BAYESIAN THINKING ABOUT MACROSOCIOLOGY

Bruce Western

Estudio/ Working Paper 2000/152
June 2000

Bruce Western is Associate Professor of Sociology, and Faculty Associate of the Office of Population Research, at Princeton University. This paper is based on a seminar that he presented at the *Center for Advanced Study in the Social Sciences,* Juan March Institute, Madrid, on 15 March 2000, under the same title.

Abstract

One question dominates methodological debate in macrosociology: In explaining large-scale social processes should we seek simple theories that apply under a range of conditions, or complex theories that are tailored to specific cases? Complex explanations are justified by their explanatory power and theoretical appeal. Simple explanations are justified by their falsifiability and their capacity to sustain strong conclusions. I weigh the merits of simple and complex explanations applying some basic ideas from Bayesian statistics. The Bayesian analysis provides support for a compromise position that is rigorously grounded in the fundamental rules of probability. These ideas are illustrated in an analysis of labor union growth in Sweden.

INTRODUCTION[*]

Recent sociological debates about general theory and historical explanation revisit one of the most enduring themes of the discipline: In explaining large-scale social processes should we seek simple theories that apply under a range of conditions, or complex theories that are tailored to specific cases? This debate tracks trends in theory and empirical focus, pitting at various times, Parsons and Durkheim against Marx and Weber, or qualitative area studies against quantitative research. Lately, lively exchanges over rational choice and the historic turn in macrosociology dominate discussion. While the controversy covers substantial ground, the choice between simple and complex explanation remains a consistent theme.

Despite a hundred years of discussion, I think the costs and benefits of simple and complex explanation in macrosociology are poorly understood. In most cases, the methodological debate is suffused with theoretical disagreement. Most recently, Kiser and Hechter (1991) sparked disputes about the role of theory in historical explanation that extended to a symposium about the status of rational choice theory and proposals for a relational sociology (Somers 1998; Kiser and Hechter 1998). Because the protagonists tie theory and method so closely, arguments for one side or another often boil down to assertions of theoretical preference. This paper tries to move the pendulum away from theory and towards methodology. Here, I examine the relative merits of simple and complex explanation using a rudimentary logic of scientific explanation.

This paper studies whether there is a principled reason to prefer simple or complex explanations, regardless of theoretical taste. This narrowly methodological objective is pursued with ideas from Bayesian statistics. Bayesian statistics clarify the logic of social explanation by providing formal rules for learning from data. These rules are coherent in the technical sense of conforming

to some basic precepts of rational choice under uncertainty. Bayesian thinking points to a strong compromise between an *a priori* preference for simple or complex explanation that has a rigorous basis in the probability calculus. This position is illustrated in an analysis of the growth of labor unions in Sweden.

COMPLEX AND SIMPLE EXPLANATIONS

The relationship between history and sociology was recently re-examined in several prominent debates involving John Goldthorpe (1991, 1997) and Edgar Kiser and Michael Hechter (1991, 1998). These were just the latest contributions to a century-long discussion of the distinctive features of a sociological account of large-scale social processes. This effort has enlisted philosophies of science, social theories, and specific techniques in arguing for parsimony or historical detail. Although the continuum between simple and complex explanation is richly populated by intermediate positions, some clear arguments describe the two approaches and their basic strengths and weaknesses.

Complexity and Historical Sensitivity

Contemporary comparative and historical sociology is distinguished by a renewed emphasis on historically-sensitive explanation (Calhoun 1996; Paige 1999). For the new macrosociology, social processes are contextual. A causal condition may have one effect in one setting, but a different effect in another. In a characteristic formulation, Tilly (1984, 79) celebrates "genuinely historical" research which shows that the "time and place in which a structure or process appears makes a difference to its character." Recently, Paige (1999) identifies this perspective with a theoretical outlook that justifies limited generalizations to well-defined

historical conditions. More idiographically, Quadagno and Knapp (1992, 502) argue that causal processes must be specified in terms of dates and place names. In this perspective, cases consist of complex combinations of characteristics and have highly differentiated identities. Although historical sensitivity can vary in degree, the many proponents share a belief in complex explanations that enlist additional conditions–specific features of the local context–to sustain causal arguments.

Complexity in macrosociological explanation is illustrated by an institutional account of Swedish union growth. Although the following account is stylized, it helps illustrate historically-sensitive explanations. Nearly all workers in Sweden are union members. While the growth of unions is often explained by business cycle fluctuations, the size of the Swedish labor movement has been traced to an unusual system of unemployment insurance, controlled by unions since the mid-1980s (Rothstein 1990, 1992). Union-controlled unemployment insurance–called a Ghent system–allowed labor officials to protect union wages from competition from the unemployed (Rothstein 1990). The Swedish Ghent system also allowed marginal workers to retain contact with unions during spells of joblessness (Western 1997).

After abortive attempts at national unemployment insurance through the 1920s, the governing Social Democrats fashioned an agreement with the Swedish Liberal Party to support a Ghent system in 1934. To secure the Liberal's cooperation, the unemployment funds were poorly funded, subject to strict government control and open to nonunion workers. Because of these compromises, the scheme was weak by European standards. Still, Social Democrats gambled that the system would ultimately increase union membership. In 1941, the levels of insurance and government contribution expanded and many unions began to join the scheme. Throughout the post-war period, unions ensured

that all members join the insurance fund while nonunion workers who joined the fund were encouraged to take up union membership. As female and white collar workers swelled the unemployment insurance rolls in the 1960s and 1970s, unionization among these workers also grew rapidly. In the 1990s, unionization rates held steady despite high unemployment. A modest prewar innovation in the administration of unemployment thus had large effects on the postwar growth of the Swedish labor movement.

In contrast to the dominant business cycle account, Sweden's experience suggests that unionization does not depend in a general way on economic conditions like the unemployment rate. Instead, union growth varies with the historical development of labor market institutions. Because unions gained some control over the labor supply and played a key role in the welfare of the unemployed under the Ghent system, they were immunized from the disorganizing effects of unemployment. This example illustrates the complexity of contextual explanations. Business cycle theories view union growth as depending mainly on the unemployment rate. In Sweden, however, we can only understand the effects of economic conditions by taking additional account of the system of unemployment insurance.

