact, some of the individual level variables provided by CSES have somewhat different answer categories. For instance, CSES collaborators fielded the variable of church attendance usually with five, but sometimes with only four response categories. For more detailed information on these differences see Appendix 3 and the CSES codebook.
Reife, Abitur, university degree). The answer category ‘high school degree’ in the USA is neither identical to the German answer category ‘Mittlere Reife’ nor ‘Abitur’ and responses to the survey question ‘education’ are therefore not fully comparable between both countries. Similar problems exist when considering other systems. ‘Church attendance’ is another variable that is notorious for its difficulties in comparative studies. Religious services are held in dissimilar formats and in different frequency across societies. ‘Union membership’ is yet another variable where response categories may have different meanings due to dissimilar organisational principles. To compensate for these calibration problems I do not investigate compositional effects of individual variables like ‘education’, ‘union membership’, ‘urban residence’, and ‘church attendance’ in this study but only (relative-) individual effects. For reasons

27 Moreover, measures of ‘church attendance’ and ‘union membership’ do not only face calibration problems in comparative research but also validity problems. Does an indicator of, for instance, ‘church attendance’ tap the same theoretical construct in all societies? It may be that ‘church attendance’ measures religiosity in one country but social conformism in another. See the third point of this subsection for validity problems.

28 An extreme case of calibration problems represents the variable of ‘political knowledge’ as analysed in this study. Different knowledge questions have been asked across countries. The cross-country variation of the scale is therefore coincidental and does not reflect any systematic differences in the degree of political knowledge across electorates. Moreover, as country-specific scaling analyses of knowledge items reveal, a single scale does not reliably represent all factual knowledge questions (see Appendix 3 for more information). Therefore, compositional variation of the knowledge scale is, just as compositional variation of variables like ‘education’ and ‘church attendance’, not modelled in the analyses.

29 As can be shown formally and as is illustrated by exemplary analyses of relative and absolute effects, parameter estimates of variables centred by group-means are identical to the ones of absolute values (Kreft & Leeuw 1995; Snijders & Bosker 1999). This is because centring is a linear transformation of variables and does not alter the estimation of effect parameters. Also, excluding compositional effects from the analysis of (relative-) individual variables does not generate omitted variable bias,
of consistency this is also not done for those compositional variables that have stable answer categories across countries like ‘sex’ or ‘age’. 

Differences in data collection across countries may diminish the reliability of compositional variation of individual-level variables. Even if election study teams across countries intend to include identical questions and identical response categories in their surveys, translation of the questions and mode of data collection are probably not completely identical. Small changes in measurement can generate differences in responses across countries, which are mistakenly treated as substantive differences if one does not correct for such measurement error. As Saris (1997) shows for Eurobarometer data, both language and the mode of data collection affect response probabilities for survey questions across countries. Preferably one would test (and subsequently correct) for different response functions in surveys across countries before variables are used for a substantive analysis of country differences. However, such an approach requires multiple responses for the same concept in each country. Since such data are not available in the CSES project, I can only make the reader aware of possible biases in the country comparative analyses due to differences in data collection.

Differences in the meaning of survey questions across countries are the third uncertainty related to the comparison of individual-level variables. Such differences may diminish the validity of variables across countries. The comparison of responses to opinion questions like the left-right scale or

---

because the correlation between (relative-) individual and compositional variables is by definition zero (because compositional variation is constant within contexts and relative individual variation is constant across contexts).

