Political Methodology
Committee on Concepts and Methods
Working Paper Series

13

February 2007

Comparative-Historical Analysis in Contemporary Political Science

James Mahoney & P. Larkin Terrie

Northwestern University

(james-mahoney@northwestern.edu &

P-Terrie@northwestern.edu)

C&M
The Committee on Concepts and Methods
www.concepts-methods.org

IPSA
International Political Science Association
www.ipsa.ca

CIDE
Teaching and Research in the Social Sciences
www.cide.edu

The C&M working paper series are published by the Committee on Concepts and Methods (C&M), the Research Committee No. 1 of the International Political Science Association (IPSA), hosted at CIDE in Mexico City. C&M working papers are meant to share work in progress in a timely way before formal publication. Authors bear full responsibility for the content of their contributions.

The Committee on Concepts and Methods (C&M) promotes conceptual and methodological discussion in political science. It provides a forum of debate between methodological schools who otherwise tend to conduct their deliberations on separate tables. It publishes two series of working papers: "Political Concepts" and "Political Methodology."

Political Concepts contains work of excellence on political concepts and political language. It seeks to include innovative contributions to concept analysis, language usage, concept operationalization, and measurement.

Political Methodology contains work of excellence on methods and methodology in the study of politics. It invites innovative work on fundamental questions of research design, the construction and evaluation of empirical evidence, theory building and theory testing. The series welcomes, and hopes to foster, contributions that cut across conventional methodological divides, as between quantitative and qualitative methods, or between interpretative and observational approaches.

Submissions. All papers are subject to review by either a member of the Editorial Board or an external reviewer. Only English-language papers can be admitted. Authors interested in including their work in the C&M Series may seek initial endorsement by one editorial board member. Alternatively, they may send their paper to workingpapers@concepts-methods.org.

The C&M webpage offers full access to past working papers. It also permits readers to comment on the papers.

www.concepts-methods.org

Forthcoming in Henry Brady, Janet M. Box-Steffensmeier,
and David Collier, eds., *Oxford Handbook of Political Methodology*
(Oxford: Oxford University Press).

Although comparative-historical analysis has roots as far back as the founders of modern social science, its place in contemporary political science can be traced to a series of successful books published in the 1960s and 1970s, such as Moore (1966), Bendix (1974), Lipset and Rokkan (1968), Tilly (1975), and Skocpol (1979). Over the last twenty years, the tradition has sustained momentum in part through the publication of scores of major new books. This scholarship includes work across many of the key substantive areas of comparative politics: social provision and welfare state development (e.g., Esping-Anderson 1990; Hicks 1999; Huber and Stephens 2001; Pierson 1994; Skocpol 1992; Steinmo 1993); state formation and state restructuring (Bensel 1990; Ekiert 1996; Ertman 1997; Tilly 1990; Waldner 1999); economic development and market-oriented adjustment (Bunce 1999; Evans 1995; Haggard 1990; Karl 1997; Kohli 2004; Sikkink 1991); racial, ethnic, and national identities (Lustick 1993; Marx 1998; Yashar 2005); revolutionary change (e.g., Goldstone 1991; Goodwin 2001; Wickham-Crowley 1992); and democratic and authoritarian regimes (Collier 1999; Collier and Collier 1991; Downing 1992; Linz and Stepan 1996; Luebbert 1991; Mahoney 2001; Rueschemeyer, Stephens, and Stephens 1992). Beyond these books, comparative-historical analysis also represents a significant portion of published work in scholarly journals on comparative politics -- a fact that we empirically demonstrate in this chapter.

While comparative-historical analysis "certainly has claimed its proud place as one of the most fruitful approaches in modern social science" (Skocpol 2003: 424), it is

also true that methodological aspects of the approach are still received skeptically in some quarters. Perhaps most notably, scholars who pursue the statistical testing of hypotheses with large numbers of cases have raised concerns about this approach. They have argued that, from the standpoint of statistical methodology, the field of comparative-historical analysis violates well-known aspects of good research design and procedure (e.g., Coppedge 2007; King, Keohane, and Verba 1994; Geddes 1990, 2003; Goldthorpe 1997; Lieberson 1991, 1994, 1998). They use these criticisms as a basis for questioning whether the influential substantive findings produced in this field are, in fact, valid.

In this chapter, we suggest that existing concerns arise from a fundamental misunderstanding of the goals and methods of comparative-historical analysis. This misunderstanding, in turn, is linked to a failure to appreciate basic differences between comparative-historical analysis and statistical analysis. We show that these two research traditions are best understood as adopting distinct research goals, using different methods to achieve these goals, and thus quite justifiably pursuing different kinds of overall research designs. Once basic differences in research orientations are recognized, it becomes clear that advice for and criticisms of comparative-historical research derived solely from a statistical template are not appropriate (see also Brady and Collier 2004; Mahoney and Goertz 2006).

Clarifying the differences between comparative-historical and statistical analysis helps to promote a more fruitful dialogue among scholars representing these two traditions. Despite their different research objectives and contrasting methodological tools, comparative-historical and statistical researchers stand to benefit from better understanding of one another's methods and research practices. There are at least two

reasons why. First, insights from one tradition often can stimulate new and useful ideas for the other tradition. For example, insights about combinatorial causation and equifinality in comparative-historical methods have led to the creation of new statistical methods (Braumoeller 2003, 2004). The same is true of recent writings on necessary and sufficient causes (Clark, Gilligan, and Golder 2006). Likewise, comparing frequentist and Bayesian statistics has stimulated new insights about process tracing in comparative-historical research (Bennett 2006). And statistical techniques have been combined with formal qualitative comparative analysis in creative ways (Ragin 2000).

Second, the proliferation of multi-method research in contemporary political science makes knowledge of a wide range of methods increasingly important. Obviously, scholars who themselves pursue multi-method research should be well schooled in all of the relevant methodological traditions. At the same time, it seems increasingly important that methodologists themselves be able to offer sound advice to scholars who seek to combine statistical and case-study methods, including comparative-historical analysis. Obviously, no one methodologist can be expected to be an expert across the board; however, methodologists should know when the limits of their expertise are reached and thus when it is time to defer to specialists in other methodological orientations.

