Whaley, Barton (1973). *Codeword Barbarossa.* Cambridge, MA: MIT Press.

White, Harrison C. (1970). *Chains of Opportunity.* Cambridge, MA: Harvard University Press.

(1991). "Theory for Control in Social Relations by Decoupling Production from Identities among Networks." Unpublished manuscript, Center for the Social Sciences, Columbia University.

(in press). *Identity and Control: A Structural Theory of Social Action.* Princeton, NJ: Princeton University Press.

Whyte, William F. (1943). *Street Corner Society: The Social Structure of an Italian Slum.* Reprinted Chicago: University of Chicago Press, 1955.

Zannetos, Zenon S. (1966). *The Theory of Oil Tankship Rates.* Cambridge, MA: MIT Press.

Zipf, George K. (1949). *Human Behavior and the Principle of Least Effort: An Introduction to Human Ecology.* Cambridge, MA: Addison-Wesley.

LIEBERSON, STANLEY 1992 "Small N's and Big Conclusi[ons"] in What is a Case? Exploring the Foundations of So[cial] Inquiry, ed. C.C Ragin and H.S Becker. New York Cambridge University Press.

4

Small *N*'s and big conclusions: an examination of the reasoning in comparative studies based on a small number of cases[1]

STANLEY LIEBERSON

This chapter evaluates an approach which is gaining in usage, especially for historical and comparative problems. Namely, we will consider the causal inferences drawn when little more than a handful of nations or organizations – sometimes even fewer – are compared with respect to the forces driving a societal outcome such as a political development or an organizational characteristic.[2] Application of this method to a small number of cases is not new to sociology, being in one form or another a variant of the method of analytical induction, described by Znaniecki (1934:236) and analyzed succinctly by Robinson (1951) and Turner (1953).[3] These conclusions rely on a formalized internal logic derived from Mill's method of agreement and his method of difference [see the discussion of Mill in Nichols (1986:170ff).] The formal rigor of this type of analysis sets it off from other small-sample procedures which also imply causality, as say in *Street Corner Society* (Whyte 1943) or in the development of the model of urban structure and growth of Burgess (1925). It is also different from case studies which seek to point out merely that a given phenomenon *exists* in some setting, as opposed to an analysis of its causes. The comments are, however, to some degree relevant for evaluating the Boolean method proposed by Ragin (1987) for dealing with somewhat larger samples used in comparative and historical research. Moreover, although the analysis is stimulated by recent developments in macrohistorical research, it is pertinent to a wide variety of other studies that use Mill's logic with a small number of cases.

One has no difficulty appreciating the goal of applying formal procedures to make causal inferences in a manner analogous to what is otherwise restricted to studies based on a much larger number of cases. If data were available with the appropriate depth and detail for a large number of cases, obviously the researcher would not be working with these few cases (assuming a minimal time-energy cost). Since the data are not available, or the time-energy cost is too great, one can only approach these efforts with considerable sympathy for their objective. We address three questions: (1) What are the assumptions underlying these studies? (2) Are these assumptions reasonable? (3) What can be done to improve such studies in those instances when they might be appropriate forms of inquiry?

Probabilistic and deterministic perspectives

Let us start by distinguishing between causal propositions that are *deterministic* as contrasted with those that are *probabilistic*. The former posits that a given factor, when present, will lead to a specified outcome. The latter is more modest in its causal claim, positing that a given factor, when present, will increase the likelihood of a specified outcome. When we say, "If X_1 then Y," we are making a deterministic statement. When we say, "the presence of X_1 increases the likelihood or frequency of Y," we are making a probabilistic statement. Obviously, if given the choice, deterministic statements are more appealing. They are cleaner, simpler, and more easily disproved than probabilistic ones. One negative case, such that Y is absent in the presence of X_1, would quickly eliminate a deterministic statement.

Alas, a probabilistic approach is often necessary to evaluate the evidence for a given theoretical perspective, even if we think in deterministic terms. This occurs for a variety of reasons, not the least being measurement errors – a serious problem in the social sciences. The existence of a measurement error means that a given data set may deviate somewhat from a hypothesized pattern without the hypothesis being wrong. In addition to this technical matter, there is an additional problem: complex multivariate causal patterns operate in the social world, such that a given outcome can occur because of the presence of more than one independent variable and, moreover, may not occur at times because the influence of one independent variable is outweighed by other influences working in the opposite direction. Under such circumstances, the influence of X_1 is only approximate (even without measurement errors), unless one can consider all of the other independent variables, through controls or otherwise.