30 Entering compositional variables in the analysis is a sensible strategy if it rests on theoretical grounds (Snijders & Bosker 1999). Can one expect the composition of age or sex across countries to be directly related to voters’ reasoning? In my view this is perhaps possible but it is probably not central for the analysis of the calculus of voting.
party loyalty may be questioned due to differences in the meaning of these concepts across countries. Do respondents in Peru understand the terms left and right in the same way as respondents in Iceland do? Is the same concept measured? It is rather unlikely that the substantive meaning of left and right is identical in all countries. Knutsen (1995b), for example, demonstrates for several Western democracies that different issues and ideologies are summarized under the scale (see also van der Brug 1997). But it is also very implausible to assume that all respondents within a single country understand all survey questions exactly in the same way. Van der Eijk and Niemöller (1983: 225-247) use open-ended questions to demonstrate that for the Netherlands respondents comprehend terms of ‘left’ and ‘right’ in different ways. But as they also show, these terms have the same instrumental meaning across individuals. Even if voters disagree on what ‘left’ exactly means they all classify a communist party as leftist, etc. The same is assumed to be the case across countries. For the aim of this study it is not necessary that respondents in different countries utilise the same substantive factors when evaluating parties, politicians, and policies. But I do assume that they use party leanings, leader evaluations, and left/right for the same purpose: to derive a vote decision from these evaluations. In the literature of the comparative approach this is referred to as functional equivalence (Przeworski & Teune 1970; Dogan & Pelassy 1984). Hence, the variables included in the vote function are assumed to be functionally comparable across countries.

---

31 As analyses of Chapter 3 will show, this assumption is supported by the data: party leanings, leader evaluations, and left-right distances significantly affect vote choice in all contexts (with one exception). See particularly the subsection on ‘The Applicability of the Vote Function’.
2.4 Choice of Statistical Model

This book investigates variation in voters' reasoning across individuals and contexts. More specifically, it studies variation in the effect of party leanings, leader evaluations, and left-right distances on vote choice. Integrating all individuals and contexts into a single model of vote choice allows an analysis of the moderating effect of individual and contextual characteristics for the weight of party, politician, and policy orientations for vote choice. Such an analysis of voters' reasoning in an integrated (or pooled) model across individuals and contexts requires two features from statistical models which traditional regression models do not account for. This refers, first, to the multinomial structure of the dependent variable 'vote choice' in multiparty contests, and second, to the multilevel data structure across individuals and contexts.

The dependent variable of 'vote choice' varies in the number of categories across political systems simply because of differences in the number of contesting parties in these elections. Moreover, these categories include quite different types of alternatives, from socialist to right wing parties, in an unordered form. Several attempts to order categories of vote choice in multiparty contests can be distinguished in the literature. Different parties could be grouped in stable categories of party families, which apply to each country. Or, one may transform vote choice into a dichotomy (for example, 'governmental' versus 'opposition' parties, 'liberal' versus 'other', etc.) or an interval-level set of party positions on a particular dimension (e.g., a left-right dimension). Such solutions, however, are unsatisfactory. They replace the dependent variable of 'vote choice' with a single aspect of it, leaving out all sorts of other considerations of vote

\[^{32}\] The single aspect would be, for example, a reasoning that concerns solely differences between party families, governmental versus opposition
choice that are also likely to be relevant for voters' reasoning. As will be elaborated below in the section on conditional logit models, observed electoral utilities and the application of discrete choice models enables us to overcome these obstacle.

The analysis of voters' reasoning in an integrated model across individuals and contexts by means of traditional regression techniques rests on the assumption that the underlying sample of individuals has been drawn randomly from all individuals (and thereby all contexts) of the world. This is not the case for data from the CSES or probably any cross-national data set. Such data are habitually collected in a two-step sampling procedure of drawing contextual units and thereafter (nested) individuals within these contexts. This hierarchical data structure impairs traditional regression methods and requires multilevel techniques, as I will argue in the section below on multilevel models.

Which statistical tools allow for the most accurate inferences on variation in the calculus of voting, given the design and available data? The following two subsections report technical adaptations of traditional regression models that account for the problem of analysing vote choice in multiparty elections (conditional logit model) and the analysis across multiple levels (multilevel model). These two sections depict the statistical model applied in the empirical analysis of this book, a multilevel conditional logit model. Appendix 2 provides the formal definition of the model.

