CURRENT ISSUES IN COMPARATIVE MACROSOCIOLOGY: A DEBATE ON METHODOLOGICAL ISSUES

John H. Goldthorpe
Comparative Social Research, Volume 16, 1997, pages 1-26.

ABSTRACT

Within comparative macrosociology, quantitative or "variable oriented" and qualitative or "case-oriented" methodologies are typically counterposed. It is, however, argued that in this way the nature of key methodological problems is often obscured. Three such problems - labeled the small N, the Galton and the black-box problems - are shown to arise with both approaches, and a critique is advanced of recent claims by exponents of case-oriented work that that they dispose of special and privileged means of by-passing or overcoming these problems.

I seek in this chapter to intervene in what is in fact a rather long-standing debate within comparative macrosociology, but one which appears of late to have acquired new vigor. The contending parties in this debate are now usually characterized as exponents of quantitative, "variable-oriented" methodologies, on the one hand, and of qualitative, "case-oriented" methodologies, on the other (see e.g., Ragin 1987; Rueschemeyer 1991; Janoski and Hicks 1994). I shall, however, argue that while the issues caught up in the protracted and complex exchanges that have occurred do include ones of major importance, the form that the debate has taken has not been especially helpful in highlighting just what these issues are, nor yet in pointing to ways in which they might be more effectively addressed.

I shall develop my position as follows. To begin with, I give a brief account of the contrast, or opposition, that has been set up between variable-oriented and case-oriented approaches. I then pursue my central argument by considering three rather well-known methodological problems that are encountered in the practice of comparative macrosociology. These problems are ones that have in fact been chiefly discussed in connection with variable-oriented research. But, I aim to show, they are present to no less a degree in case-oriented studies and, contrary to what several prominent authors have maintained or implied, the latter can claim no special advantages in dealing with them. Largely on account of misconceptions in this regard, I conclude, much recent discussion has tended to obscure, and divert attention away from, questions of method that comparative macrosociology does now need to engage with more actively - in whatever style it may be carried out.

VARIABLE-ORIENTED VERSUS CASE-ORIENTED APPROACHES

The variable-oriented approach to comparative macrosociology stems from a now famous proposal made by Przeworski and Teune (1970, chap. 1; cf. Zeiditch 1971, pp. 269-273): that is, that the ultimate aim of work in this Field should be to replace the proper names of nations (or of states, cultures etc) with the names of variables. Przeworski and Teune first illustrate the logic they would recommend by examples such as the following. Rates of heart attack are lower in Japan than in the United States. But, in seeking an explanation for this, we do not get far by treating the differing rates as simply "Japanese" or "American" phenomena. Rather, we have to drop proper names - or adjectives - and introduce generally applicable variables: that is, variables on which each nation can be given a comparable value. Thus, in the case in point, one such variable might be "per capita consumption of polysaturated fat."

Przeworski and Teune then of course go on to provide further illustrations of their position drawn from the social sciences; and, by the present day, one could in fact add to these entire research programs in sociology - and political science - that essentially follow the approach that they advocate. As a paradigm case here, one might take research that is aimed at explaining cross-national differences in the size and institutional form of welfare states (for Current Issues in Comparative reviews, see Quadagno 1987; O'Connor and Brym 1988). In such research, the names of nations are typically "replaced" by such variables as "GNP per capita," "proportion of population over age 65," "degree of trade-union centralization," "share of left-wing parties in government" etc. That is to say these are the independent variables, by reference to which the dependent variables - cross-nationally differing aspects of welfare provision - are to be "accounted for." The relationships that actually prevail between independent and dependent variables are then investigated statistically, through various techniques of multivariate analysis.

It is, for present purposes, important to recognize what Przeworski and Teune were defining their position against. Most importantly, they sought to challenge the "historicist" claim that any attempt to make macrosociological comparisons must fail in principle because different national societies are *sui generis*: that is, are entities uniquely formed by their history and culture, which can be studied only, so to speak, in their own right and on their own terms. In opposition to this, Przeworski and Teune point out that being "comparable" or "non-comparable" are not inherent properties of things: whether meaningful comparison is possible or not is entirely a matter of the analytic concepts that we have at our disposal. Thus, apples and oranges may appear to be non-comparable - but only until we have the concept of "fruit" (cf. Sartori 1994).

At the same time, though, Przeworski and Teune do insist that *if* the historicist position is accepted, then it must indeed follow that a comparative macrosociology is ruled out. If nations can only be studied as entities in themselves that will not allow of any kind of analytic decomposition- if, in other words, nations can only be studied "holistically" then comparisons cannot be undertaken. Considered as wholes, nations are unique, and "holistic comparison" is thus an impossibility. As Zeiditch (1971, p. 278) later put the point: "There is nothing else on earth quite like the United States (or the Navaho, or the Eskimo...) taken as a whole. Therefore the rule of holism [in comparative work] yields a clear and straightforward contradiction: only incomparables are comparable."

