



MACRO COMPARISONS WITHOUT THE PITFALLS
A Protocol for Comparative Research

J. Samuel Valenzuela

Working Paper #239 - April 1997

Using comparative analysis to explain historical or contemporary social phenomena has a long tradition in the social sciences. Durkheim, Gramsci, Montesquieu, Smith, Tocqueville, Weber, and others—whose classic works are still read for their insights into the defining forces of political, economic, and social institutions—were all informed by observations regarding the similarities and differences among such institutions in various settings. Juan Linz’s writings follow optimally in this social science tradition. Whether Linz focuses on the features of Spain’s regions, on the breakdown of democracies, on the distinguishing components of democratic, authoritarian, or totalitarian regimes, on the social origins of fascists, on nationalism or religion and their effects on political identities, or on the defects and virtues of presidential versus parliamentary regimes—to name a few of his contributions—his approach always draws at one point or another on comparative analysis. Even when Linz is examining a phenomenon in a single setting, such as local elites in Andalucía, the characteristics of the Franco regime, the party system of democratic Spain, or the latest survey results on political legitimacy in Chile, his interpretations are enriched by his knowledge of the characteristics of the phenomenon in other units or at other points in time.¹

Linz’ work over the past four decades has been an important component of an ever increasing postwar literature that uses macro comparisons, historical or contemporary, of a few cases (generally national societies) to develop concepts, to reach empirical generalizations, and to build middle-range social science theory. Despite the growth in the number of studies using this approach, the methodological literature devoted to it has not developed at the same pace.² This contrasts sharply with the enormous development in the number and sophistication of publications devoted to other methods, especially survey research and a considerable variety of statistical approaches. Few graduate studies curricula devote a whole semester to comparative analysis, and general social science methodology courses hardly allocate one or two sessions to

¹ I will refrain from providing extensive citations to Linz’s work here. The reader is strongly advised to examine the impressive list of his publications in Houchang E. Chehabi, ed., *Politics, Society, and Democracy: Juan J. Linz—Untranslated Writings and Complete Annotated Bibliography* (forthcoming), a companion volume to *Politics, Society, Democracy: Latin America*, S. Mainwaring and A. Valenzuela, eds. (Westview Press, forthcoming).

² A large number of publications use the term ‘comparative’ in their titles, but few among them are specifically devoted to the problem of how to draw inferences from research that requires considering many complex variables within a far smaller number of cases. In this paper the term ‘comparative analysis’ refers only to this situation. Defining comparative analysis as “the study of dissimilar social units,” as Neil J. Smelser does in his *Comparative Methods in the Social Sciences* (Englewood Cliffs, NJ: Prentice-Hall, 1976), 2, is inadequate because it makes ‘comparative analysis’ virtually indistinguishable from social scientific investigation—as Smelser notes himself, 2–3. It also places the accent on the study of differences among units, while much of comparative analysis should rest, when possible (as will be noted below), on examining their similarities. It is the use of Smelser’s and other such definitions that has led to the apparently very large number of works on comparative analysis since the 1950s.

it—if at all. In such courses comparative analysis is still usually assimilated with what Campbell and Stanley call a “quasi experiment,” in other words, a notably unreliable method—given the absence of any controls—even if it has detailed observations of a considerable number of items within a few distinct units.³ Moreover, the core of the comparative research design according to most presently available methodological statements is a very old one, as it continues to be John Stuart Mill’s 1843 “method of agreement” and “method of difference.”⁴ Notwithstanding the importance of comparative analysis in Linz’s own work, he himself has published little about the problems encountered in using this mode of research.⁵

This paper seeks to provide pointers for comparative analysis while avoiding a largely fictitious textbook version of the research process. It discusses systematically how to set up macro comparisons in ways that will avoid common error-inducing pitfalls associated with this approach to social science questions. The paper is also a homage to Juan Linz’s work, not only because it draws illustrations from it but also because his writings stand up very well to critical scrutiny from the comparative design perspective delineated here.

The protocol presented in these pages is intended for using comparative analysis to build explanations of phenomena with the ultimate aim of developing social science theory. There is a softer form of comparative analysis which aims at better analyzing or presenting certain events, settings, or epochs through a systematic comparison of two or more units. Such more descriptive use of comparisons can be very illuminating and rewarding, as they can lead the analyst to cover

³ Donald T. Campbell and Julian Stanley, *Experimental and Quasi-Experimental Designs for Research* (Chicago: Rand MacNally, 1966). Campbell himself reiterates that he and Stanley viewed “cross cultural comparison as a weak form of quasi-experimental design,” in Donald T. Campbell, “‘Degrees of Freedom’ and the Case Study,” *Comparative Political Studies*, 8, 2 (July 1975), 179.

⁴ See, for example, Adam Przeworski and Henry Teune, *The Logic of Comparative Social Inquiry* (New York: Wiley-Interscience, 1970); Arend Lijphart, “The Comparable Cases Strategy in Comparative Research,” *Comparative Political Studies*, 8, 2 (July 1975): 178–93; Smelser, *Comparative Methods*, op. cit. (n. 2); Theda Skocpol and Margaret Sommers, “The Uses of Comparative History in Macrosocial Inquiry,” *Comparative Studies in Society and History*, 22, 2 (April 1980): 174–97; and Charles C. Ragin, *The Comparative Method: Moving beyond Qualitative and Quantitative Strategies* (Berkeley: University of California Press, 1987). Mill’s terms are contained in his *A System of Logic: Ratiocinative and Inductive* (London: Longmans, Green, Reader and Dyer, 1875).

For a review essay discussing developments in comparative analysis (although it also discusses statistical and case study approaches) since the early 1970s, see David Collier, “The Comparative Method: Two Decades of Change” in Dankwart A. Rustow and Kenneth Paul Erickson, eds., *Comparative Political Dynamics: Global Research Perspectives* (New York: Harper Collins, 1991), ch. 2.

⁵ An exception within his voluminous writings is Juan J. Linz and Amando de Miguel, “Within-Nation Differences and Comparisons: The Eight Spains” in Richard L. Merritt and Stein Rokkan, eds., *Comparing Nations: The Use of Quantitative Data in Cross-National Research* (New Haven: Yale University Press, 1966). This paper stresses the convenience of using within-country comparisons, and not only cross-country ones, given that many variables are held constant by the presence of overarching national institutions as well as important historical and cultural commonalities.

facets of the subject matter—or to better appreciate the importance of some aspect or aspects—that would escape attention if the focus were only on a single unit. This latter mode of comparative analysis is more common in history, whose aim is generally to reconstruct carefully the fullest possible view of particular settings, highlighting as a result their uniqueness.⁶ As soon as systematic comparisons lead to a discussion of the sources of similarity or of variance of particular features, the comparisons come closer to the theory-building use of comparative analysis discussed here.⁷ Comparative history and comparative sociology converge as an explanation for perceived commonalities and/or differences between units becomes the agenda of research.⁸

This more demanding use of comparative analysis may be seen, following suggestions advanced by Smelser and Lijphart in the late 1960s and early 1970s, as one of four different approaches to explaining phenomena in social science—the other three being the experimental, case study, and statistical methods.⁹ However, while Lijphart notes that the comparative method is like the statistical one “in all respects except one,” namely, “that the number of cases it deals with is too small to permit systematic control by means of partial correlations,”¹⁰ my view is that the difference between these approaches is much greater. In arguing, by contrast, that “there are reasonably clear boundaries between the comparative method...and case study methods,” Lijphart underestimates the extent to which comparative analysis relies for its creative insights on

⁶ Marc Bloch, for instance, asserted that the purpose of comparisons was to “analyze and isolate the ‘originality’ of different societies,” cited by Stefan Berger, *The British Labour Party and the German Social Democrats, 1900–1931: A Comparative Study* (Oxford: Clarendon Press, 1994), 3. In an inaugural lecture delivered at Oxford University in 1991 John H. Elliot noted that “the value of comparative history lies not so much in discovering the similarities as in identifying the differences,” cited by Eduardo Posada-Carbó, “Elections before Democracy: Some Considerations on Electoral History from a Comparative Approach” in Eduardo Posada-Carbó, *Elections before Democracy: The History of Elections in Europe and Latin America* (London: Institute of Latin American Studies Series, University of London, at MacMillan Press, 1996), 4.