**Conditional Logit Model**
The first problem of the analysis of voters' reasoning that cannot be handled by traditional regression models is the multinomial structure of the dependent variable of 'vote choice'. Two ways of analysing vote choice in multiparty
elections are proposed in the literature: the analysis of observed party utilities\(^\text{33}\) (Tillie 1995; van der Eijk et al. 1996; van der Eijk 2002; van der Eijk & Kroh 2002; Kroh & van der Eijk 2003) and the application of discrete choice models (Fuchs & Kühnel 1994; Whitten & Palmer 1996; Alvarez & Nagler 1998; Quinn et al. 1999; Thurner 2000; Glasgow 2001). Both approaches follow the logic of utility theory (e.g., Fishburn 1970; Ben-Akiva & Lerman 1985: 31-57). Voters’ reasoning is regarded as a decision process in which voters, first, evaluate each available choice alternative according to a set of attributes (or, external political objects). Subsequently, voters select the most preferred alternative, or in other words, the alternative with the perceived maximum utility.

The first approach, analysing complete preference orders by observed party utilities, creates a dependent variable that varies between respondents and between parties that can be analysed straightforwardly across political systems.\(^\text{34}\) Discrete choice models on the other hand estimate party utilities instead of observing them. Based on attributes of parties and characteristics of voters these utilities are derived iteratively, assuming a choice according maximum

\[^{33}\text{In comparison with the categorical variable of vote choice, questions about the electoral utilities are asked for each party separately. The wording of a well-known example is: “We have a number of parties in <country> each of which would like to get your vote. How probable is it that you will ever vote for the following parties? Please specify your views on a 10-point scale where 1 means “not at all probable” and 10 means “very probable”. ” Analysing observed party utilities instead of the categorical variable of vote choice has several advantages. Most importantly, they allow the analysis at a complete preference order of parties (Kroh & van der Eijk 2003). Discrete choice models of the categorical variable ‘vote choice’ have to estimate (unobserved) party utilities. Whereas this is unproblematic in the two party case, it becomes increasingly difficult and inaccurate if the number of parties in a political system is high (van der Eijk & Kroh 2002).}\]

\[^{34}\text{See Oppenhuis (1995), van der Eijk and Franklin (1996), and van der Brug et al. (2003) for applications.}\]
utility (McFadden 1974; Hausman & Wise 1978; Hausman & McFadden 1984). Because the CSES data do not contain observed utilities or preferences,\(^{35}\) discrete choice models have to be used.

The analysis of moderating effects of individual and contextual characteristics for the weight of party leaning, leader evaluation, and left-right distances for vote choice impairs the estimation of single effect parameters of party, politician, and policy evaluations for each context or for each party elected. The goal should rather be to estimate single effect parameters for party, politician, and policy orientations that hold for all units of analysis as well as the increments/decrements of these values dependent upon individual and contextual variables. In other words, one does not want to have parameter estimates of, for example, party leanings for the Social Democrats in Denmark, the Democratic People’s Party in Korea, the Liberal Democratic Party in Slovenia, etc. but one would like to have one estimate for party leanings that holds across all voters, parties, and political systems only to be altered for specific individual and contextual characteristics if a compelling need

\(^{35}\) Using ‘party sympathy ratings’ that is actually included in the CSES as a substitute for electoral utilities turns out not to be feasible (Kroh 2001). A comparison of the properties of the question about electoral utilities (likelihood of ever voting for parties) and the ‘party sympathy ratings’ based on data of the Dutch parliamentary election study 1994 reveals that party sympathy is to a considerably lesser extent able to explain vote choice. Whereas 93% of the respondents vote for the party with the highest electoral utility, this is only true for 72% of the respondents in case of the sympathy ratings. Moreover, these ratings reflect different causal relationships with respect to orientations towards choice options. Evidently, sympathy for a party is neither a necessary nor a sufficient condition for vote choice. Sympathy ratings measure only part of voters’ reasoning, namely feelings/sympathy and does not sufficiently account for considerations such as party competences.
exists to add such variations. This limits alternative discrete choice models to the conditional logit model.\textsuperscript{36}

The conditional logit model has two shortcomings, which should be kept in mind. First, all discrete choice models may produce biased estimates (van der Eijk and Kroh 2002). Second, within the class of discrete choice models the most proper models for vote choice are those that allow for error correlations, such as Mixed Logit models (McFadden & Train 2000). Yet, the choice of the Conditional Logit model is a consequence of the availability of data and of the requirements of the research question. As long as there is not more adequate data or methods available—and CSES is already a major step towards studying electoral behaviour comparatively—the chosen approach is in my view most appropriate.\textsuperscript{37}