Several researchers have developed methods specifically for contextual, historically-sensitive, explanation. For instance, Charles Ragin's (1987) qualitative comparative analysis (QCA) uses Boolean algebra to formulate highly conjunctural causal accounts. QCA identifies clusters of causal conditions that are related to some outcome of interest. QCA's sensitivity to combinations of causes admits very complex explanations. In a comparative analysis of welfare states, QCA indicated that the adoption of welfare programs in 15 countries depended on one of three configurations of causal conditions. A typical configuration showed

the causal force of paternalistic state institutions combined with working class mobilization, but without Catholic government or a unitary democratic system (Hicks, Misra, and Ng 1995). Complex explanations of this kind can exhaustively account for all variation in the outcome of interest.

While QCA admits complex cross-sectional explanation, narrative methods can account for complex dynamic processes. Unlike the formal methods of QCA, the narrative approach consists of a diffuse collection of social theory, presentational styles, and specific techniques (e.g., Sewell 1996; Stone 1979; Somers 1998; although compare the formal method of Griffin 1993). Narratives are stories consisting of a cast of characters with personal traits, relationships, and motives that impel social action through time and space (Tilly 1997, 21). Narratives yield complex explanations by emphasizing the sequential and contingent character of events (Griffin 1992; Abbott 1990; Rueschemeyer and Stephens 1997). Narratives are sequential in the sense that the impact of events depends on the order in which they occur (Abbott 1990). The ordering of events provides the context for their causal power. The idea of contingency claims that "nothing in social life is ultimately immune to change" (Sewell 1996, 264). As a consequence historical events are often unexpected, capable of undoing or altering the most durable trends of history. The role of contingency in narrative repudiates the idea of directionality in the historical process (Sewell 1996, 263-64).

Complex explanations are commonly recommended for their explanatory power and theoretical appeal. The empirical power of complex explanation is a direct product of sensitivity to historical detail. Attention to detail is sometimes inspired by the work of historians or anthropologists. Thus comparative and historical researchers identify "actual historical causal forces" (Huber, Ragin, and

Stephens 1993) and are committed to "thickness" in social explanation (Somers 1998, 739; Ortner 1996, 282). QCA demonstrates the premium on explanatory completeness by attaching causal inferences to all unique combinations of causes. There is no residual, and unexplained cases are resolved by finding more complex constellations of causal conditions. Comparative case researchers take this approach too, elaborating explanations until all anomalous observations are resolved (Ragin 1987, 42-44). The empirical power of narrative is not measured by the brute facts of explained variance. Instead, narratives are viewed as intuitively attractive, capturing commonsense understanding of how social processes really work. For students of narrative, "social reality happens in sequences of actions located within constraining or enabling structures. It is a matter of particular social actors, in particular social places, at particular social times" (Abbott 1992, 428). In short, narrative is realistic, providing a close fit between theory and humanly enacted events.

In addition to their explanatory power, complex explanations claim a theoretical affinity with the classic traditions of sociology. Karl Korsch's Marxism argued for "historical specificity" in economic explanation. Weber's rejection of highly economistic Marxisms is often taken as a much broader rejection of historical generalization (Mann 1986, 523; Abbott 1992, 430). Abbott (1992) relates narrative methods to the processual orientation of the Chicago School. Some historical sociologists also treat the complex and unpredictable nature of social processes as a theoretical principle. For example, Visser and Ebbinghaus's (1999, 150) QCA study of unions in Western Europe takes a historical institutionalist perspective. This view "does not assume universal linear or cyclical processes but institutional combinations and path-dependent trajectories." Sewell's (1996, 264) historical sociology reflects a similar belief in which "social

relations are characterized by path dependency, temporally heterogeneous causalities, and global contingency." Expanding the argument beyond comparative and historical research, Portes (1999) argues that in social life generally, "goals may not be accomplished by the intended means, but by a fortuitous concatenation of events." Sensitivity to the untidiness of social life is thus seen as is a distinctive element of the sociological viewpoint.

Simplicity and Scientific Understanding

The trend to complex explanation was forcefully challenged by Kiser and Hechter (1991, 1998) and Goldthorpe (1991, 1997). Kiser and Hechter (1991) contrast the historian's commitment to descriptive accuracy with "general theory." General theories use omnitemporal laws as a source of causal propositions that describe how social processes operate under a wide variety of historical conditions. In Kiser and Hechter's analysis, rational choice offers the key example of general theory. For Goldthorpe (1991, 14), "history may serve as a 'residual category' for sociology, marking the point at which sociologists... curb their impulse to generalize..." In Goldthorpe's approach, the historical detail that excites many macrosociologists is just noise for the generalist.

Just as historicist research plies special methods, parsimonious explanation also uses distinctive techniques. The controlled experiment is the ideal method for simple explanation. Instead of identifying combinations of causal conditions, experiments narrow the focus on individual causes (King et al. 1994, 196-97; cf. Ragin 1987, 26-27). Although experimental reasoning is often inexplicit, the importance of controlled comparison for simple explanation is clear. Randomization is typically impossible in macrosociology but statistical control

has become a familiar alternative (Ragin 1987, 58-61). Statistical methods may often be infeasible because few cases are available for comparison. Simple explanation then adapts the logic of quantitative analysis. Thus macrosociologists have applied the controlled comparisons of Mill's method to isolate the impact of causal factors (Skocpol 1984; Lijphart 1971). In the boldest adaptation of statistics, King and his colleagues (1994) apply discussions of collinearity, omitted variable bias, and selectivity to the qualitative setting. Whatever the specific technique, controlled comparisons aim to pinpoint specific causes often by eliminating local context as a source of confounding variation.

Simple explanations are justified differently from complex explanations. Complex accounts are tailored to closely fit the observed data and invoke theoretical arguments for historical sensitivity. Simple explanations are often justified by a model of scientific inquiry. Kiser and Hechter (1991, 9) argue that general theory is scientifically attractive because it generates many testable implications. In comparative and historical research where data are scarce, general theories offer ample opportunity for falsification. The falsifiability of simple explanations is often contrasted with the malleability of complex accounts. Unlike parsimonious explanations, complex stories can accommodate new data by incremental modification. Such fine-tuning, however, shields complex explanations from disconfirmation. In the limit, fine-tuning may yield "a useless hodge-podge of exceptions and exclusions" (King *et al.* 1994, 104). Where complexity in historical explanation is raised to a theoretical principle, the possibility of falsification may be deflected altogether (Kiser and Hechter 1991, 9).

