

The role of ideas

- Hall, P. A. (1989) *The Political Power of Economic Ideas* (Princeton, NJ: Princeton University Press).
- Putnam, R. D. (1993) *Making Democracy Work: Civic Traditions in Modern Italy* (Princeton, NJ: Princeton University Press).

Individual theories

- Greenstein, F. I. (1987) *Personality and Politics* (Princeton, NJ: Princeton University Press).
- Helms, L. (2013) *Oxford Handbook of Political Leadership* (Oxford: Oxford University Press).



Web links

- <http://www.community.apsanet.org/comparativepolitics/ComparativePolitics/Home/>
Website of the Comparative Politics Section, American Political Science Association.
- <http://www.asu.edu/clas/polisci/cqrm>
Website of the Arizona State University Institute for Qualitative Methods.

The role of the international environment

- Cowles, M. G., and Caporaso, J. A. (2002) *Europeanization and Domestic Change* (Ithaca, NY: Cornell University Press).
- Keohane, R. O., and Milner, H. (1997) *Internationalization and Domestic Politics* (Cambridge: Cambridge University Press).

Governance

- Peters, B. G., and J. Pierre (2016) *Governance and Comparative Politics* (Cambridge: Cambridge University Press).
- Pierre, J. (2003) *Debating Governance* (Oxford: Oxford University Press).

- <http://www.upslinks.net>
Ultimate Political Science Links Page.
- <http://www.politicalresources.net>
Political Resources on the Net.



online
resource
centre

For additional material and resources, please visit the Online Resource Centre at:
www.oxfordtextbooks.co.uk/orc/caramani4e/



Endnotes

¹ This classification of research types comes from Arend Lijphart's seminal article (Lijphart 1971).

² A great deal of political science theory has been developed in reference to the US, given the size and importance of the political science profession there. However, a good deal of that theory does not appear relevant beyond the boundaries of the US (in some cases not within those boundaries either).

³ This is something of an oversimplification of the assumptions of rational choice approaches, but the central point here is not the subtlety of some approaches but rather the reliance on individual-level explanations. For a more extensive critique of the assumptions see Box 13.5 in Chapter 13.

⁴ That is, if we could reject more theories and models then we could focus on the more useful ones. As it is, we are overstocked with positive findings and theories that have credible support.

⁵ The classic example of a study that uses triangulation explicitly is Allison (1971). However, this book uses multiple theories but it does not verify the results through multiple research methods.

⁶ See, for example, Adcock and Collier (2001), who stress the need for common standards of validity for all varieties of measurement, as well as the interaction of those forms of measurement.

⁷ Lijphart (1999) has provided a slightly different conceptualization by distinguishing between majoritarian and consensual political systems (see Chapter 5 on 'Democracies'). Some parliamentary systems, such as the Westminster system, are majoritarian, designed to produce strong majority governments that alternate in office. Others, such as in the Scandinavian countries, may have alternation in office, but the need to create coalitions and an underlying consensus on many policy issues results in less alternation in policy.

⁸ These shifts are to some extent a function of changes in political culture, especially the movement towards 'post-industrial politics' (Inglehart 1990).

⁹ This is more true for American than for European political science. Discourse theory and the use of rhetorical forms of analysis have been of much greater relevance in Europe than they have in North America, and qualitative methodologies remain more at the centre of European political analysis.

CHAPTER 3

Comparative research methods

Hans Keman and Paul Pennings

Chapter contents

Introduction	50
The role of variables in linking theory to evidence	50
Comparing cases and case selection	53
The logic of comparison: relating cases to variables	56
The use of Methods of Agreement and Difference in comparative analysis	57
Constraints and limitations of the comparative method	58
Conclusion	61

Reader's guide

In this chapter the 'art of comparing' is explored by demonstrating how to relate a theoretically guided research question to a properly founded research answer by developing an adequate **research design**. First, the role of **variables** in comparative research will be highlighted. Second, the meaning of 'cases' and their selection will be discussed. These are important steps in any comparative research design. Third, the focus will turn to the 'core' of the **comparative method**: the use of the logic of comparative inquiry to analyse the relationships between variables—representing theory—and the information contained in the cases—the data. Finally, some problems common to the use of comparative methods will be discussed.

Introduction

As the Introduction to this volume stresses, both its substance and its method characterize comparative politics. The method is the 'toolkit' of what, when, and how to compare political systems. In this chapter the focus is on research methods used in comparative political science: what rules and standards should we adopt to develop a comparative research design? A research design is a crucial step for developing and testing theories and for the verification of rival theories. Hence, as Peters emphasizes, '[t]he only thing that should be universal in studying comparative politics ... is a conscious attention to explanation and research design' (Peters 1998: 26). Theory development and research design are closely interlinked in comparative politics.

Contrary to everyday practice, where most people are often *implicitly* comparing situations, in comparative politics the issue of what and how to observe reality is *explicitly* part of the comparative method. Dogan and Pelassy (1990: 3), for example, remark '[t]o compare is a common way of thinking. Nothing is more natural than to consider people, ideas, or institutions in relation to other people, ideas and institutions. We gain knowledge through reference.' Yet, the evolution of comparative politics has moved on from implicit comparisons in pre-modern times to explicit ways of comparing political systems and related processes. The major modern development in comparative political science is on linking theory to evidence by means of comparative methods. The particular method to be used depends on the research question (RQ) asked and the research answer (RA) to be given (see also Box 3.1). The actual method chosen is what we label research design (RD), and this is what this chapter is about.

A theory, in its simplest form, is a meaningful statement about the relationship between two real-world phenomena: X, the independent variable, and Y, the dependent variable. According to theory, it is expected that change in one variable will be related to change in the other. The conceptual and explanatory understanding of such a relationship is the point of departure for conducting research by comparing empirical evidence across systems (see also Brady and Collier 2004: 309; Burnham *et al.* 2004: 57). In more formal terms a theory posits the *dependent variable* in the analysis—what is to be explained? Additionally, the researcher wishes to know: what are the most likely 'causes' of the phenomenon under investigation? Again, in formal terms: which *independent variables*, or explanatory factors, can account for the variation of the dependent variable across different systems (e.g. countries) or features of political systems (e.g. parties)? The answer to this question rests heavily on the development of a 'correct' research design. Comparative methods can be considered, therefore, as a 'bridge' between the research question asked and the research answer proposed. This is what we label the 'triad' $RQ \rightarrow RD \rightarrow RA$.



BOX 3.1 ZOOM-IN The triad $RQ \rightarrow RD \rightarrow RA$

The point of departure is that all research questions are theory guided. The *theoretical* guidance is expressed in relating research questions (RQs) to research answers (RAs) in the shape of *logical* relationships between a dependent variable (Y: what is to be explained) and the independent variables (X: the most likely causes, i.e. factors serving as an explanation). The 'bridge' between RQ and RA is called a research design (RD). Therefore, the comparative method is a means *to an end*: to make choices as to which of the potentially vast mass of relevant empirical data (the evidence) and possible causes (X) explaining variations in Y are valid and reliable in arriving at a research answer.

Developing a research design in comparative politics requires careful elaboration. First, the research design should enable the researcher to *answer the question* under examination. Second, the given answer(s) ought to meet the '*standards*' set in the *social sciences*: are the results valid (authoritative), reliable (irrefutable), and generalizable (postulated) knowledge (Sartori 1994)? Third, are the research design and the methods used indeed *suitable for the research goals* set? This chapter will elaborate these issues and attempt to guide the student towards linking research questions to research answers.