\textsuperscript{36} Conditional Logit conceptualises the $i^{th}$ individuals utility of the $j^{th}$ choice as $U_{ij} = \beta x_{ij} + u_{ij}$. The choice unspecific estimate $\beta$ makes an integrated comparison of vote functions possible. Refer to Appendix 2 for a formal description and a brief discussion of the conditional logit model. Although the multinomial logit model is often advocated and accepted as appropriate way to model vote choice comparatively (Whitten & Palmer 1996), it yields at best a series of non-integrated separate country analyses. Multinomial logit/probit conceptualise the $i^{th}$ individuals utility of the $j^{th}$ choice as $U_{ij} = \beta x_{ij} + u_{ij}$. The choice specific estimate $\beta_j$ makes an integrated comparison of vote functions impossible. Other (advanced) discrete choice models that can handle violations of the assumption of Independence of Irrelevance Alternatives (IIA) are also inadequate for the pooled analysis. Their correction of error correlations between choices is based on country specific choice groupings, owing to which their solution to IIA violations is incomparable in a pooled model across contexts. Mixed logit and multinomial probit models for discrete dependent variables that cope for IIA, correct for covariance between error $u_{ij}$ and $j^{th}$ choice. The choice specific estimate correction $\gamma$ impairs an integrated comparison of vote functions (Hausman & Wise 1978; Alvarez & Nagler 1998; McFadden & Train 2000; Glasgow 2001).

\textsuperscript{37} As Coombs (1964: 5) puts it: "all knowledge is the result of theory—we buy information with assumptions—“facts” are inferences, and also are data and measurements”. In this sense, knowledge is bought by assuming that weaknesses of the statistical model do not affect substantive empirical results. As indicated before, this study is considered explorative.
**Multilevel Model**

A problem that relates to the intricacies of the comparative approach is the accurate estimation of effects across different levels of analysis. The multi-stage sampling process of first selecting countries and then individuals within countries leads to a hierarchical structure of data. Disturbance terms have to be specified for both draws, for the sample of countries and for the sample of individuals within countries. An ordinary analysis of pooled cross-sectional data in several countries does not take adequate account of the first disturbance term. Moreover, in multi-stage samples, individuals of the overall sample are not fully independent from one another. To assume that error exists only at the individual level leads to underestimation of standard errors and to subsequent biases.

In recent years, new statistical methodologies have been developed to overcome this problem by applying multilevel techniques (Bryk & Raudenbush 1992; Goldstein 1995; Snijders & Bosker 1999). These methods are also becoming standard in political science (Western 1998; Steenbergen & Jones 2002) and they are increasingly applied in studies of electoral behaviour (Charnock 1997; Lubbers et al. 2000; Pickery 2002).

Multilevel analysis is used in this study by estimating a conditional logit model of vote choice with random residuals for the slopes of explanatory variables (party leaning, leader evaluation, and left-right distance) across countries. This means that effect parameters of the vote function are allowed to differ across contextual units. A (random) sample of countries allows stochastic inferences.

Assumptions are required to obtain explorative answers. These, in turn, are useful if they trigger more appropriate measures in subsequent research and in the end improve knowledge on voters’ reasoning.
about these effect parameters across political systems.\textsuperscript{38} A formal specification is provided in Appendix 2. The multilevel model produces not only more realistic estimates of effect parameters and standard errors but also lends itself to a decomposition of variance of the dependent variable and of variance in effect parameters between levels of analysis. It is therefore possible to define how much variation in the pooled data is due to contextual heterogeneity. I will demonstrate this point in Chapter 3.

