

Ústřední knihovna FSS MU Brno



4240748307

Issues and Methods in Comparative Politics: An Introduction

■ 2nd Edition

Todd Landman

 **Routledge**
Taylor & Francis Group
LONDON AND NEW YORK

Part I

WHY, HOW, AND PROBLEMS OF COMPARISON

1	Why compare countries?	3
2	How to compare countries	23
3	Choosing countries and problems of comparison	39

The chapters in this part of the book establish the rationale for the systematic comparison of countries, demonstrate the different ways in which countries can be compared, and examine the various problems that scholars have confronted or will confront when comparing countries. Too often, both the choice of countries and the way in which they are compared are decided for reasons not related to the research question. In contrast, these chapters argue that the comparative research strategy matters. From the initial specification of the research problem, through the choice of countries and method of analysis, to the final conclusions, scholars must be attentive to the research question that is being addressed and the ways in which the comparison of countries will help provide answers.

To this end, Chapter 1 shows that the comparison of countries is useful for pure description, making classifications, hypothesis-testing, and prediction. It then shows how methods of comparison can add scientific rigour to the study of politics in helping students and scholars alike make stronger inferences about the political world they observe. This is followed by a discussion of key terms needed for a science of politics including theory and method; ontology, epistemology, and methodology; cases, units of analysis, variables, and observations; levels of analysis; and quantitative and qualitative methods. Chapter 2 delves deeper into the different ways in which countries can be compared and why these different methods matter for making inferences. It argues that scholars face a key trade-off between the level of conceptual abstraction and the scope of countries under study. It shows how comparing many countries, few countries, or single-country studies all fit under the broad umbrella of 'comparative politics', and that all have different strengths and weaknesses for the ways in which political scientists study the world.

Finally, Chapter 3 outlines the main problems that confront comparativists and suggests ways in which to overcome them. These problems include 'too many variables and too few countries', establishing equivalence between and among comparative concepts, selection bias, spuriousness, ecological and individualist fallacies, and value bias. Together, these chapters offer a synthesis of comparative methods and provide a 'toolchest' for students and scholars that can be used to approach both existing and new research questions in political science.

Chapter 1

Why compare countries?

Reasons for comparison	4
The science in political science	10
Scientific terms and concepts	15
Summary	20
Further reading	20

Making comparisons is a natural human activity. From antiquity to the present, generations of humans have sought to understand and explain the similarities and differences they perceive between themselves and others. Though historically, the discovery of new peoples was often the product of a desire to conquer them, the need to understand the similarities and differences between the conquerors and the conquered was none the less strong. At the turn of the new millennium, citizens in all countries compare their position in society to those of others in terms of their regional, ethnic, linguistic, religious, familial, and cultural allegiances and identities; material possessions; economic, social and political positions; and relative location in systems of power and authority. Students grow up worried about their types of fashion, circle of friends, collections of music, appearance and behaviour of their partners, money earned by their parents, universities they attend, and careers they may achieve.

In short, to compare is to be human. But beyond these everyday comparisons, how is the process of comparison scientific? And how does the comparison of countries help us understand the larger political world? In order to answer these important questions, this chapter is divided into four sections. The first section establishes the four main reasons for comparison, including contextual description, classification and 'typologizing', hypothesis-testing and theory-building, and prediction (Hague *et al.* 1992: 24–27; Mackie and Marsh 1995: 173–176). The second section specifies how political science and the sub-field of comparative politics can be scientific, outlining briefly the similarities and differences between political science and natural science. The third section clarifies the terms and concepts used in the preceding discussion and specifies further those terms and concepts needed for a science of politics. The fourth section summarizes these reasons, justifications, and terms for a science of comparative politics.

Reasons for comparison

Today, the activity of comparing countries centres on four main objectives, all of which co-exist and are mutually reinforcing in any systematic comparative study, but some of which receive more emphasis, depending on the aspirations of the scholar. *Contextual description* allows political scientists to know what other countries are like. *Classification* makes the world of politics less complex, effectively providing the researcher with 'data containers' into which empirical evidence is organized (Sartori 1970: 1039). The *hypothesis-testing* function of comparison allows the elimination of rival explanations about particular events, actors, structures, etc. in an effort to help build more general theories. Finally, comparison of countries and the generalizations that result from comparison allow *prediction* about the likely outcomes in other countries not included in the original comparison, or outcomes in the future given the presence of certain antecedent factors.

Contextual description

This first objective of comparative politics is the process of describing the political phenomena and events of a particular country, or group of countries. Traditionally, in political science, this objective of comparative politics was realized in countries that were different to those of the researcher. Through often highly detailed description, scholars sought to escape their own ethnocentrism by studying those countries and cultures foreign to them (Dogan and Pelassy 1990: 5–13). The comparison to the researcher's own country is either implicit or explicit, and the goal of contextual description is either more knowledge about the nation studied, more knowledge about one's own political system, or both. The comparative literature is replete with examples of this kind of research, and it is often cited to represent 'old' comparative politics as opposed to the 'new' comparative politics, which has aspirations beyond mere description (Mayer 1989; Apter 1996). But the debate about what constitutes old and new comparison often misses the important point that all systematic research begins with good description. Thus description serves as an important component to the research process and ought to precede the other three objectives of comparison. Purely descriptive studies serve as the raw data for those comparative studies that aspire to higher levels of explanation.

From the field of Latin American politics, Macauley's (1967) *Sandinista Affair* is a fine example of contextual description. The book is an exhaustive account of Augusto Sandino's guerrilla campaign to oust US marines from Nicaragua after a presidential succession crisis. It details the specific events surrounding the succession crisis, the role of US intervention, the way in which Sandino upheld his principles of non-intervention through guerrilla attacks on US marines, and the eventual death of Sandino at the hands of Anastasio Somoza. The study serves as an example of what Almond (1996: 52) calls 'evidence without inference', where the author tells the story of this remarkable political leader, but the story is not meant to make any larger statements about the struggle against imperialism. Rather, the focus is on the specific events that unfolded in Nicaragua, and the important roles played by the various characters in the historical events.

Classification

In the search for cognitive simplification, comparativists often establish different conceptual classifications in order to group vast numbers of countries, political systems, events, etc. into distinct categories with identifiable and shared characteristics. Classification can be a simple dichotomy such as between authoritarianism and democracy, or it can be a more complex 'typology' of regimes and governmental systems. Like contextual description, classification is a necessary component of systematic comparison, but in many ways it represents a higher level of comparison since it seeks to group many separate descriptive entities into simpler categories. It reduces the complexity of the world by seeking out those qualities that countries share and those that they do not share.

The process of classification is not new. The most famous effort at classification is found in Aristotle's *Politics* (Book 3, Chapters 6–7), in which he establishes six

types of rule. Based on the combination of their form of rule (good or corrupt) and the number of those who rule (one, few, or many), Aristotle derived the following six forms: monarchy, aristocracy, polity, tyranny, oligarchy, and democracy (see Hague *et al.* 1992: 26). A more recent attempt at classification is found in Finer's (1997) *The History of Government*, which claims that since antiquity (ca. 3200 BC), all forms of government have belonged to one of the following four basic types: the palace polity, the church polity, the nobility polity, and the forum polity. Each type is 'differentiated by the nature of the ruling personnel' (ibid.: 37). In the palace polity, 'decision-making rests with one individual' (ibid.: 38). In the church polity, the church has a significant if not exclusive say in decision-making (ibid.: 50). In the nobility polity, a certain pre-eminent sector of society has substantial influence on decision-making (ibid.: 47). In the forum polity, the authority is 'conferred on the rulers from below' by a 'plural headed' forum (ibid.: 51). Aristotle's classification was derived deductively and then 'matched' to actual city states, while Finer's classification scheme is based on empirical observation and inductive reasoning (see below for the distinction between these two types of reasoning). Both scholars, however, seek to describe and simplify a more complex reality by identifying key features common to each type (see Briefing Box 1.1).

Hypothesis-testing

Despite the differences between contextual description and classification, both forms of activity contribute to the next objective of comparison, hypothesis-testing. In other words, once things have been described and classified, the comparativist can then move on to search for those factors that may help explain what has been described and classified. Since the 1950s, political scientists have increasingly sought to use comparative methods to help build more complete theories of politics. Comparison of countries allows rival explanations to be ruled out and hypotheses derived from certain theoretical perspectives to be tested. Scholars using this mode of analysis, which is often seen as the *raison d'être* of the 'new' comparative politics (Mayer 1989), identify important variables, posit relationships to exist between them, and illustrate these relationships comparatively in an effort to generate and build comprehensive theories.

Arend Lijphart (1975) claims that comparison allows 'testing hypothesized empirical relationships among variables'. Similarly, Peter Katzenstein argues that 'comparative research is a focus on analytical relationships among variables validated by social science, a focus that is modified by differences in the context in which we observe and measure those variables' (in Kohli *et al.* 1995: 11). Finally, Mayer (1989: 46) argues somewhat more forcefully that 'the unique potential of comparative analysis lies in the cumulative and incremental addition of system-level attributes to existing explanatory theory, thereby making such theory progressively more complete'. The symposia on comparative politics in *World Politics* (Kohli *et al.* 1995) and the *American Political Science Review* (vol. 89, no. 2, pp. 454–481), suggest that questions of theory, explanation, and the role of comparison are at the forefront of scholars' minds.

Briefing box 1.1 Making classifications: Aristotle and Finer

Description and classification are the building blocks of comparative politics. Classification simplifies descriptions of the important objects of comparative inquiry. Good classification should have well-defined categories into which empirical evidence can be organized. Categories that make up a classification scheme can be derived inductively from careful consideration of available evidence or through a process of deduction in which 'ideal' types are generated. This briefing box contains the oldest example of regime classification and one of the most recent. Both Aristotle and Samuel Finer seek to establish simple classificatory schemes into which real societies can be placed. While Aristotle's scheme is founded on normative grounds, Finer's scheme is derived empirically.

Constitutions and their classifications

In Book 3 of *Politics*, Aristotle derives regime types which are divided on the one hand between those that are 'good' and those that are 'corrupt', and on the other, between the different number of rulers that make up the decision-making authority, namely, the one, the few, and the many. Good government rules in the common interest while corrupt government rules in the interests of those who comprise the dominant authority. The intersection between these two divisions yields six regime types; all of which appear in Figure 1.1. The figure shows that the good types include monarchy, aristocracy, and polity. The corrupt types include tyranny, oligarchy, and democracy. Each type is based on a different idea of justice (McClelland 1997: 57). Thus, monarchy is rule by the one for the common interest, while tyranny is rule by the one for the one. Aristocracy is rule by the few for the common interest, while oligarchy is rule by the few for the few. Polity is rule by the many for the common good, while democracy is rule by the many for the many, or what Aristotle called 'mob rule'.

		Those Who Rule		
		One	Few	Many
Form of Rule	Good	Monarchy (kingship)	Aristocracy	Polity
	Corrupt	Tyranny	Oligarchy	Democracy (mob rule)

Figure 1.1 Aristotle's classification scheme

Sources: Adapted from Aristotle (1958: 110–115); Hague *et al.* (1992: 26); McClelland (1997: 57)

Types of regime

Finer (1997: 37) adopts an Aristotelian approach to regime classification by identifying four 'pure' types of regime and their logical 'hybrids'. Each regime type is based on the nature of its ruling personnel. The pure types include the palace, the forum, the nobility, and the church. The hybrid types are the six possible combinations of the pure types, palace-forum, palace-nobility, palace-church, forum-nobility, forum-church, and nobility-church. These pure and hybrid types are meant to describe all the regime types that have existed in world history from 3200 BC to the modern nation state. Finer concedes that there are few instances of pure forms in history and that most polities fit one of his hybrid types. These pure forms, their hybrids, and examples from world history appear in Figure 1.2. The diagonal that results from the intersection of the first row and column in the figure represents the pure forms, while the remaining cells contain the hybrid forms. Many regime types that were originally pure became hybrid at different points in history. Of all the types, the pure palace and its variants have remained the most common through history, and despite its popularity today, the forum polity that represents modern secular democracies is a relatively rare and recent regime type (Finer 1997: 46).

	Palace	Forum	Nobility	Church
Palace	<i>Pure Palace</i> Persian, Roman, Byzantine, Chinese, and Islamic Empires; 18th-century absolutisms	<i>Palace-Forum</i> Greek tyrants, Roman dictators, Napoleonic France, modern dictatorships and totalitarian regimes	<i>Palace-Nobility</i> Court of Louis XIV, Britain 1740-60, Poland, Mamluk Regime in Egypt, and pre-1600 Japan	<i>Palace-Church</i> Traditional Thailand: the <i>sangha</i> ; European Middle Ages
Forum		<i>Pure Forum</i> Greek <i>poleis</i> , Roman republics, and medieval European city-states; modern secular democracies	<i>Forum-Nobility</i> Roman republic, Republic of Venice	<i>Forum-Church</i> Ephrata Mennonites 1725, Amish 1700-present, [†] both near Lancaster, Pennsylvania
Nobility			<i>Pure Nobility</i> 17th- and 18th-century Poland	<i>Nobility-Church</i> Teutonic Order 1198-1225
Church				<i>Pure Church</i> Vatican; Tibet 1642-1949

Figure 1.2 Pure and hybrid regime types with examples from history

Source: Adapted from Finer (1997: 34-58)

Note: [†] Author's addition

Furthermore, the publication of truly comparative books in the field continues to demonstrate the fruitfulness of this mode of analysis. For example, Luebbert (1991) compares Britain, France, Switzerland, Belgium, The Netherlands, Denmark, Norway, Sweden, Czechoslovakia, Germany, Italy, and Spain to uncover the class origins of regime type in inter-war Europe. Rueschemeyer *et al.* (1992) compare the historical experiences of the advanced industrial countries with those of the developing world to uncover the relationship between capitalist development and democracy. Wickham-Crowley (1993) compares instances of revolutionary activity in Latin America to discover the causal configuration of successful and unsuccessful social revolution in the region. Foweraker and Landman (1997) compare the authoritarian cases of Brazil, Chile, Mexico, and Spain to illustrate the relationship between citizenship rights and social movements. Finally, Inglehart (1997) compares survey data from forty-three societies to assess the mutual relationship between the process of modernization (or post-modernization) and changing value systems. In all these works, key explanatory and outcome variables are carefully defined and the relationships between them are demonstrated through comparison of empirical evidence (see Briefing Box 1.2).

Briefing box 1.2 Hypothesis-testing

Voting participation

In *Contemporary Democracies*, Powell (1982) examines a number of key hypotheses concerning voter participation in twenty-nine democratic countries. Participation is measured using voter turnout, or the percentage of the eligible voters who actually voted in national elections. He argues that voting participation ought to be higher in countries with higher levels of economic development (per capita GNP), a representational constitution, electoral laws that facilitate voting, and a party system with strong alignments to groups in society (Powell 1982: 120-121). His statistical analysis of the data from these countries reveals positive effects for all these variables on voter participation, which are depicted graphically in Figure 1.3.