Critics also claim that only simple explanations can sustain strong conclusions. Generalizing the idea of the identification of statistical parameters,

King and his colleagues (1994, 119) warn of the difficulties of supporting complex explanations with small samples: "Each observation can help us make one inference at most" (King *et al.* 1994, 119). They add that strong inferences depend on many cases, not just one. From this perspective, an intricate narrative account that culminates in a single event of interest provides weak evidence of the causes of that event. Goldthorpe (1997, 8) makes a similar point in relation to QCA. "The small N problem is not one of method but one of data... it is a problem of *insufficient information* relative to the complexity of the macrosociological questions that we seek to address" (Goldthorpe 1997, 8, original emphasis). These arguments adopt a statistical idea of evidence in which support for an explanation grows with the accumulation of independent information consistent with that explanation.

Somers (1998, 761) claims that the chief appeal of parsimonious explanation is aesthetic. The aesthetic appeal of parsimony is clear in the physical sciences. Isaac Newton argued that

> We are to admit not more causes of things than such as are both true and sufficient to explain their appearances. To this purpose the Philosophers say that Nature does nothing in vain, and more is in vain when less will serve; for Nature is pleased with simplicity, and affects not the pomp of superfluous causes. (Quoted in Beck 1943, 618-19.)

Modern physicists are also credited with the belief that the world has a simple, discoverable, structure. The statistical astronomer, Harold Jeffreys (1961, 47), describes a simplicity postulate which states that simpler laws are more likely *a priori.* Developing this idea in a statistical context, Jeffreys (1961, 342) later enlists William of Occam in support of the idea that "all variation is random until

the contrary is shown." A frankly aesthetic sensibility is revealed by the physicist Paul Dirac, who claimed that "a theory with mathematical beauty is more likely to be correct than an ugly one that fits some experimental data" (MacKay n.d., 2). This belief in the simplicity of nature contrasts strikingly with the sociological belief in the complexity of social life.

Should macrosociologists prefer simple or complex explanations? This could be treated as a purely empirical question. Because complex explanations are tuned to fit the observed facts they will tend to be more complete than simple explanations. Still, if the explanatory power of simple explanations was sufficiently close to that of a complex explanation we might prefer the simpler approach. The idea of a trade-off between explanatory power and explanatory complexity was detailed by Przeworski and Teune (1971, 211) and Heckathorn (1984). Their discussions suggest that in real analyses, an optimal point could be found that balances complexity and explanatory power. Both sides of the complexity debate in macrosociology would probably agree if such an empirical test could be devised for a given research question, the chips should fall where they may.

However the methodologists in macrosociology debate whether simple or complex explanations can be preferred in principle, before the data have been observed. The debate offers little resolution because the two camps use different criteria for evaluating social explanations. The empirical detail prized in contextual explanation is criticized as a method for evading falsification. The critics also charge that complex sociological theories fail to generate strong evidence. Historicists counter that the generalists have bad taste in theory matched only by their poor taste in austere explanation. Simple explanations fail by the standards used to judge complex explanations, and complex explanations fail by

the standards of parsimony. Despite substantial discussion there has been little systematic effort to weigh the arguments for simple and complex explanation and little compromise.

Theoretical disagreement has obstructed progress. My approach is narrowly methodological and indifferent to sociological theory. I think sociological theory offers little for understanding the merits of simple and complex explanations and the quality of theory is fundamentally an empirical question. For example, Kiser and Hechter (1991) argue that theories should always specify causal relations but Somers (1998) may disagree. My analysis assumes that if there is some advantage in causal theories (or any substantive approach), this will be reflected empirically. This does not imply that theoretical discussion is generally unimportant. The debate on macrosociological methods has usefully discussed whether some kinds of theories yield more constructive programs of research, but this question is beyond the current scope.

BAYESIAN ASSESSMENT OF EXPLANATIONS

Bayesian statistics apply elementary rules of probability to provide a method for learning from data. Bayesian analysis begins with *a priori* beliefs about a theory which are updated by data to form an *a posteriori* conclusion. Imposing the discipline of probability theory ensures that our beliefs are rational and coherent in a technical sense. Bayesian analysis yields rational beliefs which conform to commonsense axioms about the ordering of preferences (Berger 1985, 49). Complying with the rules of probability also offers a coherent way of gambling on our theories. If we took bets on our theories, the Bayesian odds are guaranteed not to lose money under some quite general conditions (Howson and Urbach 1993, 78-86).

Bayesian probability differs from the dominant, frequentist, probability concept in sociology (Berk, Western, and Weiss 1994). Frequentist probability describes the behavior of a statistic over a large number of repeated trials. Frequentist probability thus describes of a class of cases. Bayesian probability describes a researcher's degree of belief in a theory, conditional on observing particular cases. Consistent with the paradigm of complex explanation, the data have identity for the Bayesian. Bayesian conclusions refer to specific cases, and not a general class from which those cases might be drawn.

We begin a Bayesian assessment of explanations by viewing theories as conditional probability statements. A conditional probability is written, $P(D\backslash A)$, and is pronounced "the probability of D given A." We can think of $P(D\backslash A)$ as the likelihood of observing certain data, D, given that theory A is true. Equivalently, $P(D\backslash A)$ describes the data predicted by the theory. Thinking about postwar Swedish union growth, the institutional theory claims the causal importance of the Ghent system and the economic variables thought to influence unionization. In the language of conditional probability, given the Ghent system and economic conditions in postwar Sweden (theory A), we would expect to observe a high level of union organization (the data, D). Here, theory refers to any statement that allocates probabilities to observations. The framework is very general, allowing everything from regression equations to narratives to count as theories. To come under the Bayesian umbrella, researchers need only describe what they would expect to observe (D) if the theory (A) were true.

Because $P(D\backslash A)$ is a probability statement, it allows that some predictions may be more or less likely, not just possible or impossible. Theories in this approach are thus non-deterministic. As a probability, $P(D\backslash A)$ quantifies

uncertainty about the data with a number between 0 and 1. The theory of Swedish unionization may say that, given a Ghent system of unemployment insurance under postwar economic conditions, there is an 80 percent chance of a high unionization, while only a 20 percent chance of union weakness. Of course, we can cast this theory in deterministic terms by specifying a probability of 1 if the causal conditions are met and 0 otherwise.