2.5 Incomplete Data

As in most other surveys, response to stimuli in the CSES data is far from perfect. Respondents often state that they have no opinion on a survey question or they refuse to give an answer. Moreover, some variables of the CSES module are not included in all country surveys. Hence, individual level data of this study are frequently incomplete. This becomes problematic when the incidence of missing data is so high that variables of interest cannot be included in the analysis. An even more problematic situation arises when countries cannot be integrated in the pooled model because not all variables of interest have been included in the respective survey.\textsuperscript{39} Loss of efficiency in the estimation of parameters

\textsuperscript{38} Clearly, the question has to be addressed to what extent countries included in CSES can be regarded as a random sample of democracies in the world. There is for example no African democracy in the sample, and a strong bias exists towards European countries. The only criterion to include a country in the sample is the availability of data rather than a random draw. However, to conclude that multilevel techniques are not applicable seems to me untenable. Lack of a random sample of countries implies that standard errors in a multilevel model are underestimated. To reject the method and to apply ordinary tools, however, means even to increase this problem. In my opinion, the multilevel technique provides the best statistical model for the given data.

\textsuperscript{39} In the Spanish survey, for example, the church attendance question is not included. As a consequence, one could leave out Spain from the
may also be problematic when missing data are excluded, leading to a reduced sample size. It is therefore preferable to include missing data in the analysis rather than deleting or excluding them.

Selection bias can be the result of a listwise deletion of incomplete data, one of the standard procedures of handling missing data. This goes back to the likelihood of missing data. Data can be missing completely at random (MCAR) or missing at random, MAR\(^40\) (Schafer 1997; King 2001; Allison 2002). In the first case, MCAR, the occurrence of missing data is uncorrelated with any other variable in the dataset. A listwise deletion of incomplete data will approximately generate a random subsample of the original \(N\). To the extent that incomplete data are related to other variables in the dataset, MAR, listwise deletion of these missing data will generate a biased subsample. It is an empirical question whether or not correlations are evident between patterns of incomplete data and variables of the analysed dataset.

Table 2.1 below reports the proportion of complete data for two orientations of the vote function, leader evaluations and perceived left-right positions. I do not report pooled analysis across all countries, or, one could exclude the variable of church attendance from explanatory models. Both alternatives evidently are unattractive. Keeping in mind that the CSES is a collaborative project, it is little surprising that local expediency occasionally overrules shared comparative aspirations, with the consequence that not all questions are included in all surveys. The cross section of countries and variables that is aimed at in this book therefore becomes rapidly small when no solution for missing data is found.

\(^{40}\) A third category of missing data distinguished in the literature is nonignorable missing data, NI. Nonignorable missing data denotes a case in which missing data is related to the true but unobserved score. If wealthy individuals are more likely to respond with a “don’t know” to questions regarding their income, nonignorable missing data is encountered. This form of missing data requires specific treatment (Heckman (1976) models for selection bias) and is not included in the following discussion of the missing data problem.
the corresponding figures on party leanings because there are very few missing cases on the question in all analysed political systems. Whereas it seems difficult for many respondents to place parties on a left-right scale or to evaluate all political leaders of a political system, they do not have problems to report whether or not they feel close to a party. The first column of the table indicates how many parties are analysed per country and how many options respondents therefore have to evaluate. Which specific parties and political leaders this are can be obtained from Table A2.3 in Appendix 2. The second column reports the percentage of respondents found to evaluate all the leaders of all parties. For example, 94% of the Australians know and evaluate the four party leaders, which are included in the analysis of voting behaviour in this political system. Likewise, 70,7% of Australians manage to place themselves and these four parties on a left-right scale. Figures indicate that the likelihood of incomplete data clearly is different across countries and can reach up to 70% of a national sample, like in Hong Kong. In the comparison across country surveys, Norway and Denmark have the lowest numbers of missing cases on the left-right scale with about 10% in respective samples. More than half of respondents have incomplete information on the left-right scale in Hong Kong, Korea, New Zealand, Romania, Russia, Slovenia, Taiwan, and the Ukraine. The number of missing cases seems to be above average in transition democracies. Such differences may simply be explained by the fact that citizens in established democracies have had more time and experience to evaluate parties than citizens in new democracies. Another explanation would be