KEY POINTS

- The proper use and correct application of methods is essential in comparative politics.
- A correct application implies that the comparative method meets the 'standards' set in terms of validity, reliability, and its use in a wider sense, i.e. generalizability.
- The relationship between variables and cases in comparative research is crucial in order to reach empirically founded conclusions that will further knowledge in political science.

The role of variables in linking theory to evidence

Since the 1960s the comparative approach in political science has been considered highly relevant to theory development (see also the Introduction to this volume). Therefore, a research question should always either be guided by theory or itself constitute a potential answer to an existing theoretical argument. The comparative method is about observing and comparing carefully selected information (across space or time, or both) on the

basis of a meaningful, if not causal, relationship between variables. A variable is a concept that can be systematically observed (and measured) in various situations (such as in countries or over time). It allows us to understand the similarities and differences between observed phenomena. For example, we can make a difference between democracies and non-democracies or between different types of democracies (e.g. presidential, semi-presidential, and parliamentary). The extent to which the similarities and differences across systems are more or less systematic can tell us more about the plausibility of a theoretical relationship under review. For example, Linz and Stepan (1996), discussing the pros and cons of presidentialism, argued that parliamentary democracies are more enduring than presidential ones. They found that the independent variable, parliamentary vs presidential systems (a dichotomy), differed considerably in terms of political stability measured in years.

Typologies are often used as a first step in examining the theoretical association between two variables without explicitly arguing a causal relationship. The first step towards a typology is to decide what is to be classified on the basis of a research question. Take, as an example, research on the link between federalism and electoral laws. Different **electoral systems** (for example, proportional and majoritarian) are distinguished and their significance for territorial structures within a nation (for example, federal and unitary) is assessed. The major problem of this type of analysis is that miscomparing can lead to misclassification and therefore to wrongly informed conclusions. However, this can be avoided by checking that all (in this case, four) cells validly include at least a case (inclusiveness), and, further, that one case cannot be placed in more than one cell (exclusiveness). This is called an 'in-between' or hybrid case (see also Braun 2015).

In sum, the comparative method allows us to investigate hypothesized relationships among variables systematically and empirically. In contrast with the methodology of the 'exact sciences', however, the conclusions are drawn from comparisons *not* experiments. Therefore, the real world of comparative politics provides a *quasi-experimental* workplace for political scientists to examine how the complex world of politics 'turns' by demonstrating in a systematic and rigorous fashion *theoretical relationships among variables*.

An example of how the triad (recall Box 3.1) works and helps to answer a contested issue is the debate 'does politics matter?' (see Chapter 21). The dependent variable or the outcome (Y) in this example is **welfare state** development, i.e. what the researcher seeks to explain. It is called dependent because we expect that the variation in welfare state provisions across systems also depends on one or more independent variables.² As a tentative answer, the researcher comes up with a *hypothesis*. In this example the variation in welfare state development

(Y) is dependent on the relative strength of left-wing parties and trade unions in a country (X). This research answer, or hypothesis, is a conjecture about the relationship between the dependent variable and the independent variable and is supposed to explain the outcome, i.e. the development of the welfare state. In a comparative research design a theoretical relationship is elaborated to account for the differences and similarities in welfare state development (Figure 3.1).

Obviously, any type of 'X-Y' relationship in social science is an abstraction from the complexities of the real world. This is deliberate. By means of hypotheses or explanations (X) those factors are included that can account for the variation in Y. This procedure allows us to establish whether or not a meaningful relationship indeed exists, and whether or not this relationship can be qualified as 'causal' or not (i.e. it is noted as $X \rightarrow Y$).

Causality is a fraught concept in the social sciences and in strict terms is hard to establish. Yet, it is now accepted that if the variation in the dependent variable (Y—here: more or less expansion of the welfare state) is evidently and systematically related to the variation in (one of) the independent variable(s) and a theory as to why this is the case (X_1 —socio-economic; and X_2 —political indicators), then we can assume causality—at least for the cases included in the analysis. This refers to the idea of 'internal validity' (see Box 3.2).

Our ability to establish causal relationships by means of comparative research design is considered a major advantage. As we have already stated, comparative analysis is often labelled 'quasi-experimental', meaning that,



BOX 3.2 DEFINITION Internal and external validity in comparative methods

Internal validity refers to the degree to which descriptive or causal inferences from a given set of cases are indeed correct for most, if not all, the cases under inspection. *External validity* concerns the extent to which the results of the comparative research can be considered to be valid for other more or less similar cases but were not included in the research.

Both types of validity are equally important, but it should be noted that there is a trade-off (Peters 1998: 48; Pennings 2016). The more the cases included in the analysis can be considered as representative, the more 'robust' the overall result will be (*external validity*). Conversely, however, the analysis of fewer cases may well be conducive to a more coherent and solid conclusion for the set of cases that is included (*internal validity*). It should be noted that the concepts of internal and external validity are *ideal-typical* in nature: in a perfect world with complete information the standards of both internal and external validity may well be met, but in practice this is hard to achieve.

to a certain extent, we can manipulate reality, enabling the researcher to conduct *descriptive inference* (King *et al.* 1994: 34ff). This implies that the empirically founded relationship between the independent and dependent variables, based on a number of observations, allows generalization over and beyond the cases under review. Hence, the results of the analysis are considered to be relevant for all political systems where a welfare state is present or emerging (for instance, in Southern and Eastern Europe since the 1990s). In these circumstances the researcher claims that his/her results are 'externally valid' (see Box 3.2). It is obvious that this 'leap' from the empirical evidence to a more general explanation (the 'theory') is open to criticism (like the occurrence of multi-causality and conjunctural causality; see below) and drives contesting theories—in this example, politics does *not* matter but socio-economic developments do—that are developed to disprove or to enhance the theory.

Hence, socio-economic development (represented as X_1 in Figure 3.1) is considered as an important 'cause' explaining the variation in the development and level of welfare statism across OECD countries (e.g. Wilensky 1975). Political variables, such as differences between left- and right-wing parties with respect to how much welfare state is sufficient or the relative strength of these parties in government and the strength of trade unions (X_2 in Figure 3.1), were considered as less relevant (or merely coincidental) to explaining the growth of a welfare state. In other words, the research conducted appeared to prove that 'politics did *not* matter'.

Yet, a major objection concerned the finding that the *non*-political variables (X_1 in Figure 3.1) were insufficiently capable of explaining *why* cross-system differences in welfare statism *differed*, although in many instances the non-political variables X_1 were quite *similar*. The descriptive inference was not homogeneous (i.e. the assumption that a given set of variables always produces the same outcome). This criticism was supported

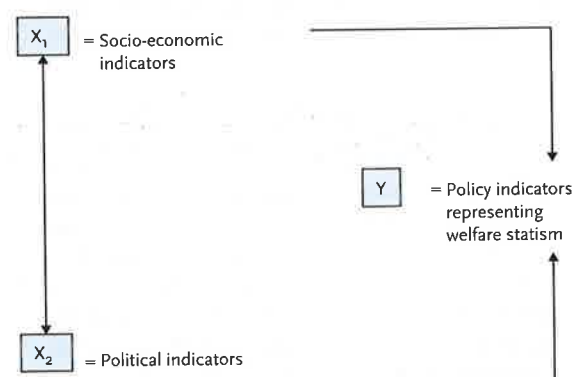


Figure 3.1 Investigating 'does politics matter?'