The idea of theory as a conditional probability statement has an eminent tradition. Weber (1949, 183) cautions that it may be impractical to assign precise numerical probabilities to predictions, but argues that the causal explanation of historical fact involves an assessment of the likelihood of various "objective possibilities." He writes that, "we can... render generally valid judgments which assert that as a result of certain situations, the occurrence of a type of reaction,... is 'favored' to a more less high degree." In a modern context, Stinchcombe (1968, 16) also takes a conditional view of the relationship between theory and data. If Durkheim's theory of egoistic suicide is true, he reasons, we would expect French Protestants to have higher rates of suicide than French Catholics, Protestant regions of German provinces to have higher rates of suicide than Catholic regions, and so on. A probabilistic element is injected into this conditional thinking by acknowledging that the data's inconsistency with theory may be due to the impact of "a large number of small unorganized causes" (Stinchcombe 1968, 23).

While a theory assigns probabilities to data that we might observe, we are really interested in how credible a theory is, having observed some data. In other words, the theory describes $P(D\backslash A)$, but our interest centers on $P(A\backslash D)$, the probability that theory A is true given some data. Using a basic rule of probability, called the conditional probability rule or Bayes rule, we

can write:

$$P(A|D) = \frac{P(D|A)P(A)}{P(D)}$$

The new terms here are $P(A)$, pronounced "the marginal probability of A," and $P(D)$, "the marginal probability of D." In Bayesian statistics these quantities describe the unconditional probability of observing some data, D, and the credence of theory A before the data are observed. The credibility of theory A before the data are observed is called the prior probability of A. The quantity describing the credibility of theory A after the data are observed, $P(A|D)$, is called the posterior probability of A. In the Bayesian model of learning, the data transform prior probabilities of theories into posterior probabilities.

Marginal probabilities can be calculated from conditional probabilities by using the sum rule of probability:

$$P(D) = \sum_A P(D|A)P(A).$$

The summation sign indicates that we are summing over different values of A. Say A is a dummy variable that can equal 1 or 0, perhaps indicating whether A is true or false, then the marginal probability of D is given by:

$$P(D) = P(D|A = 0)P(A = 0) \quad + \quad P(D|A = 1)P(A = 1).$$

This operation is called marginalization. It allows us to describe our uncertainty about one quantity, D in this case, while allowing for uncertainty about about another, A.

When will a theory A be more complicated than an alternative, theory B? In the methodological debates of macrosociology, the complexity of a theory is measured by the number of causal variables involved in the explanation.

In highly idiographic accounts where the rich detail of local context is important, a large combination of causal conditions are held to be operating. In simple explanations, attention may focus on the influence of just one or two causes. For instance, the institutional account of Swedish unions could be contrasted with a simpler economic explanation that emphasizes only the influence of unemployment. The conjunctural account says the influence of unemployment is contingent on the institutional setting (enlisting two causal variables); the simple economic account just emphasizes the impact of one variable—unemployment. We can think of the effects of causal conditions as unobserved quantities, or parameters. For the Bayesian framework the complexity of an explanation is indexed by the number of uncertain parameters implied by that explanation.

Now we are in a position to evaluate simple and complex explanations from a Bayesian perspective (see Gull 1988; Mackay n.d.). To assess the relative support for theory A compared to B we write the the posterior odds,

$$\frac{P(A|D)}{P(B|D)} = \frac{P(D|A)}{P(D|B)} \times \frac{P(A)}{P(B)},$$

where terms for the marginal probability of the data, $P(D)$, cancel out. The posterior odds consists of two terms. The right-hand term, $P(A)/P(B)$, is ratio of prior probabilities which expresses the researcher's relative belief in theory A compared to theory B before observing the data. If—perhaps on the basis of previous studies or theory—A and B seem equally likely, the prior ratio could be set to 1. Posterior support for one model or another then depends on the conditional probabilities of the data under each model, $P(D|A)/P(D|B)$. These conditional probabilities describe how well the data are predicted by each model. The ratio of conditional probabilities is called the Bayes factor.

Theory A
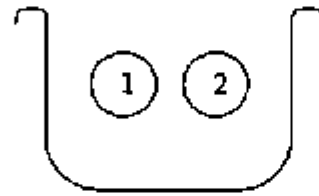
N=2

Theory B
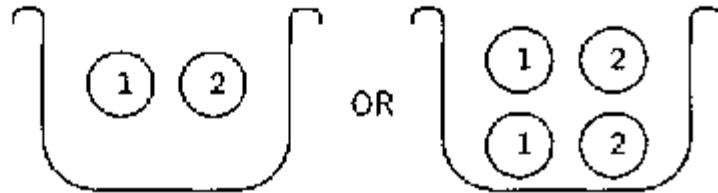
N=2   N=4

Figure 1. Simple theory A specifies there are 2 chips in the bag and complex theory B specifies that there might 2 or 4 chips in the bag.

The Bayes factor prefers simple models. To explain this preference consider the example illustrated in Figure 1. Say we have a bag with poker chips marked "1" and "2." We know that the bag contains an equal number of 1's and 2's. Two theories are proposed to explain the contents of the bag. Theory A assumes the bag contains just 2 chips. The more complex theory B proposes that the number of chips, N, may be 2 or 4. Theory B is more complicated in the sense that it has more parameters than theory A. Theory A has no parameters, but theory B is uncertain about the number of chips, N. Now come the data: Two chips are drawn from the bag, first a 1 and then a 2, which we write $D = \{1, 2\}$.

Do these data provide stronger support for the simple or the complex theory? Under the simple theory A, the probability of drawing a 1 first is one out of two or one-half. If A is true, once a 1 is drawn, only a 2 remains and this will be drawn with certainty. So the conditional probability of the data under theory A is $P(D|A) = \frac{1}{2} \times 1 = \frac{1}{2}$. Under theory B we have to consider the conditional probability of the data under the scenario that N equals two or four. For $N = 2$ we have already calculated the conditional probability of one-half. For $N = 4$, the probability of drawing a 1 first is one-half. If a 1 is drawn, 3 chips remain and the probability of drawing a 2 is two-thirds. With 4 chips, the probability of $\{1, 2\}$ is then $\frac{1}{2} \times \frac{2}{3} = \frac{1}{3}$. Thus $P(D|N = 2) = \frac{1}{2}$ and $P(D|N = 4) = \frac{1}{3}$. To calculate the probability of the data given that $N = 2$ or $N = 4$, we have to specify a probability for the parameter N. If previous research does not point strongly to either the 2-chip or the 4-chip scenario we could assign $N = 2$ and $N = 4$ the same prior probability. Applying the sum rule to obtain the marginal predictive distribution for the data,

$$
\begin{aligned}
P(D|B) &= P(D|N = 2)P(N = 2) + P(D|N = 4)P(N = 4) \\
&= \left(\tfrac{1}{2} \times \tfrac{1}{2}\right) + \left(\tfrac{1}{3} \times \tfrac{1}{2}\right) \\
&= \tfrac{5}{12}
\end{aligned}
$$

The Bayesian calculations show that the data are more probable under the simpler explanation, $P(D|A) = \frac{1}{2}$, than under the complex explanation, $P(D|B) = \frac{5}{12}$. The Bayes factor, the ratio of the evidence under each theory, equals $\frac{6}{5}$ indicating that the data are 20 percent more probable under the simple explanation than the complex. If explanations A and B are equally likely *a priori*, the simple explanation, theory A, has greater posterior probability. In the current example, theory A has posterior probability .545 and

theory B has posterior probability .455. The posterior odds depend on the assumption the $N = 2$ and $N = 4$ are equally likely *a priori* under theory B, but the general result that posterior odds favor the simple theory is insensitive to the choice of prior (see Appendix). Although this example is elementary the same idea has been shown for more complicated problems (Gull 1988).