However, if the variable-oriented approach thus developed out of a critique of holism, the case-oriented approach is usually taken to represent a revival of holism - and indeed one directed against the kind of analytic reductionism that Przeworski and Teune would favor. Thus, for example, Ragin (1991, pp. 1-2) would regard it as being the very raison d'etre of case studies that they allow a return to holism in comparative research: that is, they allow nations, or other macrosocial units, to be considered as "meaningful wholes" rather than serving simply as the basis on which "to place boundaries around the measurement of variables."

It must, though, be noted that the holism that Ragin and others thus set against multivariate analysis is not as radical as might at first appear. Case studies are indeed regarded as the only way in which macrosocial entities can be treated in their distinctive historical contexts, in their proper detail and as each constituting, as Skocpol and Somers (1980, p. 178) put it "a complex and unique socio-historical configuration." But this, it turns out, does not imply a historicism of a quite thoroughgoing kind, which would deny the validity of any concepts that are formed in order to transcend particular cases (cf. Skocpol 1994, pp. 328-329). It is still seen as permissible to "abstract" from different cases certain of their "features" or "attributes" which can then be compared for theoretical purposes. In other words, variables are identified even if sometimes behind a verbal smokescreen. Where holism enters in is with the insistence that, in any comparison, the unity of the particular cases involved should always be preserved. What is required is that, in the process of comparison, cases should always remain identifiable as such, rather than being decomposed into variables that are then interpreted only in the course of the simultaneous analysis of the entire sample of cases under investigation.

In actually pursuing holistic comparisons in this sense, exponents of the case-oriented approach appear to have found their chief methodological inspiration in the logic of John Stuart Mill (1843/1973): specifically, in Mill's "canons," or rules, of experimental induction the "Method of Agreement," the "Method of Difference," and so forth (see e.g., Skocpol 1979, pp. 36-37, 1984, pp. 378- 381; Skocpol and Somers 1980, pp. 183-184; cf. Ragin 1987, pp. 36-42). Following Mill, it is believed, each case included in a comparative enquiry can be taken as representing the presence or absence of a given phenomenon of interest - each case, that is, can be taken as a "naturally occurring" experiment relating to this phenomenon. Inferences regarding the causation of the phenomenon can then be drawn by considering which other features are concomitantly present or absent, and by in turn applying Mill's logical rules to the resulting set of comparisons. Thus, Skocpol, in her well-known study of social revolutions (1979), seeks to explain their outcomes by comparing national cases, on the one hand, in terms of whether or not revolutionary attempts succeeded and, on the other hand, in terms of the presence or absence of what she takes as likely determining factors: that is, various features of the agrarian economy and class structure, international pressures, internal political crises and so forth

It might seem that in both multivariate and logical comparisons alike the aim is in effect to "control variation" in the making of causal inferences - so that the two approaches are not, after all, so very far apart. And, indeed, the application of Mill's methods in the comparison of cases has not infrequently been represented as itself a form of multivariate analysis (e.g. Smelser 1976, chap. 7; Skocpol and Somers 1980, pp. 182-183; Dogan 1994, p. 35). However, as other commentators have pointed out (e.g. Lieberson 1992, 1994), there is one quite fundamental difference. The various forms of multivariate analysis used in quantitative work are statistical techniques, and the propositions to which they give rise are therefore probabilistic: they are based on associations or correlations that need not be perfect. In contrast, the methods proposed by Mill, being logical in character, entail propositions of a *deterministic* kind: they entail relationships that are entirely invariant. As will later be seen, this is a difference that matters, and indeed to overlook it is to neglect a major development in the history of sociological analysis: that which, in the course of the nineteenth and earlier twentieth centuries, saw sociology become part of "the probabilistic revolution" (cf. Krüger, Daston, and Heidelberger 1987; Krüger, Gigerenzer and Morgan 1987)

The distinction between variable-oriented and case-oriented approaches is not then a meaningless one. It captures an important divergence in preferred styles of comparative macrosociological research and further, one may suspect, in basic assumptions about the character of social phenomena. But, I would argue, focusing on this distinction will not in itself provide the key to an understanding of the more taxing methodological problems that arise in the conduct of such research; nor are attempts at combining or synthesizing the two approaches likely to make the main contribution to overcoming these problems, since they are in fact ones that confront both approaches alike. This I argument I now seek to sustain with reference to what may be labeled as (i) the small N problem; (ii) the Galton problem; and (iii) the black-box problem.

THE SMALL N PROBLEM

The small N problem arises in that, if nations or other macrosocial entities are taken as units of analysis, the number available for study is likely to be quite limited. Where individuals are the units, populations can be sampled so as to give Ns of several hundreds or thousands; but where nations are the units, N cannot rise much above one hundred even if all available cases are taken, and is often far less. In applying techniques of multivariate analysis, serious difficulties tend therefore to be encountered in that N is not much greater than the total number of variables involved. Statistically, this means that there are too few degrees of freedom, that models become "over-determined," that inter-correlations among independent variables cannot be adequately dealt with and that results may not be robust. Substantively, it means that competing explanations of the dependent variable may not be open to any decisive evaluation. Thus, it has been recently claimed (Huber, Ragin, and Stephens, 1993) that, for just these reasons, the research program on the determinants of state welfare provision - in which analyses based on a maximum of *c.* 20 nations have been typical - has by now reached a virtual "impasse." Theories privileging different sets of determinants can claim similar degrees of statistical support".

