



Causal Inference in the Social Sciences

Author(s): Michael E. Sobel

Source: *Journal of the American Statistical Association*, Vol. 95, No. 450 (Jun., 2000), pp. 647-651

Published by: American Statistical Association

Stable URL: <http://www.jstor.org/stable/2669410>

Accessed: 07/04/2010 03:38

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=astata>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



American Statistical Association is collaborating with JSTOR to digitize, preserve and extend access to *Journal of the American Statistical Association*.

<http://www.jstor.org>

- Clogg, C. C., and Shidadeh, E. S. (1994), *Statistical Models for Ordinal Variables*, Thousand Oaks, CA: Sage.
- Cressie, N. E., and Holland, P. W. (1983), "Characterizing the Manifest Probabilities of Latent Trait Models," *Psychometrika*, 48, 129–141.
- Christensen, R. (1996), *Plane Answers to Complex Questions: The Theory of Linear Models* (2nd ed.), New York: Springer-Verlag.
- Darroch, J. N. (1962), "Interaction in Multi-Factor Contingency Tables," *Journal of the Royal Statistical Society, Ser. B*, 24, 251–263.
- (1986), "Quasi-Symmetry," in *Encyclopedia of Statistical Sciences*, Vol. 7, eds. S. Kotz and N. L. Johnson, New York: Wiley, pp. 469–473.
- Darroch, J. N., Fienberg, S. E., Glonek, G., and Junker, B. (1993), "A Three-Sample Multiple-Recapture Approach to Census Population Estimation With Heterogeneous Catchability," *Journal of the American Statistical Association*, 88, 1137–1148.
- Darroch, J. N., Lauritzen, S. L., and Speed, T. P. (1980), "Markov Fields and Log-Linear Interaction Models for Contingency Tables," *The Annals of Statistics*, 8, 522–539.
- Diaconis, P., and Sturmfels, B. (1998), "Algebraic Algorithms for Sampling From Conditional Distributions," *The Annals of Statistics*, 26, 363–397.
- Diggle, P. J., Liang, K.-Y., and Zeger, S. L. (1996), *Analysis of Longitudinal Data*, New York: Oxford University Press.
- Duncan, O. D. (1983), "Rasch Measurement: Further Examples and Discussion," in *Survey Measurement of Subjective Phenomena*, Vol. 2, eds. C. F. Turner and E. Martin, New York: Russell Sage, pp. 367–403.
- Fienberg, S. E. (1972), "The Multiple Recapture Census for Closed Populations and Incomplete 2^k Contingency Tables," *Biometrika*, 59, 591–603.
- (1980), *The Analysis of Cross-Classified Categorical Data* (2nd ed.), Cambridge, MA: MIT Press.
- (1999), "Fréchet and Bonferroni Bounds for Multi-Way Tables of Counts With Applications to Disclosure Limitation," in *Statistical Data Protection: Proceedings of the Conference*, Luxembourg: Eurostat, 115–129.
- Fienberg, S. E., Johnson, M. A., and Junker, B. (1999), "Classical Multi-Level and Bayesian Approaches to Population Size Estimation Using Data From Multiple Lists," *Journal of the Royal Statistical Society, Ser. A*, 162, 383–406.
- Fienberg, S. E., and Meyer, M. M. (1983), "Loglinear Models and Categorical Data Analysis With Psychometric and Econometric Applications," *Journal of Econometrics*, 22, 191–214.
- Fienberg, S. E., Meyer, M. M., and Wasserman, S. S. (1985), "Statistical Analysis of Multiple Sociometric Relations," *Journal of the American Statistical Association*, 80, 51–67.
- Good, I. J. (1963), "Maximum Entropy for Hypothesis Formulation, Especially for Multidimensional Contingency Tables," *Annals of Mathematical Statistics*, 34, 911–934.
- Goodman, L. A. (1963), "On Methods for Comparing Contingency Tables," *Journal of the Royal Statistical Society, Ser. A*, 126, 94–108.
- (1979), "Simple Models for the Analysis of Association in Cross-Classifications Having Ordered Categories," *Journal of the American Statistical Association*, 74, 537–552.
- Haberman, S. J. (1974), *The Analysis of Frequency Data*, Chicago: University of Chicago Press.
- (1977), "Log-Linear Models and Frequency Tables With Small Expected Cell Counts," *The Annals of Statistics*, 5, 1148–1169.
- Holland, P. W., and Leinhardt, S. (1981), "An Exponential Family of Probability Distributions for Directed Graphs" (with discussion), *Journal of the American Statistical Association*, 76, 33–65.
- Koehler, K. J., and Larntz, K. (1980), "An Empirical Investigation of Goodness-of-Fit Statistics for Sparse Multinomials," *Journal of the American Statistical Association*, 75, 336–344.
- Lauritzen, S. L. (1996), *Graphical Models*, New York: Oxford University Press.
- McCullagh, P., and Nelder, J. A. (1989), *Generalized Linear Models* (2nd ed.), London: Chapman and Hall.
- Plackett, R. L. (1974), *The Analysis of Categorical Data*, London: Charles Griffin.
- Rasch, G. (1960), *Probabilistic Models for Some Intelligence and Attainment Tests*, The Danish Institute of Educational Research, expanded ed. (1980), Chicago: The University of Chicago Press.
- Read, T. R. C., and Cressie, N. A. C. (1988), *Goodness-of-Fit Statistics for Discrete Multivariate Data*, New York: Springer-Verlag.
- Roy, S. N., and Kastenbaum, M. A. (1956), "On the Hypothesis of No Interaction in a Multi-way Contingency Table," *Annals of Mathematical Statistics*, 27, 749–757.
- Tjurr, T. (1982), "A Connection Between Rasch's Item Analysis Model and a Multiplicative Poisson Model," *Scandinavian Journal of Statistics*, 9, 23–30.
- Wasserman, S., and Pattison, P. (1996), "Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and p^* ," *Psychometrika*, 61, 401–425.
- Whitaker, J. (1990), *Graphical Models in Applied Multivariate Statistics*, New York: Wiley.