The example provides our first main conclusion: *When two explanations account equally well for the observed data, we can prefer the simpler explanation because it has higher posterior probability.* Bayesian reasoning thus provides a rational justification for the principle of parsimony. The superior credence of the simple theory does not depend on an aesthetic commitment to the elegance of nature or scientific explanation. In the current example, the analysis was indifferent to the simple and complex explanations before the data were observed. The simple explanation was rewarded in this case because the probability calculus rewards sharp predictions that are consistent with the data.

Although this conclusion supports the parsimonious approach, simple explanations should not generally be preferred to complex explanations. Simple explanations are clearly preferred over the complex when both account for the observed data equally well. Since complex explanations claim a close fit between theory and data, there may be occasions in practice when the complex account is justified by its superior explanatory power.

Why do simple explanations have high posterior probability? Simple theories yield simple predictions (MacKay n.d., 9). Complex theories must spread their probability over a broad range of outcomes. Figure 2 illustrates this idea. Theory A, which assumes there are 2 chips in the bag, can only predict the outcomes $\{1, 2\}$ and $\{2, 1\}$. The complex theory B entertains
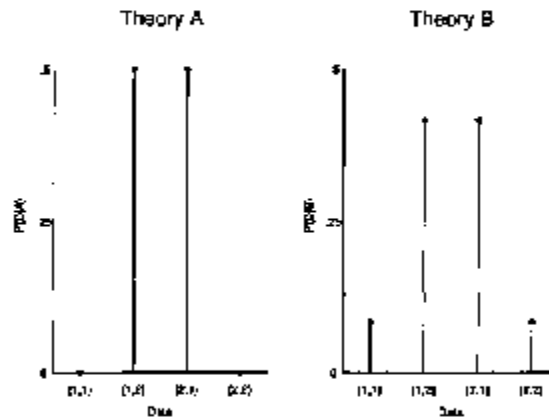
Figure 2. Marginal predictive distributions for theory A and theory B.

a richer variety of predictions, including $\{1,1\}$ and $\{2,2\}$. Drawing two chips with the same number received zero probability under theory A, but received some positive probability under theory B. Again, this can sharpen our intuitions about comparative explanation. Complex theories tend to be more open-ended, adapting to a wide variety of possible outcomes. The predictions of simple theories are much sharper, but simple theories have a higher risk of encountering inexplicable outcomes.

Marginal predictive distributions shed light on the falsifiability of simple explanations. Its often claimed that simple explanations are more testable or easier to falsify (Kiser and Hechter 1991; King et al. 1994; Goldthorpe 1991). Because their marginal predictive distributions are concentrated, simple theories sharply discriminate between supportive and unsupportive data. Observations under a simple theory either yield strong evidence or weak evidence. Data are less informative about the credibility of complex theories. The predictive distributions of complex theories are dispersed, so data often impart intermediate evidence, providing neither compelling support nor clear disconfirmation. Simple theories are strongly falsifiable in the sense of

having narrow predictive distributions. Although it is often claimed as such, falsifiability is not an unambiguous advantage. Because a simple theory predicts a narrow range of outcomes, the area of unexplained events is large. A large range of observations will have very low probability under a simple theory, but higher probability under a complex theory. This implies our second conclusion: *Because marginal predictive distributions of simple theories are concentrated, simple theories are not just "more falsifiable," they are more likely to be falsified.*

Marginal predictive distributions also sharpen thinking about the predictive power of complex explanations. Historically-oriented researchers often acknowledge that their theories are weakly predictive. The idea of contingency in historical explanation sometimes implies unpredictability, not just conditionality (Quadagno and Knapp 1992, 502). For instance, Sewell's (1996, 264) eventful sociology is "inherently unpredictable." Somer's (1998, 769) insists on the "contingent and indeterminate nature" of causal mechanisms. The elevation of unpredictability to an explanatory principle–while perhaps resonating with intuitions about historical processes–comes at a cost. Because the predictive distributions of complex theories are not strongly peaked, the most likely predictions of a complex theory will be less probable than the most likely predictions of a simple theory. Consequently, open-ended theories that entertain a broad array of possible events cannot sustain strong conclusions. The Bayesian analysis thus provides formal basis for the claim that comparative and historical data are often insufficient for supporting complicated explanations (Goldthorpe 1997; King *et al.* 1994). This implies the third conclusion of the Bayesian analysis: *Because the marginal predictive distribution of a complex theory is relatively dispersed, our most confident conclusions about a complex theory can never be as strong as our*

most our most confident conclusions about a simple theory.

How can we draw confident conclusions about complex theories? Consider the accumulation of evidence through multiple experiments. Multiple empirical tests are commonly recommended in discussions of comparative and historical methods. Multiple tests expand variation on independent and dependent variables and help control rival sources of variation (Stinchcombe 1968; Campbell 1975; Smelser 1976). Kiser and Hechter (1998) argue that scope conditions for a theory must be defined abstractly to enable a variety of empirical tests. From this viewpoint multiple tests are important because they "increase the generalizability of explanations" (Kiser and Hechter 1998, 797). Although multiple tests are important in the Bayesian analysis, theoretical generality is not an objective of Bayesian inference. Like more idiographic approaches, cases have identity in the Bayesian approach. Bayesian inference focuses on explaining a particular case, not a general class of cases. By sharpening prior information through multiple tests, the Bayesian improves understanding of particular cases.