Furthermore, we often do not know or cannot measure all of the factors that we think will influence Y. As a consequence, we are again obliged to give up on a deterministic *measurement* of the influence of X_1 on Y, even if we are prepared to make a deterministic statement about its influence. There are yet other reasons for reverting to a probabilistic

rather than a deterministic approach, namely, the role of chance in affecting outcomes. Beyond consideration here is the question of whether *chance* per se exists or is simply a residual label referring to our ignorance about additional influences and/or inadequate measures for the variables under scrutiny. In either case, some form or another of indeterminacy is clearly useful to employ in the physical sciences, let alone in the social sciences [see examples in Lieberson (1985:94–7)]. Any of these factors would lead to probabilistic statements rather than deterministic statements of outcome.

This distinction is more than merely an academic one. Rather, it is embedded in our daily thinking. Suppose we examine the influence of alcohol on automobile accidents. Even if we believe there is such an influence, we still will expect some sober drivers to have chargeable accidents and not all drunk drivers to experience accidents. If we find that some sober drivers did cause accidents and some drunk drivers did not, these deviations would not lead us to reject automatically the proposition that drunkenness causes automobile accidents.[4] Rather, we would look at a set of data and ask if the probability or frequency of accidents were greater for drunk than for sober drivers. Why is this so? Even if taking a deterministic view, we would expect several factors to influence the likelihood of an accident, alcohol being only one of them. Indeed, we would expect an interaction effect for drunkenness, such that one drunk driver might run a red light in a busy intersection and have an accident, whereas another driver might be fortunate to enter the intersection when the light was green. To be sure, we might want to take some of these additional factors into account, and we would then expect the influence of drinking to be more sharply displayed. But it is unlikely that we could isolate alcohol's influence from all of the additional conditions that either prevent drinking from causing an accident or lead a sober driver to have an accident. The net effect is that we will not totally reject our idea about alcoholism and driving if we compare a drunk driver with a sober one and find the latter has an accident and the former does not. Likewise, if we learn of one drunk driver who has an accident and a sober driver who does not, that will hardly be persuasive data that the pattern is indeed in the direction anticipated. The point is clear-cut: *a deterministic theory has deterministic outcomes, but often we can measure it only in probabilistic terms.*

Despite these facts, small-N studies operate in a deterministic manner, avoiding probabilistic thinking either in their theory or in their empirical applications. As one distinguished proponent of the small-N approach puts it, "in contrast to the probabilistic techniques of statistical analysis – techniques that are used when there are very large numbers of cases and continuously quantified variables to analyze – comparative historical analyses proceed through logical juxtapositions of aspects of small numbers of cases. They attempt to identify invariant causal configurations that necessarily (rather than probably) combine to account for outcomes of interest" (Skocpol 1984:378). One good reason for this disposition is the following principle: *except for probabilistic situations which approach 1 or 0 (in other words are almost deterministic), studies based on a small number of cases have difficulty in evaluating probabilistic theories.*

Let us draw an analogy with flying a given airline. Suppose a rude employee is encountered, or luggage is lost, or the plane is delayed. One could, after such an experience, decide to use a different airline. However, one would know that although airlines may differ in their training programs, employee relations, morale, luggage practices, airplane maintenance, and other factors affecting their desirability, a very small number of experiences is insufficient to evaluate airlines with great confidence. If airlines differ, it is in the *frequency* of unpleasant experiences rather than that one airline has only polite employees, never loses luggage, or avoids all mechanical problems. Based on a small number of experiences, one may decide to shun a certain airline, and the decision is not totally wrong, since the probability of such experiences in any given small number of events is indeed influenced by the underlying distribution of practices in different airlines. However, conclusions drawn on the basis of such practices are often wrong. We would know that passengers with small numbers of experiences will draw very different conclusions about the relative desirability of various airlines. This is because a small number of cases is a bad basis for generalizing about the process under study. Thus if we actually knew the underlying probabilities for each airline, it would be possible to calculate how often the wrong decision will occur based on a small number of experiences. The consumer errors are really of no great consequence, since making decisions on the basis of a small number of events enables the flyer to respond in some positive way to what can otherwise be a frustrating experience. Such thinking, however, is not innocuous for the research problems under consideration here; it will frequently lead to erroneous conclusions about the forces operating in society. Moreover, other samples based on a small number of different cases – when contradicting the first sample, and this is almost certain to occur – will create even more complicated sets of distortions as the researcher attempts to use deterministic models to account for all of the results. This, in my judgment, is not a step forward.