⁷ Charles Tilly, *Big Structures, Large Processes, Huge Comparisons* (New York: Russell Sage Foundation, 1984), distinguishes “individualizing comparisons,” in which the uniqueness of each case is emphasized, from “encompassing comparisons,” “universalizing comparisons,” and “variation finding comparisons,” all of which posit some element of commonality—in addition to their individual features—between the units that fall within the scope of the comparisons. The first pertains to the contrasting-descriptive approach often used by historians and the latter to the more theory-building use of comparisons. An excellent example of an ‘individualizing comparison’ is the classic work of Sir John Clapham, *The Economic Development of France and Germany, 1815–1914* (Cambridge: Cambridge University Press, 1936).

⁸ This convergence can be seen in Berger, who is a historian. His work focuses on the similarities between British and German labor movements and their related parties, and criticizes the often expressed view that they are polar opposites, given their ideological differences.

⁹ See Arend Lijphart, “Comparative Politics and the Comparative Method,” *American Political Science Review* 65, 3 (September 1971): 682–93, who acknowledges (in a footnote, 684), his debt on this point to a paper presented by Neil Smelser at the Sixth World Congress of Sociology (1966).

¹⁰ Lijphart, “Comparative Politics,” 684.

case studies, as will be noted here.¹¹ Given these disagreements, it is best to begin with a brief prefatory comment on the differences between these methods in order to establish the distinctiveness of comparative analysis. I exclude consideration of the experimental approach here because it is only remotely related to comparative analysis and is of little use to clarify the latter's procedures.¹²

Approaches to Explaining Macrophenomena

Although comparative analysis has a unique character, it shares some aspects of both the case study and statistical methods. Hence, this discussion will begin by characterizing the latter two.

Case studies have been the poor relative of social science explanation. Although scholars have pointed to their virtues in building theories by helping to generate new hypothesis and by providing testing grounds for existing theories,¹³ these are uses that view case studies as a tool or a stepping stone for eventually generating or sharpening explanations by means of comparative or statistical approaches. In the process, the distinctiveness of the case study method itself has been largely overlooked.

This distinctiveness is best expressed by the in-depth qualitative research that ethnographers and cultural anthropologists prefer, although it is also present, sometimes unwittingly, in the work of case study analysts from other disciplines, including history. The purpose of the research is to examine in detail a particular social phenomenon in a complex collectivity, the 'case' at hand, whose boundaries make it a relatively self-enclosed system of social interaction. There is a sharp awareness that the research pertains to the 'case,' but conceiving it as such actually adds little if anything to the analysis. The researcher does not presume to know beforehand which are the phenomenon's constituent elements (even if theories may point initially to them) but tries to discover and conceptualize them. The next step is to establish how these elements relate to each other, charting the phenomenon's configurations of social interactions. Once these are known, i.e., the constituent elements and their configurations, individual and collective behavior expressing the matter under research within the unit are understood to be

¹¹ Lijphart, "Comparable Cases," 160. Lijphart does note that the analysis of 'deviant cases' can be a form of comparative analysis, 160. In "Comparative Politics," 691–93 he has a useful discussion of six varieties of case studies, deviant cases among them.

¹² With experimental designs researchers introduce scientific controls by manipulating variables and by repeating the experiments. This can hardly be done with the macrophenomena that are the object of comparative analysis. Naturally, all scientific explanation requires some form of control over variables, and in this sense there is a remote relationship between comparative analysis and experimental designs. The protocol presented here shows how controls can be introduced in comparative studies.

¹³ See Collier, *op. cit.* (n. 4), 23–24, for a brief review of the literature assessing case studies.

'explained' and as such, even 'predictable,' as they flow under the constraints of a certain 'grammar' or 'logic.'¹⁴ Case study analysts have little sense of working with 'variables' and they have a highly developed sense that every aspect they examine is located within a larger context. As such, individual elements of a phenomenon can be understood fully only when the workings of the whole relevant to it are clarified. For this reason the analyst must return several times to assess the significance of each element, and these operations are themselves steps to constructing piece by piece an image of the whole. The research process seeks to put into sharp relief the distinctive, essential, defining characteristics of the case at hand, those that pervade even the phenomenon's variations. Campbell likens case study research to 'pattern-matching,' as the analyst tests and refines a theoretically informed pattern with a broad variety of observations within the case.¹⁵ (Naturally, some researchers may want to show a phenomenon's variations within the case, although when this becomes the basic object of the study it shifts either into the comparative or statistical method.) The ethnographic variant of case studies relies heavily on informants and on discovering the phenomenon's elements and configurations on the basis of their visions of it.¹⁶ Other forms of case study, for instance, the analysis of a particular national political regime such as the Spanish one in many of Linz's publications, may draw on interviews, historical documents, secondary literature, newspaper accounts, survey research, electoral data, demographic trends, and so on.

Unlike case studies, the statistical method is not concerned so much with the distinctive or essential features of a social phenomenon as with measuring and explaining the extent of its variations; in fact, grasping these variations becomes from a statistical perspective a surrogate approximation to a phenomenon's distinctive characteristics. The constitutive elements of the phenomenon are not normally an object of discovery but are furnished by a pre-existing theory, and they are expressed as a limited set of variables subject to quantification (sometimes by means of indicators that only capture an aspect of the element or stand in lieu of it). There is also a clear distinction between the variables to be explained and those that do the explaining, dependent and independent, in the statistical procedures, although the analyst may choose to invert or combine them in various ways in the course of the research. The statistical mode requires a large

¹⁴ In his generally informative discussion of the differences among case study, comparative, and statistical approaches, Charles Ragin, *Constructing Social Research: The Unity and Diversity of Method* (Thousand Oaks, London, New Dehli: Pine Forge Press, 1994), ch. 2, does not consider this form of predicting social phenomena by using case or ethnographic studies; see especially 37–39 and 51.

¹⁵ Campbell op. cit. (n. 3), 182. Given the many different observational points in case studies, Campbell argues convincingly that they in fact have an n which is much larger than one (1) and that they do not have, therefore, the negative degrees of freedom problem that is usually associated with them.

¹⁶ See, for an example of this type of work, Elijah Anderson, *Streetwise: Race, Class, and Change in an Urban Community* (Chicago: University of Chicago Press, 1990).

number of units relative to the number of variables it uses from which to derive the latter's measurement. The notion of 'case' is most often conflated with the units from which the measurements are derived, i.e., the n in the sample or in the data base. When the numerical results produce, exceptionally, distinctive clusters that the analyst can treat as separate entities, statistical approaches can approximate the notion of case held in case studies or ethnographies in the sense that they become more complex units with distinctive configurations of elements or variables. However, in statistical analysis the resulting 'cases' are constructed on the basis of the measurements presented by the variables, and there is no sense that the whole will affect in turn the significance of the elements or that the cluster constitutes necessarily a collectivity of bounded social interaction.

While case studies and statistical analysis are in their starkest form polar opposites, comparative analysis, as noted previously, shares elements of both. With the case study or ethnographic approach, comparative analysis has as its object of analysis complex collectivities, viewed as the 'cases.' It also tries to explain the characteristics of the phenomenon it examines within the cases by looking at a large number of elements and their configurations and by paying close attention to their context within the cases. However, by curious contrast with the case study approach, the notion of 'case' lies at the very center of the intellectual enterprise in comparative analysis while it is simply a given for the former. This is due to the fact that comparative analysis does not rest with the discovery of the essential elements of the phenomenon under research and its configurations within the cases but it seeks to examine the phenomenon in more than one case, attempting to construct explanations that account for the variations of the phenomenon across the cases. Hence, as with statistical analysis, the comparative approach is interested in knowing how and why phenomena vary, acquiring either different or similar forms across the units of study. Comparative analysis also shares with the statistical approach a clear-cut conception of variables, both independent and dependent. However, by contrast to the statistical approach, comparative analysis is unable to quantify its variables and expand the number of its cases to the point of satisfying the requirements of most statistical procedures.¹⁷ As a result, comparative analysis is a very complex method to use, because on the one hand it requires a deep understanding, as in case studies, of the cases it examines, and on the other it seeks to explain variations of social phenomena despite having a dearth of cases to sort out a large number of complex variables.