THE FIELD OF COMPARATIVE-HISTORICAL ANALYSIS

Several famous books are strongly associated with the comparative-historical tradition. But one might reasonably wonder about the overall commonality of the approach in contemporary political science, including in journals. Does a significant body of literature that can be called comparative-historical analysis actually exist beyond

the famous examples? How could we identify such a literature if it did? We address these questions by measuring several traits associated with comparative-historical analysis and assessing empirically the extent to which these traits are found together in published studies on comparative politics.

As with any research orientation, there are different ways of defining comparative-historical analysis. According to Mahoney and Rueschemeyer (2003), comparative-historical analysts investigate "big questions" -- substantively important and large-scale outcomes -- that take the form of puzzles about specific cases. In addressing these puzzles, comparative-historical researchers are centrally concerned with causal analysis, the examination of processes over time, and use of systematic and contextualized comparison. This understanding of the field is similar to that adopted by Collier (1998) and Skocpol (1979: 36-7; 1984: 1).

In this chapter, we are especially interested in defining the field in terms of characteristics that can be empirically measured. With this in mind, we emphasize here three core traits and two secondary traits as features that are important to most work in the field. The three core traits concern explanatory goal, conception of causation, and method of theory testing. On each of these three dimensions, comparative-historical analysis directly contrasts with statistical analysis (see Tables 1 and 2). Comparative-historical analysis adopts a causes-of-effects approach to explanation, a necessary and/or sufficient conception of causation, and process tracing to test theories. By contrast, statistical analysis uses an effects-of-causes approach to explanation, an average effects conception of causation, and regression techniques to test theories. In addition to these three main dimensions, two other attributes are often associated with comparative-

historical analysis: the use of a comparative set-theoretical logic and the analysis of temporal sequencing and/or path dependence.[1] In appendix A, we discuss the definition and measurement of each of these traits.

Tables 1 and 2 about here

With these defining traits at hand, we explored empirically whether a tradition of comparative-historical analysis could be found within the subfield of comparative politics. We gathered data from articles that recently appeared in the major comparative politics journals -- *Comparative Political Studies, Comparative Politics,* and *World Politics.* We set out to sample approximately 100 articles, evenly distributed, from these journals.[2] To make the sample representative of recent work in comparative politics, we first coded articles from 2005 for each journal. Articles from earlier years were coded if doing so was necessary to obtain a sufficiently large sample for the journal. In order to check the robustness of our results, we also analyzed approximately 40 articles on comparative politics from the main discipline-wide journals -- the *American Journal of Political Science,* the *American Political Science Review,* and the *Journal of Politics*.[3]

The data reveal that all five of the attributes associated with comparative-historical analysis commonly appear in the comparative politics literature, particularly in the articles published in the subfield journals. Table 3 reports the frequency with which

[1] We also coded studies according to whether they adopt a rational choice framework.
[2] The final sample consisted of 107 articles: 30 from *CP*, 38 from *CPS*, and 39 from *WP*. Note that descriptive, theoretical, and methodological articles were excluded from the sample.
[3] This sample consisted of 42 articles: 13 from *AJPS*, 15 from *APSR*, and 14 from *JOP*. These journals were not included in the original sample because the empirical studies they publish are almost exclusively statistical, and, as such, they are less representative than the subfield journals of the methodological diversity in comparative politics (Mahoney 2007). Nevertheless, the inclusion of these additional articles did not significantly alter the initial findings. Hence, the factor analysis results presented below in Table 4 are for the combined sample of articles from both the subfield and discipline-wide journals.

each of these attributes appeared in the subfield journals. The three core traits of the field -- a causes-of-effects approach to explanation, a necessary/sufficient conception of causation, and a process-tracing methodology -- each appear in over half of these articles. The two secondary traits -- a comparative set-theoretic logic and a concern with temporal sequencing or path-dependence -- also appear in a significant proportion of articles. What is more, attributes are not randomly distributed among the articles in the sample. They have a marked tendency to cluster together, as the factor analysis results in Table 4 show. The three core attributes of comparative-historical analysis are especially likely to hang together (as are the core attributes of statistical work). The temporal process/path dependence and comparative set-theoretic variables also tend to cluster with the three core attributes of comparative-historical analysis, though they exhibit lower factor loadings.

Tables 3, 4, and 5 here

Given the high frequency of each individual attribute and the strength of their tendency to cluster, it is clear that a body of work that can be called comparative-historical analysis is relatively common in the literature on comparative politics. Table 5 displays the number of articles from the subfield journal sample that can be classified as comparative-historical analysis according to four possible definitions (articles that have any of the three core attributes associated with statistical work are excluded). Just under half of the articles in this sample have all three of the core attributes. When more restrictive definitions are used, the proportion declines, but a non-trivial percentage of articles still qualify as comparative-historical analysis.

Comparative-historical analysis is, in short, a leading research tradition in the subfield of comparative politics based on prominence of current usage alone. Given this prominent place, it seems quite important that we assess soberly the validity of methodological concerns that have been raised about the tradition.

CONCERNS ABOUT METHODOLOGICAL PRACTICES

All observational studies in the social sciences confront important obstacles and potentially are subject to error. However, some analysts have argued that comparative-historical research faces especially grave problems that can often be easily avoided in statistical research (for a recent statement, see Coppedge 2007). The implicit or explicit implication is often that social scientists should pursue statistical research when possible (see also Lijphart 1971). In this section, by contrast, we argue that comparative-historical analysis and statistical analysis pursue different research goals, and that while they both face methodological challenges, they both play an essential role in generating knowledge in political science.

Selection Bias

Several methodologists have sounded alarm bells about the tendency of qualitative researchers to select cases based on their value on the dependent variable (Achen and Snidal 1989; Geddes 1991, 2003; King, Keohane, and Verba 1994). These stern warnings about deliberately selecting cases because they exhibit certain outcomes are especially applicable to comparative-historical studies, which quite explicitly engage in the practice. On the one hand, of course, selection on the dependent variable in this

field is hardly surprising, given that the research goal is precisely the explanation of particular outcomes. If one wishes to explain certain outcomes, it seems natural to choose cases that exhibit those outcomes. On the other hand, however, selecting cases based on their value on the dependent variable will bias findings in statistical research. From this standpoint, the practice seems to violate a basic norm of good research.