---

41 This does not mean that 70% of the sample does not have information on any political leader, but that 70% of respondents has missing information for at least one party leader. If, for example, the political leader of a small party happens to be unknown to a large part of the electorate, all these cases are neglected in a solution of listwise deletion. Figures therefore do not necessarily indicate how well informed respondents in different countries are.
that parties in transformation democracies are more volatile in their ideological positioning than in established party systems (e.g., Fish 1995; White et al. 1997). It may therefore be more difficult for voters in transition countries to locate parties’ ideological stands. Similar result can be obtained from the comparison of missing data on leader evaluations across countries. Less than 50% of the samples know all politicians in Hong Kong 2000, Mexico 2000, and the Ukraine, whereas more than 90% of Australian, Canadian, Danish, Norwegian, Spanish, and British citizens know all politicians ascertained by CSES in their political system. Again, orientations to political objects (politicians) are found more frequently in traditional Western democracies.\footnote{This descriptive observation of non-random country differences in the proportion of missing data is matched by comparisons of bivariate correlations and means. The proportion of missing data on leader evaluations is positively correlated with the proportion of missing data on left-right distances across the analysed 30 contexts ($r = 0.62; n = 30$; t-test; $p < 0.01$). Moreover, splitting the sample of countries between established and transition democracies allows the comparison of the proportion of missing data for both groups. The group of transition democracies consists of the Czech Republic, Hong Kong, Hungary, Poland, Romania, Russia, Slovenia, and the Ukraine. The proportion of incomplete data on the left-right scale in this group of countries is significantly higher than in other democracies ($n = 30$; F-test; $p < 0.05$).} Evidently, the number of missing cases across countries reflects an important aspect of these political systems.

Splitting the British sample exemplarily into subsamples with low, medium, and high education\footnote{For the operationalisation of the variable education see Appendix 3.} illustrates that systematic differences exist not only in the cross-country comparison but also within countries: less educated citizens apparently have more problems in evaluating parties in terms of leader evaluations and left-right distances. The gap in the share of missing cases on the left-right scale between low and highly educated respondents is about 27 percentage points.
Table 2.1 Percentage Complete Data on Leader Evaluations and Left-Right Distances