Source: Adapted from Pennings *et al.* (2006: 34).

by empirical observations. It appeared that levels of welfare statism tended to become more *divergent* (i.e. more different), whereas the explanatory variables (X_1) would predict otherwise: a *convergent* (i.e. more similar) development would be expected to occur. Another criticism concerned the **operationalization** of the dependent variable. By examining the various policy components of 'welfare statism' (like expenditure on social security, education, and health care), it could be demonstrated that the *design* of the welfare state showed a large cross-system variation in the distribution of what was spent on education, health care, and social security within the different countries. Table 22.2 (Chapter 22) shows that change in social expenditures differed considerably between 1980 and 2003. In short, the dependent variable 'welfare statism' could neither exclusively nor causally be linked to non-political factors alone, and nor could political factors be ignored. The comparative analysis conducted demonstrated that political variables appeared to have a considerable (and statistically significant) impact as well. Hence, the 'new' research answer became: 'yes, politics does matter'!

Note that the research design not only concerns establishing the $X_2 \rightarrow Y$ relationship, proving that political variables played their role, but is also controlling for the relative impact of 'politics' by including the original explanatory relationship $X_1 \rightarrow Y$ as a rival explanation. The message conveyed here is that research in comparative politics requires a precise and detailed elaboration of a research design—in terms of relationships examined by means of variables and developing corroborating evidence across the cases under review—connecting research questions to research answers that are conducive to causal interpretation by means of descriptive inferences.³ In the next section we turn to selecting the cases suited for **comparison**. This will direct the 'logic' of comparison implied in the research design to be used.

KEY POINTS

- Theory comes *before* method and is expressed in its simplest form as the relationship between dependent (Y) and independent (X) variables. The research method follows the research question in order to find the proper research answer.
- Research answers are (tentative) hypotheses that are interpreted by means of descriptive inference on the basis of *comparative* evidence, possibly allowing for causal interpretation.
- The research design is the toolkit to systematically link empirical evidence to theoretical relationships by means of comparative methods, *enhancing* the internal and external validity of the results.

Comparing cases and case selection

Recall that linking theory to evidence always entails the reduction of real-world complexities so as to analyse the (logical) relationship between the X and Y variables. Hence, researchers must make decisions about *what* to compare, i.e. select the cases (the carriers of relevant information), and about *how* this information can be transformed into variables. The key to the development of a proper comparative research design is to decide *which* cases are useful for comparing and *how many* can be selected (see Figure 3.2). The answers to these questions have led to many views and debates (see Brady and Collier 2004; Braun and Maggetti 2015). The thrust of this concerns the apparent trade-off between selecting many cases, but with few relevant variables available for analysis, or a few crucial (or contrasting) cases, but with many variables for use. We shall first clarify what a 'case' is.

Cross-case and within-case

The term 'case' has a general meaning in social science methodology, but in comparative methods it is used in a specific manner and is to some extent confusing (Pennings *et al.* 2006: 34ff). In comparative politics, cases denote the units of observation to be compared. Often it concerns countries, but one can also compare sub-systems (like regions) or organized entities (like governments or bureaucracies). Yet the level of measurement may be different. Take, for instance, individual voters in several countries: the country is the case compared and determines the level of *analysis*, whereas the voter is the unit of observation (within the case). Conversely, if one compares **party governments** within a country, both the case and the unit of observation—governments—are at the same level of observation. For clarity, we propose to reserve the term 'case' in comparative methods for any type of *system* included in the analysis (recall Note 1). In addition we refer to *observations* as the values (or scores) of a variable under investigation. For example, if one compares party behaviour of regional parties, then the 'case' is the regional party *within* a system. Conversely, if the welfare state is the focus of comparison, then this concerns the systems to be compared. Public expenditures on social security are the empirical value for each system (see, for instance, Table 22.4). In Box 3.3 cases and variables are discussed in terms of a data matrix (i.e. the organization of the empirical observations by case and by variable). It is important to be precise in this matter because the *number* of observations—large- or small-N—determines what type of (**quantitative** or **qualitative**) analysis is feasible in terms of descriptive inference, given the available variation across the systems, or cases, under review (Pennings *et al.* 2006: 11).

The relationship between the cases selected and the variables employed to analyse the research question is a crucial concern. As can be seen from Figure 3.2, the process of case selection is structured as a kind of scale: from one case (often including many variables) to maximizing the number of cases (often with few(er) variables). In addition, it is sometimes suggested that this choice between few or many cases is related to the type of data—quantitative or qualitative—used. This is debatable (see also Brady and Collier 2004: 246–7). For example, the study of welfare states often combines qualitative elements with statistical data (e.g. Kumlin *et al.* 2016). Or take the study on 'social capital' by Putnam (1993), who combines survey data (reporting attitudes across the population) with a historical analysis of the development of politics and society in Italy since medieval times. Hence, comparative historical analysis can be usefully combined with a cross-sectional quantitative approach.

There is an ongoing discussion on *how* to combine different types of data. This is called **triangulation**, methodological **pluralism**, or multi-method research. In its purest form it involves cross-case causal inference and within-case causal mechanism analysis and inference (Goertz 2016), but other combinations of methods are also possible. Collier and Elman (2008) distinguish between three types of multi-method research. The first type combines conventional qualitative approaches, such as case study methodologies and interviews. The second type combines quantitative and qualitative methods, for example statistical analysis and process tracing. The third type combines conventional qualitative methods with either constructivist or interpretivist approaches (for more detailed overviews, see: Giraud and Maggetti 2015; Berg-Schlosser 2012; Goertz 2016). The strength of methodological pluralism is that it helps to overcome the limitations of a single design, for example doing either interviews or statistical analysis. Multi-methods enable the researcher to both explain (using cross-case data) and interpret (using within-case data). It also allows the researcher to address a question or theoretical perspective at different levels, for example both at the individual level and at the country level. This can be useful when unexpected results arise from a prior study. In addition, multi-methods may help to generalize, to a degree, qualitative data. There are also a number of potential weaknesses or risks involved. The results may sometimes be puzzling because different types of data and designs may generate unequal evidence. As Braun and Maggetti (2015) note, there is no 'yardstick' available to judge the extent to which the different methods indeed reinforce the results of the analysis. Hence, before one decides to adopt a multi-method approach these risks should be carefully considered (Creswell 2015).

These complexities may motivate researchers to stick to one method and keep on improving it instead of combining it with other approaches. One example of

an innovative method that is used as a single method is process tracing (see Trampusch and Palier 2016). This is a qualitative method used to evaluate complex causal processes by means of historical narratives and/or within-case analysis (also Beach and Pedersen 2016). The main goal of process tracing is to make unit-level causal inferences (i.e. how a given cause affects a single unit like an international organization or a country). It focuses on a causal mechanism by examining how 'X' produces a series of conditions that come together to produce 'Y'. Process tracing is a useful tool for testing hypotheses provided that the causal mechanism under study is well theorized and not a black-box. It is particularly suited for small-n studies, for example to study why deviant cases diverge from expected trends. For example, Gheciu (2005) uses process tracing by combining interviews, participant-observation, and discourse analysis in order to demonstrate how NATO operated after the Cold War as an agent of socialization by introducing liberal-democratic norms into Central and Eastern Europe. The results refute the rationalist assumption that NATO is a military alliance which is irrelevant to processes of constructing domestic norms and institutions. The approach helps to explain how and why the national elites switched from an authoritarian to a liberal-democratic view and conduct.