The small N problem is then a real and troubling one. However, what I would wish to question are suggestions to the effect that it is a problem specific to the variable-oriented approach to comparative macrosociology, and that the case-oriented approach in some way or other allows it to be solved or circumvented. Most explicitly, Skocpol (1979, p. 36; cf. Rueschemeyer 1991, pp. 27-28, 32-34) has maintained that application of the methods "laid out" by Mill "is distinctively appropriate for developing explanations of macro-historical phenomena" when the small N problem arises; that is, "when there are too many variables and not enough cases".

This claim calls for comment in several respects. To begin with, it is unclear whether Skocpol realizes that Mill himself (1843/1974: Bk. VI, chap. 7 esp.(went to some lengths to explain that his rules of induction, being developed for use in the experimerutal sciences, were not appropriate to the study of social phenomena and that, if used, would be likely to prove inconclusive if not actually misleading. At all events, Skocpol fails to take sufficient account of certain assumptions on which Mill's methods depend but which, as various critics have followed Mill in observing (e.g. Nichols 1986; Lieberson 1992, 1994) are assumptions rarely, if at all, defensible in social research. For example, Mill's logic presupposes that, in any analysis, *all* of the relevant causal factors can be identified and includedthat is, that there are no "unmeasured variables"; and further that there is no multiple (or "plural") causation, nor again any interactions among causal factors.'

At the same time, though, Skocpol is well enough aware that Mill's canons are designed to lead to causal propositions of a deterministic kind - and does not appear much disturbed by this fact (cf. also 1984, p. 378). What, therefore, her argument would appear to come down to is this: that, in circumstances where there are too few cases for the satisfactory evaluation of probabilistic theories, deterministic ones may none the less be established. However, to accept this position, it should be noted, one must be ready to believe not just that the social world is indeed subject to deterministic theory rather than being inherently probabilistic. One must *further* believe that socio-historical *data* can be obtained that are of such a quality and completeness - that are so error-free that a probabilistic approach is not even required for the purposes of relating these data to (deterministic) theory (cf. Lieberson 1992, pp. 106-107; King, Keohane, and Verba 1994, pp. 59-60). This latter implication at least is one that, I suspect, would be found by most sociologists, on due reflection, to be far more daunting than the small N problem itself.

Various attempts have been made to develop the logical analysis of relatively small numbers of cases so as to overcome some of the more obvious limitations of Mill's methods in the context of social research. Most notable in this connection is perhaps the technique of "qualitative comparative analysis" (QCA) proposed by Ragin (1987) which is based on Boolean algebra. This technique aims to alleviate the small N problem by allowing inferences to be drawn from the maximum number of comparisons that can be made, in terms of the presence or absence of attributes of interest, across the cases under analysis. And, at the same time, it does permitindeed is primarily directed towards - the analysis of multiple causation and interaction effects. Thus, Ragin (1994b, p. 328) maintains that while a regression exercise with, say, seven independent variables and only 18 cases would be generally regarded as untrustworthy, QCA would make possible the examination of all 128 (i.e., 2*2*2*2*2*2*2) different combinations of the causal conditions involved: that is, would in fact enable the analyst to address a degree of causal complexity far beyond the reach of regression.

Given the nature of QCA, Ragin would then further argue, it allows the macrosociologist to combine analysis with holism in that the distinctive features of particular cases need never be lost sight of. However, while this may be so, it is still somewhat misleading for Ragin to represent QCA as being a *synthesis* of the case- and variable-oriented approaches, since, as he indeed recognizes (1994b, pp. 305-306), QCA remains, no less than Mill's methods, entirely logical and non-statistical in character. And it does therefore still share with the latter the major disadvantages of being unable to make any allowance either for "missing variables" or for error in the data used.

Moreover, with QCA these disadvantages combine with two other evident weaknesses of the technique: its requirement that all variables should be treated as merely two-valued; and its high degree of sensitivity to the way in which each case is coded on each variable. Thus, where essentially continuous variables are involved, such as "GNP per capita," "proportion of population over 65" and so forth, these must be reduced (with, of course, much loss of information) to more or less arbitrary dichotomies; and all subsequent results will then be strongly dependent on the way in which particular cases are allocated. If, on account of error in the original data, or in its treatment, even a single case happens to be placed on the "wrong" side of a dichotomy, the analysis could well have a quite different outcome to that which would have been reached in the absence of such error. In an application of QCA, it should be noted, the independent variables are simply shown to be causally relevant or not; no assessment of the *relative strengths* of different effects or combinations of effects is, or can be, made.