<table>
<thead>
<tr>
<th></th>
<th>Number of parties to be evaluated</th>
<th>% Valid: Leader Evaluation</th>
<th>% Valid: Left-Right Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia 1996</td>
<td>4</td>
<td>94.0</td>
<td>70.7</td>
</tr>
<tr>
<td>Canada 1997</td>
<td>4</td>
<td>94.8</td>
<td>58.6</td>
</tr>
<tr>
<td>Czech Rep. 1997</td>
<td>6</td>
<td>85.8</td>
<td>80.4</td>
</tr>
<tr>
<td>Denmark 1998</td>
<td>6</td>
<td>93.2</td>
<td>88.5</td>
</tr>
<tr>
<td>Germany 1998</td>
<td>5</td>
<td>64.5</td>
<td>76.7</td>
</tr>
<tr>
<td>Hong Kong 1998</td>
<td>6</td>
<td>65.3</td>
<td>30.8</td>
</tr>
<tr>
<td>Hong Kong 2000</td>
<td>6</td>
<td>29.5</td>
<td>28.0</td>
</tr>
<tr>
<td>Hungary 1998</td>
<td>5</td>
<td>89.2</td>
<td>65.3</td>
</tr>
<tr>
<td>Iceland 1999</td>
<td>5</td>
<td>83.7</td>
<td>61.8</td>
</tr>
<tr>
<td>Israel 1996</td>
<td>3</td>
<td>84.5</td>
<td>72.6</td>
</tr>
<tr>
<td>Japan 1996</td>
<td>5</td>
<td>72.9</td>
<td>54.6</td>
</tr>
<tr>
<td>Korea 2000</td>
<td>6</td>
<td>56.9</td>
<td>42.1</td>
</tr>
<tr>
<td>Mexico 1997</td>
<td>3</td>
<td>70.2</td>
<td>53.0</td>
</tr>
<tr>
<td>Mexico 2000</td>
<td>3</td>
<td>35.2</td>
<td>55.9</td>
</tr>
<tr>
<td>Netherlands 1998</td>
<td>6</td>
<td>52.1</td>
<td>63.5</td>
</tr>
<tr>
<td>New Zealand 1998</td>
<td>6</td>
<td>64.7</td>
<td>44.4</td>
</tr>
<tr>
<td>Norway 1997</td>
<td>6</td>
<td>94.4</td>
<td>90.9</td>
</tr>
<tr>
<td>Peru 2001</td>
<td>6</td>
<td>72.3</td>
<td>52.5</td>
</tr>
<tr>
<td>Poland 1997</td>
<td>6</td>
<td>73.3</td>
<td>64.2</td>
</tr>
<tr>
<td>Portugal 2002</td>
<td>6</td>
<td>77.3</td>
<td>62.3</td>
</tr>
<tr>
<td>Romania 1996</td>
<td>4</td>
<td>76.1</td>
<td>44.5</td>
</tr>
<tr>
<td>Russia 1999</td>
<td>6</td>
<td>75.4</td>
<td>45.8</td>
</tr>
<tr>
<td>Slovenia 1996</td>
<td>5</td>
<td>78.5</td>
<td>47.7</td>
</tr>
<tr>
<td>Spain 1996</td>
<td>3</td>
<td>94.9</td>
<td>78.9</td>
</tr>
<tr>
<td>Spain 2000</td>
<td>3</td>
<td>82.0</td>
<td>79.1</td>
</tr>
<tr>
<td>Sweden 1998</td>
<td>6</td>
<td>68.6</td>
<td>77.2</td>
</tr>
<tr>
<td>Switzerland 1999</td>
<td>4</td>
<td>75.5</td>
<td>76.1</td>
</tr>
<tr>
<td>Taiwan 1996</td>
<td>3</td>
<td>69.5</td>
<td>35.3</td>
</tr>
<tr>
<td>Ukraine 1998</td>
<td>4</td>
<td>18.9</td>
<td>38.0</td>
</tr>
<tr>
<td>UK 1997</td>
<td>3</td>
<td>93.2</td>
<td>71.3</td>
</tr>
<tr>
<td>UK, Low Education</td>
<td>3</td>
<td>90.9</td>
<td>64.4</td>
</tr>
<tr>
<td>UK, Medium Education</td>
<td>3</td>
<td>96.5</td>
<td>78.7</td>
</tr>
<tr>
<td>UK, High Education</td>
<td>3</td>
<td>98.6</td>
<td>91.6</td>
</tr>
</tbody>
</table>

Note. Incomplete information is defined as 'don't know' or 'haven't heard of' answers on a survey questions. Data Source. CSES.
Table 2.1 demonstrates that the likelihood of missing data is related to other variables in the dataset. Hence, incomplete data are not missing completely at random, MCAR, which would permit a listwise deletion of incomplete data. Rather, one encounters data being missing at random, MAR, which produces possibly biased results if one were to use listwise deletion. This MAR finding concerns comparisons between as well as within countries. In the next chapters I will demonstrate that voters’ reasoning is not identical for groups of different education or in different countries. Therefore, it is likely that listwise deletion generates selection bias. In addition to potential biases, differences in the occurrence of incomplete data may also not be ignored from a substantive point of view: they constitute an important element of variation between different democracies and between different voters.

The discussion above leads to the conclusion that there are several reasons to incorporate rather than delete missing data in the analysis. This can only be done by replacing a missing data value by a valid one, a procedure known as substitution, replacement, or imputation. Ad hoc methods for a data imputation such as mean substitution or substitutions by the midpoint of valid values are arbitrary and more importantly do not solve the problems related to data missing at random (King et al. 2001). Data augmentation techniques and related methods have been proven to replicate unobserved data most effectively (Rubin 1987; Schafer 1997; King et al. 2001). These methods aim to retrieve unobserved data using information in the observed cases. They try to arrive at a reasonable guess for missing scores. Such procedures are often regarded as the proper tools for the inclusion of missing data in statistical analysis (Allison 2002).