Say we conduct another experiment by drawing two more chips from another bag, Bayes rule can be applied again to calculate the posterior probability of the theories A and B. We approach this second calculation with the benefit of information obtained from the first experiment. For experiment 2, the prior odds of the two theories, $P(A)/P(B)$, are now given by the posterior odds obtained from experiment 1. For theory B in experiment 2, the prior probability that $N = 2$ and $N = 4$ are revised in light of evidence from the first experiment. In approaching experiment 2, theory A is now preferred to theory B a priori, and under theory B, $N = 2$ has greater prior probability than $N = 4$. Now assume that experiment 2 yields the same data as experiment 1, $D = \{1, 2\}$. Repeating the calculations of experiment

1 with these new data shows that posterior belief in A compared to B has increased from $\frac{9}{5}$ to $\frac{18}{13}$. As evidence accumulates for one theory over another, the posterior probability of that theory also increases.

Although the two experiments in this example favor the simple theory A over the complex theory B, the accumulation of data also induces a change in the predictions of theory B. The probability of each possible outcome under theory B can be updated as new data arrive. Figure 3 shows the marginal predictive distribution of theory B before experiment 1 and after experiment 7, where every experiment yields the result $\{1,2\}$ or $\{2,1\}$. As more data arrive, evidence concentrates in the center of the range of possible outcomes, and the extreme outcomes, $\{1,1\}$ and $\{2,2\}$, are becoming less likely. The marginal predictive distribution becomes more concentrated because the probability that $N = 2$ increases with each experiment. By experiment 7, the probability that $N = 2$ is 95 percent, nearly double the prior probability of 50 percent before experiment 1. The accumulation of consistent data thus leads the marginal predictive distribution of the complex theory B to increasingly resemble the marginal predictive distribution of the simple theory A. As consistent evidence accumulates, the observed data become more probable under the complex theory, and we become more sure about the values of unknown parameters. The sequential application of Bayes rule provides our fourth conclusion: *Confident conclusions about complex theories require multiple empirical tests.*

In sum, the Bayesian approach admits a strong compromise in the debate over explanatory complexity. The preference for simple explanation is strongly justified by the probability calculus in the special case where a complex alternative fits the data equally well. Simple explanations also have the advantage—consistent with Goldthorpe (1997) and King et al. (1994)—of
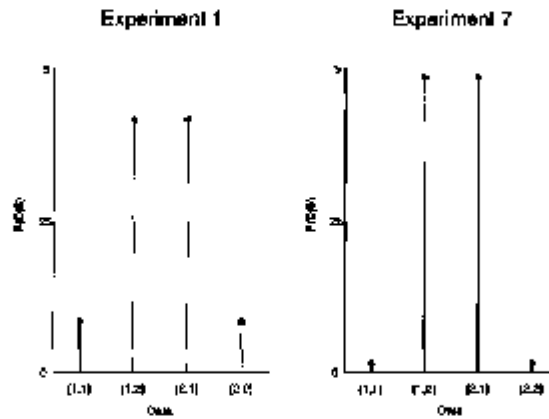
Figure 3. Marginal predictive distributions of theory D before experiment 1 and after experiment 7, where all experiments yield the data $\{1, 2\}$.

yielding potentially stronger conclusions. This is not an unambiguous advantage of parsimony however, because simple explanations also face a relatively high risk of disconfirmation. Still, the conclusion that our most confident inferences about a complex theory can never be as strong as our most confident inferences about a simple theory should be troubling for those whose sociological theory emphasizes the conjunctural and contingent character of historical events. To boost our confidence in such complex theories we must—as comparative and historical researchers are often advised—perform multiple empirical tests. Such multiple tests are important, not for establishing the generality of a theory but for drawing strong conclusions about theories of specific cases.

AN EMPIRICAL ILLUSTRATION: SWEDISH UNION GROWTH

Postwar Swedish union growth can be explained by an historical institutionalist theory that emphasizes the role of the Ghent system and a business cycle explanation that focuses on the impact of unemployment. The institutional and business cycle theories vary in their complexity. The business

cycle explanation is simple: unemployment reduces unionization because union membership provides few benefits and is costly during spells of unemployment. This simple explanation yields the sharp prediction that unemployment reduces unionization everywhere. The institutional account is relatively complex. Unionization may be low or high when unemployment is high, depending on the administrative form of unemployment insurance. The institutional and business cycle explanations appear to provide equally good accounts of Swedish unionization. Consistent with the data, both theories predict high rates of unionization in Sweden. We should then favor the business cycle account, because it has higher posterior probability.

The Bayesian analysis suggests that we can improve our confidence in the institutional explanation by a considering more evidence. The first column of Table 1 reports some estimates from a regression of the annual change in unionization on unemployment from a sample of 18 countries for the period 1950-1985. We would expect that the association between unemployment and unionization is weak in Ghent system countries but strongly negative elsewhere. The results are inconclusive. Large negative unemployment effects can only be found in countries with public unemployment insurance. Still some unemployment effects are close to zero in the public-insurance countries, and small negative unemployment effects can be found in two of the countries with Ghent systems. In this case, the institutional theory may edge the business cycle approach, but our confidence in the more complex account is rather low.

The institutional theory suggests that those at high risk of unemployment will by highly unionized in Ghent system countries. Table 1 provides some support, showing coefficients from a logistic regression of union membership on demographic and work characteristics using survey data from ten countries.

Table 1.   Unemployment coefficients, youth coefficients and unionization rates among the unemployed, 18 OECD countries.

	Unemployment Coefficients (1)	Youth Coefficients (2)	Unionization Rates	
			Employed (3)	Unemployed (4)
Ghent Countries				
Belgium	-.02	-	34.9	46.4
Denmark	.01	.06	81.2	71.5
Finland	-.02	.18	-	-
Sweden	.03	-.19	-	-
Non-Ghent Countries				
Australia	-.15	-.36	-	-
Austria	.03	-	-	-
Canada	-.14	-.19	-	-
France	-.15	-	10.6	4.9
Germany	-.23	-.41	23.7	4.9
Italy	-.18	-	26.5	2.4
Netherlands	-.11	-	30.3	15.8
Norway	-.01	-.23	-	-
Switzerland	-.13	-	-	-
United Kingdom	-.18	-.32	31.0	1.6
United States	-.16	-.18		

Note: Unemployment coefficients are from a time-series regression of the annual change in the percent unionized on unemployment and other variables, 1950–1985 (Western 1997); youth coefficients are from a logistic regression of survey data on union membership on demographic and work characteristics (Western 1997); unionization for employed and unemployed workers is reported by Scruggs (1999).