Briefly, in most social-research situations it is unlikely that the requirements of a deterministic theory will be met. When these conditions

are not met, then the empirical consequences of deterministic and probabilistic theories are similar in the sense that both will have to accept deviations: the former because of errors in measurement and controls; the latter both because of those reasons and because the theory itself incorporates some degree of indeterminacy (due to inherent problems in either the measurement or knowledge of all variables or because of some inherent indeterminacy in the phenomenon).

The implications of this are seen all the time in social research. In practice, for example, it is very difficult to reject a major theory because it appears not to operate in some specific setting. One is wary of concluding that Max Weber was wrong because of a single deviation in some inadequately understood time or place. In the same fashion, we would view an accident caused by a sober driver as failing to disprove the notion that drinking causes automobile accidents.

Suppose, for example, there is a single deviation among a small number of cases or a modest number of deviations among a larger number of cases. What are the consequences for the deterministic theory under consideration?[5] If the deterministic theory is univariate, that is, either only one variable or one specific combination of variables (an interaction) causes a given outcome, the theory can be rejected with a single deviation if one is confident that there are no measurement errors (a nontrivial consideration for either statistical or "qualitative" descriptions) and there are no other possible causes of the dependent variable.[6] As for a multivariate deterministic theory, where more than one variable or more than one combination of variables could account for the consequence, it can be rejected with a single deviation if there is confidence that there are no measurement errors – as before – *and* also that all other factors hypothesized to be affecting the outcome are known and fully taken into account.

The importance of all of this is that the formal procedures used in the small-*N* comparative, historical, and organizational analyses under consideration here are all deterministic in their conception. Indeed, small-*N* studies cannot operate effectively under probabilistic assumptions, because then they would require much larger *N*'s to have any meaningful results. This becomes clear when we watch the operation of their reasoning with the methods described by Mill.

Mill's method

As Skocpol (1986) observes, the key issue is the applicability of Mill's "method of agreement" and "method of difference" to such data. Nichols (1986) agrees, but then criticizes the application of this logic in an

Table 4.1. *Application of the method of difference*

Accident (Y)	Drunk driving (X_1)	Car entering from right-hand direction (X_2)	Driver speeding (X_3)	Runs a red light (X_4)
Yes	Yes	Yes	No	Yes
No	Yes	No	No	Yes

earlier study; for example, she shows that it assumes interaction effects but no additive influences. I will build on, and modify, this important critique here.

Let us start with the method of difference, which deals with situations in which the dependent variable (outcome) is not the same for all of the cases. Here the researcher examines all possible independent variables that might influence this outcome, looking for a pattern where all but one of the independent variables do not systematically vary along with the dependent variable. Examples of this might be where X_i is constant in all cases or varies between cases in a manner different from the dependent variable. This method is applied even with two cases, so long as only one of the independent variables differs, while the others are constant across the cases (Orloff and Skocpol 1984). Table 4.1 illustrates this type of analysis. For simplicity, let us assume that all the independent variables as well as the explanandum are dichotomies with "yes" and "no" indicating the presence and absence of the attribute under consideration. To illustrate my points as clearly as possible, I have used an illustration based on automobile accidents. The logic is that followed in Mill's methods and is identical with that employed in these deterministic studies of macrophenomena.