To evaluate this concern, when applied to comparative-historical analysis, we need to recognize that the statistical literature on bias deriving from selection on the dependent variable assumes that one wishes to generalize about average causal effects from a sample to a well-defined larger population. In comparative-historical research, however, one seeks to identify realized causal effects in particular cases; generalizing about averages from a sample to a larger population is at most a secondary goal. Insofar as comparative-historical researchers select what can be considered the entire universe of cases, therefore, standard issues of selection bias do not arise, regardless of whether the cases were chosen for their values on the dependent variable (for more extensive discussions, see Collier and Mahoney 1996; Collier, Mahoney, and Seawright 2004).[4]

Scope and Generalization

These observations raise questions about generalization in comparative-historical analysis. Obviously, comparative-historical researchers cannot simply make-up whatever definition of the universe of cases they so choose; the decision to limit a theory's applicability to a particular set of cases should not be arbitrary. One needs to ask, therefore, about the methodological basis for adopting a restrictive understanding of

[4] Problems of selection bias, as conventionally understood in the statistical literature, arise in comparative-historical studies primarily when analysts seek to generalize their theories beyond the initial cases investigated.

scope. This question is critical because the practice of restricting generalizations to a limited set of cases -- not issues of selection bias, as conventionally understood -- is often the real source of concern held by statistical methodologists about case selection in comparative-historical analysis. In particular, given the expansive definition of scope often employed in statistical research, the findings of comparative-historical researchers appear to be derived from a potentially unrepresentative sample of cases that is arbitrarily treated as the full population. Let us then explore these issues of scope and generalization.

Social scientists commonly impose scope restrictions on their findings to avoid problems associated with causal heterogeneity, which generates instability in estimates of causal effects. Indeed, primarily because of causal heterogeneity, social scientists of all traditions rarely develop theories that are intended to apply to all places and times. In addition to issues of causal heterogeneity, the need for stable concepts and measurement lead to the use of scope restrictions in which the analyst excludes cases where conceptual and measurement validity cannot be maintained.

In comparative-historical researcher, analysts adopt a narrow scope because they believe that causal and conceptual heterogeneity are the norm for their theories when assessed across large populations (Mahoney and Rueschemeyer 2003). But is this belief justified? Here we need to recognize that causal heterogeneity is not an ontological property inherent to a population of cases, but rather a feature of the *relationship* between a specific theory and a population of cases (Seawright and Collier 2004:276; Goertz and Mahoney 2007). A given population of cases may be heterogeneous vis-à-vis one theory but not another. The same is true of conceptual heterogeneity: cases may be

heterogeneous vis-à-vis some concepts but not others. One key implication is that some types of theories (or concepts) may be more likely than others to produce heterogeneity as the size of the population of cases increases.

There are very good reasons for believing that the type of theories evaluated in comparative-historical analysis are especially likely to generate causal heterogeneity in response to even modest increases in population size. To understand why, we need to compare the problem of missing variables in comparative-historical analysis and statistical analysis. This discussion, in turn, will take us back to the contrasting research goals of the two traditions.

In comparative-historical analysis, the exclusion of one or more important explanatory variables from a theory is appropriately regarded as a major problem (Ragin 2004:135-138). This is true because the very goal of comparative-historical research is to explain particular outcomes in specific cases as completely and adequately as possible. All relevant evidence pertaining to the cases should be gathered and assessed. If theories are missing key variables, or have misspecified key relationships among the variables that are included, these facts count significantly against comparative-historical arguments. The failure of previous investigators to consider one or more critical variables in fact provides a common basis for comparative-historical analysts to criticize existing work and build new explanatory theories. Missing variables are thus a constant potential source of causal heterogeneity in this field. Similar arguments can be extended to measurement error. In comparative-historical research, measurement error needs to be addressed and eliminated completely for each specific case, if possible (Ragin 2004). Otherwise, the goal of adequately explaining an outcome in particular cases is

compromised. In comparative-historical analysis, indeed, theory falsification often occurs with the change in value of one or a small number of variables. Accordingly, in this mode of research, one needs to strive to avoid measurement error for the cases analyzed, otherwise conceptual heterogeneity problems will likely arise.

There are thus good methodological reasons related to the need to avoid causal and conceptual heterogeneity that explain why comparative-historical researchers restrict the scope of their analysis to a limited number of cases. Given the kind of explanatory theory that these analysts pursue, built around the idea of realized causal effects for particular outcomes, they must quite carefully and deliberately define their population of cases to try to avoid all heterogeneity problems. Once the population is defined, even a modest increase in the number of cases runs the risk of excluding key causal factors relevant to the new cases or introducing measurement problems for the variables that are already included in the theory. Because significant modifications to the theoretical model are often required as new cases are added, the best solution may be to impose restrictive scope conditions that limit generalization.

In statistical analysis, by contrast, the goal of research is typically to estimate the average effects of one or more independent variables. Given this goal, missing variables are not necessarily a problem as long as key assumptions, especially that of conditional independence, still hold. Independent variables that are important for only a small subset of cases may be appropriately considered "unsystematic" and relegated to the error term of a regression model. Indeed, even missing independent variables that are systematically related to the outcome of interest will not necessarily bias estimates of the average effects, as long as conditional independence still applies. Likewise,

measurement error in statistical analysis does not raise the kinds of problem that it does in comparative-historical analysis. With large numbers of cases, measurement error is always present and cannot be completely eliminated. However, if one seeks to identify average effects, as statistical analysts do, measurement error is not a devastating problem as long as it is non-systematic or at least can be adequately modeled in the event that it is systematic. Unbiased estimates of average effects are quite possible in the presence of measurement error.

The fact that statistical analysis can maintain causal homogeneity even in the presence of missing variables and measurement error allows this kind of research to embrace a more expansive understanding of scope and generalization than comparative-historical analysis. For example, the inclusion of new cases with outcomes that were partially caused by idiosyncratic factors will not necessarily raise any special heterogeneity problems in statistical analysis. As long as assumptions such as conditional independence are valid and measurement error can be modeled, the extension of the scope to include new cases is usually not a problem in statistical research. Not surprisingly, therefore, statistical researchers worry less about issues of heterogeneity as they extend their arguments to new cases.