Given the many variable and small as well as complex n problem, the possibility of making the wrong inferences using the comparative approach is quite high. The protocol presented in

¹⁷ Ragin, *The Comparative Method*, op. cit. (n. 14), advocates the use of Boolean algebra for drawing inferences in comparative analysis. Obviously, this is a very different form of quantification of variables from that used in conventional statistical analysis.

this paper is intended as a basis to generate comparative research designs that will hopefully minimize the possibility of such errors.¹⁸ Comparative analysis should also be based on the most thorough understanding possible of the cases, so that its explanations can be supported by a base drawn from case study analysis. For this reason, while comparative analysis is at a disadvantage with respect to statistical analysis given the few cases it has to sort out its many variables, in my view this drawback is not as debilitating as it seems at first glance, because *the comparisons can be grounded on case studies to an extent that has no parallel in the statistical approach*. This occurs because the conception of ‘cases’ in comparative analysis is much closer to that of case studies than that of statistical approaches, and therefore the difference between the comparative and statistical approaches is not simply, as Lijphart and Ragin indicate,¹⁹ a matter of the lesser or greater number of variables relative to cases. The fact that inferences in comparative analysis may be supported by the intensive acquaintance with the material generated by case studies points to a greater imbrication of these two approaches than analysts have envisioned.²⁰

Complex research endeavors may require using more than one—even all three—approaches. However, it is usually possible to identify which form of analysis predominates in drawing the major inferences or conclusions. While case studies may use statistical methods, for example, these are often employed to better characterize the properties of the case at hand or as an additional element in its puzzle of complex configurations. Juan Linz’s work contains examples of all three approaches, although his main research mode is comparative analysis. Even when Linz writes case studies, as he has on numerous occasions focusing particularly on Spain, his material includes many comparative references. And when using a statistical approach, as he most frequently does with survey results, the numbers help him characterize aspects of a case for purposes of analyzing its similarities with or differences from others.²¹

A final prefatory point. There is a fundamental difference between cross-national analysis and comparative analysis, although they are sometimes confused, given the fact that the former

¹⁸ Lijphart, “Comparative Politics,” op. cit. (n. 9), 685–86, and “Comparable Cases,” op. cit. (n. 4), 159–63, has suggested four strategies to deal with the small *n* many variable problem. However, the first one, increasing the number of cases, is often not a realistic option unless the variables permit quantification, although if that is possible a statistical approach was preferable from the very beginning. I am skeptical of the other strategies for reasons that will become apparent in reading this paper.

¹⁹ Lijphart, “Comparative Politics,” 684; Ragin, *Constructing Social Research*, op. cit. (n. 14), 49.

²⁰ See in particular Lijphart, “Comparable Cases,” 160.

²¹ For an example that illustrates both of these observations, see Juan J. Linz, “La sociedad española: Presente, pasado y futuro” in Juan J. Linz, ed., *España: Un presente para el futuro* (Madrid: Instituto de Estudios Económicos, 1984), 63–64 and 73–95.

always draws its data from attributes of national societies while the latter often (but not exclusively) takes them as its 'cases.' However, it is best to retain the term 'cross-national analysis' for statistical approaches using national societal data. As such they do not have the small n but many variable characteristic of comparative analysis.²² If the analyst is interested, for instance, in determining the relationship between church attendance and socioeconomic development, obviously such a question is resolved with a statistical approach that may be cross-national if the data banks that are used draw from national censuses and/or surveys. But a comparative historical analysis is appropriate if the question relates, say, to the effects of Catholic, Orthodox, and Protestant variants of Christianity on the formation of party systems or national identities.

A Protocol for Comparative Research

The Question

The research question at hand determines which approach, or combination of approaches, is needed to answer it. Although this relationship between question and research procedure seems to be straightforward, there is some confusion in the literature over this matter. Stanley Lieberman, for example, in criticizing comparative analysis for drawing firm conclusions from the examination of a very small number of cases, presents a hypothetical example in which the question is what determines collisions between cars on street corners with traffic lights. He then sets up a 'test' following Mill's methods of agreement and difference, with two 'cases' per method, leading to absurd conclusions.²³ Although doing a 'case study'—an essential tool of error minimization in comparative analysis as I noted earlier—of each of these four events by interviewing the drivers would have kept Lieberman from drawing his conclusions, this is not the main problem with his simple-minded mimicry of the comparative method. The problem with this example is that the initial question calls for a statistical approach, not comparative analysis, to answer it. The 'cases' are brief events (two cars approaching a street corner at the same time), not complex collectivities of social interaction with definable boundaries. Such events are easily

²² The confusion in question appears most clearly in Przeworski and Teune, *op. cit.* (n. 4). Their preferred approach to comparison, which they label as "most different systems" (to be discussed below), generally assumes a statistical mode and should be assimilated to 'cross-national analysis,' not the comparative method. Lijphart also makes this criticism of these authors when he notes that their most different system "should be assigned to the category of statistical analysis" ("Comparable Cases," 164). Much of the discussion in Smelser, *Comparative Methods*, *op. cit.* (n. 2), chs. 6 and 7, refers also to cross-national statistical approaches rather than to comparative analysis.

²³ Stanley Lieberman, "Small N s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases" in Charles C. Ragin and Howard S. Becker, eds., *What Is a Case? Exploring the Foundations of Social Inquiry* (Cambridge: Cambridge University Press, 1992), 105-18.

quantifiable, and it therefore makes no sense at all to collect only two cases for analysis.²⁴ Whenever variables are quantifiable and the units from which their measures are derived can be expanded far beyond their number, a statistical approach should be used. An important source of error in what is sometimes billed as comparative analysis stems from the fact that another approach, usually a statistical one, is more appropriate to answer the question.

Sometimes different approaches may provide complementary answers. For instance, while Juan Linz answered his question regarding the superiority of parliamentary regimes over presidential ones with comparative analysis, the matter is also amenable to a statistical treatment that shows, albeit with a narrower set of variables, the extent to which there is a difference in terms of democratic stability and other measures between the two regimes.²⁵

A question pursued with one form of analysis may eventually lead to a question pursued with another. A comparative analyst exploring the relationship between labor movements and transitions to democracy may observe that the strike rate increases in the cases under examination; however, the extent to which this is so can only be determined by adopting a statistical approach using all cases for which there is reliable data. The statistical test may show that there is only one exception to the pattern, and the analyst may wish to focus on the reasons by doing a case study of it.²⁶ Similarly, survey results may show considerable differences among attitudes towards democratic legitimacy across recent cases of transition to democracy, and this may lead to a comparative approach in order to find the reasons for these differences. In all of what follows here I will assume that the question at hand can only be resolved satisfactorily by using comparative analysis.

While it is in the nature of a truism that all research should begin with a question, it is remarkable to note how many times scholars initiate comparisons not with a question but with what could be called a space. Thus, a comparison will be drawn between two or more national societies because they share common borders, they are located in the same area of the world, or they are

²⁴ Lieberson also argues that conclusions in the social sciences should take the form of probabilistic statements, while comparative analysis draws deterministic ones (106–9, 117–18). It is indeed very hard to state conclusions in a probabilistic manner without using a statistical approach. This does not mean that comparativists are not aware of the tentative nature of their conclusions. They hold only in so far as another analyst, after making the great effort required to understand the pertinent cases in all their complexity, is able to point to their deficiency and to generate a different, more convincing set of conclusions. In the last analysis this difference between statistical and comparative analysis is more of form than of substance.