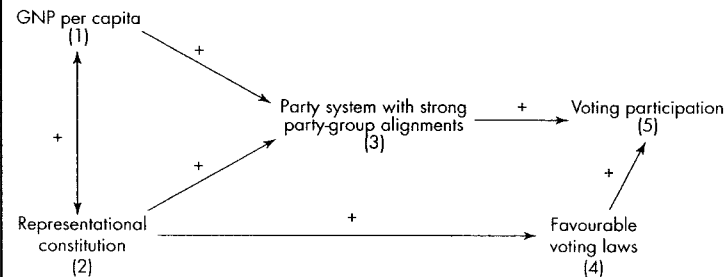


Figure 1.3 Four hypotheses on voting participation

Source: Adapted from Powell (1982: 121)

Moreover, his analysis shows that the level of economic development and constitutional structure are not directly related to voter participation, but that they 'lead to or help sustain the development of party systems and the choice of voting laws, which do get the voters to the polls' (ibid.: 120). This causal ordering is depicted in the figure with the arrows and the numbering of each variable.

Prediction

The final and most difficult objective of comparative politics is a logical extension of hypothesis-testing, namely, to make predictions about outcomes in other countries based on the generalizations from the initial comparison, or to make claims about future political outcomes. Prediction in comparative politics tends to be made in probabilistic terms, such as 'countries with systems of proportional representation are more likely to have multiple political parties'. In this example, a political scientist would know the likely effect of a nation switching its electoral system from a plurality or 'first-past-the-post' rule to a proportional one (Hague *et al.* 1992). Another predictive example involves the benefits accrued to political incumbents in contesting future elections. Based on the empirical observations of past electoral contests, political scientists could be reasonably secure in predicting that the incumbent in any given election has a higher probability of winning the election than the non-incumbent (see King *et al.* 1994).

Although prediction is less an aspiration of comparativists today than in the past, there are those who continue to couch their arguments in predictive language. For example, weak predictive arguments are found in Huntington's (1996) *The Clash of Civilizations and the Remaking of the New World Order*, and strong predictive arguments are found in Vanhanen's (1997) *The Prospect of Democracy*. Huntington (1996) identifies nine key cultural groupings which he believes currently characterize the world's population, and predicts that future conflicts will be more likely to appear in the areas where two or more of these cultures meet or 'clash'. Not only does he seek to predict future conflicts in the world, but claims that his 'civilization' approach accounts for more post-Cold War events than rival approaches. His predictions became all the more relevant after the terrorist attacks on the World Trade Center and the Pentagon on 11 September 2001, which many saw as proof of a clash between the 'Western' and 'Islamic' civilizations outlined in his book. Similarly, based on observations of the presence of economic resources and the occurrence of democracy in the world from the middle of the nineteenth century until today, Vanhanen (1997: 99–154) predicts the degree to which individual countries and regions in the world are likely to become democratic (see Briefing Box 1.3).

The science in political science

The preceding section specified the four main objectives of comparison in political science and hinted, through reference to questions of explanation, theory-building, and prediction, how comparison might be considered a science. The key term used

Briefing box 1.3 Making predictions

Democracy in East and Southeast Asia

Using similar methods as Burkhart and Lewis-Beck, Vanhanen (1997) seeks to predict the expected level of democracy in specific countries and regions of the world based on their distribution of 'power resources'. Democracy is measured by a combination of the smallest parties' share of the vote and the percentage turnout (ibid.: 35). The distribution of power resources is measured by an index that combines the urban population, the non-agricultural population, proportion of students, the size of the literate population, the number of family farms, and the degree of decentralization of non-agricultural economic resources (Vanhanen 1997: 59–60). By examining the relationship between the level of democracy and the distribution of power resources from 1850–1993, Vanhanen compares the actual 1993 values of democracy to those that were predicted using regression analysis. Figure 1.4 shows the actual and predicted values of democracy for sixteen

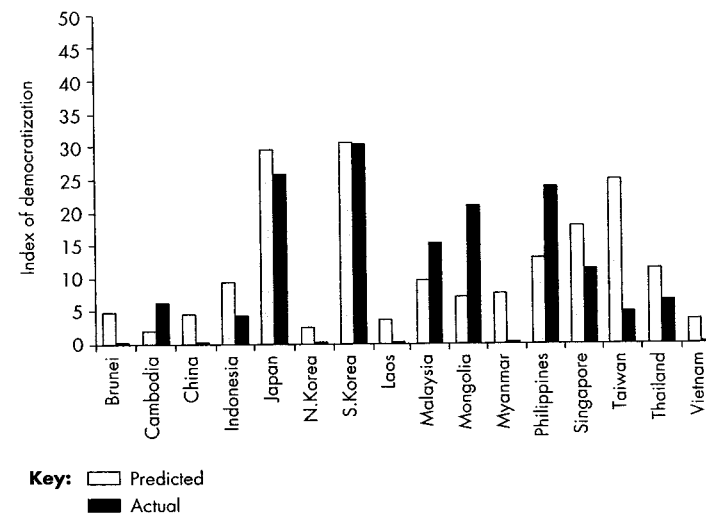


Figure 1.4 Predicting democracy in East and Southeast Asia

Source: Adapted from Vanhanen (1997: 88–89)

countries from East and Southeast Asia. The sixteen countries are listed along the horizontal axis and the values of the index of democratization are listed on the vertical axis. The predicted scores of democracy represent the level of democracy that each country ought to have obtained by 1993, given its corresponding

distribution of power resources. The actual level is the score for 1993. The difference between the two values is known as the residual. Japan and South Korea appear to have obtained the levels of democracy that were predicted, while Malaysia, Mongolia, and the Philippines have higher levels of democracy than expected and Brunei, China, and Taiwan have lower scores than were expected. These varied results have several implications. First, the discrepancy between the actual and the predicted values may mean that something other than the distribution of power resources accounts for the level of democracy (see Chapter 3). Second, the deviant cases whose level of democracy is unexpected for 1993 may be temporary exceptions to the overall pattern. Third, the indicators that were used may not accurately reflect the concepts Vanhanen seeks to measure (see Chapter 3). Overall, however, the process of making predictions can raise new research questions and identify the need to focus on those cases that do not 'fit' the pattern (see Chapter 2).

throughout the discussion was *inference*. Simply put, making an inference is 'using facts we know to learn something about facts we do not know' (King *et al.* 1994: 119 after Mill; see also Couvalis 1997). Gabriel Almond (1996) observes that 'the object of political science . . . is the creation of knowledge, defined as inferences or generalizations about politics drawn from evidence'; and Mayer (1989: 56) claims that 'comparative analysis . . . [is] a method that plays a central role in the explanatory mission of political science itself'. Thus, comparative politics seeks to achieve the goal of inference about politics through comparing countries. This section of the chapter clarifies how the process of making inferences is the underlying principle of comparative politics, and how the methodological assumptions of natural science are important to a science of politics.

For the purposes of this volume, science is defined as the gradual accumulation of knowledge about the empirical world through systematic practices of inquiry, including the collection of evidence, the generation and testing of hypotheses, and the drawing of substantive inferences.¹ But beyond this basic definition, what are the parallels between political science and natural science? What are the main differences between the two? And how does comparison help resolve these differences? The strong case for a science of politics suggests that both (comparative) political science and natural science share the same basic goals, namely, description, classification, hypothesis-testing, and prediction. Both activities require the systematic collection of evidence; an ordering of the evidence and the search for discernible patterns; the formulation and testing of contending explanations for the occurrence of the patterns; and the building of more general theories. Thus, a science of politics always contains this 'evidence-inference methodological core' (Almond 1996: 52), or the 'customary pair' of theory and observation (Feyerabend 1993: 23; see also Gordon 1991: 589–634).

Two examples from the natural sciences may help make these points clearer. Both the theory of evolution and the theory of gravity are based on the systematic collection of evidence. Charles Darwin sought to document the entirety of the Earth's flora and fauna. Originally in an effort to demonstrate the glory of God's creation,

Darwin soon discovered a pattern in what he was observing for which an alternative explanation was possible. The theory of evolution, buttressed later by the theory of natural selection, emerged as the new explanation for the variety of species found in the natural world. Similarly, Isaac Newton formulated the theory of gravity based on the collection of evidence (the falling apple!). Neither scientist had actually seen evolution or gravity but merely *observed its effects*. In this way, evolution and gravity are mental constructs, whose repeated empirical verification has given them a law-like status.

Political scientists also collect evidence systematically (e.g. archival records, interviews, official statistics, histories, or surveys), search for discernible patterns in the evidence, and formulate theories to account for those patterns. In comparative politics, the political scientist compares countries in an effort to verify the theories that have been formulated. Thus, both the natural and political sciences seek to make inferences based on the empirical world they observe, and both seek to maximize the certainty of these inferences. Despite these general similarities between natural science and political science, there remain two important (albeit not absolute) differences: experimentation and the generation of scientific 'laws'. These differences are discussed in turn.

The first difference between natural science and political science is the role of experimentation. While for some areas of natural scientific research, such as astronomy and seismology, experimentation is not possible, the advances in natural science are generally supported by evidence gathered through experimentation, which involves the controlled manipulation of the subject under study in an effort to isolate causal factors. Evidence in political science, on the other hand, tends not to be gathered through experimentation, even though some political scientists use experiments in their research (e.g. those who work on game theory, focus groups, and 'citizen-juries'). Comparative politics, in particular, cannot use experimentation for both practical and ethical reasons. For example, it would be practically impossible to re-run the same election in the same country with a different electoral system to observe the differences in the outcome of the two systems. Ethically, it would be impossible to redistribute income intentionally in a developing country to see if civil strife erupts. Both these examples demonstrate the use of *counterfactuals*, or situations in which the researcher imagines a state of affairs where the antecedent factors to a given event are absent and where an alternative course of events or outcomes is considered (Ferguson 1997b).

Whether it is different electoral systems, different distributions of income, different levels of economic development, or the absence of particular revolutionary groups, political scientists implicitly suggest a counterfactual situation when making claims about important explanatory factors. The claim that 'single-member district electoral systems tend to produce two-party systems' is in effect also claiming that countries *without* such electoral systems will necessarily have different political party systems. While some historians may construct alternative historical scenarios based on 'calculations about the relative probability of plausible outcomes' (*ibid.*: 85), political scientists compare countries that differ in ways that supply the counterfactual situation. For example, by comparing the political party systems across countries with different electoral systems, the comparativist seeks to demonstrate that the type of electoral system has some bearing on the type of party system. In this way,

comparative research 'simulates' experimentation (Lieberson 1987: 45; Ferguson 1997b; see also Tetlock and Lebow 2001).

The second difference between natural science and political science involves the law-like status that is given to certain scientific theories. Experimentation and repeated empirical verification give theories in the natural sciences the status of laws (e.g. the law of conservation of energy, or Boyle's Law of Gases); however, the nature of evidence marshalled in support of theories of political science is such that law-like generalizations are rare. Three famous 'laws' of political science are well known. Michels' 'Iron Law of Oligarchy' suggests that the natural processes observable in the dynamics of organizations and small groups are such that over time, all groups and organizations develop a hierarchical structure of authority with a small elite at their head. In an example from the comparative literature, this law has been tested in the examination of social movement organizations, where evidence suggests that the most successful and longstanding social movement organizations tend to have formal bureaucratic structures and authoritative bodies composed of elites from the movement (see Tarrow 1994). The second law, called 'Duverger's Law', states that electoral systems based on single-member districts tend to produce two parties while systems with proportional representation tend to produce multiple parties. This law has been repeatedly tested in comparative studies on electoral systems and on balance, is supported by the evidence (see Rae 1971; Lijphart 1994a).

The third law on 'the democratic peace' states that democracies do not go to war with each other (Babst 1964). Repeated comparative studies in international relations of war 'dyads' (i.e. pairs of countries that engage in war with each other), demonstrate that '[t]he number of wars between the democracies during the past two centuries ranges from zero to less than a handful depending on precisely how democracy is defined' (Levy 1989: 87–88). Scholars argue that this 'absence of war between democracies comes as close to anything we have to an empirical law in international relations' (ibid.: 88). Moreover, combined with the process of democratization, which has become more pronounced since 1974 (see Chapter 7), the law of democratic peace offers optimism about future conflict in the world, since a larger proportion of democracies in the world means fewer inter-state wars (see Ward and Gleditsch 1998; Przeworski *et al.* 2000; Gelpi and Griesdorf 2001).

Aside from these three 'laws' of political science, the bulk of comparative research eschews making such strong claims. What then are the main conclusions about comparative politics that can be drawn from this cursory comparison to natural science? First, for practical and ethical reasons, comparative politics relaxes some of the rigours of natural science, but still employs the same logic of inference. Second, comparative politics is a non-experimental (or quasi-experimental) social science that seeks to make generalizations based on the best available evidence (Campbell and Stanley 1963; Lijphart 1975: 162; Lieberson 1987). Third, as a substitute for experimentation, comparison allows for control (Sartori 1994: 16), holding certain things constant while examining and accounting for observed differences (see Chapter 2). Fourth, while not seeking ironclad laws, comparative politics seeks clarity, understanding, and explanation of political phenomena about which it can be reasonably certain. The goal of this book therefore, is to provide the necessary tools for students of politics to achieve this clarity, understanding, and explanation while avoiding the pitfalls and obstacles that limit such an enterprise.

Scientific terms and concepts

Before concluding this chapter, it is necessary to define and clarify terms that have been used thus far, as well as terms that will be encountered throughout the book. These are general terms used throughout the social sciences that all students of politics ought to know if they aspire to a more scientific approach to understanding the political world. These terms include theory and method; ontology, epistemology, and methodology; cases (or countries), units of analysis, variables, and observations; levels of analysis; and quantitative and qualitative methods. Throughout the discussion every effort is made to show how the book uses these terms and concepts of social science.

Theory and method

There are two basic types of theory in political science, normative and empirical. Normative theory specifies how things in society *ought to be*, given a desired set of outcomes and philosophical position. From the Greeks and Romans to Rawls, normative political theorists establish frameworks for realizing the common good and address key problems of society through theoretical argumentation. For example, Rawls (1971) carries on the tradition of liberal contract theory found in Locke, Rousseau and Kant, by deriving principles of justice from an idealized thought experiment. In contrast, empirical theory seeks to establish causal relationships between two or more concepts in an effort to explain the occurrence of observed political phenomena. For example, an empirical theory of social revolution may posit a series of socio-economic factors that account for revolutionary behaviour in certain types of people, which would then be tested using evidence (see Chapter 5). In addition, theories in political science can be deductive or inductive. Deductive theories arrive at their conclusions by applying reason to a given set of premises (Stoker 1995: 17; Lawson 1997: 16–19; Couvalis 1997). For example, the rational choice perspective in political science assumes that all political actors maximize their own personal utility, or self-interest, when choosing between alternatives. From that basic assumption, the scholar logically deduces the range of possible outcomes (Ward 1995: 79; Levi 1997). Inductive theories, on the other hand, arrive at their conclusions through observation of known facts (Couvalis 1997). For example, a scholar observing higher instances of peasant rebellion in geographical areas with higher levels of land and income inequality will arrive inductively at the conclusion that inequality is related to rebellion. Comparison of evidence from other countries or geographical regions would seek to confirm this generalization.

Method, on the other hand, is the means by which a theory is derived and tested, including the collection of evidence, formulation and testing of hypotheses, and the arrival at substantive conclusions. Evidence can be collected, for example, through the examination of historical records, the collation and analysis of open-ended interviews of political activists, the systematic reporting of the participant observation of social movement activities, or the construction and analysis of mass surveys of a sample of the population. In formulating and testing hypotheses, method makes the decision rules and the rejection of rival hypotheses explicit. Finally,

substantive conclusions are drawn from the theories and the evidence. As the preceding discussion in this chapter suggests, this book, although not primarily concerned with different theories of comparative politics, seeks to demonstrate the different ways in which comparative methods can be used to test deductive and inductive empirical theories of politics.