Applying the method of difference to the hypothetical data in Table 4.1, we would conclude that the auto accident was caused by X_2, because in one case a car entered the intersection whereas in the other case no car did. We would also conclude that the accident was *not* caused by drunk driving or the running of a red light, because the variables (respectively X_1 and X_4) were the same for both drivers, yet only one had an accident. Such conclusions are reached only by making a very demanding assumption that is rarely examined. The method's logic assumes no interaction effects are operating (i.e., that the influence of each independent variable on Y is unaffected by the level of some other independent variable). The procedure cannot deal with interaction effects; the procedure cannot distinguish between the influence of inebriation or running a red light and the influence of another constant, such as the benign fact that both drivers were not exceeding the speed limit. Since X_1 and X_4 are constant, under this logic it would follow that neither inebriation nor running a red light had anything to do with the accident occurring. *The procedure does not empirically or logically eliminate interaction effects. Rather, it arbitrarily assumes that they do not operate and that therefore constants cannot influence the dependent variable.*[7] Unless interactions are automatically ruled out a priori, this means that the results in Table 4.1 (and all other small-N applications of the method) *fail* to provide any determination of the influence of variables X_1, X_3, and X_4 on the phenomenon under consideration.[8] Just to make the point very clear, consider another example of the same sort: ten people apply for a job; five are blacks and five are whites. One of the five blacks and all the five whites are hired. Applying the method of difference, one would conclude that race did not affect employment. Rather, it would have to be some variable that separates all of the employed persons from the four who did not get a job. Using a small N with the method of difference, it is not possible to examine interaction effects or multiple causes. Their absence is assumed.

The reader should also note how this method has a certain limited generality unless one assumes, a priori, that only one variable causes the phenomenon under study. For variables that are constant, it is impossible to rule out their influence under different levels simply because there are no measurements. From Table 4.1, for example, we know that an accident occurs although X_3 is constant. Even ignoring the question of interaction effects, it is impossible to conclude that X_3 does not cause accidents unless one assumes there is only one cause of accidents. In this case, and this asymmetry is common in small-N studies, we only know about situations where drivers are *not* speeding. Note again the assumptions that are introduced: if there is any generality to the results, it means that only a single causal variable is operating, otherwise, under the logic used in such studies, the influences of constants are not really taken into account in the method of difference.[9] This has a great bearing on the generality of such small-N comparative studies.

In Table 4.2, we have a new situation in which two drivers both experience accidents. As before, the two drivers are drunk, both cars run red lights, and again in only one instance another car was appropriately entering the intersection, whereas in the other instance there was none. This time, however, the second person was driving at a high speed, whereas the first driver was not. Intuitively, it is not unreasonable that high speed driving could affect the chances of an accident, say causing

Table 4.2. *Application of the method of agreement*

Accident (Y)	Drunk driving (X_1)	Car entering from right-hand direction (X_2)	Driver speeding (X_3)	Runs a red light (X_4)
Yes	Yes	Yes	No	Yes
Yes	Yes	No	Yes	Yes

a skid, or the car could have failed to make a turn in the intersection. At any rate, since both drivers have accidents, the logic generated by Mill's method of agreement is applied here, where presumably the causal variable is isolated by being the only constant across the two instances, whereas all of the other attributes vary. However, notice what happens under that logic here. The previous cause, X_2, is now eliminated since it varies between two drivers who both have accidents. Previously, X_1 and X_4 could not have caused an accident, but are now the only two contenders as a possible cause. Since only one driver is going at a high speed now and both drivers have accidents, it follows that the addition of this factor could not have caused an accident, an extraordinary conclusion, too. What has gone wrong? This is an example of how Mill's method cannot work when more than one causal variable is a determinant and there is a small number of cases. Comparison between the two tables shows how volatile the conclusions are about whether variables cause or do not cause accidents. Every fact remains the same regarding the first driver in both cases, but the fact that the second driver was speeding and therefore had an accident completely alters our understanding of what caused the first driver to have an accident. Another shortcoming to such data analyses is that the conclusions are extremely volatile if it turns out that a multideterministic model is appropriate. Moreover, with a small-*N* study, although it is possible to obtain data which would lead one to reject the assumption of a single-variable deterministic model (assuming no measurement error), it is impossible for the data to provide reasonable assurance that a single-variable deterministic model is correct, even if the observed data fit such a model.

These comparisons suggest more than the inability of Mill's methods to use a small number of cases to deal with a multivariate set of causes. As Nichols points out, Mill had intended these methods as "certain only where we are sure we have been able to correctly and exhaustively analyze all possible causal factors" (1986:172). Nichols goes on to observe that Mill rejects this method when causality is complex or when more than one cause is operating. Beyond these considerations, important as they clearly are, the foregoing analysis also shows how exceptionally vulnerable the procedure is to the exclusion of relevant variables. In Table 4.2, had we left out X_4, inebriation clearly would have been the causal factor, but it is not clear because X_4 is included. This is always a danger; large-*N* studies also face the potential danger that omission of variables will radically alter the observed relations, but the susceptibility to spurious findings is much greater here.