An important issue arises at this point: if comparative-historical explanations are highly fragile when new cases are introduced, but statistical explanations are much less fragile, does it not also follow that statistical explanations are "superior"? There are two reasons why this conclusion is not correct. The first is that the ability of statistical analysis to adopt a wide scope of generalization is dependent on the validity of key assumptions, especially conditional independence. In contemporary political science,

many empirical researchers feel quite comfortable making this assumption with little elaboration. Yet methodologists and statisticians often suggest that the assumption is an unrealistic leap of faith in much of the observational research pursued in the social sciences (Lieberson 1985; Freedman 1991). Insofar as the assumption of conditional independence cannot be sustained, statistical research suffers from unrecognized and unmodeled causal heterogeneity. In other words, it is possible that the expansive understanding of scope adopted in statistical analysis is often not appropriate.

Second, and more important for our purposes, it is essential to remember that comparative-historical researchers and statistical researchers have distinct research goals. If one wishes to explain particular outcomes in specific cases, as comparative-historical researchers do, then one must formulate theories in which it is not possible to easily extend the scope of generalization. The alternative is to reject the research goals of comparative-historical research; that is, to prohibit studies that seek to explain particular outcomes in specific cases and encourage scholars only to ask questions about average effects across large populations. For reasons that we discuss below, this kind of prohibition against asking comparative-historical questions would be extremely costly for social science knowledge. In short, if one is going to remain open to different forms of knowledge accumulation and permit the asking of comparative-historical research questions, then one must be willing to live with the restricted scope that accompanies comparative-historical analysis.

Assessing Causation with a Small N

Even if the limited scope of comparative-historical inquiry makes good sense, some analysts are still concerned that the small number of cases that fall within this scope

does not permit the scientific testing of hypotheses. From a statistical standpoint, a small population poses a degrees of freedom problem and insurmountable obstacles for hypothesis testing. How can comparative-historical researchers ever hope to adjudicate among rival explanations if they select so few cases?

The answer to this question again requires appreciating differences between statistical analysis and comparative-historical analysis. In statistical research, where the goal is to estimate average causal effects, one needs to have enough cases to control for relevant variables and still achieve specified confidence levels. However, with comparative-historical research, the goal is not to generalize about typical effects for a large population. Rather, the goal is to determine whether a given variable *did exert* a causal effect on an outcome in a particular set of cases. Given this goal, comparative-historical researchers need to embrace a distinct understanding of causation and indeed of explanation, which -- as we shall now see -- obviates the need for a large number of cases to achieve valid causal assessment.

Comparative-historical researchers ask the following question about any potential causal factor: did it exert an effect (alone or in combination with other variables) on the specific outcomes of interest in the particular set of cases that comprise the population? Sometimes, even a cursory examination will allow one to dispose of certain causal factors that might be generally relevant across a large population of cases. For example, when explaining the emergence of democracy in economically poor India or Costa Rica, the variable of development is clearly not useful (at least in the usual way), even though it is positively related to democracy in a large sample of cases. In other instances, however, plausible causal factors cannot be so quickly dismissed. Many potential causal factors

are “correlated” with the specific outcome of interest. How do researchers adjudicate among these rival explanations that are matched with the outcome of interest?

Comparative-historical researchers use the method of process tracing -- which involves marshalling “within-case” data -- to pass judgment on the validity of rival explanations emphasizing factors that cannot be eliminated through comparative matching techniques. Although here is not the place to discuss at length the mechanics of process tracing (see George and Bennett 2005), a few words are in order. Most basically, process tracing helps one to assess whether a posited causal factor actually exerts a causal effect on a specific outcome. This is done by exploring the mechanisms through which the potential causal factor is hypothesized to contribute to the outcome. If intervening mechanisms cannot be located, then doubt is cast upon the causal efficacy of the factor in question. By contrast, if appropriate intervening mechanisms are found, then one has grounds for believing that the factor in question did exert the effect. Beyond this, process tracing allows one to evaluate hypotheses by considering “sub-hypotheses” that do not necessarily refer to intervening mechanisms but that should be true if the main hypothesis of interest is valid (Mahoney and Villegas 2007).

It bears emphasis that this mode of hypothesis assessment does not require a large number of cases. Rather, like a detective solving a crime, the comparative-historical researcher who uses process tracing draws on particularly important facts from individual cases (see Goldstone 1997; McKeown 1999). Not all pieces of evidence count equally. Some forms of evidence are “smoking guns” that strongly suggest a theory is correct; others are “air-tight alibis” that strongly suggest a theory is not correct (Collier, Brady, Seawright 2004). For comparative-historical researchers, a theory is often only one key

observation away from being falsified. Yet, they may have certain kinds of evidence that suggest that the likelihood of theory falsification ever occurring is small.

Another relevant consideration concerns the conception of causation that is used in comparative-historical explanation. The various small-N comparative methods adopted by these researchers -- Mill's methods of agreement and difference, explanatory typologies, and qualitative comparative methods -- all assume understandings of causation built around necessary and/or sufficient causes (Ragin 1987, 2000; Mahoney 2000; Goertz and Starr 2003; Elman 2005; George and Bennett 2005).[5] By contrast, mainstream statistical methods assume forms of symmetrical causation that are not consistent with necessary and/or sufficient causation.

To assess hypotheses about necessary and sufficient causes, including combinations of causes that are jointly sufficient, a large number of cases usually is not needed. One or two cases may be enough for the simple purpose of eliminating (though not confirming) an explanation about necessary and sufficient causation. A medium number of cases is normally needed to achieve statistical confidence about the validity of an explanation that invokes necessary and/or sufficient causation solely by using cross-case matching techniques.[6] In some comparative-historical studies, this medium number of cases is analyzed. However, in small-N studies (e.g., N = 3), cross-case analysis is generally combined with process tracing. Because the N needed for necessary and

[5] Quite often, researchers treat individual causes as parts of a larger combination of causes that are together *jointly sufficient* for the outcome of interest (Mackie 1980). In fact, in this field, distinct combinations of causes may each be sufficient, such that there are multiple causal paths to the same outcome (see Ragin 1987).