Ontology, epistemology, and methodology

Ontology, epistemology, and methodology are terms that occur in the discussion of the philosophy of science and the distinctions between them often become blurred in the comparative literature. Ontology is, quite literally, the study of being, or the metaphysical concern with the essence of things, including the 'nature, constitution, and structure of the objects' of comparative inquiry (Lawson 1997: 15). It concerns what can be studied, what can be compared, and what constitutes the political. In other words, for comparative politics, ontology concerns the countries, events, actors, institutions, and processes among other things that are observable and in need of explanation. Epistemology is the study of the nature of knowledge, or how scholars come to know the world, both through a priori means and through a posteriori means of observation, sense impression, and experience. In contrast to ontology, it concerns what knowledge of the political world is possible and what rules of inquiry scholars follow in knowing the political world. In the history and philosophy of science, epistemology has moved from the strong claim made by positivists that a unity of the natural and social sciences is possible to one that recognizes a certain plurality of approaches grounded in the link between evidence and inference of the kind that this book advocates (see Gordon 1991: 589–668). In contrast to ontology and epistemology, methodology concerns the ways in which knowledge of the political world is acquired. As its name suggests, methodology is the study of different methods or systems of methods in a given field of inquiry. There are thus rules of inquiry specific to qualitative and quantitative methods, even though both strive to provide explanation and understanding of observed political phenomena. These three concepts also have 'directional dependence' such that ontology establishes what is knowable, epistemology how it is knowable, and methodology how it is acquired systematically (Hay 2002: 61–66).

Having defined these terms, it is helpful for the reader to know how the discussions throughout the rest of this book are grounded in certain ontological, epistemological, and methodological assumptions. Without entering a philosophical debate, this book is grounded in the ontological belief that animate and inanimate objects in the world exist in and of themselves, and by extension observable events exist in and of themselves. The object of political science is to account for and understand these events in terms of why they happened, how they happened, and the likelihood of them happening again in the future, as well as in different parts of the world. While adhering to the notion that history is 'open ended' (Popper 1997), this book accepts that there are certain 'event regularities' (Lawson 1997) in the world that political science seeks to describe and explain.

Epistemologically, comparative politics inhabits a broad spectrum. One end of the spectrum contends that all things political and social are knowable through

the process of deduction based on indisputable assumptions about human nature. Typically labelled nomological-deductivism, such an epistemological position adheres to the positivist quest for law like generalizations about political behaviour. The other end of the spectrum claims that all knowledge is culturally bound and relative, suggesting that it is impossible to know anything beyond the strict confines of the local cultural context (Kohli *et al.* 1995). Such a position suggests that a science of comparative politics is not possible, since political concepts would not 'travel' across different cultural contexts and there would be fundamental differences in their meaning (see Macintyre 1971).

In an effort to be inclusive of different methods of comparison, this book is located somewhere in between these two extremes. On the one hand, it accepts that certain deductive theories of politics can be tested in the real world and that generalizations about the world of politics are possible given the proper adherence to rules of inquiry. On the other hand, it recognizes that knowledge of the political world cannot be 'value-free' and that the processes of theory generation and observation may not be mutually exclusive (Feyerabend 1993: 27; Sanders 1995: 67–68; Couvalis 1997). It therefore accepts that certain kinds of cross-cultural comparisons and cross-national comparisons can be made if certain procedures are adopted (see Chapter 3 in this volume). Methodologically, the book is concerned with the application of comparative methods to real research problems in comparative politics in an effort to help students make more valid generalizations about the political world they observe. These different methods of comparison, as well as their advantages and disadvantages are outlined in Chapter 2.

Cases, units of analysis, variables, and observations

These four terms are vital aspects of systematic research in comparative politics. Cases are those countries that feature in the comparative analysis. For example, in *States and Social Revolutions* (1979), Theda Skocpol examines the cases of France, Russia, and China. Units of analysis are the objects on which the scholar collects data, such as individual people, countries, electoral systems, social movements, etc. Variables are those concepts whose values change over a given set of units, such as income, political party identification, propensity to join a protest movement, etc. Observations are the values of the variables for each unit, which can be numeric, verbal, or even visual. For example, a hypothetical study of social movements in Britain, France, The Netherlands, and Germany may have a variable entitled 'strategy', which has categories denoted 'political lobbying', 'peaceful demonstration', 'violent direct action', 'grass-roots organizing', and 'consciousness-raising'. In this hypothetical study, the cases are the countries, the units of analysis are the movements, the variable is 'strategy', and the observation is the value of the strategy variable for a given movement in a given country.

In addition to the different values that variables assume, they can either be *dependent* or *independent*. Dependent variables (alternatively referred to as outcome variables, endogenous variables, or the explanandum) are those political outcomes that the research is trying to explain. An independent variable, on the other hand, is that which explains the dependent variable (and is alternatively labelled a causal

variable, an explanatory variable, an exogenous variable, or the explicandum). The distinction between dependent and independent variables is derived from the specific research question of a comparative project and the particular theoretical perspective that has been adopted. Since most political events have multiple explanations, it is possible to have more than one independent variable for a given dependent variable. In formal models of politics, the dependent variable is often depicted by a *y*, and the independent variable is often depicted by an *x*.

For example, a dependent variable may include votes for a leftist party, military coups, revolutions, or transitions to democracy. Independent variables to account for each of these dependent variables may include, respectively, social class, economic crisis, the commercialization of agriculture, or elite bargaining. In his study of guerrillas and revolution in Latin America, Wickham-Crowley (1993) seeks to explain the occurrence of successful social revolutions. In this case, successful social revolution is the dependent variable. The independent variables include the presence of a guerrilla group, the support of workers and peasants, sufficient guerrilla military strength, the presence of a traditional patrimonial regime, and the withdrawal of US military and economic support for the incumbent regime (Wickham-Crowley 1993: 312; see Chapter 5 in this volume).

Levels of analysis

Levels of analysis in political science are divided between the micro, or individual level, and the macro, or system level. Micro-political analysis examines the political activity of individuals, such as respondents in a mass survey, elite members of a political party or government, or activists in a protest movement. Macro-political analysis focuses on groups of individuals, structures of power, social classes, economic processes, and the interaction of nation states. As in other divisions in political science, there are those who believe all of politics can be explained by focusing on micro-level processes, and there are those who believe that all of politics can be explained by a focus on macro-level processes. This is sometimes called the 'structure-agency' problem of politics (see Hay 1995, 2002). Micro-analysts believe that the world of politics is shaped by the actions of 'structureless agents', while macro-analysts believe that that world is shaped by the unstoppable processes of 'agentless-structures'.

The comparative politics literature is rich with examples of these different levels of analysis. In *The Rational Peasant*, Samuel Popkin (1979) argues that revolutionary movements are best understood by focusing on the preferences and actions of individual peasants (a micro-level analysis). Support for this assertion comes from his intense study of peasant activity in Vietnam. In contrast to Popkin, Jeffrey Paige (1975) in *Agrarian Revolution*, demonstrates that revolutions are most likely in countries with a particular structural combination of owners and cultivators. This macro-level analysis is carried out through comparing many countries at once, and then verifying the findings in the three countries of Vietnam, Angola, and Peru (see Chapter 2). In *Liberalism, Fascism, or Social Democracy*, Gregory Luebbert (1991) claims that the types of regime that emerged in inter-war Europe had nothing to do with 'leadership and meaningful choice' (ibid.: 306), but were determined

structurally by mass material interests, social classes, and political parties (a macro-level analysis). Finally, in the *Breakdown of Democratic Regimes*, Stepan (1978) finds the middle ground in accounting for the 1964 breakdown of democracy in Brazil, where he suggests that macro-political conditions at the time of breakdown certainly limited but did not determine the actions of individual leaders. This present book does not privilege one level of analysis over another. Rather, it demonstrates the ways in which different levels of analysis fit into different comparative methods.

Quantitative and qualitative methods

Simply put, quantitative methods seek to show differences in number between certain objects of analysis and qualitative methods seek to show differences in kind. Quantitative analysis answers the simple question, 'How many of them are there?' (Miller 1995: 154), where the 'them' represents any object of comparison that can either be counted or assigned a numerical value. For example, it is possible to count the number of protest events or assign values to different social movement strategies (see above, p. 17), the degree to which human rights are protected (see Chapter 9), and an individual's identification with political parties. Quantitative data can be official aggregate data published by governments on growth rates, revenues and expenditures, levels of agricultural and industrial production, crime rates and prison populations, or the number of hectares of land devoted to agrarian reform. Quantitative data can also be individual, such as that found in the numerous market research surveys and public opinion polls. Quantitative methods are based on the distributions these data exhibit and the relationships that can be established between numeric variables using simple and advanced statistical methods.

Qualitative methods seek to identify and understand the attributes, characteristics, and traits of the objects of inquiry, and the nature of the method necessarily requires a focus on a small number of countries. In comparative politics, there are three types of qualitative methods: macro-historical comparison (and its three subtypes) (Skocpol and Somers 1980; Ragin *et al.* 1996); in-depth interviews and participant observation (Devine 1995); and what is variously called interpretivism, hermeneutics, and 'thick description' (Geertz 1973; Fay 1975). In none of these types of method is there an attempt to give numerical expression to the objects of inquiry, and in all of them, the goal is to provide well-rounded and complete discursive accounts. These more complete accounts are often referred to as 'ideographic' or 'configurative', since they seek to identify all the elements important in accounting for the outcome.

Through focus on a small number of countries, comparative macro-history allows for the 'parallel demonstration of theory', the 'contrast of contexts', or 'macro-causal' explanation (Skocpol and Somers 1980). Parallel demonstration of theory tests the fruitfulness of theory across a range of countries. The contrast of contexts helps to identify unique features of countries in an effort to show their effect on social processes, while bringing out the richness of the individual countries and aspiring to 'descriptive holism'. Macro-causal analysis seeks to explain observed political phenomena through the identification and analysis of 'master' variables (Luebbert 1991: 5). In-depth interviews and participant observation strive to uncover

a deeper level of information in order to capture meaning, process, and context, where explanation 'involves describing and understanding people as conscious and social human beings' (Devine 1995: 140). Similarly, interpretivism, hermeneutics, and 'thick description' are concerned with interpretation, understanding, and the deeper structures of meanings associated with the objects of inquiry.

Over the years a division in political science has developed between those who use quantitative methods and those who use qualitative methods; however, it seems that this division is a false one if both methods adhere to the goal of making inferences from available evidence (Foweraker and Landman 1997: 48–49). In other words, this book is grounded in the belief that the same logic of inference ought to apply equally to quantitative and qualitative methods (see King *et al.* 1994). Perhaps more importantly, the qualitative distinction made among categories in comparative classification schemes necessarily precedes the process of quantification (Sartori 1970, 1994). And, as the ensuing chapters will demonstrate, it is clear that the field of comparative politics is richly populated with studies that use quantitative and qualitative methods (or both) at all levels of analysis, as well as across all methods of comparison.

Summary

This chapter has outlined the four main objectives of comparative politics and argued further that all co-exist and are necessary for systematic research. Predictions cannot be made without well-founded theories; theories cannot be made without proper classification; and classification cannot be made without good description. The chapter has shown how comparative politics is scientific if it aspires to making inferences about the political world based on the best available evidence. Finally, it defined the key terms that will be used throughout the book. The next chapter examines the different methods of comparison that are available to students, all of which can be used to make larger inferences about the political world that we observe.

Note

- 1 A slightly more cumbersome definition is offered by Goodin and Klingemann (1996a: 9): 'science . . . [is] systematic enquiry, building toward an ever more highly differentiated set of ordered propositions about the empirical world.'

Further reading

Chilcote, R.H. (1994) *Theories of Comparative Politics: The Search for a Paradigm Reconsidered*, 2nd edn, Boulder, CO: Westview.

An overview of the main theories of comparative politics, including system theory, state theory, political culture theory, modernization theory, dependency theory, and class theory.

Dogan, M. and Pelassy, D. (1990) *How to Compare Nations: Strategies in Comparative Politics*, 2nd edn, Chatham, NJ: Chatham House.

A review of why and how to compare countries as well as a brief overview of popular concepts in comparative politics.

Hague, R., Harrop, M., and Breslin, S. (1992) *Political Science: A Comparative Introduction*, New York: St Martin's Press.

A textbook on political science written from a comparative perspective.

Hay, C. (2002) *Political Analysis*, London: Palgrave.

Excellent summary and exposition on the purpose and meaning of conducting political analysis.

King, G., Keohane, R.O., and Verba, S. (1994) *Designing Social Inquiry: Scientific Inference in Qualitative Research*, Princeton: Princeton University Press.

Demanding but worthwhile effort to unify qualitative and quantitative research methods under one logic of inference.

Lichbach, M. and Zuckerman, A. (eds) (1997) *Comparative Politics: Rationality, Culture, and Structure*, Cambridge: Cambridge University Press.

Advanced text on rational, cultural, and structural theories and how they are used in comparative politics.

Marsh, D. and Stoker, G. (eds) (1995) *Theories and Methods in Political Science*, London: Macmillan.

Excellent reader on the main theories and methods of political science.

How to compare countries

Introduction	24
Methods of comparison	24
Comparing many countries	26
Comparing few countries	29
Single-country studies as comparison	34
Conclusion	35
Further reading	36

Introduction

As the last chapter made clear, there are different strategies of comparative research in political science, including comparing many countries, comparing few countries, and single-country studies. In contrast to some comparativists (Lijphart 1971; Peters 1998) and in agreement with Mackie and Marsh (1995: 177), this book argues that all three of these strategies of research are subsumed under the broader umbrella of comparative politics, which can be unified under one logic of inference. The comparative literature is replete with examples of all these methods, but why have they come about and what are the advantages associated with each? This chapter demonstrates that these methods are a function both of the explanatory aspirations of the researcher and the level of conceptual abstraction contained within a given study. The chapter outlines each method and discusses how each is useful for drawing inferences. In no way is one method privileged over another, as each has different advantages and disadvantages.

Methods of comparison

The distinction between different comparative methods should be seen as a function of the particular research question, the time and resources of the researcher, the method with which the researcher is comfortable, as well as the epistemological position he or she adopts. Different research questions require different methods. For example, someone wanting to know why Tony Blair and New Labour won the 1997 General Election in the United Kingdom after eighteen years of Conservative government will necessarily focus on that one country. But someone interested in the electoral support for reformed left-of-centre political parties may choose all the countries in the European Union. Second, the time and resources of researchers are often constrained, which limits the number of countries that can be feasibly researched in any one project. Third, some are comfortable using quantitative methods while others are not. Some enjoy large comparisons while others enjoy researching the fine details of particular countries. Finally, researchers who adhere to deductive theory may use different methods to those adhering to inductive theory. Those seeking more universal generalizations may use different methods from those that seek more contextually specific levels of explanation.