Comparative historical analysis has returned to the comparative method of late. 'Comparative historical analysis aims at the explanation of substantively important outcomes by describing processes over time using systematic and contextualized comparisons' (Mahoney and Rueschemeyer 2003: 6). Thus, this type of historical analysis is meant to be explanatory and its mode of analysis is to use time (i.e. change) as the major operationalization of a variable. Processes are studied within the context of historical developments, not in isolation, and historical sequences can be employed to explain the meaning of change. For example, some students employ so-called 'critical junctures' (e.g. the World Wars or the End of the Cold War) that have transformed the relationship between state and society. Finally, there is the notion of '**path dependence**', meaning that certain political choices made in the past can explain certain policy outcomes at present (see Chapter 22). The explanation rests, then, on the idea that alternative options for choice were not open any more, or, given the time a policy exists, the 'point of no return' has been definitively passed (Pierson 2000). In short, comparative historical analysis has a lot to offer to the comparative student, either in combination with other approaches or on its own (Keman 2013).

In sum, the selection of cases and variables depends on a deliberate choice in relation to the research question and on consideration of the type of approach chosen in view of the explanatory goals set (see also Ragin 2008). Further, the set of cases and variables in use also affects the so-called *ceteris paribus* condition (all other things



BOX 3.3 DEFINITION Cases and variables in a comparative data set

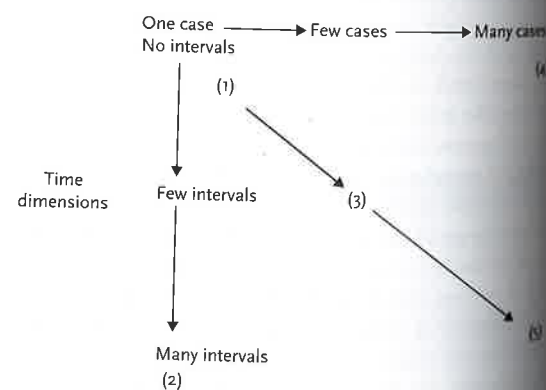
In comparative research the term 'case' is reserved for units of observation that are comparable at a certain level of measurement, be it micro (e.g. individual attitudes), meso (e.g. regional parties), or macro (e.g. national government). The information in a data set is a two-dimensional rectangular matrix: variables in columns (vertical) and cases in rows (horizontal). Each cell contains values for each variable, i.e. observations (e.g. levels of GDP, types of governments, or votes for parties), for each case (e.g. countries, regions, or parties). Likewise, variables may represent information over time (e.g. points in time, such as years, and are then cases; see Tables 22.1 and 22.2).

are considered to be constant for all cases). Therefore, case selection is a crucial step.

Case selection

Cases are the building blocks for the theoretical argument underlying the research design. The number of cases selected in the research design directs the type and format of comparison. This is illustrated in Figure 3.2.

Figure 3.2 shows that there are different options for selection depending on how many cases and how many variables are involved. *Intensive strategies* are those with many variables and few cases. An example is the analysis of the few consociational democracies that exist. These democracies are Austria, Belgium, the Netherlands, and



- (1) Case study (at one time point);
- (2) Time series (one case over time);
- (3) Closed universe (relevant cases in relevant periods);
- (4) Cross-section (all cases at one time point);
- (5) Pooled analysis (maximizing cases across time and space).

Figure 3.2 Types of research design

Source: Adapted from Pennings *et al.* (2006: 21).

Switzerland. They are characterized by major internal divisions along religious and/or linguistic lines that divide the population into more or less equally sized minorities. Despite this societal segmentation, which is often regarded as a source of conflict and instability, these countries managed to remain stable due to cooperation between the elites of the segments. *Extensive strategies* are those with few variables and many cases. An example is the analysis of welfare states, as discussed in Chapters 21 and 22. Here, many cases, if not all (e.g. all established democracies), are selected, whereas only a few variables are included. If N (number of cases included) is less than ten or fifteen, the strategy is intense. In addition, whether or not 'time' is a relevant factor needs to be taken into account (Pennings *et al.* 2006: 40–1). This is often the case, in particular when change (or a process development) is a crucial element of the research question (e.g. explaining welfare state developments). This is called *longitudinal analysis* if it is quantitatively organized or *historical analysis* if it is based on qualitative sources. Finally, if statistical analysis is used, as many cases as feasible are required to allow for tests of significance (King *et al.* 1994: 24; Burnham *et al.* 2004: 74). Five options for case selection are presented in Figure 3.2.

The single-case study

A single-case study may be part of a comparative research design. But as it stands alone it is at best implicitly comparative and its external validity is low or absent (see Landman 2003: 34–5). However, it can be used for post hoc validation to inspect whether or not the general findings hold up in a more detailed analysis or to study a *deviant case* (i.e. a case that appears to be an 'exception to the rule'), and can also be used as a *critical* or *crucial* case study (Seha and Müller-Rommel 2016). Another use of a single-case study is as a pilot for generating hypotheses, or confirming or invalidating extant theories (Lijphart 1971: 691).

Time series

Time series or longitudinal analysis can be useful in two ways: first, to compare a specific configuration within a few cases in order to inspect comparative change (which was discussed earlier in association with comparative historical analysis). An example is the analysis of new parties that were electorally more successful after the 1990s than before (see Table 12.2). Second, to analyse which factors are (or have become) relevant over time as causes. An example is the comparison of the different waves of **democratization** over time from the nineteenth century up to the present (see Chapter 12). Another use is to replicate a cross-sectional study by time-series analysis to observe differences in the outcomes (King *et al.* 1994: 223).

Closed universe

The third option in Figure 3.2 concerns the 'few' cases for comparison at different points of time, taking into account change by defining periodic intervals based on external events (or 'critical junctures'), for instance after a discrete event like war or an economic crisis. A good example is the developments during the inter-war period when in some European countries democracy gave way to dictatorship, whereas in other countries democracy was maintained (see Berg-Schlosser and de Meur 1996). A few(er) cases research design is often called a 'focused comparison' derived from the research question under review.

Cross-section

The fourth option in Figure 3.2 implies that several cases are compared simultaneously. This research design is frequently used. It is based on a selection of those cases that resemble each other more than they differ and thereby reduce variance caused by other (unmeasured) variables. It implies that the 'circumstances' of the cases under review are assumed to be constant, whereas the included variables vary. This enhances the internal validity of the analytical results. For example, if the focus is on formation of coalition governments, then it follows that we only take into account those democracies where multiparty systems exist.

Pooled analysis

The final option is disputed among comparativists. Although the number of cases can be maximized by pooling cases across time and systems (e.g. 20 rows and 20 columns taken together implies that the N of cases is 400 instead of 20), whereas in a time series data set the years (or other points in time, e.g. periods) are exclusively the cases and in a cross-sectional data set the cases are exclusively the political systems to be compared, the pitfall of pooling is that the impact of time is held constant across all cases (or, at least, changes across cases do not vary (Kittel 1999)). A possible fallacy is that comparative analysis suffers from the fact that the cases are 'too much' alike and therefore there are no meaningful differences from which to draw conclusions (see King *et al.* 1994: 202–3). To avoid this, one can include control cases (e.g. analysing EU members in conjunction with non-member states). Another remedy would be to include a 'rival' explanatory variable (as we showed in Figure 3.1: politics vs economics). Pooled analysis is mainly used in sophisticated quantitative approaches and it requires skills in statistical methods at a more advanced level.