The table reports coefficients for young workers, aged 35 and under, who have a relatively high risk of unemployment. Consistent with the Ghent system hypothesis, small logistic regression coefficients indicate that unionization rates for young workers are similar to those for middle-aged workers in Ghent system countries but not elsewhere. The final column of Table 1 provides additional support, showing that unionization rates are much higher among the unemployed in Ghent system countries. All these results are consistent with the Ghent system hypothesis.

Quasi-experiments are also useful. We can contrast the historical experience of Sweden and Norway. The two countries share structurally similar union movements, a similar cultural heritage and ethnic homogeneity. The Norwegian Ghent system was replaced by state unemployment insurance in 1938, and since then Swedish union density has continuously exceeded the Norwegian by 20 to 30 percentage points. Belgium and the Netherlands provide another paired comparison. Industrial relations are similarly structured in the two countries, but the Netherlands abandoned its Ghent system shortly after the war. Unionization rates in the two countries diverged from the early 1950s. Finally, we can also examine differences in unionization across industries within the Dutch labor market. Although compulsory state unemployment insurance was established in 1952, construction unions retained control over unemployment insurance and unionization rates in the building trades remained relatively high (Western 1997, 58). In sum, econometric evidence, the Sweden-Norway and Belgium-Netherlands comparisons, and inter-industry trends in Dutch unionization support the theory that the Ghent system nullifies the negative impact of unemployment and promotes unionism. With this additional information, we can be more confident that strong growth in Swedish reflects the influence of the Ghent system

 rather than low unemployment.

    It may be objected that the Ghent system theory is simple compared to many of the narrative accounts of historical sociology. As explanations become highly particularistic, multiple empirical tests may be impossible. Two counter-examples challenge this argument. Paige's (1999) recent discussion of the effect of communist leadership on union radicalism is more complex and may be closer to the explanatory norms of historical sociology. In addition to communist union leadership, Paige emphasizes the influence of worker insurgency, employer intransigence and severe economic recession. His account joins a case study of U.S. longshoremen, an analysis of 38 CIO unions in the 1930s, and case studies of Costa Rica and El Salvador. The Bayesian perspective shows that results from each one of these studies helps reinforce our confidence in the others. The limiting example might be provided by the historian, Simon Schama (1991). Schama's *Dead Certainties* recounts the deaths of two different individuals in two different episodes. Each episode consists of several narratives, each written from the viewpoint of a different protagonist. Although this case is unambiguously idiographic, each narrative provides a separate empirical test that influences our credence in the other accounts. Schama's narratives frequently diverge, leaving readers uncertain about the sequence of events leading to the deaths in question.

CONCLUSION

This analysis says nothing about the specific merits of general theory or conjunctural explanation, narrative or regression analysis, QCA or Mill's method. Instead, the discussion identifies the costs associated with simple and complex explanations that are generally unacknowledged by their proponents. Complex explanations–perhaps based on narrative analysis and historically conditional

theory–are relatively uncertain. Simple explanations–perhaps based on regression analysis of general theory–runs a high risk of falsification.

Is it better to propose complex or simple explanations? The Bayesian approach offers no specific guidance. Parsimony should be preferred over complexity when two explanations account equally well for the observed data. This preference is not due to a taste for elegant explanations. When simple and complex explanations provide equally good accounts of the observed data, simple explanations have higher posterior probability. However, in most real empirical analysis, two explanations will not provide identical accounts of the observed data. If the simple explanation provides the better account, it will definitely have higher posterior probability. More commonly, however, the complex explanation provides the better account, and the explanatory gain must be sufficiently large to justify the additional complexity.

This explanatory gain may be difficult to assess in a qualitative setting. To be more confident of the complex explanation in this situation it is necessary to sharpen prior information. We can do this by conducting multiple tests, studying a range of empirical implications of a given theory. The importance of multiple tests has an important place in the canon of comparative methods. In many cases, the appeal of multiple tests is essentially negative; multiple tests allow many opportunities for disconfirmation of a theory (Stinchcombe 1968, 19; Lijphart 1971, 686; Smelser 1976, 200-02; King *et al.* 1994, 19). The Bayesian justification for multiple tests comes from the other direction. Because our most confident inferences about complex explanations can never be as strong as our most confident inferences about simple explanations, we need multiple tests to strengthen our belief in complex accounts.

While the Bayesian analysis offers something to both sides in the debate over macrosociological methods, a strongly historicist perspective falls outside the current approach. If we believe that the union growth in Sweden depends intimately on conditions that can only be defined by dates and place names, and other social mobilizations cannot help us understand the Swedish events, we have no way of improving our prior information. This type of historicism leads us unavoidably to relatively weak conclusions. In practice I think the comparative impulse in macrosociology is strong and there are few examples of a purely idiographic approach. Still, the Bayesian outlook underlines the fundamental rationale of comparative research: to learn about one setting, we must necessarily examine others.

APPENDIX. PROOF THAT THE BAYES FACTOR FAVORS PARSIMONY

To prove that the Bayes factor favors the simple theory A regardless of the priors for $N = 2$ and $N = 4$ under theory B, write the priors for N as α and $(1 - \alpha)$. The priors sum to one because they are probabilities. The marginal predictive distribution for the complex theory is,

$$
\begin{aligned}
P(D|B) &= P(D|N=2)P(N=2) + P(D|N=4)P(N=4) \\
&= \frac{\alpha}{2} + \frac{1 - \alpha}{3} \\
&= \frac{\alpha + 2}{6}.
\end{aligned}
$$

Theory B gives positive probability to $N = 2$ and $N = 4$, so $0 < \alpha < 1$ and $\frac{1}{3} < P(D|B) < \frac{1}{2}$. Because $P(D|A) = \frac{1}{2}$, the Bayes factor $P(D|A)/P(D|B) > 1$, favoring the simple theory regardless of prior.

REFERENCES

Abbott, Andrew. 1990. "Conceptions of Time and Events in Social Science Methods: Causal and Narrative Approaches." *Historical Methods* 23: 140-50.

Abbott, Andrew. 1992. "From Causes to Events: Notes on Narrative Positivism." *Sociological Methods and Research* 20: 428-455.

Beck, Lewis White. 1943. "The Principle of Parsimony in Empirical Science." *Journal of Philosophy* 40: 617-33.

Berger, James O. 1985. *Statistical Decision Theory and Bayesian Analysis.* Second edition. New York: Springer.