Suppose a researcher has a sufficient number of cases such that there are several drivers who have accidents and several who do not. Would the deterministic model based on a small number of cases now be facilitated? In my opinion, it is unlikely. If drinking increases the probability of an accident but does not always lead to one, and if sobriety does not necessarily enable a driver to avoid causing an accident, then it follows that some drunk drivers will not experience an accident, and some accidents will be experienced by sober drivers. Under the circumstances, there will be no agreement for these variables among all drivers experiencing an accident, and there will be no agreement among those not experiencing an accident. This means that neither of Mill's methods will work. A difference in the frequency of accidents linked to drinking will show up, but this of course is ruled out (and more or less has to be) in the deterministic practices involving small-*N* studies. Multicausal probabilistic statements are simply unmanageable with the procedures under consideration here.[10]

One way of thinking about this small-*N* methodology is to visualize a very small sample taken from a larger population. Let us say we have a small sample of nations or of political developments drawn randomly from the universe of nations or the universe of political developments.[11] What is the likelihood that the application of Mill's methods to this small sample will reproduce the patterns observed for the larger universe? Rarely, in my estimation, do we encounter big-*N* studies in which all of the relevant causal variables are determined and there are no measurement errors such that all cases are found so neatly as is assumed here with small-*N* studies. Yet in order to draw a conclusion, the small-*N* study assumes that if all cases were equally well known, the patterns observed with the small sample would be duplicated without exception. Is this reasonable? Also ask yourself how often in large-*N* studies would restrictions to a deterministic univariate theory make sense.

It is also impossible for this type of analysis to guard against the influence of chance associations. Indeed the assumption is that "chance" cannot operate to generate the observed data. Because it is relatively easy to develop a theoretical fit for small-*N* data, researchers are unable to guard against a small-*N* version of the ad hoc curve fitting that can be employed in large-*N* studies [see the discussion of Taylor's theorem in Lieberson (1985: 93)]. Ironically, small-*N* deterministic analyses actually have the same goal as some types of large-scale empirical research, namely, explaining all of the variance. The former is just another version of this, subject to the same dangers (Lieberson 1985: chap. 5), along with special ones due to their very demanding assumptions necessary when using a small *N*.

Theoretical concerns

Two implications follow from this review; one is theoretical and the other deals with empirical procedures.

Dealing with the theoretical questions first, obviously the small-*N* applications of Mill's methods cannot be casually used with all macro-societal data sets. The method requires very strong assumptions: a deterministic set of forces; the existence of only one cause; the absence of interaction effects; confidence that all possible causes are measured; the absence of measurement errors; and the assumption that the same "clean" pattern would occur if data were obtained for all cases in the universe of relevant cases.

At the very least, users must recognize that these assumptions are mandatory in this procedure. The issue then becomes this: Under what conditions is it reasonable to make these assumptions ("reasonable" in the sense that they have a strong likelihood of being correct)? Keep in mind that the empirical data themselves cannot be used to test whether the assumptions are correct or not; for example, the empirical data gathered in the typical small-*N* study cannot tell us if a univariate deterministic cause is operating or if there are no interaction effects. Theories of large-scale organizations, "qualitative" or not, must direct themselves to these questions before the data analyses begin. Moreover, the theories have to develop ways of thinking about these problems so the researcher can decide if they are reasonable. Admittedly, this is vague advice, and hopefully those dealing with this type of research will come up with solutions. Certainly, the Boolean method proposed by Ragin (1987) is a step in the right direction, although it does require a relatively larger *N* than the type of small-*N* studies under consideration here.[12]

The quality of qualitative data

It should be clear how critical it is that small-*N* studies take extraordinary care in the design and measurement of the variables, whether or

not it is a so-called qualitative study. Care is always appropriate, but the impact of error or imprecision is even greater when the number of cases is small. Keep in mind that the deterministic model used in these studies requires error-free measurement. The choice of cases for study is itself critical, requiring great thought about the appropriate procedure for choosing them. Presumably, these are self-evident facts to practitioners of this approach, and the intense scrutiny of a small number of cases should mean exceptional care with the descriptions.