[6] Using bayesian assumptions, for example, Dion (1998) shows that only five cases may be enough to yield 95 percent confidence about necessary causes. Using a simple binomial probability test, Ragin (2000, pp. 113-15) shows that if one works with "usually necessary" or "usually sufficient" causes, seven consistent cases are enough to meet this level of significance. Braumoeller and Goertz (2000) offer many examples of case-oriented studies that pass such significance tests.

sufficient causation is relatively modest, the "burden" that process tracing must carry in such studies is not overwhelming. Rather, the small N comparison does some of the work, with process tracing contributing the rest.

IMPLICATIONS OF DIFFERENCES

Our discussion has called attention to fundamental differences between comparative-historical analysis and statistical analysis. On the one hand, an awareness of these differences provides a basis for appreciating their distinctive contributions in political science. On the other hand, these differences raise questions about the extent to which the two research traditions might be meaningfully combined. By way of conclusion, we address these implications.

The kinds of knowledge generated by comparative-historical research and statistical research are clearly different. Comparative-historical studies tell us why particular outcomes happened in specific cases -- this is one important sense in which these studies are "historical," though there are others (see Mahoney and Rueschemeyer 2003; Pierson 2004; Skocpol 1984). The historical knowledge generated in comparative-historical studies, in turn, is relevant for policy and practical reasons. By teaching us about the genesis of outcomes in certain specific cases, comparative-historical research provides a critical foundation for hypothesizing about the effects of subsequent developments in these cases. Here a comparison with physicians who seek the medical history of their patients is useful. A cardiologist can offer better advice to a patient if the causes of the patient's earlier heart attack are well understood. Analogously, policy makers can pursue better interventions and offer more helpful suggestions if they

understand well the causes of prior relevant outcomes in the cases of interest. Indeed, if one understands a particular pattern of causation in a given case, one would seem especially well situated to explore whether the causal pattern might apply to another similar case. These points will be obvious to some, but the tendency for many in the discipline is nevertheless to assume that comparative-historical studies are of mostly historical relevance alone.

The strengths and pay-offs of statistical research are different. Whereas comparative-historical analysis is excellent at engaging complex theories with fine-grained over-time evidence, statistical research has the virtue of allowing for the testing of hypotheses about the average effects of particular variables (or specified interactions of variables) within large populations in a way that mimics aspects of a controlled experiment. Findings from large populations may or may not be relevant for thinking about particular cases. For example, a causal variable that promotes a given outcome in the population as a whole might have the opposite effect in a particular case of interest. But statistical findings certainly are relevant for generalizing. Indeed, if one wishes to offer policy advice or recommendations that are intended to -- on average -- make changes across a large population, the findings generated from statistical methods would seem especially appropriate.

This discussion is not intended to suggest that statistical work is irrelevant for thinking about particular cases. Nor is it meant to suggest that comparative-historical works cannot arrive at quite general findings. Rather, the point is that comparative-historical and statistical studies have different goals, produce different kinds of information, and are thus *tend to be* useful for different (though equally valid) purposes.

Given that each tradition has its own distinctive contributions to make, it is not surprising that there would be interest in combining the two, which perhaps could allow for a "best of both worlds" synthesis. While contemporary political scientists often value multi-method research, we nevertheless wish to raise here some cautionary notes about combining comparative-historical analysis and statistical analysis. We believe that the combination is more difficult to achieve than is sometimes suggested, and that multi-method research is not always an improvement over work that is exclusively comparative-historical or exclusively statistical.

When they engage in multi-method research, most analysts still pursue either a causes-of-effects approach or an effects-of-causes approach. In this sense, much multi-method research can be considered *primarily* comparative-historical or *primarily* statistical in orientation. With multi-method work that is primarily comparative-historical, the main goal remains the explanation of specific outcomes in particular cases. The statistical analysis is subservient to this goal. By contrast, with multi-method work that is primarily statistical, the main goal is to estimate average causal effects for a large population. Here one or more case studies are used to service this larger goal. Occasionally, of course, some studies will pursue both goals equally and thus truly cross the divide. However, in our sample, this kind of multi-method research characterized only 8.7 percent of all journal articles.

How is statistical analysis used in multi-method studies that are primarily comparative-historical in orientation? In the most basic way, generalizations from prior statistical research represent background knowledge that comparative-historical analysts must consider as they formulate their own explanatory hypotheses for their case studies.

All comparative-historical analysts react to prior general theories relevant to their outcomes, which often entails situating one's argument in relationship to existing statistical knowledge. Beyond this, comparative-historical researchers also may use statistical findings -- including findings they generate themselves -- in conjunction with process tracing. Much as a detective draws on knowledge of general causal principles to establish a link between suspect and crime, so too a comparative-historical researcher may use existing or newly discovered statistical findings when attempting to establish the mechanisms that connect cause and effect. For example, one might hypothesize that slow increases in grain prices in eighteenth-century France contributed to peasant revolts by deflating rural wages (i.e., the impact of declining grain prices on overall revolts worked through lower wages at the individual level). To develop this idea, a comparative-historical researcher might wish to carry out regression analysis to assess the effects of prices on wages in France -- to make sure that the two are, in fact, statistically linked net of other factors (see Goldstone 1991: 188-89). In doing this, the researcher collects a large number of observations from what is, given the perspective of the comparative-historical research design, a single case. Comparative-historical researchers thus may be especially likely to turn to statistical analysis when macro hypotheses in the small-N research design suggest mechanisms that work at lower levels of analysis. The statistical confirmation of these hypotheses serves the larger goal of validating the small-N argument.

For their part, statistical researchers may draw on the findings from comparative-historical analysis to develop their own hypotheses; comparative-historical work can inspire new ideas about causally relevant factors that can be tested in a statistical model.