The central distinction between different comparative methods depends on the key trade-off between the *level of abstraction* and the *scope of countries* under study (Mair 1996). In general, the higher the level of conceptual abstraction, the more potential there is for the inclusion of a large number of countries in a study, where political science concepts 'travel' across different contexts (Sartori 1970, 1994). Alternatively, focus on one country or a few countries means that the researcher can use less abstract concepts that are more grounded in the specific contexts under scrutiny. For example, in the study of democratic institutions, a comparison of many countries may use a simple dichotomy between 'presidential' or 'parliamentary' political systems (Stepan and Skach 1993). A comparison of Latin American political systems, however, would have to adopt more refined categories of presidentialism since all the countries in the region are presidential (Jones 1995;

Foweraker 1998). Finally, further refinements of the concept of presidentialism could be made in order to fit the nuances of a particular country, such as the United States.

Figure 2.1 summarizes these methods of comparison by showing this trade-off between the level of abstraction and the scope of countries. The cells identifying each method are determined by the intersection between the level of abstraction (high, middle, and low) and the scope of countries (one, few, and many). The figure is a heuristic device to illustrate this trade-off in stark terms. In reality, the lines of distinction between the various methods are more blurred, and there are studies that use several different methods at once. For example, Paige's (1975) *Agrarian Revolution* compares many countries at once to uncover the structural determinants of revolution in the world, and then compares the specific countries of Angola, Vietnam, and Peru to see if the cross-national findings hold at the local level.

This representation of comparative methods differs slightly from that outlined in previous work on comparative politics (Lijphart 1971; Collier 1991: 9–12). First, it includes all three methods under the comparative umbrella. In the past, Lijphart (1971) called comparing many countries using quantitative analysis the 'statistical' method and comparing few countries using qualitative analysis the 'comparative' method. For many, single-country studies are by their nature not comparative but may have comparative merit. Many such studies either use concepts that are applicable in other countries, develop new concepts that may become applicable in other countries, and/or embed their studies in a comparative context (Sartori 1994: 15). This book argues that if the research strives to make larger inferences about politics through some form of comparison and uses concepts applicable to more than the country under study, then it is comparative (Lichbach and Zuckerman 1997: 4). Thus, all three methods are deemed comparative.

Second, comparing many countries is commonly referred to as 'large-*n*' comparison, and comparing few countries 'small-*n*' comparison, where *n* is the number of countries. It is important not to confuse the usage of *n* when carrying out a comparative study, since it can also refer to the overall number of observations. As Eckstein (1975: 85) rightly observes, it is possible to have a single-country study with many observations, such as six general elections, or 2,000 respondents in a national survey (see also Ragin 2000: 67–69). Putnam's (1993) *Making Democracy*

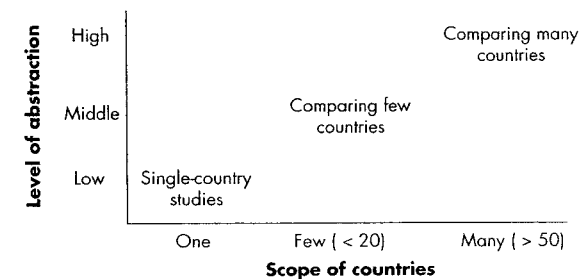


Figure 2.1 Methods of comparison

Sources: Based on Sartori (1970) and Mair (1996)

Work compares many regions within Italy, which, in this case, is a single-country study drawing inferences from a large- n . To prevent confusion in this book, n is always used to denote the number of observations (King *et al.* 1994: 51–52). For example, Burkhart and Lewis-Beck (1994) compare 131 countries from 1972–1989 ($n = 2, 358$), and Foweraker and Landman (1997) compare Brazil (1964–1990), Chile (1973–1990), Mexico (1963–1990), and Spain (1958–1983), producing $n = 99$ (four countries times the total number of years compared). While the former study compared many countries and the latter a few countries, both could be considered ‘large- n ’ comparative studies. Thus, this book divides the three methods into comparing many countries, comparing few countries, and single-country studies.

Comparing many countries

Comparing many countries most closely approximates the experimental method of science, since it is particularly suited to quantitative analysis through measurement and analysis of aggregate data collected on many countries (Lijphart 1971). Although there are examples of qualitative comparisons of many countries, such as Huntington’s (1996) *The Clash of Civilizations* and Finer’s (1997) *History of Government*, the majority of studies that compare many countries simultaneously use quantitative methods. This method of comparison requires a higher level of abstraction in its specification of concepts in order to include as many countries as possible. Its main advantages include statistical control to rule out rival explanations, extensive coverage of countries, the ability to make strong inferences, and the identification of ‘deviant’ countries or ‘outliers’.

Comparing many countries is referred to as ‘variable-oriented’, since its primary focus is on ‘general dimensions of macro-social variation’ (Ragin 1994: 300) and the relationship between variables at a global level of analysis. The extensive coverage of countries allows for stronger inferences and theory-building, since a given relationship can be demonstrated to exist with a greater degree of certainty. For example, Gurr (1968: 1015) demonstrates that levels of civil strife across 114 countries are positively related to the presence of economic, political, short-term, and long-term deprivation. His analysis also explains that this relationship holds for roughly 65 per cent of the countries (see Chapter 5 and Sanders 1995: 69–73). More recently, Helliwell (1994) has shown that for 125 countries from 1960–1985 there is a positive relationship between per capita levels of income and democracy. After controlling for the differences between OECD countries, Middle Eastern oil-producing countries, Africa, and Latin America, this relationship is demonstrated to hold for about 60 per cent of the countries.

A second advantage of comparing many countries lies in the ability to identify so-called ‘deviant’ countries or ‘outliers’. These are countries whose values on the dependent variable (levels of civil strife or democracy in the examples above) are different than expected, given the values on the independent variables (levels of deprivation or per capita income). In testing for the positive relationship between income inequality and political violence in sixty countries, Muller and Seligson (1987: 436) use a simple scatter plot to identify which countries fit their theory and which

do not. For example, Brazil, Panama, and Gabon were found to have a lower level of political violence than was expected for the relatively high level of income inequality. On the other hand, the UK was found to have a particularly high level of political violence given its relatively low level of income inequality. By identifying these ‘outliers’, scholars can look for other explanations that account for their deviance, and they can remove them from their analysis to make more accurate predictions for the remaining countries. Thus, in this case the unexpected level of political violence observed for the UK was due to the Northern Ireland conflict. Such deeper analysis of outliers is also known as conducting ‘crucial’ case study (see below, p. 35).

Quantitative studies of many countries help in building general theories of politics since they allow other scholars to replicate their findings. The data sets for these studies can be read and analysed by a variety of statistical software packages. Scholars doing this kind of research often deposit their data in national data archives, such as the UK Data Archive at the University of Essex, the Roper Center at the University of Connecticut, the Inter-university Consortium for Political and Social Research (ICPSR) at the University of Michigan, and the Human Relations Area Files at Yale University. More recently, these data sets have been made available in files that can be downloaded from the Internet. In this way, new measures and new methods of analysis can be applied to these data to test the same theories or develop new theories. As a general rule, all scholars should strive to make their data public in an effort to keep a record of the progress of research, as well as help develop new understandings of politics.

Qualitative comparison of many countries is more difficult for two reasons. First, qualitative analysis generally requires a richer level of information, such as deep history of all the countries, which is often difficult to collect and synthesize. Indeed, Finer’s (1997) attempt to compare regime types over 5,000 years and across the globe represents a monumental task that occupied all the years of his retirement and produced a three-volume study with 1,700 pages. Second, it is more difficult to draw strong inferences from these data since they cannot be subjected to statistical analysis. Thus, Finer is able to describe and analyse different regime types as they have appeared in history to show how those in existence today are products of innovations from the past, but he is unable (or unwilling) to make any larger causal inferences. Even though he ‘privilege[s] those governmental innovations that are still relevant today’, he is adamant in stating that these regime types are not the product of a process of ‘linear evolution’ (*ibid.*: 88–89).

Despite the advantages of comparing many countries, there are some distinct disadvantages, including the availability of data, the validity of measures, and the mathematical and computing skills needed to analyse data. First, collecting relevant data on the independent nation states of the world can be difficult and time-consuming. Aggregate data are often published only for selected years or selected countries, making comprehensive comparison difficult. In the past, students had to rely on statistical abstracts and yearbooks produced by governments and international organizations, but the advent of the Internet has made the search for data much easier. By using careful search terms on any of the search ‘engines’ on the Internet (e.g. Lycos, Excite, Magellan, Google, Metacrawler), students can locate official statistics produced all over the world that can be downloaded quickly.

Second, measuring concepts from political science is difficult and can affect the validity of the measures. Valid measures closely approximate the true meaning of a concept, or what the researcher thinks he or she is measuring (King *et al.* 1994: 25). For example, the literature on economic development and democracy (see Chapter 4) tends to measure economic development with a country's level of per capita gross domestic product. But some argue that this measure does not take into account the distribution of income, which is also needed in order to capture the nature of a country's level of development. Democracy is also measured in a variety of ways. Freedom House (e.g. 1995) uses abstract scales that measure the degree to which political and civil liberties are protected. Vanhanen (1997: 35) measures democracy with an index that combines the vote share of the smallest party with the level of electoral turnout. Banks (1994) measures the presence of democratic institutions, including the competitiveness of the nomination process, executive effectiveness, legislative effectiveness, legislative selection, and party legitimacy. Many argue that this plethora of democratic measures highlights problems of validity.

Many students eschew quantitative comparison of many countries since it requires mathematical and computing skills. Statistical analysis of data requires an understanding of basic four-figure mathematics, algebra, probability theory, and calculus. It also requires knowledge of computers, spreadsheets, and statistical software packages. In response to these worries of students, there are several important things to consider. First, many undergraduate and most graduate programmes in political science require their students to take courses in statistics and political explanation, and some universities offer intensive data analysis training. Second, the development of computer technology combined with the availability of data makes this type of analysis much easier than in the past, and it is not unreasonable to assume that it will continue to do so. Third, a large portion of published literature in comparative politics uses quantitative analysis. Students who avoid learning even the basics can shut themselves off from important sources in the field. Thus, all students of comparative politics ought to achieve a basic understanding of the principles of quantitative analysis in order to evaluate studies that use it and employ it when appropriate (Collier 1991: 25).

An underlying assumption of statistical analysis is that events and facts in the world exhibit certain distributions, which can be described, compared, and analysed. But the comparison and analysis of these distributions of data is done from a collected sample of countries during specific periods of time. The comparison of the distributions is carried out in an effort to see if a relationship exists between them for the sample, and whether this relationship would hold for all countries in all periods of time. This basic practice of making inferences from a sample (some countries over one period) to a population (all countries in all time) lies at the heart of statistical analysis in comparative politics. This basic principle of statistical analysis can be demonstrated using a deck of playing cards (see Knapp 1996). A deck of playing cards has a known population of fifty-two cards. Each card has known characteristics, including the four suits (clubs, hearts, spades, and diamonds), the two colours (red and black), and the different values (Ace through King). There is thus a distribution of suits (thirteen cards in each), colours (twenty-six red cards and twenty-six black cards), and values (four cards of each value). Assuming that the entire deck of cards represents all countries for all time, it is possible to see how the examination of a

sample of cards from the deck could tell us much about the whole population. Using a sample of twenty cards from a well-shuffled deck, a student could get a first approximation of any of the distributions of a deck's attributes (suits, colours, and values). Replacing the sample, drawing repeated samples, and noting the distributions of the various characteristics would allow the student to get a more accurate picture of the whole deck. This process of sampling and inference is precisely what comparativists are trying to do when they collect and compare aggregate statistics from many countries.

Comparing few countries

Variouly called the comparative method, the 'comparable cases strategy' (Lijphart 1975), or 'focused comparison' (Hague *et al.* 1992), comparing few countries achieves control through the careful selection of countries that are analysed using a middle level of conceptual abstraction. Studies using this method are more intensive and less extensive since they encompass more of the nuances specific to each country. The political outcomes that feature in this type of comparison are often seen to be 'configurative', i.e. the product of multiple causal factors acting together. This type of comparison is thus referred to as 'case-oriented' (Ragin 1994), since the country is often the unit of analysis, and the focus tends to be on the similarities and differences among countries rather than the analytical relationships between variables. Comparison of the similarities and differences is meant to uncover what is common to each country that accounts for the observed political outcome.

The method of comparing few countries is divided primarily into two types of system design: 'most similar systems design' and 'most different systems design' (Przeworski and Teune 1970; Faure 1994). Most similar systems design (MSSD) seeks to compare political systems that share a host of common features in an effort to neutralize some differences while highlighting others. Based on J.S. Mill's (1843) method of difference, MSSD seeks to identify the key features that are different among similar countries and which account for the observed political outcome. Most different systems design (MDSD), on the other hand, compares countries that do not share any common features apart from the political outcome to be explained and one or two of the explanatory factors seen to be important for that outcome. This system is based on Mill's method of agreement, which seeks to identify those features that are the same among different countries in an effort to account for a particular outcome. In this way, MDSD allows the researcher to distil out the common elements from a diverse set of countries that have greater explanatory power (Collier 1993: 112).

Table 2.1 clarifies the distinction between these two systems and shows to which of Mill's methods they adhere. For MSSD on the left-hand side of the figure, the countries share the same basic characteristics (*a*, *b*, and *c*), and some share the same key explanatory factor (*x*), but those without this key factor also lack the outcome which is to be explained (*y*). Thus, the presence or absence of the key explanatory factor is seen to account for this outcome, a state of affairs that complies with Mill's method of difference. For MDSD on the right-hand side of the figure, the countries have inherently different features (*a* through *i*), but share the same key

Table 2.1 Most similar systems design (MSSD) and most different systems design (MDSD)

	MSSD Difference [†]			MDSD Agreement [†]		
	Country 1	Country 2	Country Φ	Country 1	Country 2	Country Φ
Features	a b c	a b c	a b c	a b c	d e f	g h i
Key explanatory factor(s)	x	x	not x	x	x	x
Outcome to be explained	y	y	not y	y	y	y

Source: Adapted from Skocpol and Somers (1980: 184)

Note: [†] Based on J. S. Mill's (1843) method

explanatory factor (x) as well as the presence of the outcome to be explained (y). In this system, the outcome to be explained is due to the presence of the key explanatory factor in all the countries (x), and thus adheres to Mill's method of agreement. In both systems, the presence of x is associated with the presence of y , and some would argue that x actually causes y . The difference between the two systems resides in the choice of countries.

Most similar systems design is particularly well suited for those engaged in area studies (Przeworski and Teune 1970: 33). The intellectual and theoretical justification for area studies is that there is something inherently similar about countries that make up a particular geographical region of the world, such as Europe, Asia, Africa, and Latin America. Whether it is common history, language, religion, politics, or culture, researchers working in area studies are essentially employing most similar systems design, and the focus on countries from these regions effectively controls for those features that are common to them while looking for those features that are not. For example, Jones (1995) compares the institutional arrangements of Latin American countries, which not only share the same cultural and historical Iberian legacies, but also share the same basic form of presidentialism. Similarly, Collier and Collier (1991) compare the experiences of eight Latin American countries to uncover the 'critical junctures' during which labour movements were incorporated into the political system.