However, exceptionally rigorous practices are necessary to avoid some methodological pitfalls. If a small number of cases are selected using reasonably rigorous criteria, then it makes a great difference whether the outcomes are the same or not in each case. If the same, then the method of agreement is used such that a solution occurs only if one variable is constant in all cases; if different, then the only solution occurs when all but one of the variables are constant across all the cases. All of this is nothing more than a repetition of procedures dating back to Mill. Less obvious, at least as far as I can tell, are the implications this has for the delineation of each independent variable. If an independent variable consists of nominal categories, there should be little difficulty, since presumably trained observers would agree on the classification of each measure. The researcher uses the same checks as would be performed in any large-scale study (e.g., content analysis). But if the independent variable is even ordinal, there is a certain arbitrariness in the way an ordered variable is dichotomized or otherwise divided (polytomized).

To simplify the point, just consider dichotomies. The method of agreement will work only if all the cases for one causal variable fall in the same category *and* if no other variable has such uniformity. This means that the cutoffs are critical. The same holds for the method of difference, but here the results must be such that the results are uniform for all but one variable, with the one critical exception being associated with differences in the dependent variable. Under the circumstances, the delineation of the dichotomies or polytomies is critical and has to be done as rigorously as possible since the boundaries will influence the results enormously. All of this means that rigor is mandatory when locating the variables if they are nominal, and even more so when they are ordinal, for example, careful driver versus careless driver, etc.

With the method of difference, where there is an inverse relationship between the number of cases and the difficulty of finding all but one variable constant across cases, researchers have to guard against using such broad categories as to make it relatively easy for cases to fall under the same rubric. With the method of agreement, where it is vital that all but one variable be different across the cases, the danger is in constructing narrow categories within each variable so that it will be relatively hard for cases to fall under the same rubric. In short, because of the subtle pressure to obtain only one variable that is homogeneous (in the case of agreement in the dependent variable) or only one variable that is heterogeneous (in the case of disagreement in the dependent variable), one must also guard against the bracketing of attributes in the former case, and decomposition in the latter. For this method to work at all, researchers must introduce formal criteria for these decisions which can be followed in advance of a given research project. To my knowledge, they do not exist at this time. (It would be an interesting study in both the sociology of knowledge and research methodology to see if the breadth of categories used in recent studies is related to whether the study involves cases calling for one or the other method.)

Because of the small *N*'s and the reasoning this method requires, it is vital to include all possible causal variables. Yet this will tend to lead to inconclusive results if carried out in a serious way, since the method of agreement will probably turn up with more than one variable that is constant for all the cases and, likewise, the method of difference will have more than one independent variable that is associated with the difference in the dependent variable. Suppose, for example, we find that a drunk driver has no automobile accident, but the sober driver experiences one. In such a case, using the small-*N* methods practiced in historical sociology, the investigator is in danger of concluding that sobriety causes automobile accidents, or at the very least is the cause in the observed situation. At best, and only if the correct causal factor is included, the study will conclude that either sobriety or some other factor causes automobile accidents. At worst, if the correct causal factor is excluded, sobriety will be the cause. So there is a kind of dilemma here; a "clean" result will tend to occur only with a modest number of independent variables, but this very step is likely to increase the chances of an erroneous conclusion.

Also, the relationship between the independent variables and the dependent variable is distorted if the cases are selected so as to have agreement or disagreement with respect to the dependent variable (rather than simply sampling from all of the cases). It can be shown that sampling in order to obtain a certain distribution with respect to the dependent variable ends up distorting the explanandum's association with the independent variables (unless the ratio of odds is used). Obviously not all cases are equally good, since the quality of the data presumably varies between them, as does the researcher's access to and knowledge about the relevant information. However, this distortion is beyond that problem and makes it even harder to assume that one small

sample and another small sample by a second researcher can be combined to generate a more accurate model of the forces under consideration.