All in all, the message is that the range of choice with regard to case selection is larger than is often thought. First, the options available (in Figure 3.2, options 3 and 5 are often used in combination). Second, the options in

developing a research design can be used sequentially. For instance, one could follow up a cross-sectional analysis (option 4) with a critical or a crucial case study (option 1) as an in-depth elaboration of the comparative findings. However, it should be noted that the options for choice as depicted in Figure 3.2 are not completely free. For instance, if industrialization is seen as a *process*, it must be investigated *over time* in order to answer the research question of whether or not this results in a change towards welfare statism.

The main point of this section is not only that case selection is important for how many cases can or should be included in the analysis, but also that the choice is neither (completely) free nor (completely) determined. First, the choice of cases depends on the theoretical relationship under review ($X-Y$), which defines what type of political system can be selected. Obviously, if one researches the working of democracies, non-democratic systems cannot be included. Second, the type of empirical data available can limit the choice of cases. Third, one should bear in mind that the relationship between cases and variables also determines what type of technique can be used: statistical analysis can only be used if the N is sufficient for tests of significance, whereas a small- N allows for including contextual information or multi-causal analysis (e.g. by means of Fuzzy-Set logic; see Section: The use of Methods of Agreement and Difference in comparative analysis). Finally, if the research question involves a specific phenomenon—like Federalism or Semi-Presidentialism—the N of cases is obviously limited by definition.

Relating the cases and associated information (i.e. data) is the next step in performing comparative analysis. This stage of the research design concerns establishing and assessing the relationship between the evidence (data) collected across the selected cases for the independent and dependent variables in search of a (causal) relationship.

KEY POINTS

- In comparative research case selection is a central concern for the research design. It is important to keep in mind that the level of inquiry as derived from the research question is related to the type of system under investigation. The comparative variation across systems is empirically observed by means of indicators representing the variables that are in use.
- The balance between (many or fewer) cases and variables is an important option for case selection and the organization of the data set (see Box 3.3).
- Figure 3.2 shows the options for case selection. The selection of cases depends on the research question and the hypotheses that direct the research design. The choice of cases can be limited due to lack of data and therefore can impair the chosen research design.

The logic of comparison: relating cases to variables

In comparative methods there are two well-known research designs that employ a different type of logic: the **Most Different Systems Design** (MDSD) and the **Most Similar Systems Design** (MSSD). These designs relate directly to the type and number of cases under review and to the selection of variables by the researcher in view of the research question and the related (hypothetical) answers. Both have been developed following John Stuart Mill's dictum: *maximize experimental variance—minimize error variance—control extraneous variance* (Peters 1998: 30). In fact, they are 'ideal types'—something to strive for.

Experimental variance

This points to the observed differences or changes in the dependent variable (Y) of the research question, which is supposed to be a function of the independent variable (X). Figure 3.1 is an example of the basic structure of modelling the relationship between a research question and a research answer. The question at stake was whether or not 'politics matters'. A crucial requirement for answering this question and attempting to settle this debate is that the dependent variable (Y = 'welfare statism') indeed varies across cases or over time (or both). Where there is *no* experimental variance we cannot tell whether or not the independent variables make a difference or not. Hence, the research design would lead to insignificant results because we cannot tell whether the effect-producing variables (X) account for the observed outcomes (in Y).

Error variance

This is the occurrence of random effects of unmeasured variables. These effects are almost impossible to avoid in the social sciences, given its quasi-experimental nature which always implies a reduction of 'real-life' circumstances. Even in a single-case study or comparing a few cases, a 'thick' **descriptive analysis** cannot provide full information. However, error variance should be minimized as much as is feasible (in statistical terms, the error term in the equation is then constant or close to zero). One way to minimize error variance would be to increase the number of cases. However, this is not always feasible, as mentioned earlier in the discussion on case selection (see also 'Conceptual stretching').

Extraneous variance

The final requirement in Mill's dictum is controlling for extraneous variance. If there is no control for other possible influences, the hypothetical relation $X-Y$ may in part be produced by another (unknown) cause. One

example of an unknown cause, also called a confounding variable, is the favourable impact of consensus democracy on performance, which could also be caused by economic growth. If the latter is excluded, the analysis may suffer from underspecification, which leads to erroneous results. This is often due to omitted variables and can lead to a *spurious relationship* (a third variable affects both the independent and dependent variables under investigation). There is no 'best' remedy to prevent extraneous variance exercising an influence other than by having formulated a fully specified theory or statistical significance tests and control variables.⁴ One approach is to apply the principles of the Methods of Agreement and Difference. Using these methods we are in a position to draw causal conclusions by means of logically ordering the *differences* and *similarities* between the dependent and independent variables, based on the empirical evidence available.

The use of Methods of Agreement and Difference in comparative analysis

The logic of comparative enquiry is obviously meant to assess the relationship between the independent variables and the dependent variable in light of the number of cases (many, few, or one) selected for comparison. As we have already seen, case selection has implications for the use of the logics of comparison. Two logics are distinguished:

- Method of Difference;
- Method of Agreement.

The Methods of Difference and Agreement originate from John Stuart Mill's *A System of Logic* (1843). The basic idea is that comparing cases is used to interpret commonalities and differences between cases and variables. Hence, these 'logics' refer to the type of descriptive inference used to examine whether or not there is indeed a causal relationship between X and Y . This assessment is inferred from the empirical evidence (data) collected.

The **Method of Difference** focuses on comparing cases that *differ* with respect to either the dependent variable (Y) or the independent variable (X) but do *not* differ across comparable cases with respect to other variables (the *ceteris paribus* clause). Hence, covariation between the dependent and independent variables is considered crucial under the assumption that the context remains constant. This is the MSSD: locating variables, in particular the dependent variable, that differ across similar systems and accounting for the observed outcomes. An example is the debate on the role of 'politics' as regards the welfare state. We look at the political differences between systems that are similar in terms of their institutional

design and examine the extent to which party differences (X) match differences in welfare state provisions (Y). The stronger the match between, for instance, the strength of the left in parliament and government and the 'generosity' of welfare entitlements, on the one hand, and its absence in cases where parties of the right are dominant, the more likely it appears that 'politics matters'. Alternatively, the **Method of Agreement** consists of comparing cases (systems) in order to detect those relationships between X and Y that remain *similar* notwithstanding the differences in other features of the cases compared. Hence, other variables may be different across the cases except for those relationships that are considered to be causal (or effect-productive). This is the so-called MDSD. An example is Luebbert's analysis investigating the possible causes of regime types during the inter-war period (1919–39). He distinguishes three regime types: liberalism, social democracy, and fascism (Luebbert 1991). The explanatory variable (X) is 'class cooperation' (between the middle class, farmers, and the working class) and regime type is the dependent variable (Y). Luebbert finds that only specific patterns of class cooperation consistently match the same regime type across twelve European countries. Most other variables considered (as possible causes) in the comparative analysis do not match the outcome (regime type) in the same way.