Metodologia e tecnica della ricerca sociale a.a. 2010-2011

Causal Inference in the Social Sciences

Michael E. SOBEL

The subject of causation is controversial. For several hundred years, since Hume, who essentially equated causation with constant conjunction ("same cause, same effect"), until recently, theories that view causation as a form of empirical regularity (predictability) have dominated philosophical discussions of the causal relation. Many writers who subscribe to this view have also argued that not all regular sequences evidence causation. Thus Hume's account is regarded as incomplete (nor would this account be completed by including Hume's other criteria, temporal prior-

ity and contiguity), and it is important to understand just what separates causal from noncausal sequences. One idea that has been put forth (it can even be found in Hume, though he appears not to have realized its implications) is that whereas both types of sequences exhibit de facto regularity, only causal sequences are counterfactually regular. That is, in both cases, one can state that when the object or event X is present ($X = 1$), it is followed by the effect Y ($Y = 1$), but in the case of a causal sequence, one can also state that if X is not present, then Y also is not be present.

Michael E. Sobel is Professor, Department of Sociology, Columbia University, New York, NY 10027 (E-mail: mes105@columbia.edu). The author thanks Sandler Greenland, Donald Rubin, and an anonymous reviewer for helpful comments.

© 2000 American Statistical Association
Journal of the American Statistical Association
June 2000, Vol. 95, No. 450, Vignettes

These notions raise epistemological problems. First, the relationships that scientists study rarely hold in every instance. Thus it seems reasonable to entertain notions of causation that allow us to speak of causation in individual instances and/or that feature so-called “probabilistic causation.” Second, if a causal relationship must sustain a counterfactually conditional statement, then additional problems arise because one can observe the actual states of X and Y , but not the state of Y that would occur were X in a state other than its actual state. Thus it is possible that $Y = 1$ in those instances where $X = 1$, but this might be true even if $X = 0$. Consider the case of weight loss among motivated persons who take diet pills ($X = 1$). Intuition suggests that by controlling who receives (does not receive) diet pills, the possibility of encountering this situation is reduced. In other words, experiments might be useful for inferring causation.