Conclusions

A number of assumptions made in these small-*N* macrocomparative studies are not only very demanding, but to my knowledge they are normally not made explicit or seriously examined. Yet they entail assumptions that are usually indefensible in social research. This leads to a certain curiosity. One possibility is that these assumptions occur because they are the only way of proceeding with such data sets, not because the investigators commonly believe they are correct. In that circumstance, the same assumptions will collapse when studies based on large *N*'s are attempted. Another possibility is that such assumptions are appropriate for certain subject matters such as major institutions, nations, and the like. If that is the case, then a very important step is missing, since these assumptions are rarely justified with empirical data based on a larger number of cases. (That is, as a test, by sampling an extremely small number of cases from large macrosocietal data sets it should be possible to show that the same conclusions would occur with Mill's method as by studying the universe of cases.) At the moment, however, it appears that Mill's procedures cannot be applied to small-*N* studies. There are strong grounds for questioning the assumptions essential to causal analyses generated by such procedures.

As matters now stand it appears that the methodological needs are generating the theory, rather than vice versa. Put bluntly, application of Mill's methods to small-*N* situations does not allow for probabilistic theories, interaction effects, measurement errors, or even the presence of more than one cause.[13] For example, in the application shown earlier, the method cannot consider the possibility that more than one factor causes automobile accidents or that there is an interaction effect between two variables.[14] Indeed, if two drivers are drunk, but one does not have an accident, the procedure will conclude that the state of inebriation could not have been a cause of the accident that did occur.

I have selected the automobile-accident example because it should be patently clear that the special deterministic logic does not operate in that instance. Perhaps one may counter that nations and major institutions are neither persons nor roulette wheels; surely their determination is less haphazard, and therefore deterministic thinking is appropriate for these cases. Hence, one might argue, the points made are true for automobile accidents but not for major social institutions or other macrosocietal phenomena. This sounds plausible, but is it true? It turns out that many deep and profound processes are somewhat haphazard too, not so easily relegated to a simple determinism. Elsewhere, I have cited a wide variety of important phenomena which appear to involve chance processes, or processes that are best viewed that way. These include race riots, disease, subatomic physics, molecules of gas, star systems, geology, and biological evolution (Lieberson 1985:94–9, 225–7). One must take a very cautious stance about whether the methods used in these small-*N* studies are appropriate for institutional and macrosocietal events. At the very least, advocates of such studies must learn how to estimate if the probabilistic level is sufficiently high that a quasi-deterministic model will not do too much damage.

References

Burgess, Ernest W. (1925). "The Growth of the City: An Introduction to a Research Project," pp. 47–62 in Robert E. Park, Ernest W. Burgess, and Roderick D. McKenzie (eds.), *The City*. Chicago: University of Chicago Press.

Isaac, Larry W., and Larry J. Griffin (1989). "Ahistorism in Time-Series Analyses of Historical Process: Critique, Redirection, and Illustrations from U.S. Labor History." *American Sociological Review* 54:873–90.

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

Marini, Margaret Mooney, and Burton Singer (1988). "Causality in the Social Sciences," pp. 347–409, in Clifford C. Clogg (ed.), *Sociological Methodology, 1988, Vol. 18*. Washington, DC: American Sociological Association.

Nichols, Elizabeth (1986). "Skocpol and Revolution: Comparative Analysis vs. Historical Conjuncture." *Comparative Social Research* 9:163–86.

Orloff, Ann S., and Theda Skocpol (1984). "Why Not Equal Protection? Explaining the Politics of Public Social Spending in Britain, 1890–1911, and the United States, 1880s-1920." *American Sociological Review* 49:726–50.

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

Robinson, W. S. (1951). "The Logical Structure of Analytic Induction." *American Sociological Review* 16:812–18.

Skocpol, Theda (1984). "Emerging Agendas and Recurrent Strategies in Historical Sociology," pp. 356–91 in Theda Skocpol (ed.), *Vision and Method in Historical Sociology*. Cambridge University Press.

(1986). "Analyzing Causal Configurations in History: A Rejoinder to Nichols." *Comparative Social Research* 9:187–94.

Turner, Ralph H. (1953). "The Quest for Universals in Sociological Research." *American Sociological Review* 18:604–11.

Whyte, William F. (1943). *Street Corner Society: The Social Structure of an Italian Slum*. Reprinted Chicago: University of Chicago Press, 1955.

Znaniecki, Florian (1934). *The Method of Sociology*. New York: Holt, Rinehart & Winston.

Part II
Analyses of research experiences