In sum, the fact that QCA remains a logical technique means that its results are far more exposed to major distortion, both by difficulties in the selection of independent variables (cf. Amenta and Poulson 1994) and by the occurrence of error in data, than are results derived from statistical techniques. And whether, then, QCA does actually mark any significant advance in the treatment of the small N problem - as, for example, Skocpol has recently claimed (1994, p. 309) - must remain open to very serious doubt.

What, I would argue, it is above all else necessary to recognize here is that *au fond* the small N problem is not one of method at all, but rather of data: more specifically, it is a problem of insufficient information relative to the complexity of the macrosociological questions that we seek to address. Thus, in so far as exponents of the case-oriented approach in effect choose to restrict ' themselves to small Ns, they are unlikely ever to avoid the difficulties of "too many variables and not enough cases" or, as King, Keohane and Verba (1994, p. 119) put it, "more inferences than implications observed" no matter what resorts to Millian logic. Boolean algebra or other technical devices they may attempt. Conversely, what is vital to overcoming the small N problem is in principle easy to state, albeit in practice toilsome, even where possible, to achieve: that is, simply to increase the information that we have available for analysis.

One way in which this can sometimes be achieved is by exploiting more fully the experience of those nations (or other macrosocial units) for which we do have good data sources. Thus, in comparative welfare state research various investigators (e.g. O'Connor and Brym 1988; Korpi 1989; Pampel and Williamson 1989; Huber, Ragin, and Stephens 1993; O'Connell 1994) have by now taken up the lead given by econometricians and demographers and have "pooled" data for the same set of nations for several different time-points. Observations - and degrees of freedom - are in this way increased, and appropriate checks and corrections can be introduced into analyses in order to allow for the fact that the successive "waves" of information thus acquired are not, of course, entirely new and independent (see e.g., Stimson 1985; Hicks 1994a). Such a "pooling" strategy can then be reckoned as a valuable resource for macrosociologists following a variable-oriented approach; and King, Keohane and Verba (1994, pp. 221-223) have recently suggested various analogous procedures that might profitably be followed in qualitative studies. More important, though, for the variable- and case-oriented approaches alike, is to increase the number of units to which comparisons extend; and further (cf. Przeworski 1987) to widen their geographical and sociocultural range, so that the greater variation thus obtained ,in supposed causal factors can improve the chances of deciding between competing theories. This will often mean bringing Third World nations into the analysis, and problems of data quality, which must always be of central concern in comparative work, may on this account be accentuated (cf. Dogan 1994, pp. 40-41). However, the challenge thus posed should not be shirked. Two recent authors, Bradshaw and Wallace (1991, p. 166) have argued for the particular appropriateness of case studies in the Third World, since, they maintain, calls for rigorous quantitative research must be biased against poor nations that lack adequate data or even computers. While this view is clearly well-intentioned, I would still regard it as quite wrong-headed. Either the assumption is being made that case studies are, in some mysterious way, immune to problems of the reliability and validity of data with which quantitative researchers have to struggle or else case studies are being recommended for Third World use as some kind of "inferior good." It would surely be, from all points of view, a better strategy for First World social scientists to seek to help their Third World colleagues to collect whatever kinds of data, and to undertake whatever kinds of analysis, are in fact demanded by the nature of the substantive problems that they wish to pursue.

THE "GALTON" PROBLEM

The "Galton" problem is named after the nineteenth-century British polymath, Francis Galton. In 1889 Galton famously criticized a pioneering comparative analysis by the anthropologist, Edward Tylor. Tylor (1889) claimed to show complex correlations among economic and familial institutions across a wide range of societies, past and present. These correlations he then sought to explain from what we would now think of as a functionalist standpoint. Galton (1889) however, questioned the extent to which Tylor's observations were independent ones, and pointed out that "institutional" correlations might arise not only under the pressure of functional exigencies, or through other processes operating within societies; they might also be the result of processes of what we would now call cultural diffusion among societies.

The problem of distinguishing between processes of these two kinds has subsequently plagued cross-cultural anthropology (Naroll 1973; Hammel, 1980), and it obviously arises in comparative macrosociology to no less a degree. Thus, to revert to the investigation of welfare state development, it would be rather implausible to suppose that this development has proceeded quite autonomously in each national case, and free of such external influence as might have been exerted by the examples of, say, Bismarckian social policy in the nineteenth century, or the Beveridge Plan for post-war Britain, or, more recently, the "Scandinavian Model" (cf. Therborn 1993). Moreover, the Galton problem could be regarded as potentially more damaging at the present time than ever before. Claims that the treatment of nations as independent units of analysis has been untenable ever since the emergence of a "world system" in the seventeenth century (Hopkins 1978; Hopkins and Wallerstein 1981) or that there now exists "a highly institutionalised world polity" (Meyer 1987, p. 42) might well be thought exaggerated. But it could hardly be denied that, by the late twentieth century the independence of "national" observations is likely to be compromised, and not merely by the acceleration and intensification of cultural diffusion but further through the quite purposive actions of a whole range of international or multinational political and economic organizations. In this way, as Przeworski (1987) has recognized, the threat is created that the small N and Galton problems run together, as we do indeed enter into a world in which N= 1.