When applying the (quasi-)experimental method to causal problems one should keep in mind that in social science the degree of control of the independent variables is always limited (Kellstedt and Whitten 2013: 70–88). We cannot control exposure to them because we cannot 'assign' a country a type of regime or a level of social expenditures. In addition, whereas the internal validity of experiments is often high, the external validity (i.e. generalizability) is low. Due to these drawbacks, genuine experiments are rare in comparative politics. Instead, many comparative studies are observational by taking reality as it is without randomly assigning units of analysis to treatment groups. The major disadvantage of this strategy is that, since we do not control for all possible causes of Y , there is always a chance that some third (confounding) variable is causing Y . This potential problem weakens the internal validity of observational studies. The only way to cope with this problem is to try to identify alternative causes of Y and to present any causal claim in a tentative way. The MSSD- and MDSD-designs are in between the experimental and observational methods and run the risk that they combine the weaknesses of both methods. As a consequence, if one adopts a MSSD- or MDSD-design one should be alert on both the internal and the external validity.

Recently, an alternative approach—*Qualitative Comparative Analysis* (QCA)—has been developed which attempts to cater for 'multiple causalities' (one of the limits of Mill's logic of comparison). This type of analysis allows for the handling of many variables in

combination with a relatively high number of cases simultaneously (recall option 3 in Figure 3.2). Ragin (2008) claims that this type of research design is a way of circumventing the trade-off between many cases/few variables vs few cases/many variables. The logic of comparison employed is based on Boolean algebra in which qualitative and quantitative information is ordered in terms of *necessary and sufficient* conditions as regards the relationship under investigation. This approach also appears to be well suited to focusing on the variation of comparative variables within cases. Instead of aiming at detecting one (at best) effect-producing circumstance (X) by means of a variable-oriented approach, the *homogeneity* of comparable cases directs the process of descriptive inference. An example is the search for the conditions under which economic development is more or less promoted by public policy (Vis *et al.* 2007). Instead of searching for the strongest or a single relationship, the researcher attempts to find out which *combination* of factors is connected with cases in view of their economic development. This procedure and concomitant logic of comparison has been developed into a 'fuzzy-set logic'.

The many applications and cross-validations of QCA have led to much support, but more recently also to criticisms. Most critics agree that QCA is useful if the goal is to test deterministic hypotheses under the assumption of error-free measures of the employed variables. However, critics argue that most theories are not sufficiently advanced to allow for deterministic hypotheses because causality is a complex phenomenon. In addition, the assumption of error-free measures is an illusion because often scientists have to use proxies or indicators that do not fully represent the original meaning (Hug 2013). Furthermore, as QCA is often used to analyse a relatively small number of cases, it has a problematic ratio of cases-to-variables which may negatively affect the stability of findings (Krogslund and Michel 2014). Several researchers have found that the results are susceptible to minor model specification changes. They argue that the identification of sufficient causal conditions by QCA strongly depends on the values of the key parameters selected by the researcher. They also show that QCA results are subject to confirmation bias because they tend towards finding complex connections between variables even if they are randomly generated. Others have argued that there is no way to assess the probability that the causal patterns are the result of chance because QCA methods are not designed as statistical techniques. The implication is that even very strong QCA results may plausibly be the result of chance (Braumoeller 2015). If we overlook these criticisms, we can conclude that the main critique of QCA is that it is not robust and often yields erroneous causal connections. On the one hand, these critiques should be taken seriously because they highlight weak spots that should not be overlooked by any researcher that uses QCA. On the other hand, doing causal analysis on a

small number of cases by means of QCA should be regarded as a qualitative approach in which interpretation is more important than quantification. If it is purely used as a statistical technique it may well lead to misleading outcomes because the researcher fully relies on the software to generate causal connections. Applicants of QCA should never do that and instead account for all the relevant steps taken during the research and cross-validate the results with alternative sources and interpretations.

KEY POINTS

- The point of departure is a hypothesis concerning the relationship between two or more variables (X–Y) whose empirical validity is to be verified by means of real-world data across a number of cases.
- The Method of Agreement uses MDSD to allow descriptive causal inference. Conversely, the Method of Difference derives its explanatory capacity from MSSD. The shared goal is to eliminate those variables that exemplify no systematic association between X and Y across the cases selected.
- An alternative logic of comparison has recently been developed: QCA/fuzzy-set logic. This approach allows scrutiny of multiple causality across various cases and variables.

Constraints and limitations of the comparative method

Although the comparative approach in political science is considered to be advantageous in linking theory to evidence, enhancing it as a 'scientific' discipline, there are a number of constraints that limit its possibilities and can impair its usefulness. In this section we discuss some of these and offer possible solutions. While it is important to be aware of them, it is often difficult for students to find appropriate solutions. Hence, it is wise to seek advice from an experienced researcher or lecturer.

One major concern is that we often have too many different theories that fit the same data. This means that collecting *valid and reliable* data for the cases we have selected to test theoretical relations can turn out to be a daunting task. If this problem is insufficiently solved, it will undermine the quality of the results. More often than not we are forced to stretch our *concepts* so that they can *travel* to other contexts and increase the number of observations across more cases. However, this may create too large a distance between the stretched concept and the original theoretical concept (Sartori 1970, 1994). If this problem occurs (and it often does), it may well affect the internal *and* external validity of the results (see Box 3.2). In conjunction with this hazard, it has been noted

that reliability problems may arise as a result of including (functional) *equivalents* that are used to widen the case selection (and thus increase the number of cases). An example is the concept of the 'federal state'. What defines a federation, and how can such a concept be transformed into measurable entities? Thus, the problem is to what extent a concept transformed into an empirical indicator has the same meaning across different settings or cultures (Van Deth 1998). Finally, some caveats need to be taken into account when interpreting the comparative data available: (1) Galton's problem, (2) individual and ecological fallacies, and (3) **over-determination**.

Conceptual stretching

Conceptual stretching is the distortion that occurs when a concept developed for one set of cases is extended to additional cases to which the features of the concept do not apply in the same manner. Sartori (1970) illustrated this problem by means of the 'ladder of generality'. Enhancing a wider use of a theoretical concept by *extension* (of its initial meaning, i.e. moving from A to B in Figure 3.3) involves a loss of intension (where the observations reflect the original features of the concept, i.e. remain close to B). *Intension* will obviously reduce the applicability of a concept in comparative research across more cases, but it enhances the internal validity of the cases compared. Extension will have the opposite effect, and the question is then whether or not the wider use (i.e. in a higher number of cases to be compared) impairs the claim for external validity of the analytical results. In Figure 3.3 this choice is visualized: the more the meaning of a concept moves due to the process of operationalization from position A to B, the

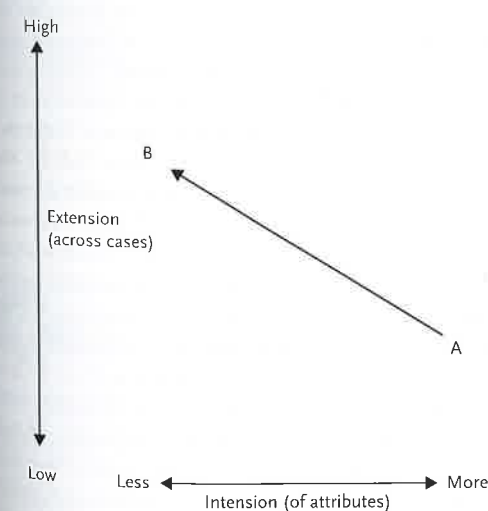


Figure 3.3 Sartori's ladder of generality
Source: Pennings *et al.* (2006: 49).

less equivalent the information collected for each case may be and therefore the validity of the results is negatively affected.