Statistical researchers may also turn to case studies to determine whether findings make sense when assessed in light of an intensive analysis of specific cases. Through such analyses, statistical researchers can evaluate whether the statistical model is adequate, needs refining and retesting, or is deeply problematic and cannot be salvaged. Although in the course of the case analyses the researcher could potentially seek to develop fully adequate explanations of the particular cases, the overarching goal typically remains estimating the average effects of independent variables of interest for the population as a whole. For instance, in Lieberman's (2005) nested analysis approach, cases are selected not because their outcomes are inherently interesting, but rather because their location with respect to the regression line makes them good candidates for further assessing the validity of the statistical model. The goal of the nested analysis is generating valid knowledge about effects of causes; the comparative-historical evidence is mostly subordinated to the larger statistical design.

Our purpose in noting that one approach typically is subordinated to the other in multi-method research is not intended as a criticism. Rather, we emphasize the point to make it clear that most multi-method research is not equal parts quantitative and qualitative -- it is, rather, driven by primarily the goals and orientations of one side or the other. When this point is acknowledged, it becomes clear that multi-method research is an advantage only to the extent that the use of the secondary method actually and effectively supplements the main method of investigation. Statistical studies that offer superficial case studies as supporting evidence do not contribute to the explanation of particular outcomes in those cases. And if the case studies are carried out without attention to good methodological practice, they will not provide a reliable basis for

evaluating the statistical model either. By the same token, comparative-historical studies that use regression analysis in the course of process tracing are not necessarily more powerful than comparative-historical studies that do not use any statistical testing. The value added by statistical testing simply depends on what kind of evidence is needed for successful process tracing to be carried out. And the use of regression analysis with process tracing will not be fruitful if the regression analysis is poorly executed.

The message of this discussion is that there is nothing inherently wrong with conducting comparative-historical work that does not include a statistical component (and vice versa). Indeed, for many research projects, an additional secondary analysis using an alternative methodology is unnecessary or inappropriate. Hence, as political science increasingly moves toward and celebrates multi-method research, we believe that some of the best work that is produced in the discipline will eschew this trend and remain squarely centered in the field of comparative-historical analysis.

Table 1:
Attributes associated with comparative-historical analysis

Attribute	Definition
Causes-of-effects approach#	Research goal is to provide complete explanations of specific outcomes in particular cases.
Necessary/sufficient conception of causation#	Study treats individual causal factors or sets of multiple causal factors as necessary/sufficient for the outcome of interest.
Process-tracing method#	Study explores the mechanisms, within particular cases, through which potential causal factors are hypothesized to affect an outcome.
Comparative set-theoretic methods	Study tests theory using set-theoretical methods (e.g., Mill's method of agreement and difference, Boolean algebra).
Temporal processes modeled	Explanation emphasizes the sequencing of independent variables and/or path-dependent processes of change.

Core attribute of comparative-historical analysis

Table 2:
Core attributes associated with statistical analysis

Attribute	Definition
Effects-of-causes approach	Research goal is to estimate the effects of one or more independent variables on a dependent variable/s across a large number of cases.
Average effects conception of causation	Study treats independent variables as parameters whose average effects can be estimated across the full population of cases.
Regression methods	Regressions are used for theory testing.

Table 3:
Frequency of attributes associated with comparative-historical analysis*

CHA Attribute	Percentage of Articles
Causes-of-effects approach	55.1
Necessary/sufficient conception of causation	57.9
Process-tracing method	58.9
Comparative set-theoretic methods	43.9
Temporal processes modeled	21.5

*Percentage of studies from sample of comparative politics journals (n=107).

Table 4:
Factor analysis of methodological attributes*

	Factor
	1
Causes of effects approach	0.9473
Necessary/sufficient conception of causation	0.9428
Process tracing method	0.8854
Comparative set theoretic methods	0.7169
Temporal processes modeled	0.4644
Effects of causes approach	-0.8654
Average effects conception of causation	-0.9321
Regression methods	-0.9149
Rational choice theory	-0.4490

*Results are from principal factors analysis performed on the entire sample of articles (n=149) from comparative politics and discipline journals. Reported is the first factor extracted, which explains 97.48 percent of the total variance in these nine variables (Eigenvalue = 5.95605).

Table 5:
Frequency of comparative-historical articles*

Definition of CHA	**Percentage of articles**
Causes-of-effects * necessary/sufficient * process-tracing	49.5
Causes-of-effects * necessary/sufficient * process-tracing * comparative set-theoretic methods	36.4
Causes-of-effects * necessary/sufficient * process-tracing * temporal processes	18.7
Causes-of-effects * necessary/sufficient * process-tracing * comparative set-theoretic methods * temporal processes	14.0

Percentage of studies in sample from comparative politics journals (n=107). The "" symbol denotes the logical AND.

APPENDIX A:
A NOTE ON CODING PROCEDURES FOR EMPIRICAL ANALYSIS

All nine attributes are measured dichotomously as present or absent in a given article. We allow that, in principle, any study could possess any combination of attributes. Here we offer brief operational definitions for the nine attributes.

(1) *Causes-of-effect approach:* present if a central goal of analysis is to explain one or more specific outcomes of interest in particular cases.

(2) *Effects-of-causes approach*: present if a central goal of analysis is to estimate the extent to which particular independent variables account for variation in the dependent variable for a population as a whole rather than any particular case.

(3) *Necessary and sufficient conception of causation*: present if individual causal factors or sets of multiple causal factors are treated as necessary and/or sufficient for the outcome of interest.

(4) *Average effects conception of causation*: present if individual causes are assumed to exert average symmetrical effects that operate within the population as a whole.

(5) *Process tracing method*: present if specific pieces of data are used to test whether a posited causal factor exerts a causal effect in one or more particular cases by exploring the mechanisms through which the potential causal factor is hypothesized to contribute to the outcome of interest.

(6) *Regression method*: present if regression techniques are used to test hypotheses.

(7) *Comparative set-theoretic methods*: present if two or more cases are compared across causal and outcome variables to assess whether potential causal variables can be logically eliminated.

(8) *Temporal sequencing or path dependence:* present if the timing or sequencing of independent variables is hypothesized to affect the outcome, or if path-dependent processes are assumed to be present.

(9) *Rational choice framework*: present if theories are either formally or informally deduced from the assumption of rational, goal-oriented actors.