Lack of independence in observations, as well as limits on their number, does then undoubtedly create serious difficulties for cross-national research. However, just as with the small N problem, what I would wish first of all to stress is that while the difficulties in question may be most apparent with the variable-oriented approach, they are by no means restricted to it; the case-oriented approach enjoys no special immunity.

Thus, the assumption that nations can be treated as units of analysis, unrelated to each other in time and space, is one required by the logical methods of comparison that are favored in case-oriented research no less than by statistical methods. And indeed where historical cases are involved, the Galton problem is then likely to be encountered in a particularly troublesome form. The scarcely disputable fact that situations and events occurring at one time tend to have been influenced by situations and events occurring earlier clearly breaches the assumption of the independence of cases - as built into Mill's or any other logical method - and in a way that is not easily remedied. Thus, for example, one finds that

Skocpol, in her study of revolutions has obviously to recognize that the course of the Chinese revolution up to 1949 was in various ways influenced by events in Russia in 1917 and subsequently. But this recognition has then to be kept quite apart from her logical analyses of the factors that determine revolutionary success, which it threatens to compromise (cf. Burawoy 1989). In other words, "analytic induction" and narrative accounts that crucially rely on temporality cannot be integrated, but have to be left to play separate, and ultimately incompatible, explanatory roles (see further Kiser and Hechter 1991, pp. 12-13; Griffin 1992, pp. 412-413; cf. also Skocpol 1994, p. 338).

Despite this, the Galton problem has in fact met with only a rather limited appreciation - and response - among exponents of case-oriented research. McMichael (1990) has proposed a solution through what he calls "incorporated comparisons," which is apparently intended to take over the insights, while avoiding the "rigidity," of a world-system perspective. But since he presents his approach as an "interpretative" one that can proceed "without recourse to formal methodological procedures or a formal theory" (1990, p. 388), it is not easy to evaluate (nor, I would have to say, to understand). Another reaction is that of Sztompka (1988) which, however, is less an attempt to grapple with the Galton problem than a capitulation to it - and one that might be seen as somewhat opportunistic. The severity of the problem in the modern world, Sztompka maintains, is such that the whole agenda of comparative macrosociology needs to be changed towards in fact a concentration on case-studies! "Globalization" has, in Sztompka's view, already made societal homogeneity and uniformity the norm. Thus, the central aim should no longer be to establish cross-national similarities or regularities of variation, using
"hard," quantitative techniques; rather, comparative work should now focus on the description and interpretation of "enclaves of uniqueness"that is, those the establishment of the International Labor Office in 1919, and by its subsequent world-wide activities.

The further large potential of attempts at thus modelling interdependence may be brought out by reference to the recent work of Castles and others (Castles 1993), who have introduced into comparative policy research the idea of "families of nations." Instead of attention centring on nations as "unattached singles," they argue, more account should be taken of the affinities that exist among groupings of nations, as a result of shared histories and cultural traditions. Castles has in fact suggested (1993, pp. xv-xvi) that recognition of such affinities may indicate the "outer limits" of the Przeworski-Teune program of replacing the proper names of nations with the names of variables. For policy similarities and differences among nations "may be attributable as much to history and culture and their transmission and diffusion amongst nations as to the immediate impact of the economic, political and social variables that figure almost exclusively in the contemporary public policy literature." And, Castles believes, the former kinds of effect are difficult to accommodate within the "prevailing intellectual paradigm," as represented by the variable-oriented approach.

Now, as regards his substantive point on the importance of historically- formed cultural patterns that transcend national boundaries, Castles may well be right. And, as will be apparent later, I share his concern with determining just where the theoretical limits of macrosociology, in whatever style it may be conducted, must in the end be drawn. But I do not see why the variables that replace the names of nations in quantitative analyses of comparative public policy need be only variables thought likely to have an "immediate impact"; nor why one cannot, in principle at least, also include variables that do indeed seek to capture nations' historical affinities and the longer-term influences that derive from them. Indeed, I would argue that to attempt to do precisely this is the obvious way to explore further the idea of families of nations. In other words, there seems no reason why the insights provided by Castles and his associates should not serve as the starting point for appropriate quantitative analyses that would enable us to form more reliable judgements on what is, after all, crucially at issue: that is, the relative importance, in regard to policy developments and repertoires, of inter- as opposed to intra-societal, and of "historical" as opposed to "contemporary" effects.

In sum, we should not be led into believing that claims regarding globalization or the existence of a world system or of families of nations necessitate some quite radical transformation of cross-national comparative macrosociology, and least of all one that would entail its restriction to case studies. In dealing with the Galton problem where there are good grounds for supposing that it does indeed exist - the variable-oriented approach at all events has resources that are in fact only beginning to be exploited.