²⁵ A statistical approach to the question can be seen in Arend Lijphart, "The Virtues of Parliamentarism: But Which Kind of Parliamentarism?" in H. E. Chehabi and Alfred Stepan, eds., *Politics, Society, Democracy: Comparative Studies* (Boulder: Westview Press, 1995). Juan Linz's basic articles on this question are "The Perils of Presidentialism," *Journal of Democracy* 1 (winter 1990): 51–69; and "The Virtues of Parliamentarism," *Journal of Democracy* 1 (fall 1990): 84–91.

²⁶ My appreciation to Nancy Bermeo for her talk on this topic at Nuffield College, Oxford University, 14 May 1996, in which she noted that strike data from Ecuador constituted such an exception.

accessible to the comparativist for reasons such as knowledge of the languages needed to study them or access to funding, family support, institutional contacts, and so on, that make the research possible. When these considerations propel the initiation of comparisons, successful ones must still formulate a question that calls for applying comparative analysis—not other methods—to cases that have been selected in this manner—and not to other cases. I will return to this latter issue below.

However, the apparent truism of beginning research with a question also has its complications. As occurs in any scientific endeavor, the initial question may not be one that is well formulated given the inadequacy of the state of knowledge over the matter to which it pertains. Consequently, the research process must often return to reformulating the question, sharpening it. With comparative analysis it is also sometimes necessary to redress the question for reasons that are unique to it, again a point to be taken up later in this paper.

After formulating the question, at least in preliminary form, the analyst should turn to the main approaches that treat the subject in order to develop a list of the causal determinants that have been suggested in the literature for the problem at hand. During the course of the research the analyst should weigh the validity of all these explanations. While in statistical approaches the hypotheses drawn from different theoretical views often must be tested sequentially, in comparative analysis the evidentiary material (historical documents, interviews, statistical series, reports, and/or secondary literature, etc.) must be approached while keeping in mind what are many times conflicting notions—and the more the better—regarding what explains the phenomenon at hand. Sometimes there are no preexisting theoretically based explanations covering the research problem. The analyst should still try to develop some notions based on his or her intuition or experience with related questions that will explain the phenomenon. It is better to stipulate such preconceptions explicitly rather than to jump into the research material, because, more often than not, they are present anyway in the back of the analyst's mind and condition his or her perceptions of the material.

In this initial stage the analyst must also pay close attention to the definitions of key concepts according to the various theoretical traditions, as they often do not agree. It is essential to generate clear definitions of the terms and to apply them consistently across all the cases. While this should be done as soon as possible, it is often not wise to adopt thoroughly fixed definitions until the empirical research has reviewed preliminarily a range of disparate cases that are relevant to the question being asked. An examination of the evidence helps to clarify the meaning of concepts, and it may even lead to creating new ones. Holding to inadequate definitions can distort considerably the interpretations of the empirical material, leading to faulty conclusions. Induction plays a central role in comparative analysis and for it to operate properly, in other words, for the evidence to contribute its part in shaping the analyst's views of it, the

definitions of key concepts should be open to revision and clarification in the early stages of the research.²⁷ Juan Linz's work in this respect offers a model to follow given its clear and concise definitions which have been crafted after a considerable review of various historical settings. For example, Linz's careful delimitation of the notion of political legitimacy and its facets resulted from his experience in dealing with the history of the delegitimation of democracies.²⁸ And his definition of democracy, minimalist and procedural, has been crafted to include in its scope regimes that nonetheless have considerable shortcomings, especially in terms of the extension of suffrage.²⁹ This is due to the fact that their overall dynamics corresponded more to democratic than to nondemocratic regimes.

The Universe

A fundamental step in comparative analysis is to determine which is the universe of cases that must be examined in order to answer properly the question at hand. Ignoring this step can lead to the well-known problem of case selection bias.³⁰ In statistical analysis different universes (and samples) may contain the necessary properties to test hypothesis, and each is arguably as valid as the next. Comparative analysis does not have the same latitude or flexibility. Working with few cases and many variables requires ascertaining exactly which are the cases that apply to the question. This means that a considerable amount of time must be devoted to an at least rapid review of a broad array of cases that may be suspected of having attributes of the problem under examination. And yet, comparativists often do not spend much time thinking about the universe of cases to which their question applies, and a recent methodological statement even ignores the significance of this issue entirely.³¹

A further difficulty presented here is that the universe in comparative analysis has not one but two components. The first its core, consisting of all cases (national often, or subnational such

²⁷ These considerations are applicable whether one subscribes to an essentialist or a minimalist view of definitions.

²⁸ See Juan J. Linz, *The Breakdown of Democratic Regimes: Crisis, Breakdown, and Reequilibration* (Baltimore: Johns Hopkins University Press, 1978), 16–23.

²⁹ Linz, *Breakdown*, 5, where even regimes with *suffrage censitaire* are included among democracies.

³⁰ See Smelser, *Comparative Methods*, op. cit. (n. 2), 221–12; and Barbara Geddes, "How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics," *Political Analysis* 2 (1990): 131–50.

³¹ Ragin, *Constructing Social Research*, op. cit. (n. 14), 113. Ragin simply notes that the cases selected for comparison are those the analysts or their intended audiences find interesting, and that they should be "comparable," i.e., they should "belong to the same category."

Geddes, op. cit. (n. 30), 134, does realize the importance of searching for the proper universe. However, she does not discuss the specific problems confronted in determining the universe when there are many variables but far fewer cases. In fact, her definition of the universe as "the cases to which the hypothesis should apply," and her discussion of sampling procedures that should be "uncorrelated with the placement of cases on the dependent variable," simply apply the standard procedures suited to statistical approaches (134–35).

as city governments, parishes, firms, etc.) where the phenomenon under investigation, in whatever variant if there is more than one, has occurred. The second component can be called the ancillary one: it includes cases where the phenomenon to be explained has not occurred, even if it may nearly have occurred, but that contain the most closely parallel or kindred phenomena to the one being investigated because they are nearest different species or subspecies of a common genus. Thus, rising up the genus-species ladder to the next generic rung and then coming down on the nearest parallel line should permit the investigator to discover which are the most kindred cases.³² For instance, if the question pertains to the possibility of democratic outcomes out of political crisis in sultanistic regimes, the core universe will include cases where the crises of such regimes have generated democratic transitions as well as where they have not; it will therefore include settings with both major variations that correspond to the question. Sultanistic regimes fall under the genus of authoritarian regime as defined by Linz,³³ and the problem being researched refers to political crisis and not to other political phenomena. The ancillary universe will therefore be composed of cases of political crisis with and without democratic outcomes in other types of authoritarian regime, i.e., other species of the same genus.

Most of the research will focus on cases in the core universe, while knowledge of the ancillary cases, which may be needed as a check for the explanations drawn from the core universe—as will be elaborated below—need not be as profound. The larger the question, i.e., the more generic and abstract it is, the more cases will be included in the core universe, to the point that the ancillary one may be reduced to zero. Thus, if the question is how collective identities are formed, there is no use wondering where this phenomenon did not take place; but if the question is how political identities influencing voter behavior are formed in democratized post-Communist regimes—a question that is certainly much lower down on the genus-species ladder—then the ancillary universe contains a large number of cases, the most proximate ones being those of political identity formation as it affects voter choices in newly democratized regimes stemming from other nondemocratic origins.

Sometimes apparently small changes in the research question can lead to considering certain cases either as part of the core or of the ancillary universes. For example, in Linz's study of the breakdown of democracies, the core universe is composed certainly of the cases where such breakdowns occurred, but he refers as well to a case, namely France at the fall of the Fourth and

³² This procedure borrows from Giovanni Sartori's classic article on the use of concepts in comparative research, "Concept Misformation in Comparative Politics," *American Political Science Review* 54, 4 (December 1970): 1033–53.