REFERENCES

Achen, Christopher H. and Duncan Snidal. 1989. “Rational Deterrence Theory and Comparative Case Studies.” *World Politics* 41: 143-69.

Bendix, Reinhard. 1974. *Work and Authority in Industry: Ideologies of Management in the Course of Industrialization.* Berkeley: University of California Press.

Bennett, Andrew. 2006. "Stirring the Frequentist Pot with a Dash of Bayes." *Political Analysis* 14: 339-44.

Bensel, Richard Franklin. 1990. *Yankee Leviathan: The Origins of Central State Authority in America, 1859-1877*. New York: Cambridge University Press.

Brady, Henry E. and David Collier, eds. 2004. *Rethinking Social Inquiry: Diverse Tools, Shared Standards*. Lanham and New York: Rowman & Littlefield.

Braumoeller, Bear F. 2004.“Hypothesis Testing and Multiplicative Interaction Terms.” *International Organization* 58: 807-20.

Braumoeller, Bear F. 2003. “Causal Complexity and the Study of Politics.” *Political Analysis* 11: 209-33.

Braumoeller, Bear F. and Gary Goertz. 2000. “The Methodology of Necessary Conditions.” *America Journal of Political Science* 44: 844-58.

Bunce, Valerie. 1999. *Subversive Institutions: The Design and the Destruction of Socialism and the State*. Cambridge: Cambridge University Press.

Clark, William Roberts, Michael J. Gilligan, and Matt Golder. 2006. “A Simple Multivariate Test for Asymmetric Hypotheses.” *Political Analysis* 14: 311-31.

Collier, David. 1998. “Comparative-Historical Analysis: Where Do We Stand?” *APSA-CP: Newsletter of the Organized Section in Comparative Politics* 9 (Winter): 1-2, 4-5.

Collier, David, Henry E. Brady, and Jason Seawright. “Sources of Leverage in Causal Inference: Toward an Alternative View of Methodology.” Pp. 229-66 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards,* edited by Henry E. Brady and David Collier.

Collier, David and James Mahoney. 1996. “Insights and Pitfalls: Selection Bias in Qualitative Research.” *World Politics* 49: 56-91.

Collier, David, James Mahoney, and Jason Seawright. 2004. “Claiming Too Much: Warnings About Selection Bias.” Pp. 85-102 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier.

Collier, Ruth Berins. 1999. *Paths Toward Democracy*. New York: Cambridge University Press.

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

Coppedge, Michael. 2007. *Approaching Democracy: Research Methods in Comparative Politics*. New York: Cambridge University Press.

Dion, Douglas. 1998. “Evidence and Inference in Comparative Case Study.” *Comparative Politics* 30: 127-46.

Downing, Brian M. 1992. *The Military Revolution and Political Change: Origins of Democracy and Autocracy in Early Modern Europe*. Princeton, NJ: Princeton University Press.

Elman, C. 2005. “Explanatory Typologies in Qualitative Studies of International Politics.” *International Organization* 59: 293-326.

Ekiert, Grzegorz. 1996. *The State Against Society: Political Crises and Their Aftermath in East Central Europe*. Princeton, NJ: Princeton University Press.

Ertman, Thomas. 1997. *Birth of the Leviathan: Building States and Regimes in Medieval and Early Modern Europe*. Cambridge: Cambridge University Press.

Esping-Anderson, Gøsta. 1990. *The Three Worlds of Welfare Capitalism*. Princeton, NJ: Princeton University Press.

Evans, Peter. 1995. *Embedded Autonomy: States and Industrial Transformation*. Princeton, NJ: Princeton University Press.

Freedman, David A. 1991. “Statistical Models and Shoe Leather.” In Peter Marsden, ed., *Sociological Methodology*. San Francisco: Jossey-Bass.

Geddes, Barbara. 1990. “How the Cases You Choose Affect the Answers You Get: Selection Bias in Comparative Politics.” Pp. 131-50 in *Political Analysis*, vol. 2, edited by James A Stimson. Ann Arbor: University of Michigan Press.

Geddes, Barbara. 1991. “Paradigms and Sand Castles in Comparative Politics of Developing Areas.” Pp. 45-75 in *Comparative Politics, Policy, and International Relations*, edited by William Crotty. Evanston, IL: Northwestern University Press.

Geddes, Barbara. 2003. *Paradigms and Sand Castles: Theory Building in Comparative Politics*. Ann Arbor: University of Michigan Press.

George, Alexander L. and Andrew Bennett. 2005. *Case Studies and Theory Development in the Social Sciences*. Cambridge, MA: MIT Press.

Goertz, Gary and H. Starr, eds. 2003. *Necessary Conditions: Theory, Methodology, and Applications*. Lanham, MD: Rowman & Littlefield.

Goertz, Gary, and James Mahoney. 2007. "Scope in Case Study Research." Manuscript.

Goldstone, Jack A. 1991. *Revolution and Rebellion in the Early Modern World*. Berkeley: University of California Press.

Goldstone, Jack A. 1997. "Methodological Issues in Comparative Macrosociology." *Comparative Social Research* 16: 107-20.

Goldthorpe, John H. 1997. "Current Issues in Comparative Macrosociology: A Debate on Methodological Issues." *Comparative Social Research* 16: 1-26.

Goodwin, Jeff. 2001. *No Other Way Out: States and Revolutionary Movements, 1945-1991*. Cambridge: Cambridge University Press.

Haggard, Stephan. 1990. *Pathways from the Periphery: The Politics of Growth in the Newly Industrializing Countries*. Princeton, NJ: Princeton University Press.

Hicks, Alexander. 1999. *Social Democracy and Welfare Capitalism: A Century of Income Security Politics*. Ithaca, NY: Cornell University Press.

Huber, Evelyne and John D. Stephens. 2001. *Development and Crisis of the Welfare State: Parties and Policies in Global Markets*. Chicago: University of Chicago Press.

Karl, Terry Lynn. 1997. *The Paradox of Plenty: Oil Booms and Petro-States*. Berkeley: University of California Press.

Kohli, Atul. 2004. *State-Directed Development: Political Power and Industrialization in the Global Periphery*. New York: Cambridge University Press.