THE BLACK BOX PROBLEM

The black box problem, even more than the small N or Galton problems, has been linked with the variable-oriented approach (see e.g., Rueschemeyer 1991, p. 26; Abbott 1992a, pp. 54-62; Esping-Andersen 1993, p. 8). A quantitative analysis may be undertaken which is successful in "accounting for" a significant part of the variation in the phenomenon of interest - let us say, the sizes of welfare states. But such an analysis, it can be objected, still tells us rather little about just what is going on at the level of the social processes and action that underlie, as it were, the interplay of the variables that have been distinguished.

We know the "inputs" to the analysis and we know the "outputs" from it; but we do not know much about why it should be that, within the black box of the statistical model that is applied, the one is transformed into the other. The problem is of course mitigated if "intervening" as well as independent variables are included in the analysis, so as to give it a more finely-grained character; and further if both independent and intervening variables are chosen on theoretical grounds, so that certain causal processes may at least be implied.

None the less, it can still be maintained that the black box problem is seriously addressed only to the extent that such processes are spelt out quite explicitly, so as to provide a "causally adequate" account of the actual generation of the regularities that are empirically demonstrated (cf. Elster 1989, chap. 1). The black box problem, thus understood, has been seized upon by exponents of case studies in order to make the claim that the results of quantitative analyses must in effect be dependent upon case studies for their interpretation. Thus, Huber, Ragin, and Stephens (1993) have argued that the problem of conflicting explanations of the growth of welfare states can only be solved through a "dialogue" between variable- and case-oriented research, and that it is case studies that must play the crucial part in identifying "actual historical causal forces." Likewise, Rueschemeyer (1991, p. 28; cf. also Rueschemeyer, Stephens, and Stephens 1992, ch. 2) has maintained, with reference to comparative research into capitalist development and democracy, that in this area the tradition of historical case studies is "far richer in theoretical argument and analysis" than is that of quantitative work. Rueschemeyer accepts that quantitative studies have established a clear positive association between capitalist development and democracy; but, he is convinced, the "key to the black box" that mediates this association will only be found in theory inspired by case studies and, especially, in "explanatory ideas grappling with historical sequences."

Again, however, I would wish to call into question the privileged status that is thus accorded to the case-oriented approach. To begin with, it should be recognized that while, just as with the small N and Galton problems, the black box problem may be most apparent in quantitative work, it does in fact arise I equally with the case-oriented approach where logical methods of comparison are applied. Contrary to what Rueschemeyer suggests (1991, pp. 32-33), logical methods too can only establish empirical regularities which may, at most, point to causal relations: they do not, in themselves, provide an account of the actual processes involved (cf. Burawoy 1989). And if, to this end, "analytical induction" is accompanied by some narrative of historical sequences, then this, for reasons earlier noted, cannot be part of the logical method itself but only in fact a rather awkward appendage to it.

I would, furthermore, argue that the theoretical achievement of case studies is, in any event, a good deal less impressive than the authors cited above attempt to make out. Where the unity of cases is preserved - where cases are studied "holistically," rather than being decomposed into variables - it is indeed possible, at least in principle, to provide detailed descriptions of "what happened" in each case, and with due regard for the specific contextual features involved. But to have a narrative account of a sequence of historical events is not the same thing as having a theoretical account, and even if one accepts as I would be ready to do - that a historical narrative can itself constitute a form of explanation. Most crucially, perhaps, such a narrative need not extend beyond the particular instance to which it is applied, or comparative narratives beyond the set of cases compared (cf. Skocpol and Somers 1980, p. 195). In contrast, a theoretical account must have some claim to generality. The explanation it provides of what is going on within the black box of a statistical or of a logical analysis is not one that is simply "extracted" from the actual events involved in the instances covered by the analysis but one that is, rather, derived from a theory that could, indeed should, apply to other instances falling within its intended scope or domain.

It might of course be suggested - and I would Find it unexceptionable - that specific narratives may serve as a valuable resource for theory development: that is, by prompting attempts to conceive of some more general ideas that would allow the accounts given in different cases to be fitted into a deductive structure of argument. In other words, detailed case studies could play a heuristic role in the "context of discovery," prior to the testing of any resulting theory against further, independent cases in the "context of validation." However, the distinction here involved is one that proponents of the case-oriented approach appear to find uncongenial, and that Rueschemeyer, for example (1991, pp. 32-33; cf. Rueschemeyer, Stephens, and Stephens 1992, p. 36; Skocpol 1994, p. 330), flatly rejects. The view that seems rather to be favored is that the process of theory development should be advanced by successive inductions from particular cases - so that it becomes in effect essentially merged with the process of theory testing. The matching of developing theory against new inductions and its modification where it is found not to hold go on as one, seamless activity.