Berk, Richard A., Bruce Western, and Robert Weiss. 1995. "Statistical Inference for Apparent Populations." Pp. 421-58 in *Sociological Methodology,* edited Peter Marsden. Oxford: Blackwell.

Calhoun, Craig. 1996. "The Rise and Domestication of Historical Sociology." Pp. 305-37 in *The Historic Turn in the Human Sciences,* edited by Terence McDonald. Ann Arbor, MI: University of Michigan Press.

Campbell, Donald T. 1975. " 'Degrees of Freedom' and the Case Study." *Comparative Political Studies* 8: 178-193.

Goldthorpe, John. 1991. "The Uses of History in Sociology: A Reflection on Some Recent Tendencies." *British Journal of Sociology* 42: 211-230.

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

Griffin, Larry J. 1992. "Temporality, Events, and Explanation in Historical Sociology." *Sociological Methods and Research* 20: 403-27.

Griffin, Larry J. 1993. "Narrative, Event-Structure Analysis, and Causal Interpretation in Historical Sociology." *American Journal of Sociology* 98: 1094-1133.

Gull, Stephen F. 1988. "Bayesian Inductive Inference and Maximum Entropy." *Maximum-Entropy and Bayesian Methods in Science and Engineering* 1: 53-74.

Heckathorn, Douglas. 1984. "Mathematical Theory Construction in Sociology: Analytic Power, Scope, and Descriptive Accuracy as Trade-Offs." *Journal of Mathematical Sociology* 10: 295-323.

Hicks, Alexander, Joya Misra, and T.N. Ng. 1995. "The Programmatic Emergence of the Social Security State." *American Sociological Review* 60: 329-49.

Howson, Colin and Peter Urbach. 1993. *Scientific Reasoning: The Bayesian Approach.* Second edition. Chicago: Open Court.

Huber, Evelyne, Charles Ragin, and John D. Stephens. 1993. "Social Democracy, Christian Democracy, constitutional Structure and the Welfare State." *American Journal of Sociology* 99: 711-49.

Jeffreys, Harold. 1961. *Theory of Probability.* Third Edition. Oxford: Oxford University Press.

King, Gary, Robert O. Keohane, and Sidney Verba. 1994. *Designing Social Inquiry.* Princeton: Princeton University Press.

Kiser, Edgar and Michael Hechter. 1991. "The Role of General Theory in Comparative-Historical Sociology." *American Journal of Sociology* 97: 1-30.

Kiser, Edgar and Michael Hechter. 1998. "Rational Choice Theory and Its Critics." *American Journal of Sociology* 104: 785-816.

Lijphart, Arend. 1971. "Comparative Politics and the Comparative Method." *American Political Science Review* 65: 682-93.

MacKay, David J.C. N.d. "Probable Networks and Plausible Predictions− A Review of Practical Bayesian Methods for Supervised Neural Networks." Unpublished manuscript. Cambridge, England: Cavendish Laboratory.

Mann, Michael. 1986. *The Sources of Social Power, Volume I, A History of Power from the Beginning to A.D. 1760.* New York: Cambridge University Press.

Ortner, Sherry B. 1996. "Resistance and the Problem of Ethnographic Refusal." Pp. 281-304 in *The Historic Turn in the Human Sciences,* edited by Terence J. McDonald. Ann Arbor: University of Michigan Press.

Paige, Jeffrey M. 1999. "Conjuncture, Comparison, and Conditional Theory in Macrosocial Inquiry." *American Journal of Sociology* 105: 781-800.

Portes, Alejandro. 1999. "The Hidden Abode: Sociology as Analysis of the Unexpected." Presidential address, American Sociological Association meetings. Chicago.

Przeworski, Adam and Henry Teune. 1970. *The Logic of Comparative Social Inquiry.* New York: John Wiley.

Quadagno, Jill and Stan J. Knapp. 1992. "Have Historical Sociologists Foresaken Theory? Thoughts on the History/Theory Relationship." *Sociological Methods and Research* 20: 481-507.

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

Rothstein, Bo. 1990. "Marxism, Institutional Analysis, and Working Class Power." *Politics and Society* 18: 317-45.

Rothstein, Bo. 1992. "Labor Market Institutions and Working Class Strength" Pp. 33-56 in *Structuring Politics: Historical Institutionalism in Comparative Analysis,* edited by Sven Steinmo, Kathleen Thelen, and Frank Longstreth. New York: Cambridge University Press.

Rueschemeyer, Dietrich and John D. Stephens. 1997. "Comparing Historical Sequences−A Powerful Tool for Causal Analysis." *Comparative Social Research* 16: 55-72.

Schama, Simon. 1991. *Dead Certainties: Unwarranted Speculations.* New York: Vintage.

Scruggs, Lyle. 1999. "Divergent Fates? The Ghent System and Union Membership in Europe, 1970-1995." Unpublished manuscript. University of Connecticut.

Sewell, William. 1996. "Three Temporalities: Toward and Eventful Sociology." Pp. 245-280 in *The Historic Turn in the Human Sciences,* edited by Terence McDonald. Ann Arbor: University of Michigan Press.

Skocpol, Theda. 1984. "Emerging Agendas and Recurrent Strategies in Historical Sociology." Pp. 356-391 in *Vision and Method in Historical Sociology,* edited by Theda Skocpol. New York: Cambridge University Press.

Smelser, Neil J. 1976. *Comparative Methods in the Social Sciences.* Englewood Cliffs, NJ: Prentice-Hall.

Somers, Margaret R. 1998. " 'We're No Angels': Realism, Rational Choice, and Relationality in Social Science." *American Journal of Sociology* 104: 722-784.

Stinchcombe, Arthur L. 1968. *Constructing Social Theories.* Chicago: University of Chicago Press.

Stone, Lawrence. 1979. "The Revival of Narrative." *Past and Present* 85: 3-24.

Tilly, Charles. 1984. *Big Structures Large Processes Huge Comparisons.* New York: Russell Sage Foundation.

Tilly, Charles. 1997. *Roads from Past to Future.* Lanham, MD: Rowman and Littlefield.

Visser, Jelle and Bernhard Ebbinghaus. 1999. "When Institutions Matter: Union Growth and Decline in Western Europe, 1950-1995." *European Sociological Review* 135-58.

Weber, Max. 1949. *Methodology of the Social Sciences.* Edited and translated by Edward A. Shils and Henry A. Finch. New York: Free Press.

Western, Bruce. 1997. *Between Class and Market: Postwar Unionization in the Capitalist Democracies.* Princeton, NJ: Princeton University Press.