Lieberman, E. S. 2005. " Nested Analysis as a Mixed Method Strategy for Comparative Research." *American Political Science Review* 99: 435-52.

Lieberson, Stanley. 1985. *Making It Count: The Improvement of Social Research and Theory*. Berkeley: University of California Press.

Lieberson, Stanley. 1991. “Small N’s and Big Conclusions: An Examination of the Reasoning in Comparative Studies Based on a Small Number of Cases.” *Social Forces* 70: 307-20.

Lieberson, Stanley. 1994. “More on the Uneasy Case for Using Mill-Type Methods in Small-N Comparative Studies.” *Social Forces* 72: 1225-37.

Lieberson, Stanley. 1998. “Causal Analysis and Comparative Research: What Can We Learn from Studies Based on a Small Number of Cases.” Pp. 129-45 in *Rational Choice Theory and Large-Scale Data Analysis*, edited by Hans-Peter Blossfeld and Gerald Prein. Boulder, CO: Westview.

Lipjhart, Arend. 1971. “Comparative Politics and the Comparative Method.” *American Political Science Review* 65: 682-93.

Lipset, Seymour Martin, and Stein Rokkan, eds. 1968. *Party Systems and Voter Alignments: Cross-National Perspectives*. New York: Free Press.

Linz, Juan J. and Alfred Stepan. 1996. *Problems of Democratic Transition and Consolidation: Southern Europe, South America, and Post-Communist Europe*. Baltimore: Johns Hopkins University Press.

Luebbert, Gregory M. 1991. *Liberalism, Fascism, or Social Democracy: Social Classes and the Political Origins of Regimes in Interwar Europe.* New York: Oxford University Press.

Lustick, Ian. 1993. *Unsettled States, Disputed Lands: Britain and Ireland, France and Algeria, Israel and the West Bank-Gaza*. Ithaca, NY: Cornell University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry: Scientific Inference in Qualitative Research*. Princeton, NJ: Princeton University Press.

Mackie, John Leslie. 1980. *The Cement of the Universe: A Study of Causation*. Oxford: Oxford University Press.

Mahoney, James. 2000. “Strategies of Causal Inference in Small-*N* Analysis.” *Sociological Methods and Research* 28: 387-424.

Mahoney, James. 2001. *The Legacies of Liberalism: Path Dependence and Political Regimes in Central America*. Baltimore: Johns Hopkins University Press.

Mahoney, James. 2007. “Debating the State of Comparative Politics: Views from Qualitative Research.” *Comparative Political Studies* 40: 32-8.

Mahoney, James, and Celso Villegas. 2007. "Historical Enquiry and Comparative Politics." Pp. 73-89 in *The Oxford Handbook of Comparative Politics*, edited by Carles Boix and Susan C. Stokes. Oxford: Oxford University Press.

Mahoney, James and Dietrich Rueschemeyer. 2003. "Comparative Historical Analysis" Achievements and Agendas." Pp. 3-38 in *Comparative Historical Analysis in the Social Sciences* edited by James Mahoney and Dietrich Rueschemeyer. New York: Cambridge University Press.

Mahoney, James and Gary Goertz. 2006. "A Tale of Two Cultures: Contrasting Qualitative and Quantitative Research." *Political Analysis* 14: 227-49.

Marx, Anthony W. 1998. *Making Race and Nation: A Comparison of South Africa, the United States, and Brazil*. Cambridge: Cambridge University Press.

McKeown, Timothy J. 1999. "Case Studies and the Statistical Worldview." *International Organization* 53: 161-90.

Moore, Barrington, Jr. 1966. *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*. Boston: Beacon Press.

Pierson, Paul. 1994. *Dismantling the Welfare State? Reagan, Thatcher, and the Politics of Retrenchment*. Cambridge: Cambridge University Press.

Pierson, Paul. 2004. *Politics in Time: History, Institutions, and Social Analysis*. Princeton, NJ: Princeton University Press.

Ragin, Charles C. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. Berkeley: University of California Press.

Ragin, Charles C. 2000. *Fuzzy-Set Social Science*. Chicago: University of Chicago Press.

Ragin, Charles C. 2004. "Turning the Tables: How Case-Oriented Research Challenges Variable-Oriented Research." Pp. 123-138 in in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier.

Rueschemeyer, Dietrich, Evelyne Huber Stephens, and John D. Stephens. 1992. *Capitalist Development and Democracy*. Chicago: University of Chicago Press.

Seawright, Jason and David Collier. 2004. "Glossary." Pp. 273-313 in *Rethinking Social Inquiry: Diverse Tools, Shared Standards*, edited by Henry E. Brady and David Collier.

Sikkink, Kathryn. 1991. *Ideas and Institutions: Developmentalism in Brazil and Argentina*. Ithaca, NY: Cornell University Press.

Skocpol, Theda. 1979. *States and Social Revolutions: A Comparative Analysis of France, Russia, and China*. Cambridge: Cambridge University Press.

Skocpol, Theda. 1992. *Protecting Soldiers and Mothers: The Political Origins of Social Policy in the United States*. Cambridge, MA: Belknap Press of Harvard University Press.

Skocpol, Theda. 1984. "Sociology's Historical Imagination." Pp. 1-21 in *Vision and Method in Historical Sociology*, edited by Theda Skocpol. Cambridge: Cambridge University Press.

Skocpol, Theda. 2003. "Doubly Engaged Social Science: The Promise of Comparative Historical Analysis." Pp. 407-428 in *Comparative-Historical Analysis in the Social Sciences*, edited by James Mahoney and Dietrich Rueschemeyer.

Steinmo, Sven. 1993. *Taxation and Democracy: Swedish, British and America Approaches to Financing the Modern State*. New Haven, CT: Yale University Press.

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

Tilly, Charles. 1990. *Coercion, Capital, and European States, AD 990-1990*. Cambridge, MA: Basil Blackwell.

Waldner, David. 1999. *State-Building and Late Development*. Ithaca, NY: Cornell University Press.

Wickham-Crowley, Timothy. 1992. *Guerrillas and Revolution in Latin America: A Comparative Study of Insurgents and Regimes since 1956*. Princeton, NJ: Princeton University Press.