It is, however, in just this regard that the case for case studies becomes least convincing. The crucial point is that if a theory is formed in such an essentially inductive way- without, so to speak, any deductive backbone - then it is hard to see how it can be genuinely tested at all. As it stands, such a theory does no more than recapitulate observations; and it is, moreover, difficult to know exactly how it would be properly extended beyond the particular circumstances from which it has been obtained so that an independent test might be attempted. Or, to put the matter the other way around, if a theory amounts to no more than an assemblage of inductions, the possibilities for "saving" or "patching" it in the face of contrary evidence are virtually unlimited. Generality can be claimed for so long as such a theory appears to Fit the cases to which it is applied; but when it fails to fit, it can then be maintained that "causal homogeneity" no longer holds, and that a somewhat different theory is required; and, in all of this, analysts can congratulate themselves on their "sensitivity to context"! However, the arbitrary delimitation of the scope of a theory that is, a delimitation that the theory does not itself provide foris an evident weakness. Thus, in the context of welfare state research, Korpi (1989, p. 324) has critically remarked that theories of "state autonomy," as advanced by Skocpol and others (e.g. Orlof and Skocpol 1984; Weir and Skocpol 1985) on the basis of qualitative case studies "leave ample room for flexible ad hoc explanation," and has urged the need for such theories to be formulated in a way that would expose them to more stringent empirical critique. And yet more prominently, the charge of arbitrariness has been levelled against Skocpol's treatment of the Iranian revolution (1982), when taken in relation to her previous analyses (1979) of the French, Russian and Chinese revolutions (Nichols 1986; Burawoy 1989; Kiser and Hechter 1991), since in the Iranian case a significant, yet seemingly quite ad hoc, theoretical shift is introduced: that is, popular urban demonstrations become a "functional substitute" for peasant revolts and guerrilla activity (cf. Skocpol 1994, pp. 313-314.(

Finally in this connection, I would also question whether the account offered by Rueschemeyer, Stephens and Stephens of the association between capitalist development and democracy does in fact bear out their contention that case studies afford a privileged ground for the development of theory capable of overcoming the black box problem. Their account fails in this respect, I would suggest, precisely because of the degree to which the analysis of their cases leads them to hedge about their central argument on power struggles among social classes with exceptions and qualifications - relating to cross-national differences in the social construction of class interests, in the possibilities for class alliances, in the form of civil society, in the role of the state, in the impact of transnational relations, and so on (1992, pp. 269-281 esp.). Not only does the ratio of explaining to "explaining away" thus seem rather low but, further, it is notable that when these authors come to address the key issue of the "generalizability" of their theory beyond the cases they have examined (1992, p. 285) - to, say, east Asian or east-central European nations of the present day - what they have to offer is not a series of derived hypotheses that would be testable against such new cases but yet more discussion of additional factors to be considered." Now it may be that the awareness that Rueschemeyer and
his colleagues here display of complexity and "causal heterogeneity" is empirically warranted. But, if so, what they have provided is a demonstration of the inherent difficulty of forming a theory of the relationship between capitalist development and democracy, and not that theory itself.

For macrosociologists seeking to treat black box problems more effectively, I would then argue, case studies, whether "historical" or otherwise, have no distinctive value, and an absorption in their specificities may indeed divert attention away from what is in fact crucially required: that is, theory that is as general as it is possible to make it. As Kiser and Hechter (1991) have maintained, in a strong critique of the quality of theory in comparative historical sociology, to illuminate the black boxes represented by mere empirical regularities, we need more than just a redescription of the latter within a "theoretical (sc. conceptual) framework" which appears indefinitely modifiable as our data-base expands. Rather, theory must be sought that is general in that it permits the specification of causal processes which, if operative, would be capable of producing the regularities in question and would have a range of further implications of at least a potentially observable kind. To the extent that theory is general in this sense, it can then claim both greater explanatory power, which theory must always seek, and greater openness to empirical test, which it must never evade. I would, moreover, add that such a concern with generality in theory might help macrosociologists to see the relevance of history to their enterprise in a different - and, I believe, more appropriate - way to that which appears currently in mode among exponents of the case-oriented approach. Instead of a recourse to history being regarded as essential to the development of theory, it might be better understood as marking the limits o/theory: that is, the point at which what is causally important in regard to certain empirical findings is recognized not in recurrent social situations and processes that might be the subject of theory but rather in contingencies, distinctive conjunctures of events or other singularities that theory cannot comprehend.

Since the foregoing is put somewhat abstractly, I may try to illustrate with reference to the (primarily quantitative) work that I have undertaken with Robert Erikson on comparative social mobility (Erikson and Goldthorpe 1992a). Perhaps the most notable finding of this work was that when intergenerational class mobility was considered net of all structural influences - or, that is, as "social fluidity" - rates and patterns showed high stability over time within nations and, further, a large measure of similarity across nations. Such a degree of invariance clearly underlines the need for general theory. For hypotheses on the causal processes capable of producing temporal constancy and cross-national commonality of the kind that our quantitative analyses revealed will have to be derived from a theory of Current Issues in considerable scope: that is, from a theory which is precisely not "sensitive to context" - unlike the theories of national "exceptionalism" in regard to mobility which our results called into doubt - but applicable to societal contexts widely separated over both time and space. And in this respect, I should say, Erikson and I were able to make only a very modest beginning."