The choice to be made and the matter of dispute among comparativists is how broadly or extensively (i.e. from A to B) we can define and measure variables without a serious loss of meaning. There are different opinions on the degree of flexibility that is allowed when 'stretching' concepts to make variables 'travel' across (more) cases. Sartori remarks that *overstretching* is dangerous, and not all concepts can travel all over the world and through all time, like a political party, whereas constitutional design may (Sartori 1994). However, attempts have been made to develop methods to cope with the problem of overstretching and travelling (Braun and Maggetti 2015).

Family resemblance

Some comparativists have suggested another solution by means of 'family resemblance' (Collier and Mahon 1993: 846–8). In its simplest fashion this method extends the initial concept by adding features which share some of the attributes of the original concept. How far this type of extension can go depends on what the research question is. For example, if we are investigating the behaviour of political parties and define these as any actor that is vote-seeking (= A), office-seeking (= B), and policy-seeking (= C), then the concept of a party can be used in a wider sense. Instead of requiring that all three characteristics are present *simultaneously*, we allow the inclusion of parties that have fewer in common. Examples are electoral democracies, where people can vote (= A), but parties are not allowed to govern (= B), let alone to make policy decisions (= C). This latter type of party often occurs in emerging democracies (see Merkel 2014).

Radial categories

The second option of going up the 'ladder of generality' is the use of *radial* categories. Here the underlying idea is that each step of extension, thus including new comparable cases, is defined by a *hierarchy* of attributes belonging to the initial concept. In Figure 3.4 this is made visible by defining A as the essential attribute, whereas B and C are considered as secondary. Figure 3.4 demonstrates these two strategies for extension through which the number of cases is to be increased. Family resemblance requires a degree of commonality and this produces three cases in comparison, instead of one under the initial categorization, by sharing two out of the three defining features (AC, AB, BC). The radial method requires that the primary attribute (A) is always included. In Figure 3.4, this means two cases instead of the original single case (A + B and A + C). It is advisable to develop a *typology* first to see empirically how this transformation of a concept may work out.

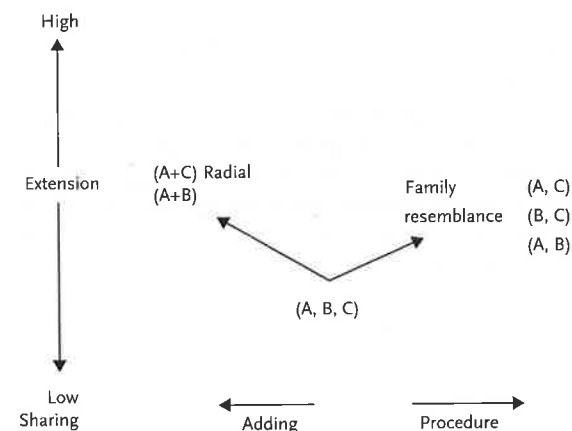


Figure 3.4 Radial categorization and family resemblance

Source: Pennings *et al.* (2006: 50).

Equivalence

A related problem of transforming concepts into empirically based indicators concerns the question of whether or not the meaning of a concept stays constant across time and space. Landman (2003: 43–6) argues that this problem is less a matter of whether or not a concept is measured with *identical* results (which is a matter of reliability regarding the measurement). Whatever solution is chosen, at the end of the day it is up to the researcher to convince us whether or not the degree of **equivalence** between measured phenomena is acceptable.

Interpreting results

Galton's problem, ecological and individual fallacies, and over-determination are all hazards that are related to the *interpretation* of the results of the comparative analysis.

Galton's problem

Galton's problem refers to the situation where the observed differences and similarities may well be caused by exogenous factors that are common to all the cases selected for comparison, such as comparing fiscal policy-making across states in Europe after the introduction of the EMU requirements, or the choice for a Westminster-style of parliamentary governance in former British colonies (Burnham *et al.* 2004: 74). Another example is the process of '**globalization**' (see Kübler 2016). Obviously, diffusion will affect the process of descriptive inference because the explanation is corrupted by a common cause that is of similar influence in each case, but is difficult to single out (Lijphart 1975: 17). A possible way to detect such a common cause is either by triangulation or by applying comparative historical analysis.

Individual and ecological fallacies

These fallacies are likewise problematic vis-à-vis the causal interpretation of evidence. An ecological fallacy occurs when data measured on an aggregated level (e.g. at the country level) are used to make inferences about individual- or group-level behaviour. Conversely, individual fallacy is the result of interpreting data measured at the individual or group level as if they represent the 'whole' (e.g. using electoral surveys for party behaviour or national attitudes; see also the Introduction to this volume). This type of fallacy occurs regularly in comparative politics and shows the need for developing a proper level of measurement.

Over-determination and selection biases

Over-determination and **selection biases** are risks that emanate from case selection. In particular, when MSSD is used, the chances are high that the dependent variable is over-determined by another difference that is not actually catered for in the research design (Przeworski and Teune 1970: 34). Conversely, if the cases included in the analysis are fairly homogeneous there is a chance that a selection bias will go unnoticed. As King *et al.* (1994: 141–2) note, if the similarities among the cases affect the degree of comparative variation of the independent and dependent variables we cannot draw valid conclusions.

In addition to these constraints and limitations, comparative methods have been criticized for being a-theoretical, empiricist, and solely country-oriented (Newton and van Deth 2016). We may argue that the comparative methods do not offer solutions to all and sundry types of research problems, but help to formulate them and find ways to arrive at plausible answers. Comparison is required to make research insightful and scientifically relevant. The alternative is a single case. However, the comparative approach is developing and learning from these criticisms (Seha and Müller-Rommel 2016). Several new approaches that are related to the 'spatial turn' in comparative politics can help to advance new insights.

One example is multi-level governance (MLG) that criticizes methodological nationalism and other forms of centrism (like Europe or democracies) that often characterize comparative politics. The MLG-approach analyses how decision-making competences are shared by policy-relevant actors at different levels of government (i.e. supranational, national, and subnational levels) (Kübler 2015: 63–80). It focuses on the dynamics of cross-level interactions between these actors in one or more policy areas. These interactions often involve forms of **soft power** such as information exchange and persuasion. A second example examines interdependencies, policy diffusion, and policy transfer (Obinger *et al.* 2013; Jahn and Stephan 2015). This type of research argues that not only developments within nation-states matter, but also

relations between nation-states and inter- and **supra-national** forms of cooperation and bargaining. Such a relational approach means that policy outcomes are accounted for by multiple interdependencies in conjunction with domestic factors. Policy diffusion and transfer are explained by causal mechanisms such as learning and competition. A related approach is Social Network Analysis (SNA). A network consists of nodes (the agents) and edges (the links or relations between the nodes). The relationships between them represent ways in which actors influence and control each other. Network analysis seeks to describe these (evolving) relationships in order to learn how power structures constrain behaviour and change (for a recent example, see Paterson *et al.* 2014).

Although these new developments are promising, we have to take into account an important caveat. The existing applications are often rather descriptive and lack explanatory power and generalizability. Hence, they should be regarded as helpful additions to conventional methodologies and not as replacements.