Where quantitative analysis requires mathematical and computer skills, area studies require language training and extensive field research. Thus, some see these requirements as distinct disadvantages to comparing countries from a given region. It can take years to learn the languages needed to compare countries in Asia or Africa. Even within Latin America, students must learn Spanish and Portuguese, let alone the various dialects of each that are spoken in different parts of the region. Extensive

Briefing box 2.1 Most similar and most different systems design

Both system designs are used in comparative politics, particularly by those who compare few countries. Both these examples show how Mill's methods of agreement and difference can be applied to research questions. The first example shows how the most similar systems design is applied to six Latin American countries in an effort to uncover the sources of peasant support for revolutionary activity. The second example shows how the most different systems design is used to account for different regime types in fourteen European countries during the inter-war period.

Most similar systems design (MSSD): sources of peasant support for guerrillas

As part of a more comprehensive effort to account for revolutionary activity in Latin America between 1956 and 1970, Wickham-Crowley (1993: 92-117) uses the most similar systems design to examine the type of peasants that are most likely to support guerrillas in the region. Drawing on the work of Jeffery Paige (1975), he argues that guerrilla strongholds and support for revolutionary behaviour ought to be higher in rural areas in which there are peasants whose livelihood is the most vulnerable to negative influences from the structure of the agricultural system of production. His hypothesis is stated as follows:

If the guerrillas gain support in an area with a relatively high prevalence of sharecroppers, squatters, or perhaps tenants, my working assumption is that there is an 'elective affinity' between the two, and that guerrillas would not have received such support in more ordinary agricultural regions.

(Wickham-Crowley 1993: 95)

To test the hypothesis, he compares the regional breakdown of Cuba, Venezuela, Guatemala, Colombia, Peru, and Bolivia to determine whether such a relationship

Table 2.2 Most similar systems design

Case	Cuba	Venezuela	Guatemala	Colombia	Peru	Bolivia
Key peasant groups	Squatters	Share-croppers	Tenants	Share-croppers	Serfs	Small-holders
Outcome to be explained	Guerrilla support	Guerrilla support	Guerrilla support	Guerrilla support	Guerrilla support	No guerrilla support

Source: Adapted from Wickham-Crowley (1993: 92-117)

Note: Cases cover the period 1956-1970

Table 2.3 Most different systems design

		Group 1			
		Britain	France	Switzerland	Belgium
Cases				The Netherlands	
Class alliance		Middle class vs. working class Liberalism	Middle class vs. working class Liberalism	Middle class vs. working class Liberalism	Middle class vs. working class Liberalism
Outcome					
		Group 2			
		Denmark	Norway	Sweden	Czechoslovakia
Cases					
Class alliance		Working class + middle peasantry Social democracy	Working class + middle peasantry Social democracy	Working class + middle peasantry Social democracy	Working class + middle peasantry Social democracy
Outcome					
		Group 3			
		Germany	Italy	Spain	
Cases					
Class alliance		Middle class + middle peasantry Fascism	Middle class + middle peasantry Fascism	Middle class + middle peasantry Fascism	Middle class + middle peasantry Fascism
Outcome					

Source: Adapted from Luebbert (1991)

exists. Table 2.2 summarizes the comparison and shows that in all the cases except Bolivia, there is the presence of both the specified types of peasants and the outcome to be explained. Bolivia has a prevalence of smallholders, who according to the theory are not likely to support guerrilla activity, and in this case, do not. Thus, across similar cases, the presence of the key explanatory factor is associated with the presence of the outcome to be explained.

Most different systems design: the origins of regimes in inter-war Europe

In seeking to account for the different regime types that emerged in twelve countries in Europe during the inter-war period, Luebbert (1991) claims that the key explanatory variable is the particular class alliance that formed within these countries. The three regime types include liberalism, social democracy, and fascism. The twelve countries are grouped according to these three outcomes and within each group, the countries share few features in common apart from the same class alliance and the same outcome. Thus, Luebbert matches the presence of a particular class alliance to a particular regime type. Table 2.3 summarizes this analysis, and shows that liberalism is the product of a strong middle class versus a weak working class. Social democracy is seen to be a product of an alliance between the working class and the middle peasantry. And fascism is seen to be a product of an alliance between the middle class and the middle peasantry. In this example, the most different systems design is applied to each group of countries.

field research can mean long periods living under adverse conditions to which the researcher is unaccustomed. Moreover, funding organizations may be less inclined to support projects that envision long periods of field research. These problems represent the practical considerations that all researchers confront, and they highlight the different ways in which comparative methods can be seen to be a function of the training and disposition of the researcher.

Most different systems design is typical of comparative studies that identify a particular outcome that is to be explained, such as revolutions, military coups, transitions to democracy, or 'economic miracles' in newly industrialized countries (Geddes 1990: 134-141). The countries that comprise these types of comparative studies are all instances in which the outcome occurs. For example, Wolf (1969) compares instances of revolutionary movements that had significant peasant participation in Mexico, Russia, China, North Vietnam, Algeria, and Cuba. Though these countries share few common features, Wolf argues that the penetration of capitalist agriculture is the key explanatory factor common to each that accounts for the appearance of the revolutionary movements and their broad base of peasant support. As the next chapter will show, this kind of intentional choice of countries based on the presence of the same outcome constitutes one form of 'selection bias' (Geddes 1990; King *et al.* 1994), which necessarily limits the types of inferences that can be drawn from comparison.

Some comparativists use both system designs. In *Problems of Democratic Transition and Consolidation*, Linz and Stepan (1996) use MSSD to compare the

experiences of democratic consolidation within the separate regions of South America, Southern Europe, and Eastern Europe; and then use MDSD to compare across these three regions. Similarly, Rueschemeyer *et al.* (1992) use MSSD to examine the relationship between capitalist development and democracy within Latin America, and MDSD to compare Latin America and the advanced industrial world. De Meur and Berg-Schlosser (1994) employ both designs to analyse the conditions of survival or breakdown of democratic systems in inter-war Europe. What remains important to all these methods of comparing few countries is the proper specification of the outcome that is to be explained, the reason for adopting either system design, as well as the choice of the particular countries under scrutiny (see Chapter 3).¹

Single-country studies as comparison

As outlined above, a single-country study is considered comparative if it uses concepts applicable to other countries, develops concepts applicable to other countries, and/or seeks to make larger inferences. What should be recognized is that inferences made from single-country studies are necessarily less secure than those made from the comparison of several or many countries. Nevertheless, such studies are useful for examining a whole range of comparative issues. For Eckstein (1975), single-country studies are the equivalent of clinical studies from medicine, where the effects of certain treatments are examined intensively. Beyond this, however, single-country studies provide contextual description, develop new classifications, generate hypotheses, confirm and infirm theories, and explain the presence of deviant countries identified through cross-national comparison. This section of the chapter will consider these in turn.

As outlined in Chapter 1, one of the goals of comparison is contextual description. Single-country studies that merely describe or interpret political phenomena have been variously referred to as 'atheoretical' and 'interpretative' (Lijphart 1971: 691), or configurative-idiographic (Eckstein 1975: 96). Strictly speaking, these types of studies are not comparative but are useful for comparison purely for their information. But single-country studies that provide new classifications are useful for comparison. For example, in describing the Franco regime in Spain, Juan Linz (1964) identified a new form of authoritarianism that was different from personalistic dictatorships and totalitarian states. The regime institutionalized representation of the military, the Catholic Church, and the Falange, as well as the Franco loyalists, monarchists, and technocrats. Unlike totalitarian states, the regime relied on passive mass acceptance rather than popular support (Linz 1964; Carr and Fusi 1979: 31–35; Foweraker and Landman 1997: xxiii). Similarly, Guillermo O'Donnell (1973) established the concept of the 'bureaucratic-authoritarian state' in his examination of Argentine politics, a concept which would later be applied not only to other authoritarian regimes in Latin America but also to those in Southeast Asia.

Single-country studies are also useful for generating hypotheses for theories that have yet to be specified fully. As 'plausibility probes' (Eckstein 1975: 108), they either explicitly or implicitly suggest that the generated hypothesis be tested in a larger selection of countries (Lijphart 1971: 692). Again, O'Donnell's (1973) work on authoritarianism is illustrative. To account for the 1966 military coup and

subsequent authoritarian regime in Argentina, O'Donnell posited a relationship between a particular stage of dependent capitalist development and the advent of the bureaucratic-authoritarian state. This hypothesis was subsequently tested in other Latin American countries and was found wanting on many grounds (see Collier 1979). The point remains, however, that the hypothesis generated from the Argentine case was stated in such a way that other scholars could test it for other countries, and its rejection led to the search for rival explanations (see Cohen 1987, 1994).

When someone gives a lecture using comparative evidence from many countries, a member of the audience may exclaim, 'But in my country, things are different!' This is undoubtedly true, but more importantly the comment illustrates how single-country studies can be used to confirm and infirm existing theories, or illuminate known deviant countries. Theory-confirming and theory-infirming studies are conducted within the confines of known generalizations (Lijphart 1971: 692) and they often adopt the 'least likely' or 'most likely' method of comparison (Eckstein 1975: 118). Least likely studies find a country where the theory suggests the outcome is not likely to occur. If the outcome is not observed, then the theory is confirmed. Most likely studies are conducted in countries where the theory suggests the outcome is definitely meant to occur. If the outcome is not observed, then the theory is infirmed. These crucial country studies do not definitively prove or disprove a theory, but merely confirm or infirm its applicability to other countries.

Finally, deviant country studies are particularly useful for theory generation. As outlined above, comparison of many countries often reveals a host of deviant countries that do not conform to the theoretical expectations of the researcher. This deviance invites further research of the countries to establish which rival explanations had not been considered, and it forces the re-evaluation of how the key variables of the study were originally operationalized. Deviant country studies can weaken existing theories as well as further refine the concepts and measures used in the original comparative analysis (Lijphart 1971: 692). The United States, China, and Brazil represent excellent examples of deviant countries for different research questions. For the United States, comparativists seek to explain the absence of a large socialist party (Lipset and Marks 2000); for China, the survival of the communist regime after the 1989 'velvet revolutions' in Central and Eastern Europe (Hague *et al.* 1992: 37–38); and for Brazil, the absence of a social revolution given its poor distribution of income. All three countries represent a state of affairs that defies predominant theories in comparative politics.

Conclusion

This chapter has shown that all three methods – comparing many countries, comparing few countries, and single-country studies – should be grouped under the umbrella of comparative politics if they seek to make generalizations through explicit comparison, or if they use and develop concepts applicable to other countries through implicit comparison. Comparing many countries is the best method for drawing inferences that have more global applicability. Through use of the method of difference and method of agreement, comparing few countries can lead to inferences that are better informed by the contextual specificities of the countries under scrutiny.

Single-country studies can provide contextual description, generate hypotheses, confirm and infirm theories, and enrich our understanding of deviant countries identified through other comparisons. Finally, the chapter has made it clear that different strategies of comparison should be seen as the product of the trade-off between the level of conceptual abstraction and the scope of countries, as well as the arbitrary and practical factors surrounding any comparative research project. The next chapter examines the process of choosing countries, the main problems associated with comparison, and summarizes the main arguments of the first three chapters.

Note

- 1 Despite the prevalence of such comparisons in the field, there are four underlying assumptions to these two methods which, if violated, reduce their ability to make valid inferences. The research must assume: (1) a deterministic explanation, rather than a probabilistic one, (2) no errors in measurement, (3) the existence of one cause, and (4) the absence of interaction effects (see Liebersohn 1991, 1994; Savolainen 1994).

Further reading

Collier, D. (1991) 'New Perspectives on the Comparative Method', in D.A. Rustow and K.P. Erickson (eds) *Comparative Political Dynamics: Global Research Perspectives*, New York: Harper Collins, 7–31.

An excellent review of comparative methods.

Eckstein, H. (1975) 'Case-study and Theory in Political Science', in F.I. Greenstein and N.S. Polsby (eds) *Handbook of Political Science*, Vol. 7: Strategies of Inquiry, Reading, MA: Addison-Wesley, 79–137.

The most comprehensive review of the scientific value of single-country studies.

Faure, A.M. (1994) 'Some Methodological Problems in Comparative Politics', *Journal of Theoretical Politics*, 6(3): 307–322.

This essay outlines the most similar and most different systems design as well as their 'mirror images'.

Lijphart, A. (1971) 'Comparative Politics and Comparative Method', *The American Political Science Review*, 65(3): 682–693.

The original statement about comparative method, locating it as a non-experimental and non-statistical social science.

— (1975) 'The Comparable Cases Strategy in Comparative Research', *Comparative Political Studies*, 8(2): 158–177.

This essay presents further reflections on comparative method.

Mackie, T. and Marsh, D. (1995) 'The Comparative Method', in D. Marsh and G. Stoker (eds) *Theory and Methods in Political Science*, London: Macmillan, 173–188.

A brief overview of comparative methods.

Ragin, C. C. (1994) 'Introduction to Qualitative Comparative Analysis', in T. Janoski and A. Hicks (eds) *The Comparative Political Economy of the Welfare State*, Cambridge: Cambridge University Press, 299–320.

This essay distinguishes between 'variable-oriented' and 'case-oriented' approaches and proposes a way to unify them.

Sartori, G. (1970) 'Concept Misinformation in Comparative Politics', *American Political Science Review*, 64: 1033–1053.

The classic statement on 'conceptual stretching' and the 'ladder of abstraction'.

— (1994) 'Compare Why and How: Comparing, Miscomparing and the Comparative Method', in M. Dogan and A. Kazancigil (eds) *Comparing Nations: Concepts, Strategies, Substance*, London: Basil Blackwell, 14–34.

A restatement of the main claims in 1970 and that 'to compare is to control'.

Skocpol, T. and Somers, M. (1980) 'The Uses of Comparative History in Macrosocial Inquiry', *Comparative Studies in Society and History*, 22: 174–197.

This essay outlines the uses of comparative history as well as Mill's methods of agreement and difference.

Choosing countries and problems of comparison

Too many variables and too few countries	40
Establishing equivalence	43
Selection bias	46
Spuriousness	51
Ecological and individualist fallacies	53
Value bias	55
Conclusion	56
Further reading	58

The preceding two chapters made it clear why and how to compare countries. Scholars compare to provide context, make classifications, test hypotheses, and make predictions. They do this by comparing many countries, few countries, or they provide in-depth studies of single countries. As there are many trade-offs associated with these different goals and methods of comparison, there are also important fundamental problems, which if not addressed explicitly can limit the types of generalizations that can be drawn from any study. While not representing insurmountable obstacles to comparison, it is important to address these problems and outline the strategies for overcoming them. This chapter discusses six complementary problems of comparison, which are associated with the choice of countries, the manner in which they are compared, the structure of the research design, and the nature of the evidence.

The first is the problem of too many variables and not enough countries (Collier 1991; Dogan and Pelassy 1990; Hague *et al.* 1992), also known more generally as 'too many inferences and not enough observations' (King *et al.* 1994: 119). This problem arises when more factors of explanation for the observed outcome have been identified than there are countries (or observations) in the study, leading to an indeterminate research design. Clearly this problem tends to be associated more often with single-case studies and those that compare few countries than with those studies that compare many countries. The second problem is one of establishing equivalence both in the theoretical concepts that are used and the operational indicators of those concepts as they are applied in multiple contexts (Sartori 1970; Macintyre 1971; Mayer 1989). For example, the concept of political participation may mean very different things across different contexts, such as voting in one country, or mobilizing activists against nuclear power in another.