³³ Juan J. Linz, "Totalitarian and Authoritarian Regimes" in Fred I. Greenstein and Nelson W. Polsby, eds., *Handbook of Political Science*, vol. 3, *Macro Political Theory* (Reading, MA: Addison-Wesley, 1975), 179 et passim.

inception of the Fifth Republics, in which democracy was 'reequilibrated' after a crisis, generating rapidly a reassertion of the democratic regime.³⁴ This was a case that had the advantage for Linz's argument of having had a disruption of normal constitutional procedures, thereby eliminating the guesswork regarding the severity of its political crisis. If it is viewed for this reason as a case of democratic collapse, then it contains a variation of the phenomenon under study, i.e., a breakdown that nonetheless led to a rapid reassertion of democracy instead of a long lasting nondemocratic regime, and in this sense it would belong in the core universe. However, if it is viewed (as Linz in fact saw it) as a case of survival of democracy despite having undergone a very severe crisis, then it is a 'near miss' situation that belongs in the ancillary universe because it does not contain the outcome under study, i.e., a breakdown of democracy leading to the instauration of a long-lasting nondemocratic regime. Consequently, this latter qualification determines whether France in 1958 (and many other cases) is part of the core or ancillary universes. The research sometimes has to be at a relatively advanced stage before such judgements can be made and the full composition of the ancillary universe of possibly useful non-x cases (in a sense to be explained later) comes to light.

One of the most common of all pitfalls in comparative analysis is to try to reach conclusions for theory-building purposes without examining the full range of variation of the phenomenon at hand in the core universe. It is to avoid this mistake that it is fundamental, as indicated above, to know which cases compose this universe. If such cases are not known or are simply dismissed from consideration, some variations of the phenomenon may not enter the analysis. The result may be inferences that have a high probability of being wrong or inadequate because they stem from a narrow expression of the phenomenon. Hans Daalder has noted, for example, that models of European state building have erroneously been based mainly on state building in France.³⁵ Similarly, for many questions it is not possible to restrict the analysis simply to a geographical area such as Scandinavia, Western Europe, the Southern Cone of Latin America, or East Asia, and yet this is done more often than not. Again, Juan Linz's work points the way in terms of its refusal to restrict its boundaries to a single area of national experiences, as shown by his taxonomic survey of the varieties of nondemocratic regimes or his studies of fascists.³⁶ There is no such thing as a

³⁴ Linz, *Breakdown*, op. cit. (n. 28), 87–88 especially.

³⁵ Hans Daalder, "Paths toward State Formation in Europe: Democratization, Bureaucratization, and Politicization" in Chehabi and Stepan, eds., op. cit. (n. 25). In Daalder's words, 115: "writings on European political development in general, but most notably writings on the development of the state, are generally couched in terms of a dominant paradigm that on closer look is clearly derived from a stylized analysis of the French case."

³⁶ Juan J. Linz, "Totalitarian and Authoritarian Regimes," op. cit. (n. 33); Linz, "Some Notes toward a Comparative Study of Fascism in Sociological Historical Perspective" in Walter Lacqueur, ed., *Fascism: A Reader's Guide* (Berkeley and Los Angeles: University of California Press, 1976); and Linz, "Political Space and Fascism as a Late-Comer" in Stein Ugelvik Larsen, Bernt Hagtvet, and Jan Petter Myklebust, eds., *Who Were the Fascists?* (Oslo: Universitetsforlaget, 1980).

'theory' valid for only a limited range of cases, and yet the comparative literature in the social sciences is unfortunately full of such attempts.³⁷

A second common error relating to the universe is the opposite of the first, namely, the attempt to treat too many cases. No analyst working alone, not even Juan Linz, can master with all the required nuances and contextualizations a long list of cases. Attempts to do so result not only in a cursory understanding of the cases but also, quite inevitably, in a form of comparative analysis that is excessively deductive, consisting basically of a set of theoretical premises and their logical derivations which are then illustrated by drawing selectively from the case materials. Such analyses therefore contain little, if anything, in terms of discovery, and will simply confirm most if not all preconceived ideas. The complexity of the cases is glossed over, as the analyst's theoretical framework overpowers everything. With too many cases comparative analysis goes too far astray from the case study approach, thereby losing the latter's contribution to error minimization through the intimate understanding of the configurations presented by the phenomenon under examination in the specific cases. It is the insider's knowledge of the material of strategic cases (given the question that is asked) that leads to the original discoveries that make comparative analysis an important tool. Consequently, it is impossible to be a good comparativist without knowing a small number of cases very, very well.

Again Juan Linz exemplifies this point. His initial studies led him to very thorough research on early to mid-twentieth-century Germany, Spain, and Italy. And it is his, again, very thorough understanding of these cases that led him to develop the authoritarian regime model that we now recognize, given his writings, as a key concept in the study of politics. Franco's Spain was simply of a different species under the nondemocratic regime genus from Nazi Germany and even Fascist Italy.³⁸ Researchers who attempt comparative analysis without in addition being thorough students of case studies not only have little to contribute beyond their theoretically derived models but often leave those who know particular cases well baffled by the use that they make of them. Case study students should, on the contrary, recognize the 'fit' between the contributions of comparative analysis and the cases with which they are acquainted and should

³⁷ It may be argued that it is possible to draw empirical generalizations from comparing cases that do not contain the full range of variation of the phenomenon at hand. But the conclusions of most such studies do not make such a limited claim. And these generalizations will be of little use in any event. What sense would it make to conclude, after studying the origins of social security systems in Brazil and in Italy, that such systems originate with fascist or protofascist authoritarian regimes?

³⁸ The authoritarian regime model was presented initially in Juan J. Linz, "An Authoritarian Regime: The Case of Spain" in Erik Allard and Yrjo Littunen, eds., *Cleavages, Ideologies and Party Systems* (Helsinki: Westmark Society, 1964).

derive insights for their own work from those contributions. The comparative analyst's theoretical imagination should not exceed his or her knowledge of the cases.³⁹

These two common inadequacies related to the number of cases—having too few or too many of them—are like the Scylla and Charybdis of comparative analysis, and it is necessary to devise strategies to navigate safely between both dangers. These will be addressed below.

Typologizing

Comparative analysis requires abstracting or highlighting a set of attributes relevant to the question at hand from the cases—attributes that are then judged similar or different in various ways to be noted later—and as a result it is impossible to do this analysis without creating what are in effect ideal types. Hence, it is useful to develop explicitly what becomes a typology as early as permitted by the minimal necessary knowledge of the cases and to formalize the attributes of each type as clearly as possible. It may be necessary to have both types and subtypes.

The typology should be organized around *the outcomes* shown by the phenomenon under investigation (or the dependent variable or variables), and it should cover *the full range of variation of such outcomes* in the core universe. Both points are important. The latter has already been noted above: covering the full range of variation of the phenomenon is essential to reach proper theory-building generalizations. Examining briefly the ancillary cases, especially if the question leads to a very circumscribed set of core cases, may add kindred manifestations of the phenomenon to the typology for later use. Again, this will be discussed below; it is the former point that requires additional comment.

³⁹ What I have called here the deductivist and illustrative approach to comparative analysis does have its advocates. One of the most extreme is Angelo Panebianco, *Political Parties: Organization and Power* (New York: Cambridge University Press, 1988), who writes that in comparative analysis “one’s research *is based on a predetermined analytic picture*, [and] an investigation of this type will inevitably do an injustice to...historiographic interpretations (filtering them through different theoretical lenses) and to historiographic material in general. Comparative historical research almost always leaves historians (specialists in individual case studies) perplexed and unsatisfied. And this is virtually inevitable because the comparativist...can only be highly selective in his choice of historical literature, *having to discard those aspects of the historiographical debate which are not compatible with his theoretical perspective*... An ever-present risk is that of doing an overly superficial analysis of the different case studies. But the alternative is even riskier; in fear of doing an injustice to history, the researcher...loses sight of his goal: to isolate the similarities and differences between the various cases (which, in turn, is possible *only if the predetermined theoretical perspective is not abandoned in the process*),” xiv–xv, my emphasis. Needless to say, I reject the substance of the comment categorically. It makes a virtue out of forcing the case study material into a preconceived conceptual and analytical straightjacket!