We also found, though, that in so far as variation in social fluidity did occur cross-nationally, we could not account for it, to any large extent, in terms of other generalizable attributes of societies, in the way that the Przeworski-Teune program would require. Our analyses pointed here to the far greater importance of historically formed cultural or institutional features or political circumstances which could not be expressed as variable values except in a quite artificial way. For example, levels of social fluidity were not highly responsive to the overall degree of educational inequality within nations, but patterns of fluidity did often reflect the distinctive, institutionally shaped character (cf. p. 11 above) of such inequality in particular nations, such as Germany or Japan. Or again, fluidity was affected less by the presence of a state socialist regime *per se* than by the significantly differing policies actually pursued by the Polish, Hungarian or Czechoslovak regimes on such matters as the collectivization of agriculture or the recruitment of the intelligentsia. In such instances, then, it seemed to us that the retention of proper names and adjectives in our explanatory accounts was as unavoidable as it was desirable, and that little was to be gained in seeking to bring such historically specific effects within the scope of theory of any kind

.

In sum, black box problems - essentially problems of "making sense" of empirical findings - are unlikely to be alleviated by comparative macrosociologists striving in effect to transcend the distinction between theory and history. For such attempts tend to lead merely to a weakening of our understanding of theory and of historicity alike, and in turn to a blurring of crucial differences in the nature of theoretical and historical explanations. A strategy of greater long-term promise would be to continue to pursue sociological theory that amounts to more than just the elaboration of concepts and aspires to generality in the sense indicated above, but at the same time to show due modesty in accepting that, for any kind of macrosociology, and no matter how theoretically accomplished it may eventually become, "history" will always remain as a necessary residual category." It may, furthermore, be a consequence of such a strategy that certain phenomena that macrosociologists have sought to study - including, perhaps, revolutions or other kinds of "regime transition" - turn out to be ones on which theory can give relatively little cognitive grasp at all. That is to say, while it may be of interest to write the comparative history of these phenomena - their history as viewed within a common conceptual framework - they appear just too few, too interdependent and too causally heterogeneous for anything of much use to be said in theoretical terms. In instances where the indications accumulate that this is indeed the case, then the course of wisdom must surely be to accept the situation with good grace. Macrosociologists will still be left with a very great deal to do, and there have not, after all, ever been any guarantees that a sociology of everything should be possible.

CONCLUSIONS

I have argued that, while a divergence can certainly be observed between variable-oriented and case-oriented approaches to comparative macrosociology, to concentrate attention on this divergence - or even on ways of overcoming it - does not provide the best focus for understanding and addressing major methodological issues that are encountered in this field. As King, Keohane and Verba have emphasized (1994, chap. 1), we may distinguish between quantitative and qualitative styles of research in the social sciences, but each must still strive to meet the exigencies of the same underlying "logic of inference" and contend with the problems to which this common requirement gives rise. Through an examination of three such problems, recurrent within comparative macrosociology, I have tried to show how each can, and does, occur in the context of variable-oriented and case-oriented work alike. These problems are not in fact ones on which alternatives in research styles have much bearing, but are of a more elementary, which is not to say easier, kind. Thus, the small N problem is essentially a problem of insufficient information on which to base analyses - or, that is, on which to draw in making inferences; and it can be resolved, or mitigated, only by more extensive data collection, aided by techniques for exploiting to the full the information that is at any time available. The Galton problem, where it arises, is one of observations lacking a property-independence - that we would like to assume in our analyses; and, to the extent that interdependence among our units of observation is simply a feature of the way the world is, we must deal with this situation by seeking to represent the interdependence (or, better, the processes creating it) within our analyses, so that we can not just recognize its presence but also assess its importance. And finally the black box problem is one of how we move from descriptive to causal inferences or, that is, go beyond our empirical findings and the regularities they allow us to establish to an understanding of how these regularities are generated. Here what is crucial is to construct theory in a way that maximizes both its explanatory power and its openness to test against further empirical research and that also allows us to see as clearly as possible where the limits to theoretical explanation are reached.

In this chapter, criticism has been more often directed against case-oriented than against variable-oriented research. This is not an expression of hostility on my part to qualitative research as such, whether in macrosociology or more generally, and especially not to such research of a historical character. Rather, it reflects my view that it is proponents of the variable-oriented approach who have, at all events, better appreciated and responded to the problems I have considered, while proponents of the case-oriented approach have sometimes failed to recognize that they too need to address these problems or, as I suggested at the start, have made claims to the effect that they dispose of special and privileged means of by-passing or overcoming them. My critical comments have then been chiefly directed against such claims. I have sought to show that they do not, at least as so far presented, have any very secure basis, and that, if they are to be maintained, they will need to demonstrated far more cogently than hitherto. I would doubt if this will prove possible, since I see no reason at all to believe in such special and privileged means. The small N, Galton
and black box problems pertain to quite basic issues that are likely to arise in any instance of comparative macrosociological research, whatever the style in which it is conducted. Whether investigators choose to work quantitatively or qualitatively, with variables or with cases, the inherent logic of these issues remains the same, and so too therefore will that of any solutions that may be achieved.