Mean substitution of incomplete data approximately yields the same (possibly biased) estimates as a procedure of listwise deletion. For a proof see Appendix 3.
Yet, existing imputation techniques rest on the assumption that behind the missing data true, but unobserved scores exist. Incomplete data in this view are a problem of measurement or of observation. In many cases this is a plausible assumption. If, for example, respondents state not to know their age or their income, one can safely assume that they nevertheless have a specific age or income. This is even so when respondents have not even been asked to state their age or income. The approximation of missing data by some kind of multivariate ‘guess’ is therefore a sensible strategy. For that reason, missing data on variables of the socio demographic background of respondents, such as their age, their education, their residence, their church attendance, and their union membership are replaced by estimates of the true scores using an EM (Expectation Maximization) procedure.45

But measurement problems are not the only cause for incomplete data. They may also originate from insufficient information by respondents. If respondents are asked to state their opinion on a stimulus, this requires from them certain information, namely, that they understand the stimulus and that they have an opinion on it. A ‘don’t know’ answer on an opinion question may be a true answer, which, however, cannot be located on the response scale. In this case existing imputation techniques lack justification and are therefore not applicable. It does not make sense to retrieve a score that does not exist. Since there is no existing technique to handle

45 The EM algorithm generates maximum likelihood estimates for missing data in the means and covariance matrix in two steps. The first E step (expectation) imputes missing data with a linear prediction that results from regressing all variables on each other (starting values are mean scores). The second M step (maximization) uses the imputed data to generate new predicted scores for the means and covariance matrix. Additional to the linear prediction (which tends to underestimate the true variance) the method incorporates the residual variance of the prediction in the imputation procedure. These two steps are repeated until the estimation converges. See Appendix 3 for variables included in the EM imputation of the variables ‘age’, ‘education’, ‘church attendance’, ‘rural and urban residence’, and ‘union membership’.
this kind of missing data, I propose a new method of imputation. This method of ‘multiple white-noise imputation’ is a solution to the problem of incomplete data when these are caused by insufficient information. This method will be applied in this study to impute incomplete data on perceived leader sympathies and on self-placements and party placements on the left-right scale. If respondents state not to know where a party can be placed on a left-right continuum, this answer is not treated as a measurement problem but as a lack of sufficient information regarding the stimulus of the survey question. The same is true for the evaluation of political leaders. If respondents state not to know a certain politician, it is regarded the true answer since it is unlikely that all voters are familiar with the entire political elite of a system.

Readers interested in the method of multiple white noise imputation, its theoretical justification, formal basis, numerical properties, and empirical application are referred to Appendix 1 of this book. For readers not interested in the detailed description of the procedure it may be noted that the method basically substitutes missing data by random values in the metric of the underlying scale. These random values keep the empirical distribution in the valid data intact. The main consequence of such a white-noise imputation is that estimated effect parameters based on purely random-imputed data equal zero. Effect parameters in the whole sample are therefore reduced in magnitude according to the proportion of incomplete data. To give an example, suppose that the proportion of incomplete data on the leader sympathy scale is higher in country A (60% valid response) than in country B (80% valid response). Suppose furthermore that a regression model explaining vote choice by leader evaluation based on listwise deleted missing data would produce in both countries an effect parameter of $\beta_{AB} = 0.5$. Now, instead of deleting missing data, all incomplete information on the sympathy rating in both countries is randomly substituted. The same regression model explaining vote choice by leader sympathy
will now produce an effect parameter that equals zero for those respondents who do not know respective politicians. It is therefore assumed post hoc that respondents did not take sympathy towards a political leader into account in their vote decision if they never heard of this leader. Since this occurs more often in country A than in country B, effect parameters will be 'dragged' closer to zero in A (approximately $\beta_A = 0.5 \times 0.6 = 0.3$) than in B (approximately $\beta_B = 0.5 \times 0.8 = 0.4$). Apart from its theoretical justification, this method has the advantage over existing data augmentation methods that it avoids artificial inflation of effect parameters and of fit measures (such as $R^2$).