Lijphart’s fourth suggestion to deal with the small *n* many variable problem, that of “restricting the analysis to the key variables and omitting those of marginal importance,” (“Comparable Cases,” op. cit. [n. 4], 159) also can lead to this problem. The analyst may restrict him or herself too much to a predetermined conceptual scheme and minimize the possibility of discovery from cases studies.

Building the typology around outcomes organizes the comparisons among the cases around the similarities and differences they contain in terms of the phenomenon to be explained. The search for what determines the phenomenon can then proceed without having the typology prejudice the results. This is what occurs if the typology is built in the other possible ways of doing so, namely, on the basis of what are understood to be, given the working hypothesis, the determinants (or independent variables) of the phenomenon under investigation or on the basis of supposed background characteristics of the cases. The first option leads the analyst to run the risk, given the complexity of the cases, of simply confirming the hypothesized determinants as they become a powerful lens or filter that presorts the evidence. The supposed explanations organize the analysis to such an extent that it is hard to escape from them, and the research becomes unwittingly tautological or self-fulfilling because it is then difficult to examine the validity of alternative perspectives or even for new ones to emerge from reviewing the evidence.⁴⁰ This is not to say that the researcher should not have working hypotheses; the point here is that the hypotheses should be assessed by looking at material organized on the basis of what they purport to explain, i.e., the outcomes of the phenomenon in question. It is possible, nonetheless, for the analyst to choose to report the research findings by presenting a typology based on the determinants of the phenomenon; but this is a matter of choice for the sake of the clarity of exposition once the conclusions of the research have been reached.⁴¹

The second alternative option is to organize the typology around certain background characteristics of the cases, but it is no better than the previous one. It generates the risk of skewing the analysis in ways that make the background attributes of the cases, i.e., those that are unconnected to either the phenomenon to be explained or to variables the theories suggest as their causes, acquire a determinative effect, even though the analyst does not have this intention. For instance, if the question is what determines the formation of welfare regimes, and the cases are classed into Buddhist, Catholic, Confucian, Hindu, Jewish, Muslim, Orthodox, and Protestant types, it is quite likely that the answers to the question will be affected by this overarching typology of the religious and/or moral basis of national cultures. A very common source of background based typologies is the tendency to classify cases according to their levels of

⁴⁰ This procedure is, again, similar to the deductive and illustrative form of comparative analysis which does not open itself to discovery. For an example of this inadequate form of typologizing in comparative analysis, see Clark Kerr, John T. Dunlop, Frederick Harbison, and Charles A. Myers, *Industrialism and Industrial Man: The Problems of Labor and Management in Economic Growth* (New York: Oxford University Press, 1964), ch. 2, in which the typology of industrializing elites also explains the course of industrialization and labor management.

⁴¹ In my "Labor Movements in Transitions to Democracy: A Framework for Analysis," *Comparative Politics* 21, 4 (July 1989): 445–72, I chose to organize the presentation of comparative research findings with a typology based on the determinants of the phenomenon I analyzed. In my "Labor Movements and Political Systems: Some Variations" in Marino Regini, ed., *The Future of Labor Movements* (London: Sage Publications, 1992), I chose to retain the working typology based on outcomes.

development or, more crudely, into whether they are part of 'advanced industrial societies' or of the 'Third World,' a distinction that never made much sense but is now clearly obsolete.

Having developed the typology, the relevant cases should then be listed into the various types. Obviously, there will generally be a disparate number of cases in each one. It would indeed be a very rare coincidence if each and every variation of the phenomenon were to be found in an equal number of cases, and such symmetry should not be forced into the interpretation of the material.

Coping with the Size of the Universe

If the research question calls for a manageable set of cases to be analyzed, then the size of the universe does not present any problems and the researcher can sail safely through the Scylla of having too few and the Charybdis of having too many cases. But often the question's relevant universe of cases is excessively large for any individual scholar to analyze with the required depth.⁴² Moreover, the language abilities of the researcher may prevent serious consideration of all the necessary cases. In such situations (in addition to learning more languages or hiring translators) three strategies can be followed.

The first is collaborative work. A researcher who has done extensive work on a question may recruit others to discuss further cases, as exemplified with Juan Linz's model of the breakdown of democratic regimes which led to a conference and to a collective set of volumes edited with Alfred Stepan.⁴³ Several researchers who know their cases well may also join forces to study a common issue. This requires optimally a unified team to discuss every step of the research and for the group's organizer or organizers to be actively engaged in studying some of the case materials in depth, otherwise the latter will be too detached intellectually to be able to pull the various strands together effectively in concluding observations. Perhaps the best example of this model is the group directed by Charles Tilly for the Social Science Research Committee which examined the issue of the formation of national states in Europe.⁴⁴ But in most situations this form of collaborative research will be impossible to organize, as its funding is hard to secure and a compatible team is rarely assembled.

⁴² Having too few cases, even one, in the core universe is not a problem as long as it is indeed true that there are no other cases with the phenomenon under investigation. It is also possible to do the comparative analysis equivalent of computer simulations, that is, to invent cases on the basis of speculation of the 'what if' sort. While historians rightly shy away from such exercises, they can be fruitful from a conceptual and theoretical point of view.

⁴³ Juan J. Linz and Alfred Stepan, eds., *The Breakdown of Democratic Regimes* (Baltimore: Johns Hopkins University Press, 1978).

⁴⁴ Charles Tilly, ed., *The Formation of National States in Western Europe* (Princeton: Princeton University Press, 1975).

A second strategy is to reformulate the question by dropping to a lower level of abstraction along the genus-species ladder, which will necessarily reduce the size of the universe—particularly the core universe. For example, if the general question is ‘what are the effects of religion on the formation of political parties in democracies?’ it calls for an examination of all cases where democracies have been established. Reducing the question to ‘what are the effects of religion on the formation of political parties in those democracies that developed before 1920?’ would, of course, drastically cut down the size of the core universe only to those cases that were earliest in developing democratic regimes; how much so would depend on the definition of democracy. It is also possible to reduce the relevant universe by following the genus-species line along the religious dimension and investigating, for instance, only the impact of Catholicism on the formation of political parties in democracies. In this case the core universe would be composed of countries with democratic regimes and a significant percentage (to be defined) of Catholics in the population.⁴⁵ Or the universe can be reduced even further by focusing on ‘what are the effects of Catholicism on the formation of political parties in well-established democracies (according to a clear cut definition) where only a minority (again, definition required) of the national population is of Catholic background or tradition?’ In this case the core universe would be composed of Germany, The Netherlands, Switzerland, the United States, the United Kingdom, and perhaps some other cases depending on the definitions.

While this strategy is highly effective in reducing the size of the core universe, it has two drawbacks. The first is that every additional qualification to an initially grand question also reduces the breadth and, possibly, the theoretical pay-off of the conclusions. Although all of these questions fall into the domain of middle-range theorizing, their scope becomes smaller and smaller with each reduction. Although it is better to have solid research on more limited questions than poorly grounded work on larger ones, at an extreme the question and the resulting theoretical pay-off may become trivial or insignificant. Hence, in doing this type of reduction of the universe, it is important not to eliminate the case or cases that present the most theoretically meaningful variations of the problem at hand, those that stand out in the literature given, for instance, their ‘deviant’ nature. For example, with the last reduction to the question in the previous paragraph the universe of cases no longer contains any where Catholicism is or was the established Church or where Catholics were the main state builders. Eliminating such cases means that the phenomenon of the effects of Catholicism on the formation of political parties in a democracy can no longer be examined, because an important variation to this phenomenon has been lost. However, a researcher may well be interested in looking specifically at the effects of

⁴⁵ For an important recent contribution to the study of the varieties of Catholic political party formation in Europe, see Jean-Dominique Durand, *L’Europe de la Démocratie Chrétienne* (Brussels: Éditions Complexe, 1995).

sizeable Catholic minorities on party formation in countries where Protestants were the main state or democracy builders, given that this matter has received less attention in the literature.