The third problem of selection bias arises from the intentional choice of countries (Liebersohn 1987; Geddes 1990; Collier 1995; King *et al.* 1994), as well as the use of historical accounts and sources that favour the particular theoretical position of the comparatist (Lustick 1996). The fourth problem is spuriousness, or the omission of key variables that may account for both the outcome and other explanatory factors already identified. The fifth problem – ecological and individualist fallacies – arises when a study seeks to make inferences about one level of analysis using evidence from another (Robinson 1950; Scheuch 1966, 1969; Miller 1995). For example, a theory of revolution may concentrate on individual psychological factors that account for rebellious behaviour, but the comparison to test the theory may use aggregate statistics across countries on levels of inequality and instances of political violence. The final problem for all comparatists to consider is that of value bias, where the particular cultural, political, and philosophical predisposition of the researcher necessarily biases the conduct and conclusions of the enquiry.

Too many variables and too few countries

This problem of comparison is illustrated initially with two simple examples, one from simple algebra and one from introductory economics. It is then illustrated using a hypothetical example from political science. Algebra courses often present simple equations that take the following form:

$$x + 5 = 10 \quad [1]$$

In this equation, x is some unknown, whose value is solved by subtracting 5 from 10. A slightly more complicated problem would include two unknowns and takes the following form:

$$y = x + 10 \quad [2]$$

$$2x = y + 35 \quad [3]$$

In equations [2] and [3], the values of x and y are not immediately known; however, by combining the two equations through substitution, it is possible to solve for both x and y . Once the value of x has been determined, the value of y can be determined. The steps for this process are as follows:

$$2x = (x + 10) + 35 \quad [4]$$

$$2x = x + 45 \quad [5]$$

$$x = 45; y = 55$$

Similarly in economics, the price and quantity of any good in a market at equilibrium is a function of its supply and demand. Goods in short supply fetch a higher price than goods in abundance, and goods in high demand are more expensive than goods in low demand. If there is an upward shift in demand for a product, then a firm raises the price until it can produce more. Similarly, if a firm produces too much of a good, it is forced to lower its price until the excess supply is sold. Knowing only the supply or demand function for a particular good could not allow the market price or quantity to be determined. As in the algebra example above, the supply and demand curves can be approximated using equations for straight lines. The market price and quantity of a good are determined by setting the two equations equal to one another, which is the same thing as saying that they intersect. Thus, given specific demand and supply equations, the market price and quantity can be derived.

In both the algebra and economic examples, the idea of a system of two equations is similar to the problem of two many variables (or inferences) and not enough countries (or observations). On its own, equation [2] above is meaningless, and x and y can have any number of values that would satisfy it. Similarly, a demand equation without its complementary supply equation is equally meaningless if one wants to know both the quantity and price at market equilibrium. In comparative politics, if a study has too many unknowns (i.e. inferences or possible explanations) and not enough equations (i.e. countries or observations), then solving for the unknowns is problematic. Consider the following hypothetical example from political science. A scholar wants to know which factors are crucial for explaining high public expenditure. After reviewing the relevant literature, it is posited that public expenditure is high in wealthy countries controlled by left-of-centre governments. In this example, there is one dependent variable, public expenditure, and two independent variables, partisan control of government and wealth of the country. Logically, there are four possible combinations of the two independent variables (Figure 3.1). It would be impossible for a scholar to know the effects of

		Wealth of country	
		Poor	Rich
Partisan control of government	Left	Country A Left-poor	Country B Left-rich
	Right	Country C Right-poor	Country D Right-rich

Figure 3.1 Logical combination of two variables in four countries

these variables on the level of public expenditure if the comparison only looked at two countries or less. For example, if a left-poor country is compared to a left-rich country, partisanship is not allowed to vary. Similarly, if a left-rich country is compared to a right-rich country, then wealth is not allowed to vary. Adding a third case to either comparison (e.g. a right-poor country), allows both variables to vary, and the hypothesis can be tested with a determinate research design.

In extending this logic to an example from the last chapter (Table 2.2), Wickham-Crowley (1993) could not know the explanatory relevance of the type of peasant if he only looks at peasants in one country. Similarly, Luebbert (1991) could not know the likely outcome of class alliances if he limited his study to Britain. In general, a study that has too many variables and not enough countries makes explanation of the outcome problematic. Although this problem is more frequent in single-case studies and those studies that compare few countries, it can also arise in those that compare many countries since there is a relatively small and finite number of them in the world (Hague *et al.* 1992: 27).

There are three solutions to the problem of too many variables and not enough countries, all of which are based on the principle that the number of variables (or inferences) must be less than the number of countries (or observations) (King *et al.* 1994: 119–122). The first solution is to raise the number of observations to allow the key factors of the study greater overall variation, sometimes referred to as ‘degrees of freedom’. This can be achieved by comparing instances of the political phenomenon and its hypothesized explanatory factors over time, by adding more countries to the study, or by comparing sub-units of the nation under scrutiny. Recent work in comparative politics has sought to compare many countries over many years using techniques in so-called ‘pooled cross-section time series analysis’ (see Stimson 1985; Beck and Katz 1995). Such analysis ‘pools’ repeated observations of countries by collecting country data for long periods of time. As the discussion in Chapter 4 shows, Burkhart and Lewis-Beck (1994) compare 131 countries from 1972 to 1989, yielding a total of 2,358 observations, while Landman (1999) compares seventeen Latin American countries from 1972 to 1995 for total sample size of 408 observations. In each example, pooling the comparison of countries over time raises the number of observations. In studies that compare few countries, more instances of the phenomenon are drawn from history to increase the number of observations, and in single-case studies, sub-units or regions within the nation are compared, such as Putnam’s (1993) study of democratic performance across the regions of Italy or Hagopian’s (1996) study of patrimonial politics in Brazil.

The second solution to the problem is to use the most similar systems design (MSSD) to achieve focused comparison of few countries. As was outlined in the last chapter, the MSSD framework seeks to control for those factors that are similar across the countries in the study, while focusing on only those factors that are different that account for the outcome. Again, this strategy of comparison underlies the justification for area studies, but some argue that the MSSD framework simply provides ‘overdetermined’ outcomes (Przeworski and Teune 1970; Collier 1991: 17), where many rival explanations are never truly eliminated. Another criticism of the MSSD framework involves one of perspective, in that similarities for one researcher may be differences for another, effectively lending little value to the approach (Collier 1991; King *et al.* 1994). Despite these criticisms, area studies continue to be carried out with the implicit or explicit reference to the MSSD framework.

The third solution is to reduce the number of variables by focusing on the key explanatory factors that are hypothesized as important for explaining the outcome. This can be achieved either by using the most different systems design (MDSD) or by having stronger theoretical specifications. Recall that the MDSD framework intentionally compares a diverse set of countries, while concentrating on their key similarities. For example, Opp *et al.* (1995) compare the relationship between left–right ideological positions and support for social protest in Germany, Peru, and Israel. For them, the comparison of such different countries allows for a rigorous test of their main theoretical propositions (*ibid.*: 71–72). In applying a variation on MDSD, Parsa (2000) compares the social revolutions in Iran, Nicaragua, and the Philippines. All three countries shared ‘similar experiences and structural features’ (economic development, authoritarian rule, strong states, and US support), yet ‘unlikely challengers were able to seize power’ and the immediate outcomes of each revolution were different (*ibid.*: 3–4, see Chapter 6 this volume). In addition to comparing most different countries as in these two examples, a strong theory can highlight a parsimonious set of explanatory factors that can travel across space and over time. For example, the ‘rational choice’ perspective examines the role that ‘selective incentives’ play in the motivations of individuals to become involved in collective action. Such attention to selective incentives has been used to account for the actions of revolutionary peasants across the globe and over the centuries (see Lichbach 1994, 1995).

Establishing equivalence

The second problem confronting comparativists is the equivalence of both their theoretical concepts and the indicators for those concepts across multiple contexts. Mayer (1989: 57) argues that ‘the contextual relativity of the meaning or the measures of indicators constitutes the most serious impediment to the cross-contextual validity of empirically testable explanatory theory’. In other words, is it possible to specify concepts and indicators that have shared meanings to allow valid comparisons? For example, does the concept of class apply equally in all societies? Does the idea of ‘civic culture’ (Almond and Verba 1963) mean the same thing in Brazil as it does in France? Is it possible to have ‘new’ social movements in Latin

America (Fuentes and Frank 1989; Escobar and Alvarez 1992)? Does it mean the same thing when a British MP votes against his or her party as when a US Senator votes against his party (Hague *et al.* 1992: 29)? The crux of the problem is not specifying identical, or even similar concepts, but *equivalent* ones so that their comparison is meaningful (Dogan and Pelassy 1990; Sartori 1994).

There are three intellectual positions that offer insight into this problem: (1) the universalist position, (2) the relativist position, and (3) the middle position. The universalist position argues that if theoretical concepts and their indicators are to have any explanatory power, they must be able to travel to all parts of the globe. For example, rationalist, functionalist, and structuralist approaches take such a position. Rationalists argue that all individuals maximize their own personal utility given a set of preferences and confronting a range of choices (Ward 1995). Functionalists argue that 'certain vital functions', such as interest articulation and interest aggregation, are 'fulfilled everywhere' (Dogan and Pelassy 1990: 42). Structuralists argue that macro-structures such as the state, economic development, and social classes are omnipresent, but exist in varying degrees and are responsible for determining political outcomes.

The relativist position argues that all meaning is locally determined, and that a general 'science' of comparative politics is necessarily limited if not impossible (Macintyre 1971). Ethnographic, interpretivist, and anthropological approaches tend to take this position (see Geertz 1973; Scott in Kohli *et al.* 1995). In a critique of Almond and Verba's (1963) study of political culture in Italy, Germany, Britain, the United States, and Mexico, Macintyre (1971: 173) argues that indicators of commitment to government were never sufficiently examined to account for their cross-cultural differences in meaning. Thus, substantive comparison of these countries and the generalizations about civic culture must be treated with suspicion. Although not an extreme relativist, Sartori (1970, 1994) argues that 'stretching' a concept too far dilutes its meaning and precision, suggesting, that once defined and operationalized, certain concepts can only travel so far.

The middle position argues that comparativists must not abandon all their concepts, but should modify them to be more sensitive to the cultural specificities of the contexts they are studying. In *Theorizing Social Movements*, Foweraker (1995) seeks to modify the North American rationalist and European culturalist perspectives on social movements to explain the patterns of social mobilization in Latin America. Key factors of explanation from the rationalist perspective (interests, strategies, micro-mobilization, and political opportunity structure) are combined with culturalist concerns of identity and expression in discussing the various origins, trajectories, and outcomes of Latin American social movements. Some comparativists consider themselves 'opportunists' as they modify, combine, and reconstitute concepts to fit the cases under study (Przeworski in Kohli *et al.* 1995: 16), and argue that wilful sacrifice of insights from different perspectives may obscure important explanatory factors (Katzenstein in Kohli *et al.* 1995: 15).

Since the relativist position obviates the reason for comparative politics, this chapter provides common solutions for those seeking to make larger inferences through comparison (i.e. those adhering to the universal and middle positions). These solutions include raising the level of abstraction (Sartori 1970), focusing on smaller numbers of countries for which the comparativist has thorough substantive

knowledge (Sanders 1994), using 'specialist teams' in compiling cross-national data sets (*ibid.*), and specifying the functional equivalence between concepts or indicators (Dogan and Pelassy 1990). As in the solutions to the problem of 'too many variables not enough countries', there are important trade-offs associated with each of these solutions. The key to all is careful specification of concepts, thoughtful construction of indicators that operationalize them, careful application of them to multiple contexts, and recognition of their limitations.

In returning to the distinctions made in Chapter 1, raising the level of abstraction allows a study to be more inclusive, while lowering the level of abstraction makes it more exclusive. For example, in the comparative study of public administration, Sartori (1970: 1042) argues that the term 'staff' is abstract enough to travel universally, 'administration' to all societies that have the presence of some form of bureaucracy, and 'civil service' to all societies with a fully developed modern state. In this way, as the level of abstraction decreases from 'staff' to 'civil service', the number of eligible countries for comparison necessarily decreases as well. Finer (1997: 78) adopts terms that will travel through space and over time. His 'master variables' for classifying the world's regimes include territory (city, country, or empire), type of regime (palace, forum, nobility, church, and hybrids), the presence or absence of a standing military or civil bureaucracy, and the substantive and procedural limitations on the activities of rulers. Inglehart (1997) seeks to apply two-value continua to forty-three countries, which range on the one hand from citizens' concerns with 'survival' vs. 'well-being' to, on the other, their concerns with 'traditional' vs. 'legal-rational' forms of authority. In the latter two examples, important concepts are specified in such a way as to incorporate many countries.

The second solution – focusing on a small set of countries for which the comparativist has thorough substantive knowledge – suggests that the analyst be 'extremely cautious about engaging in cross-national comparative research' (Sanders 1994: 43). The explanatory power of concepts can be enhanced if they are applied in contexts with which the comparativist is most familiar. Thus, those who engage in area studies spend many years studying the history, economics, politics, and culture of a regional sub-set of countries in an effort to make more meaningful explanations of political phenomena. This 'local' knowledge can identify gaps between theoretical concepts and their application, and result in more meaningful comparison. Sanders (*ibid.*: 48) summarizes this point precisely:

It is only with detailed substantive knowledge that analysts can make informed judgements either about the relevance of the characterizations that they make of particular systems or about the identity of meaning attached to the questions that they pose to people living in different countries.

The third solution necessarily follows from the second. If truly informed comparison of many countries is limited, then those seeking to compare many countries 'should venture out of the security of the familiar if they are prepared to collaborate with other scholars' who possess specialist knowledge of the countries under scrutiny (Sanders 1994). This solution was used by Fitzgibbon (1967), who sought to measure democratic change in Latin America by using a questionnaire to measure general social and political factors he believed were both preconditions and

manifestations of democracy. The questionnaire was sent to leading academics working in specific countries and regions in Latin America and was repeated at five-year intervals between 1945 and 1985. The resulting 'image index' is highly correlated with similar such measures (Foweraker and Landman 1997: 61 fn. 14; Chapter 4). Another example that follows Sanders' prescription is Inglehart's (1997) World Values Survey, which uses local specialist teams to implement a similar survey in forty-three countries. It is also common practice in the human rights community to produce world reports on human rights protection such as the Amnesty International *Annual Reports*, the US State Department *Country Reports*, or Human Rights Watch *World Report*. These reports can then be used for secondary analysis, such as Poe and Tate's (1994) global analysis of the repression of human rights (see Chapter 9 in this volume).

The final solution is the identification of 'functional equivalence' of concepts and indicators. This solution does not envisage concepts as identical or even similar, but functionally equivalent. If two entities share exactly the same qualities, properties, and characteristics, they are considered identical (apples are apples). If they share some qualities, properties, or characteristics, then they are said to be similar (apples and pears are fruit). If they share the same function, however, they are said to be functionally equivalent. For example, leaders of countries can serve three functions: symbolic representation of the nation, chief executive of state authority, and party leader. The French president embodies all three while the British monarch embodies the symbolic role and the British prime minister embodies the executive and party leader roles (see Dogan and Pelassy 1990: 37). Depending on the functional focus and political systems of the comparison, the study may include an examination of one, two, or three individuals. Thus, functional equivalence allows entities with seemingly dissimilar characteristics to be grouped into useful and exclusive categories. In general, the analyst must specify clearly in which respect the concept is comparable.