To conclude this section, the constraints and limitations of the comparative method need permanent attention. However, it would be wrong to conclude that—given the complexities and criticisms discussed in this chapter—the comparative approach to politics is therefore misdirected or fallacious. If we accept the fact that most political science is comparative, even if not explicitly so, then it is one of the *strengths* of the comparative method that both the advantages and disadvantages are recognized and discussed in terms of its methodology.

KEY POINTS

- There are many hazards and pitfalls in comparative methods that ought to be taken into account to link theory and evidence in a plausible fashion.
- Conceptual travelling is a sensitive instrument to widen the case selection as long as overstretching is avoided.

? Questions

Knowledge based

1. Why is the 'art of comparing' not only useful for *explicit* comparison, but also an *implicit* part of the toolkit of any political scientist?
2. Can you explain why the comparative method is often called 'quasi-experimental'? Can you argue *why* this would justify the use of statistics in comparative politics?

The use of 'Radial categories' and 'Family resemblance' to extend the number of cases can remedy this.

- Interpretation problems are often due to biases like Galton's problem and over-determination, as well as to individual and ecological fallacies. Avoiding these problems as far as possible reduces the probability of drawing invalid conclusions.

Conclusion

Some time ago, Gabriel Almond lamented the lack of progress in political science at large (Almond 1990). His main complaint concerned the lack of constructive collaboration among the practitioners. However, he made an exception as regards the field of comparative politics:

Mainstream comparative studies, rather than being in a crisis, are richly and variedly productive ... In the four decades after World War II, the level of rigor has been significantly increased in quantitative, analytical, and historical-sociological work. (Almond 1990: 253)

Much of the credit should go to those involved in the further development of the methodology of comparative politics by means of debates on difficult issues in the comparative method. However, new developments, like QCA, process tracing, and the re-emergence of comparative historical analysis, take place and are welcome. This chapter has attempted to demonstrate this. Throughout we have maintained that comparative politics is a (sub) discipline of political science where theory development is explicitly linked to empirical evidence by means of a rigorous application of the comparative method. Even if not all the problems—and they do exist—can be solved at this stage, we hold the view that the comparative method is the best way to go forward to further comparative politics within political science at large.

3. What exactly is the difference between internal and external validity? Why is this distinction important? Can you give an example of each type of validity?
4. If you examine the debate on 'does politics matter?' can you describe the research design used? Are you able to develop an alternative one—in terms of variables and cases—to test the same issue of this debate?

5. What is a comparative variable and what is the relation between a concept and an indicator?

Critical thinking

1. What is a case? Can you elaborate what the case is in terms of unit of observation and level of measurement if it concerns a comparative investigation of party government?
2. There are different options as regards the type and number of cases needed to develop a research design. Can you think of a research question that would justify the choice of a single-case study where 'time' is relevant and 'inter'-system references are necessary?



Further reading

General literature on methods in political science

Brady, H. D., and Collier, D. (2010) *Rethinking Social Inquiry: Diverse Tools, Shared Standards* (2nd edn) (Lanham, MD: Rowman & Littlefield). This edited volume discusses a wide variety of methodological concerns that are relevant for comparative methods.

King, G., Keohane, R. D., and Verba, S. (1994) *Designing Social Inquiry* (Princeton, NJ: Princeton University Press). This is a contemporary classic in social science methods, written by three political scientists. It is an introduction and uses much material taken from comparative politics.

Kellstedt, P. M., and Whitten, G. D. (2013) *The Fundamentals of Political Science Research* (Cambridge: Cambridge University Press). This book presents an accessible approach to research design and empirical analyses in which researchers can develop and test causal theories.

Keman, H., and Woldendorp, J. (2016) *Handbook of Research Methods and Applications in Political Science* (Cheltenham: Edgar Elgar Publishing). Representing an up-to-date state of affairs for approaches and methodologies in use, including much hands-on information, discussing different approaches and applying methods in political science.

Landman, T. (2008) *Issues and Methods in Comparative Politics: An Introduction* (3rd edn) (London: Routledge). This introductory text discusses various fields within comparative politics, focusing on different research designs by means of one, few, and many cases.

Specific literature on comparative methods

Beach, D., and Pedersen, R. B. (2016) *Causal Case Studies: Foundations and Guidelines for Comparing, Matching and*

3. Globalization is considered not only to grow but also to produce biased results due to diffusion. Can you discuss this problem in relation to the relationship between national policy formation of membership states and the EU?

4. Describe the basic differences between the Methods of Agreement and Difference. Give an example of each, demonstrating this difference.

5. A constraint of the comparative method is 'conceptual stretching' and the solution may lie in extending the number of cases by means of 'family resemblance' or 'radial categories'. Can you think of an example of each to extend the number of valid cases?

Tracing (Ann Arbor: University of Michigan Press). This book delineates the ontological and epistemological differences among causal case study methods, offers suggestions for determining the appropriate methods for a given research project, and explains the step-by-step application of selected methods.

Braun, D., and Maggetti, M. (eds) (2015) *Comparative Politics: Theoretical and Methodological Challenges* (Edward Elgar Publishing). A profound and thorough discussion of the contemporary challenges for comparative politics regarding multi-level analysis, concept formation, multiple causation, and complexities of descriptive inference.

Mahoney, J., and Rueschemeyer, D. (eds) (2003) *Comparative Historical Analysis in the Social Sciences* (Cambridge: Cambridge University Press). This reader contains many different views on developing qualitative types of comparative research with an emphasis on history and the use of case studies.

Pennings, P., Keman, H., and Kleinnijenhuis, J. (2006) *Doing Research in Political Science: An Introduction to Comparative Methods and Statistics* (2nd edn) (London: Sage). This is a course book intended for students. It is an introduction to the use of statistics in comparative research and contains many examples of published research.

Ragin, C. (2008) *Redesigning Social Inquiry. Fuzzy Sets and Beyond* (Chicago, IL: University of Chicago Press). The book centres on the 'fuzzy-set' approach as an alternative to other comparative methods as discussed in this chapter (see also Box 3.3) and discusses the advantages of this approach in comparison with extant practices.



Web links

<http://www.politicalresources.net>

Political Resources on the Net: provides links to a number of useful data sets for use in comparative politics.

<http://www.sourceoecd.org>

OECD statistics on the economy, infrastructure, and social expenditures of most developed nation-states and Central and Eastern Europe.

<http://www.ec.europa.eu/eurostat/>

Eurostat: aggregated statistical and individual level survey data on EU and member states.

<http://www.compass.org>

Website devoted to qualitative comparative analysis and Boolean analysis as used in Box 3.3 with access to literature and various software, as well as links to all relevant websites in the field.



For additional material and resources, please visit the Online Resource Centre at:

www.oxfordtextbooks.co.uk/orc/caramanica/



Endnotes

¹ In comparative politics the term 'cross-national research' is often used to depict the cases for comparative analysis (e.g. Landman 2003). Here I use the term 'cross-system analysis' to avoid the idea that only 'nations' are cases in a comparative politics (Lijphart 1975: 166).

² There are many names in use for independent variables: exogenous, effect-producing, antecedent variables, etc. They all have in common that a change in X affects Y. In this

chapter we use the term dependent (Y) and independent (X) variables.

³ Another way of developing a causal argument is counterfactual analysis, asking what happens if a variable is omitted from the equation. In this example: what if 'politics' does not play a role?

⁴ Figure 3.1 is a way to handle spuriousness: by controlling for social and economic factor X_1 , one could estimate the relative influence of politics X_2 in terms of a direct relationship.