The second drawback is that such reductions of the core universe increase the size and potential significance of the ancillary one, and hence beyond a certain point the number of cases ceases to become much smaller. While the research need not go into depth with ancillary cases, if the core universe becomes so narrow that it does not contain a sufficient number of cases for the analyst to undertake a proper check of the inferences he or she draws from its comparisons, then the ancillary ones will become much more important to the research. I will return to this point in the discussion of the process of drawing inferences from comparisons.

The third strategy to cope with the excessive size of the universe is to take a cue from statistical analysis and to draw a sample from it. This is the preferable procedure for navigating between the problem of having too many or too few cases when the analyst does not want to limit the theoretical scope of the question. Naturally, the results depend on having an adequate sample, and this kind of sampling cannot be done at random. The cases the analyst already knows best, or those in which the phenomenon under study first came to his or her attention, can of course be included in the sample. But the sample has to follow strictly the typology (which in turn has to be correctly designed), and it should include at least one case, but preferably two or three, in each type. The aim is to have a manageable number of cases but for that number to contain all the significant manifestations of the phenomenon being researched, thereby preventing the common mistake of trying to reach generalizations on less than their full range of variation. Cases from the ancillary universe, which need not reflect all possible kindred variations, should also be selected into separate types if need be.

As noted above, every question points to a set of variables pertinent to it, a set that can be determined preliminarily by examining the theories that address it. These independent variables, together with the dependent one or ones, form part of the 'active' variables the analyst focuses on during comparisons. There are also many 'background' variables, i.e., those that are known to the analyst but that express attributes of the cases that are considered to be irrelevant to researching the question at hand.

The 'background' variables do play a part, however, during the sampling phase of the research, for it is best to select cases for each type in the sample that have overall affinities in terms of such background variables. This is preferable because the research process may uncover that the active variables that were originally considered to be determinative, given the theoretical perspectives, cannot account adequately for the phenomenon at hand. In such circumstances the researcher may want to bring background variables to the foreground (making them part of the active ones) in an effort to discover new explanations, and this operation will be easier with cases that are, by and large, closer to each other in ways that were originally

considered irrelevant. Hence, if the choice is, for example, between including material from France or from Zambia in a certain type, it is best to opt for the case that has greater affinities with the one or more cases already included in the type in terms of the seemingly irrelevant political, economic, cultural, demographic, and so on, background variables.

FIGURE 1

STANDARD METHODS OF COMPARISON BASED ON MILL			
Method I		Method II	
Of Agreement (Mill) or of Difference (Przeworski and Teune)		Of Difference (Mill), of Similarity (Przeworski and Teune), or of Comparable Cases (Lijphart)	
Case 1 x present	Case 2 x present	Case 3 x present	Case 4 x absent
Other variables:		Other variables:	
a	b	a	a
c	d	b	b
e	f	c	c
g	h	d	d
i	j	e	e
k	l	f	f
Explanatory variables:		Explanatory variables:	
o	o	r	s
p	p	t	u
q	q	v	w
x results from o, p, q, variables that are matched in both cases.		x results from r, t, v, variables contained in case 3 but not in case 4.	

Setting up the Comparisons

The literature on comparative analysis still presents, as indicated previously, Mill's methods of 'agreement' and 'difference' as its main tools for working towards research conclusions. Figure 1 exemplifies these methods in a simple form with two cases and three variables to explain phenomenon *x*. With Method I, to be used when *x* is present in both cases, the researcher tries to find a few similar attributes among the many differences between the cases

with the presumption that such similarities account for the similar outcome. With Method II, to be applied when x is present in only one case, the researcher seeks to match the attributes of what are presumed to be similar cases in a search for those that are different, as the latter would then explain the x and *non-x* outcomes.

There is no reason why a researcher should use only one of the methods in Figure 1, as both can be used by comparing more than two cases, and in fact it is best to do so. Berins Collier and Collier even apply both schemes at the same time: after constructing a typology of labor 'incorporation' outcomes in Latin America, they analyze pairs of cases that presumably share the basic similarity of belonging to a single type but are as different as possible within that framework.⁴⁶

However, the methods in Figure 1 contain very simple-minded renditions of comparative research designs. First, they do not contemplate all the possibilities that present themselves in comparisons. Cases that have the outcome to be explained may have arrived there by more than one route, contrary to what is implied in Method I. There may be, in other words, coincidental determinants for similar outcomes. Moreover, Method II assumes that cases that do not show similar outcomes are nonetheless matched in every other respect! This is a tall order in comparative analysis, because the cases are of such complexity that it is impossible for them to be so extensively similar that the only differences will be those that explain the different outcomes. This points to the second deficiency in this rendition of comparative designs, namely, that it is insufficiently complex in its depiction of the variables in so far as it makes no distinction between the 'active' and 'background' ones, as defined above. Thus, Method II depicts a possible comparison between two cases if and only if all the variables, from a to w , are part of the *active set* of variables, while in Method I, for all intents and purposes, the variables from a to l are part of the irrelevant *background set*.

Figure 2 presents the missing possibilities in comparisons by placing determinants and outcomes into a fourfold table and by adding to its visual scheme the distinction between active and background variables.

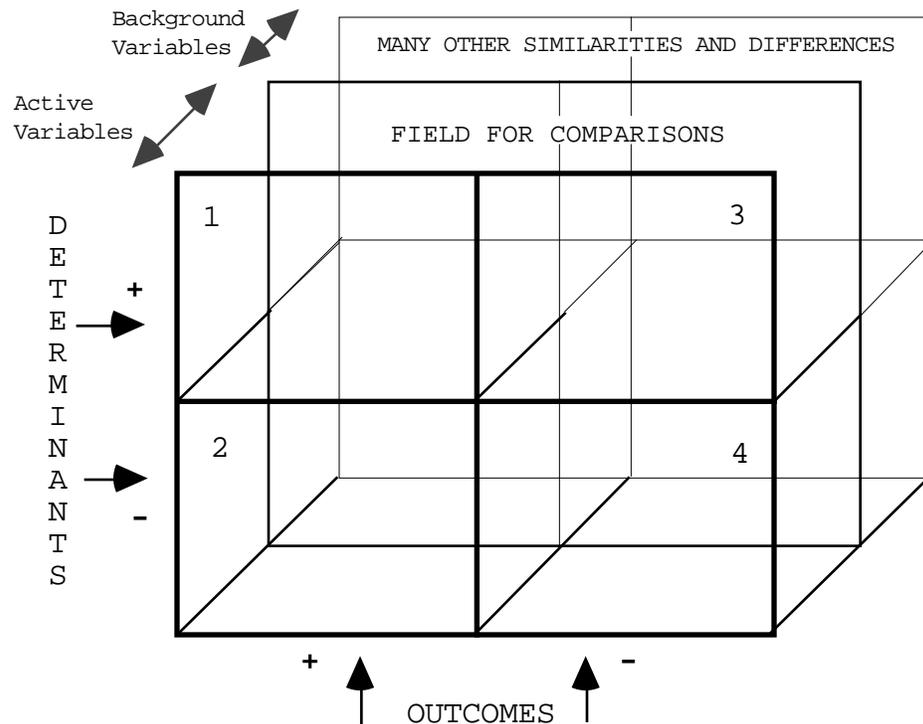
The foreground boxes in Figure 2, which contain only the active variables, illustrate the varieties of possible comparisons. The first box pertains to those cases that are seen as generating *similar cases comparisons* because they have the same outcomes in terms of the phenomenon to be explained, and matching determinants can be located to explain them. It is like Method I in the previous figure, except that box 1 in Figure 2 would contain only variables x , o ,

⁴⁶ Ruth Berins Collier and David Collier, *Shaping the Political Arena: Critical Junctures, the Labor Movement, and Regime Dynamics in Latin America* (Princeton, NJ: Princeton University Press, 1991), 16–18.

p , q , of Figure 1 (i.e., the active ones) while the rest (including some similar variables not contemplated in Method I) would fall in the background variable space behind box 1.