Selection bias

A crucial scientific principle frequently violated by comparative politics is the principle of selection. Comparison seeks to achieve experimental simulation, but experiments and mass attitudinal surveys in political science use *random selection* of individual respondents, while the essence of much of comparative politics is the *intentional selection* of countries. The basic experimental form has an experimental group and a control group. The experimental group receives the 'treatment' (stimulus, drug, or exposure to some independent factor), and the control group does not. The outcome of both groups after treatment is then compared. If the experimental group exhibits a different outcome than the control group, it is attributed to the treatment, given that all else is equal (known as the *ceteris paribus* condition). In mass attitudinal surveys, a completely random sample of individuals is selected and the subsequent data analysis of responses yields substantive inferences about the whole population from which the sample is drawn (see de Vaus 1991). In studies of electoral behaviour, a frequent finding is that those from a lower social class tend to vote for left-of-centre political parties while those of higher social class tend to vote for right-of-centre parties. The analysis of the survey data *compares* groups of

individuals from each social class and determines the effects of that difference on their preference for particular political parties.

In both these examples, *the selection of individuals or units of analysis is not related to the outcome to be explained*. Selection bias in comparative politics occurs through the non-random choice of countries for comparison, or the deliberate selection by the comparativist (Collier 1995: 462). Though selection of countries lies at the heart of comparison, selection without reflection may lead to serious problems of inference. The most blatant form of selection occurs when a study includes only those cases that support the theory. More subtle forms of selection bias, however, occur when the choice of countries relies on values of the dependent variable (Geddes 1990; King *et al.* 1994) and for qualitative studies, both the use of certain historical sources (Lustick 1996) as well as exclusive focus on contemporary political systems.

The problem of selection does not affect studies that compare many countries as much as those compare few countries, and it is a major problem for single-case studies. Studies that compare many countries usually have a sufficient number of observations to avoid the problem of selection, and quantitative studies of many countries can use a number of statistical techniques to eliminate the problem (see Gujarati 1988; Fox 1997). For studies that compare few countries and single-case studies, however, selection can seriously affect the type of inferences that are drawn. Frequently in these types of studies countries are chosen because they exhibit only the outcome the comparativist seeks to explain, such as a social revolution, a military coup, a transition to democracy, the failure of deterrence, or high economic growth rates (Geddes 1990; Collier 1995). Selecting on the dependent variable in this way can lead either to an overestimation of effects that do not exist, or to an underestimation of effects that do exist (Geddes 1990: 132–133). In other words, a study may claim that a set of explanatory variables is either more important in accounting for an outcome, or may neglect the importance of other explanatory variables. Both problems mean that the analysis is drawing false inferences.

Recall O'Donnell's (1973) study of the bureaucratic-authoritarian state (see Chapter 2 in this volume). O'Donnell tried to explain the advent of the bureaucratic-authoritarian state based on the case of Argentina in 1966. He argued that the presence of key independent factors – a collapse of a certain mode of dependent capitalist industrialization, economic stagnation, and an increase in popular demands – led the military to overthrow the democratic government, implement economic plans for recovery, and repress popular mobilization against the Argentine state. Subsequent research tested this theory both in Latin American countries that had similar experiences of authoritarianism and in countries that did not (Collier 1979). These studies showed that countries with similar authoritarian experiences did not share the same antecedent factors, while countries that sustained democracy did share these factors. Moreover, when the Latin American economy took another downturn in the early 1980s, no new instances of the bureaucratic-authoritarian state arose. Thus, the comparison across cases and time revealed that the strong connection between these independent factors and authoritarianism could not be upheld (Cohen 1987). O'Donnell's single case study overestimated the effect of the antecedent factors on the political outcome he observed (see Briefing Box 3.1). His results led him to issue a robust refutation of the thesis that economic development

Briefing box 3.1 The problem of selection bias*The rise of the bureaucratic-authoritarian state*

In explaining the rise of the bureaucratic-authoritarian state in Argentina, O'Donnell (1973) focused on two key explanatory factors: (1) the stagnation of the economy as measured by balance of payments deficits, low growth rates, rising inflation, and (2) the inability of the country to make the necessary transition from the 'easy phase' to the 'hard phase' of import-substitution-industrialization (ISI).

Under the easy phase of ISI, the state provided protection of the local economy with high tariffs and import quotas to allow new industries to develop the capacity to produce domestically what used to be imported from abroad. The policy included credit at concessionary rates, high wages for labour, and artificially high prices for traditional exports through manipulation of exchange rates. The hard phase of ISI, on the other hand, saw a shift to the domestic production of all intermediate goods necessary for finished capital goods, which was known as 'deepening' or 'vertical integration'. This phase required the attraction of foreign investment from multinational corporations, the loosening of tariff and quota restrictions, a reduction in wages, and a readjustment of exchange rates.

In the Argentine case, economic stagnation preceded the military overthrow in 1966 and 'deepening' of the economy occurred after the coup. From this chain of events, O'Donnell theorized a connection between the antecedent factors, the advent of the bureaucratic state, and the subsequent economic policy of deepening. This reasoning is depicted in column three of Table 3.1. Subsequent comparison to the cases of Brazil, Colombia, and Venezuela revealed that while all three experienced economic stagnation, two (Colombia and Venezuela) *did not* experience military coups, and one (Brazil) had *already* started a process of deepening *before* the military overthrew the democratic government in 1964. These contrasting cases are listed in columns four, five, and six of Table 3.1. Thus, by relying on only the case of Argentina, O'Donnell's theoretical conceptualization and explanation suffered from selection bias.

Table 3.1 Explaining the bureaucratic-authoritarian state in Latin America

		Argentina	Brazil	Venezuela	Colombia
Explanatory factor 1	Economic stagnation	Yes	Yes	Yes	Yes
Explanatory factor 2	Failure to make transition to hard phase of ISI	Yes	No	Yes	Yes
Outcome 1	Military coup and implementation of BA state	Yes	Yes	No	No
Outcome 2	Deepening of domestic economy	Yes	Yes (pre-coup)	Yes	Yes

Source: Adapted from O'Donnell (1973); Serra (1979)

causes democracy, while the case of Argentina turned out to be a serious outlier (see Chapter 4 this volume and Przeworski *et al.* 2000: 99–100).

In a less obvious but equally problematic example of selection bias, Skocpol (1979) compares countries that experienced social revolutions (Russia, China, and France) to contrasting countries where revolution did not occur (Japan, Prussia, and Britain) in an effort to demonstrate the explanatory relevance of certain structural factors to these revolutions. These structural factors include external military threats, regime reform, dominant class opposition, and state collapse (see Chapter 5). The contrasting cases did not share these factors and did not experience social revolutions. Geddes (1990) argues that the comparison to these contrasting cases is good but still limited, since these countries represent the other extreme of her dependent variable. The comparison confirms Skocpol's theory, but Geddes (1990: 143) asks, 'would a differently selected set of cases do so?' Comparison to the cases of Mexico, Guatemala, El Salvador, Honduras, Nicaragua, Ecuador, Peru, Bolivia, and Paraguay, which have similar structural factors and varying experiences with social revolution, would reveal the limits to the inferences about structures that Skocpol draws (*ibid.*: 144–145).

Collier (1999) compares seventeen historical cases and ten contemporary cases of democratization to examine the importance of working-class mobilization on the process of democratic reform. In all twenty-seven cases, the period of democratic reform pushed the countries 'across a threshold consistent with conceptualizing the political regime as democratic' (*ibid.*: 23), marked by an election and the installation of a new government. Collier is not concerned with the overall durability of the new regime, as many of the cases experience democratic breakdown later on, but she is interested in determining the role of labour mobilization in the reform process. The study is a curious example of selection bias since the dependent variable does not vary (all cases in the sample experienced democratic reform), the choice of countries depends on the outcome that is to be explained (historical and recent cases of democratic reform), and labour mobilization was present in some cases and absent in others. Collier (1999: 167) argues that based on these comparisons labour mobilization is not a 'decisive or even necessary, no less sufficient, factor in democratization'. But her study is an example of how an attempt to raise the number of observations by comparing many instances of democratic reform still yields an indeterminate research design. Like the problem outlined in Figure 3.1, Collier's problem can be depicted by a 2×2 matrix that is the product of the intersection between her two main variables: (1) labour mobilization (yes or no), and (2) democratization (yes or no). Her observations only cover half of all the possible combinations in the matrix (i.e. cases of democratic reform with or without labour mobilization). For a definitive rejection of the hypothesis that labour mobilization matters for democratization, she would ideally have to add cases to her sample that either (1) did not experience democratic reform and had labour mobilization, or (2) did not experience democratic reform and did not have labour mobilization. It could be that labour mobilization has a *negative* impact on democratic reform. Without adding examples of either of these two combinations of variables, her analysis suffers from indeterminacy stemming from a selection of cases on the dependent variable (see above discussion on too many variables and too few countries).

In both the O'Donnell and Skocpol examples, selection on the dependent variable led to an overestimation of the importance of certain explanatory factors, while in the Collier example, selection bias may have led to the underestimation of effects that do exist. In general, there are three solutions to the problem of choosing on the dependent variable. The first solution is to have a dependent variable that varies: i.e. countries in which the outcome has occurred and those in which it has not. Only by comparing across the presence and absence of outcomes can the importance of explanatory factors be determined. Second, when comparing few countries, the choice of countries ought to reflect substantive knowledge of parallel cases (Laitin 1995: 456). Third, stronger theory may specify more accurately a range of countries in which certain outcomes and their explanations would obtain (*ibid.*). Fourth, and related to the third solution, strong theory will also identify which countries represent 'least likely' (cf. Chapter 2) instances of the phenomenon under investigation (Caporaso 1995: 458). All four solutions demand close attention to the types of inferences that are being drawn when intentionally choosing countries for comparison.

A second form of selection bias arises in qualitative studies that rely on historical sources, where the analyst chooses historical accounts either intentionally or unintentionally whose description of events fits the particular theory being tested. As Lustick has pointed out, 'the work of historians cannot be legitimately treated as an unproblematic background narrative from which theoretically neutral data can be elicited for the framing of problems and the testing of theories' (Lustick 1996: 605). Historiography varies in its description of how the past actually unfolded, which events receive emphasis, as well as the different theoretical dispositions of the historians themselves. Thus, inferences drawn from studies using descriptive historical accounts that 'are organized and presented according to the categories and propositions of theories they are testing' will necessarily be biased (*ibid.*: 610). Solutions to this form of selection bias include using multiple sources to arrive at a 'mean' account of the events and identifying the tendencies within each source to acknowledge possible sources of bias.

A third form of selection bias can occur from the time periods that are used in the comparison, especially for those studies seeking to analyse social behaviour that has a very long history, such as warfare, trade, and the emergence of states and regimes. Selecting contemporary time periods (even those throughout the twentieth century) and drawing inferences about longer-term processes is a form of historical selection bias. In this sense, the selection is taking place at a particular time or at an arbitrary end to a time-line of events, and inferences drawn from such a comparison will necessarily be less secure (Geddes 1990). There are examples of studies in comparative politics and international relations that avoid such a problem of selection. As noted above, Finer (1997) compares ancient, medieval, and modern forms of government. Arrighi (1994) examines the relationship between capital accumulation and state formation over a 700-year period. Cioffi-Revilla and Landman (1999) analyse the rise and fall of Mayan city-states in ancient Mesoamerica from 2000 BC to AD 1521. Midlarsky (1999) examines the effects of inequality on state formation and warfare in ancient and modern societies. Finally, the work on the 'democratic peace' (see Chapter 2 in this volume) compares warfare 'dyads' from the middle of the nineteenth century to the late twentieth century. In each of these examples, there

is an attempt to provide generalizations about an important aspect of politics by comparing whole systems over long periods of time.

Spuriousness

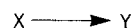
A spurious explanation is one in which some unidentified factor is responsible for the outcome, while the identified factor is mistakenly attributed to having an effect on the outcome. Also known as omitted variable bias (King *et al.* 1994: 168), this problem frequently arises in comparative politics and is related to selection bias since the choice of cases may overlook an important underlying factor that accounts for the outcome. Consider the following example. An industrious graduate student spends the summer holidays working in resorts around the United States. Over the years, the student recognizes that wherever he works, there appears to be both a high number of flamingos and retired people. He decides to spend his leisure time collecting data on the geographical distribution of flamingos and retired people. Cognizant of the problem of selection bias, the student extends the collection of data to include all the states in the US. After the data are collected, the student finds a positive correlation between the number of flamingos and the number of retired people. From these robust statistical results, he concludes that flamingos cause retired people. It is clear that the unidentified factor in this example is climate. On balance, both flamingos and retired people in the United States 'flock' to those areas with warmer climates. Thus, the mistaken connection between the two is due to the unidentified factor (see Briefing Box 3.2). By omitting the variable of climate, the student mistakenly concluded that flamingos *cause* retired people. If the student had only collected data in Florida, he may have reached the same conclusion, but one that was additionally influenced by selection bias.

In comparative politics, it has been frequently asserted that authoritarian regimes are better at promoting economic development than democratic regimes, since their 'relative autonomy' from society allows them to control more easily instances of political dissent. Global analysis of the relationship compares indicators of authoritarianism and economic performance and finds a strong positive association between the two. What these studies fail to identify, however, is that authoritarian governments tend to fall during periods of economic downturn, since much of their legitimacy rests on their ability to deliver economic benefits (Przeworski *et al.* 2000). Once discredited in economic terms, authoritarian regimes tend to lose their grip. Democracies, on the other hand, endure through periods of thick and thin. In terms of the overall relationship, this fact means that authoritarian regimes are only in power during times of good economic performance. Thus, by ignoring the important factor of regime 'attrition', the original finding in support of the connection between authoritarian regimes and economic performance is spurious (Przeworski and Limongi 1993, 1997).

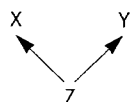
As seen above, the solutions to the problem of spuriousness are related to the number of countries in a comparative study; moreover the trade-offs associated with these solutions can often be a source of frustration. The easiest solution for spuriousness is to specify all the relevant variables that may account for the observed outcome. This solution is fine if the comparison is across many countries or many

Briefing box 3.2 Spuriousness

Simple explanations of events often take the form 'if event x then event y' (Sanders 1994, 1995; Lawson 1997), which can be depicted graphically as follows:

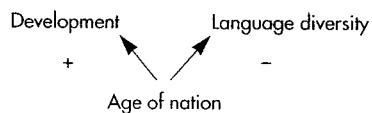


In this example, x and y are the only variables that have been identified. Suppose data collected on the occurrence of x and y shows that whenever (or wherever) x occurs, y also occurs. The regular and concomitant occurrence of both would lead to either the weak conclusion that x and y are *associated* with each other, or the strong conclusion that x actually causes y. But what if some other factor z also occurs regularly with x and y? The analyst risks specifying a relationship between x and y that may actually be the result of z acting on x and y independently. This situation is depicted as follows:



In this case, there is no direct relationship between x and y, but a common underlying factor to both, which explains their occurrence. Failure to specify this third variable and its effects on x and y constitutes the problem of spuriousness. The assertion that authoritarian regimes (x) are better at promoting economic development (y), failed to identify that authoritarian regimes tend to collapse in times of economic hardship (z).

In another example, Lieberson and Hansen (1974) found a negative relationship between language diversity (x) and development (y), when they compared a sample of countries at one point in time. Had they stopped there, they would have concluded that language diversity inhibits development. Further analysis showed, however, that for a given nation *over time*, there was no relationship between language diversity and development. What they did find, however, was that the *age of a nation* (the previously unspecified z) was negatively related to language diversity and positively related to development. Thus, the original relationship between language diversity and development was spurious (see Firebaugh 1980). This example of spuriousness is summarized as follows:



In both these examples, failure to identify the common underlying factor can lead to a false inference regarding the relationship between the two variables specified originally.

observations, but if the study is one of few countries or one country, specifying additional variables can overlap with the first problem identified in this chapter (too many variables, not enough countries). It is important not to specify irrelevant variables as they may simply cloud the analysis. The second solution is to select countries that fit the criteria of the theory that has been specified, but this solution overlaps with the problem of selection bias. Thus, the comparativist is forced to recognize these various trade-offs while maximizing the types of inferences that can be made given the countries and the evidence in the study.

Ecological and individualist fallacies

There are two types of data in the social sciences: *individual* data and *ecological* data. Individual data, as the name suggests, comprise information on individual people. Ecological data comprise information that has been aggregated for territorial units, such as voting districts, municipalities, counties, states, and countries (Scheuch 1969: 136). Individual data are collected through the use of periodic censuses carried out on the whole of a particular population, through other 'official' means, or through surveys carried out on a representative sample of the population. The twin problems of ecological and individualist fallacies occur when inferences are drawn about one level of analysis using evidence from another. An ecological fallacy occurs when results obtained through the analysis of aggregate-level data are used to make inferences about individual-level behaviour. Alternatively, an individualist fallacy occurs when results obtained through analysis of individual-level data are used to make inferences about aggregate-level phenomena. For example, claiming that women support the right to abortion by correlating the percentage of women in electoral districts with votes in support of an abortion measure is an ecological fallacy. Claiming that Germany is a more 'authoritarian' society than Britain by comparing responses to standardized survey questions is an individualist fallacy.

Both fallacies are a problem since analysis carried at one level may overestimate relationships at another level (Robinson 1950: 353), and both fallacies originate from the same sources, namely, the ontological predispositions of the researcher and data availability. In the first case, some scholars may assume that data at one level represent a higher degree of reality than data at another level. As Scheuch (1969: 134) argues, 'individual behaviour may be treated as being the only real phenomenon, while system properties are abstractions, or individual behaviour may be viewed as mere reflection of the only reality, namely structural properties'. In either case, the source of the fallacy is due to a certain ontological predisposition that serves as the starting point of the inquiry. As outlined in the previous chapter, rationalist explanations see collective behaviour as having no particular status other than the individuals who comprise it (Lichbach 1997: 245). Structuralist explanations, on the other hand, focus on the political, social, and economic connections among people, such as '[h]istorically rooted and materially based processes of distribution, conflict, power, and domination, thought to drive social order and social change' (ibid.: 247-248). Thus, a rationalist may collect information on individuals to make larger claims about groups, while a structuralist may collect information on groups of people to make larger statements about individuals.

Data availability is the second source of ecological and individualist fallacies, since scholars may be forced to substitute data from one level to examine a research question specified at another level. The first example of such a problem appears in a study of voting behaviour of newly enfranchised women in the US State of Oregon. In trying to count women voters, Ogburn and Goltra (1919) correlated the percentage of women in electoral districts with the percentage of people who voted 'no' on selected referenda in the same districts. They assumed that women would have been more likely to vote 'no' on this select set of referenda and this could therefore indirectly estimate the number of women voting in each district. Ogburn and Goltra were aware that there may be a problem drawing inferences about women voters by combining aggregate data on the female population with individual data on referenda votes (see King 1997: 3–4). Similar problems have been encountered in studies that try to estimate the socio-economic characteristics of people who voted for the Nazi Party during the Weimar Republic. The same problem is evident in Birch (2000) who examines social cleavages and party systems in the newly democratized Ukraine by combining aggregate and individual level data.

Examples of ecological and individualist fallacies are not only isolated to single-country studies. Gurr (1968) posited that a sense of relative deprivation was the prime motivating force behind rebellious activity. Relative deprivation is a psychological condition that obtains when individuals perceive that those 'goods and conditions of life to which they are rightfully entitled' fall short of those they are actually capable of achieving, given the social means available to them (Gurr 1970: 13). He posited that high levels of relative deprivation ought to be related to high levels of political violence. Since individual-level data on relative deprivation were unavailable, Gurr tested this hypothesis using aggregate data on 114 countries (see Chapters 2 and 5), which showed a positive association between his measures of relative deprivation and political violence. In this case, aggregate data were used to falsify a hypothesis at the individual level (Sanders 1981: 30–31).

In *Modernization and Postmodernization*, Inglehart (1997) commits an individualist fallacy in his study of values in forty-three societies. Using a standard battery of questions ranging from the importance of God to protection of the environment, Inglehart constructs 'clusters' of values that cohere into distinct geographical patterns. These patterns, Inglehart argues, are meaningfully distributed around the globe according to general cultural groups, including Latin America, Northern Europe, Eastern Europe, Catholic Europe, South Asia, Africa, and North America. In this study, Inglehart is aggregating individual-level responses to questions to establish simplified classifications of countries based on culture. Grouping percentages of individuals who responded similarly to a battery of survey questions and ascribing cultural 'types' to them is a clear illustration of the individualist fallacy, which confuses systemic properties with individual characteristics. Whiteley's (2000) examination of the relationship between social capital and economic growth also commits an individualist fallacy. He compares thirty-four countries using individual level measures of social capital and aggregate measures of economic growth. His analysis includes a scatterplot of the percentage of respondents in the *World Values Survey* who claimed they trust people against GDP per capita (Whiteley 2000: 455). His multivariate regression analysis includes individual level data on social capital alongside aggregate measures of investment, education, population growth, among

other control variables in order to account for changes in the average growth rate of the countries in his sample. While he does find a statistically significant relationship between high levels of trust and economic growth, the inference that social capital matters for growth is insecure owing to the problem of individualist fallacy.

In both the Inglehart (1997) and Whiteley (2000) examples, individual level data drawn from national surveys are aggregated into country scores and then included in analytical models alongside aggregate data. Each study assumes that countries can be grouped into different cultural clusters, or classified into groups that have strong and weak social capital on the basis of individual level data. Scheuch (1966: 158–159) shows that making these types of inferences is not possible. For example, a democratic system may be comprised of many individuals who respond positively to a series of questions probing their authoritarian tendencies, none the less the system is still democratic. Similarly, an authoritarian system may be comprised of individuals who respond positively to a series of questions probing their democratic tendencies or 'civic culture' (Almond and Verba 1963), but nevertheless remains authoritarian. In short, to ascribe a certain cultural or systemic trait to a country based on a sample of the population is to draw an incorrect inference about that system based on an incorrect level of analysis.

The solution for avoiding both fallacies is straightforward. The data used in any research ought to minimize the chain of inference between the theoretical concepts that are specified and the measures of those concepts that are ultimately adopted in the analysis. Known as the 'principle of direct measurement' (Scheuch 1969: 137), the solution means that research that specifies questions at the individual level ought to use individual data, and vice versa for research questions that specify systemic relationships. For quantitative analysis, Miller (1995: 155–156) argues that 'analysis of individuals can only lead to precise quantitative conclusions about individuals; an analysis of places to precise conclusions about places; and analysis of times only to conclusions about times'.¹ The pragmatic aspects of research may not allow the direct measurement of the phenomena, but the overall point remains that this measurement must be as close to the level of the phenomena being examined as possible.

Value bias

The final problem of comparison is one of value bias, a problem which depends upon the perspective from which one sees the world. Over the course of the last century, social science has come to recognize that knowledge is not 'value-free'. Classification, analysis, and substantive interpretation are all subject to the particular perspective of the researcher. Modern empirical analysis accepts that to some degree 'what is observed is in part a consequence of the theoretical position that the analyst adopts in the first place' (Sanders 1995: 67), but the quest to 'separate fact and value' is still considered worthwhile (Hague *et al.* 1992: 30). The key to making valid comparisons is *to be as public as possible* (King *et al.* 1994: 8) in terms of the judgements that have been made in the overall construction of the comparative study. These judgements include the theoretical perspective upon which the study is based, the identification of its key variables, the specification of its research design, and the limits to the type of inferences that can be drawn from it.

Conclusion

This chapter has identified six key problems of comparative method. It has also made clear that these problems are embedded in the overall trade-offs between the various methods of comparison. Specifying too many inferences without having enough observations constitutes an indeterminate research design that often affects single-case studies and those that compare few countries. Establishing cross-cultural equivalence in terms of theoretical concepts and their operational indicators is a constant worry for studies that compare many countries, since the global travel of concepts may undermine the precision of their meaning. The intentional selection of countries that support the theory being tested and that represent one or opposite values on the dependent variables can lead to an overestimation of a relationship that does not exist or the underestimation of a relationship that does. Failing to specify important 'control' or other relevant variables can lead to the overestimation of relationships. Transcending different levels of analysis can also affect the type of inferences. Finally, ignorance of the cultural and theoretical perspective that underlies a study can colour its substantive conclusions.

These problems were outlined not to paralyse comparative researchers, but to highlight possible sources of bias in drawing valid inferences. Careful attention to these problems at the outset of any comparative inquiry will maximize the types of inferences that can be drawn. Acceptance of the natural limits of comparative inquiry is a healthy step along the winding road to the production of knowledge. Taken together, the three chapters in Part I have identified why political scientists compare countries, how they compare countries, and the types of problems they frequently encounter along the way. Table 3.2 summarizes the methods of comparison and assesses their strengths and weaknesses both in terms of their ability to arrive at valid inferences and the trade-offs for the researcher that are associated with each.

Comparing many countries is susceptible to statistical analysis, which helps eliminate possible sources of selection bias and spuriousness. The large number of observations means that these types of studies are good at making strong inferences, which in turn contribute to theory-building. The comparison of many countries is good for identifying deviant cases that invite closer scrutiny both of the cases as well as of the theory that is being tested. On the other hand, comparison of many countries can rely on measures that are invalid owing to the limitations of available data. The connections established between variables may be considered too abstract and simplistic. The collection and analysis of the data may be time-consuming and may require mathematical and computing training which many comparativists are not willing to undertake.

Comparing few countries achieves control through the careful selection of countries that fit within either the most similar systems design (MSSD) or the most different systems design (MDSD). These types of studies are intensive and are good for theory generation. They avoid conceptual stretching since they rely on specialist knowledge of a few cases. These studies tend to see their objects of analysis as a configuration of multiple explanatory factors that depend on the careful comparison of history of the chosen countries. Alongside these benefits, studies that compare few countries are not able to draw strong inferences owing to problems of selection bias both in terms of the choice of countries and the choice of the historical accounts

Table 3.2 Comparative methods: an assessment

<i>Method</i>	<i>Strengths</i>	<i>Weaknesses</i>
Comparing many countries	Statistical control Limited selection bias Extensive scope Strong inferences and good for theory-building Identify deviant countries	Invalid measures Data availability Too abstract/high level of generality Time-consuming Mathematical and computer training
Comparing few countries	Control by selecting 1 Most similar systems design (MSSD) 2 Most different systems design (MDSD) Good for theory-building Intensive, less variable-oriented Avoid 'conceptual stretching' Thick description Areas studies Configurative analysis Macro-history	Less secure inferences Selection bias: 1 Choice of countries 2 Choice of historical account Language training Field research
Case study	Intensive, ideographic, path-dependent, and configurative analysis Six types: 1 Atheoretical 2 Interpretive 3 Hypothesis-generating 4 Theory-confirming 5 Theory-infirming 6 Deviant countries	Insecure inferences Selection bias: 1 Choice of countries 2 Choice of historical account Language training Field research

used for evidence. Finally, many comparativists who consider themselves 'generalists' do not want to spend their time and energy learning the languages and conducting the field research in the countries that comprise these types of studies.

Studies of single countries constitute the most intensive of the comparative methods and still make up a large proportion of research in the field of comparative politics. Single-case studies useful for comparison are those that generate hypotheses, confirm and infirm theories, and elucidate deviant cases identified through other modes of comparison. Since they are the least extensive, single-case studies are most susceptible to problems of selection bias, too many variables and not enough observations, and indeterminate research designs that yield less secure inferences than the other modes of comparison. As with area specialists, 'country specialists'

invest a tremendous amount of their time learning the local language and culture of their particular country, a commitment that other comparativists may find too demanding. Having outlined these methods of comparison, the logic that underpins them, and the problems that are associated with them, the chapters in Part II interrogate popular topics of comparative politics using the 'architecture' established in these first three chapters. Part II is primarily concerned with how different methods have been applied to different research questions, and whether these methods have produced consistent answers to these problems. In this sense, Part II *compares comparisons* in an effort to illustrate the practical implications of different comparative methods as they are applied to real research problems. The topics in Part II include economic development and democracy (Chapter 4), violent political dissent and revolution (Chapter 5), non-violent political dissent and social movements (Chapter 6), transitions to democracy (Chapter 7), institutional design and democratic performance (Chapter 8), and the comparative study of human rights (Chapter 9).

Note

- 1 Recent work in this area claims to have resolved the problem of ecological fallacy using advanced statistical techniques and the creation of specific software (King 1997), which is available for those wanting to pursue this line of research. Thus far, the new technique has been applied to voting rights cases in the United States in which aggregate data is used to make inferences about individual voting behaviour based on categories of race and social class. The extension of the method to aggregate data on nation states will certainly follow, but will involve more complicated techniques. For those not willing to pursue this line of work, however, theories that posit relationships to exist at the individual levels ought to be tested with data at the individual level, and the same rule of thumb should apply for theories that posit relationships at the aggregate level.

Further reading

Collier, D. (1991) 'New Perspectives on the Comparative Method', in D.A. Rustow and K.P. Erickson (eds) *Comparative Political Dynamics: Global Research Perspectives*, New York: Harper Collins, 7–31.

This essay addresses the problems of too many variables and not enough countries.

Dogan, M. and Pelassy, D. (1990) *How to Compare Nations: Strategies in Comparative Politics*, 2nd edn, Chatham, NJ: Chatham House.

This book provides a good discussion of establishing functional equivalence in comparative politics.

Geddes, B. (1990) 'How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics', *Political Analysis*, 2: 131–150.

An excellent article on selection bias using real examples from the comparative literature.

Hague, R., Harrop, M., and Breslin, S. (1992) *Political Science: A Comparative Introduction*, New York: St Martin's Press.

Chapter 2 makes brief statements about value bias, too many variables not enough countries, and equivalence.

Lieberson, S. (1987) *Making It Count: The Improvement of Social Research and Theory*, Berkeley: University of California Press.

Chapter 2 provides an exhaustive review of problems with selection bias.

Lustick, I. (1996) 'History, Historiography, and Political Science: Multiple Historical Records and the Problem of Selection Bias', *American Political Science Review*, 90(3): 605–618.

A good review of historical sources of selection bias.

Sanders, D. (1994) 'Methodological Considerations in Comparative Cross-national Research', *International Social Science Journal*, 46.

This article presents a strong argument in favour of using only those countries with which the comparativist has good substantive knowledge.