FIGURE 2

**BEYOND THE 'METHODS OF SIMILARITY AND DIFFERENCE':
ACTIVE AND BACKGROUND VARIABLES IN COMPARATIVE SITUATIONS**



The second box corresponds to *coincidental similarities*. The outcomes are similar but the determinants for them are different. There is no way to know a priori if a comparison is of a similar case or of a coincidental variety, and yet in the course of the research the analyst may well discover that there are different routes to the phenomenon under investigation. (Such results can lead the analyst to prefer creating a typology of determinants in order to report them.)

The third box represents comparisons in which the outcomes are different, but the determinants are very nearly similar; these are *intervening variable comparisons*, for a particular

variable (or a very limited set of them) can be clearly identified as having created the difference.⁴⁷ Box 3 is, therefore, like Method II in Figure 1, if all the latter's variables from *a* to *w* are considered active ones.

Finally, the fourth box pertains to *different cases comparisons*. The outcomes are different, and so are the determinants. Box 4 is like Method 2 in Figure 1 if the explanatory variables *r* through *w* are considered (in addition to *x*) to be the active ones, with the rest (*a* to *f*) being background variables.

The cases placed within each type of the typology created for the research establish similar cases comparisons (box 1), although on further examination they may turn out to be of a coincidental nature (box 2). Comparing cases from different types generates different cases comparisons (box 4), although a deeper analysis may reveal that they are in fact intervening variable comparisons (box 3).

Drawing Inferences

Similar cases comparisons are the best anchor for drawing inferences from comparisons and the research should always begin with them if at all possible, i.e., if there is more than one case per type. It is much better to search for the determinants of a phenomenon by studying cases where it is commonly present. It is riskier to do so with different cases comparisons because so many differences are bound to be present even within the active set of variables that it is much more difficult to determine which ones actually explain the phenomenon under examination; in such circumstances the analyst will be tempted to simply choose as explanatory variables those that correspond most closely to his or her theoretical preferences or preconceptions.

The similar cases comparisons may lead to the discovery that there are coincidental similarities at work. This conclusion should be reached only after considering the possibility that some of the background variables may provide a common explanation. If the type under examination does indeed have coincidental similarities, this means that it contains two or more subtypes of different causal determinants leading to the same outcome. It may be possible to continue doing similarity comparisons if the resulting subtypes still contain at least two cases. If no two such cases are found and the cases in the type resulted from sampling out of a larger universe, then it may be useful to go back to the original list of cases to add one or more to the respective subtypes, thereby altering the sample. At that point the similarity comparisons can be resumed.

The similarity comparisons within each type generate, however, only tentative conclusions which then must be checked by comparing cases across different types. If similar determinants are uncovered in what are supposed to be cases with different outcomes, then there is obviously

⁴⁷ Smelser uses the term 'near cases' to refer to these comparisons; Smelser, *Comparative Methods*, op. cit. (n. 2), 215–20. Linz and de Miguel's paper, op. cit. (n. 5), refers to them as well.

something wrong with the original conclusions. These different cases comparisons can be done exclusively with cases in the core universe if it contains a sufficient variety of different outcomes to the phenomenon under investigation (or, to put it differently, if the question is sufficiently general so as to produce a core universe containing several—at least three—species that can be included in a typology). If the question leads, for example, to a core universe which is then organized into four types given the main variations manifested by the phenomenon to be explained, and if each type has at least two cases (or if sampling has left two cases in each type), then there will be four similar cases comparisons (one per type) and six different cases comparisons as a check for the conclusions drawn from the former ones. However, if the core universe contains only cases with similar outcomes given a highly focused question, or even if it contains only two types (which would only permit one different cases comparison), it will be necessary to draw on cases in the ancillary universe to check with different cases the solidity of the explanations developed with the similar cases comparison. As the purpose of the different cases comparisons is to verify the solidity of the explanations derived from those that are similar, it is not necessary to do extensive research on the ancillary cases; it suffices to examine them enough to ascertain that they indeed do not have the same configuration of causal variables that have been seen to operate in the similar cases within the core universe. After all, the research question focuses on the core universe of cases and does not aim to explain the characteristics of the ancillary ones. If the analyst is driven to do so, then the research question in fact changes, and the previous steps of the research as noted above would have to be retraced.

The best support for the similarity analysis comes from difference comparisons of the intervening variable form (box 3 in Figure 2). Such box 3 comparisons may occur both with core as well as with ancillary cases. However, they are especially useful when trying to answer a question that is sufficiently generic so as to produce several outcomes in the core universe. With the intervening variable form of comparison the research is able not only to check the solidity of the similar cases analysis (i.e., by ascertaining that the causal variables do not appear in the same form), but it is also able to pinpoint much more specifically the likely reason or reasons for the difference. With a different cases contrast (box 4 in Figure 2) this latter conclusion does not go beyond showing that the cases are different indeed, and hence the similarity analysis internal to each type still bears the full weight of the explanation.

When there is only one case within a type, there is no alternative but to rely, in addition to the conclusions of the case study analysis, on different cases comparisons. Again, in such circumstances it is best if these comparisons prove to be of an intervening variable or variables kind (box 3 in Figure 2).

In conducting the comparisons, the analyst should not consider each variable as having only binary values, as if they could be coded simply yes/no or 0/1. Such treatment steers the

researcher towards viewing variables as having absolute values and one-to-one relations among them, rather than seeing them as having variously weak or strong expressions, as operating in contexts where they may acquire different significances and as forming part of complex sequences. Thus, in comparing party systems, the terms 'Conservative,' 'Liberal,' and 'Socialist' parties do not mean the same thing in different countries nor across time in the same country. Similarly, large landowners since the 1850s cannot be presumed to have everywhere the same political and economic interests. Land ownership can be used not only to produce a very broad variety of products from milk and wine to wood but also to secure credit for investment elsewhere in the national or international economy, to block other people's access to potential mineral deposits or to water, to have the necessary local influence to build a political career, and so on. Hence, Ragin's suggestion of using Boolean algebra as a shorthand mechanism to help annotate the effects of the variables can be useful,⁴⁸ as long as the researcher assigns the 0/1 values to variables after assessing their strength, placing them in their context, and viewing whether they operate as part of a sequence of variables when they exert their effects only through peculiar interactions among them.

Conclusion

The protocol contained in these pages is designed to steer the scholar who wishes to use comparative analysis for causal and theory-building purposes away from the most common pitfalls that are associated with this method. To restate these pitfalls in a nutshell, they are the following: 1) beginning comparative analysis focusing on a 'space' rather than with a clear-cut question; 2) seeking to answer with comparative analysis questions that should be addressed by using other approaches, particularly a statistical one; 3) not anchoring the analysis on clear definitions that are applied consistently across all cases; 4) not spending any time thinking about the universe of cases containing manifestations of the phenomenon under research; 5) not creating a typology explicitly, or devising it on a basis other than the variants shown by the phenomenon that is being investigated; 6) comparing fewer cases than the number of variations shown by the phenomenon, or sampling the universe without including at least two cases per variant, if at all possible, or at least one, if not; 7) trying to examine too many cases, thereby losing the ability to ground comparative analysis explanations on thorough case studies; 8) adopting a research design that leads only to examining different cases when it is possible to do similar cases comparisons; 9) neglecting to check the conclusions derived from the analysis of similar cases by looking at those that have different outcomes, referring, if necessary, to the ancillary universe.

⁴⁸ Ragin, *The Comparative Method*, op. cit. (n 14), chs. 6–8.

Comparative analysis is not an easy method to use properly. In addition to containing many possible sources of error, it often demands careful historical research, linguistic abilities, and the capacity to sift through a great deal of information. Juan Linz's work over several decades provides an excellent model to emulate. With his erudition, his knowledge of languages, his limpid questions and definitions, his attention to the universe of settings to which his questions apply, his in-depth studies of specific cases which can be seen as his 'sample' from the larger set of cases, and his careful judgements of the evidence without losing sight of its contextual configurations, Linz has basically followed, *avant la lettre*, the protocol presented here.