In the social sciences, where many of the questions asked do not appear amenable to experimentation, approaches to inferring causation have been proposed that differed from those that emerged in the statistical literature on experimental design. Prior to the development of simultaneous equation models by econometricians and the related work by Simon (1954) on spurious correlation (which suggested the possibility of distinguishing between causal and noncausal associations), many quantitative social scientists eschewed causal language. Bernert (1983) described the situation in American sociology. Subsequently, sociologists (e.g., Duncan 1966) discovered path analysis and disseminated it to other social sciences, promoting its use for drawing causal inferences and more generally, promoting the use of causal language. Later, in an important generalization, Jöreskog (1977) embedded factor analysis from psychology and simultaneous equation models from economics into the covariance structure model, seemingly allowing causal inference with latent causes and effects. (These models are now commonly called structural equation models or causal models.) More recently, a literature on graphical models (e.g., Whittaker 1990) has emerged. Use of these models for drawing causal inferences (e.g., Pearl 1998; Spirtes, Glymour, and Scheines 1993) is another outgrowth of the foregoing approach and the literature on path analysis. However, some authors (e.g., Cox and Wermuth 1993) have deliberately avoided using causal language in conjunction with such models. Other extensions of Simon’s work are the methods that economists have proposed for assessing so-called “spurious causation” in time series (Geweke 1984; Granger 1969).

Although the earlier developers of these methods tended, in tandem with the prevailing currents in philosophy in their day, to equate causality with prediction, subsequent writers, in tune with later currents, have tended toward the view that causal relationships sustain counterfactual conditionals. Thus in the literature on simultaneous equation models, exogenous variables (errors are defined so as to be uncorrelated with these exogenous variables) came to be regarded as potentially manipulable inputs, and model parameters were interpreted as indicative of what would hap-

pen under interventions (see, e.g., Goldberger 1964, p. 375). Similarly, sociologists, psychologists, and others speak of parameters in structural equation models as effects, interpreting these as indicative of what would happen to a response if the independent variables were manipulated. (For references to this literature, see Sobel 1990.) Similar remarks apply to some treatments of Granger causality. Although it is recognized that such interpretations are extramathematical and not always warranted, it is widely believed that these interpretations are licensed when the variables in the model are ordered properly and the “causal theory” underlying the model is “valid,” irrespective of study design.

It seems evident that if causation is equated with predictability, then causal inference can be approached using nonexperimental studies and statistical methods for studying association (as above). In retrospect, it also seems evident that under a counterfactual account of causation, which involves comparing an observed response with a response that would have occurred under conditions that were not observed, an analysis of the causal relationship needs to begin with the use of something like the notation invented by Neyman (1923) to discuss potential yields in an agricultural experiment. This notation allows for unambiguous definitions of effects, enabling evaluation of methods for estimating these.

To illustrate, suppose that the response (Y) is to be compared under treatment and no treatment, and there is no interference among units; that is, the response of a unit does not depend on the treatment received by other units (Cox 1958). Then either y_{ic} , the response of unit i in the absence of treatment, or y_{it} , the response of unit i under treatment, can be observed. The unit (causal) effect is defined as a function of the two responses; for example, $y_{it} - y_{ic}$. Although this cannot be observed, Neyman showed (under a model of assignment equivalent to randomization) that the difference between treatment and control group means is an unbiased estimator of the average of the unit effects—the average treatment effect (ATE). Notice that here, study design does enter the picture.

Although the potential outcomes notation was subsequently used in the statistical literature on experiments, Rubin (1974, 1977, 1978, 1980, 1991) applied it to observational studies to define effects that accorded with a counterfactual account of causation and to state sufficient conditions (ignorability and strong ignorability) for estimating these effects consistently (Rosenbaum and Rubin 1983a). Building on this work, Rubin and coworkers clarified a number of issues important to social and behavioral scientists, including Lord’s paradox (Holland and Rubin 1983), causal inference in case-control studies (Holland and Rubin 1988) and causal inference in recursive structural equation models (Holland 1988).

Increasingly, social statisticians are finding that use of the foregoing framework above results in greater clarity, enabling precise definitions of causal estimands of interest and evaluation of methods traditionally used to make causal inferences. An excellent example of this is the work on instrumental variables by Angrist and Imbens (1995) and

Angrist, Imbens, and Rubin (1996). (For related material, see also Robins and Greenland 1992.) The idea behind the instrumental variables estimator, which goes back at least to Wright (1928), is (in econometric parlance) to estimate the effect of a variable correlated with the error term in a regression (i.e., the variable is not “exogenous”) by using another variable that is correlated with the response, but that does not directly affect the response. Angrist et al. (1996) studied this estimator in a randomized study with non-compliance, showing, under assumptions that they clearly spelled out, that this estimates a parameter they called the local average treatment effect (LATE). LATE is the average treatment effect among compliers (those who both comply with their actual assignment and who would comply with the assignment not assigned). Although the subpopulation of compliers is not observable, and other effects may often be of greater interest (e.g., the average effect among the treated), it nevertheless is of great value to know what the instrumental variables estimator is actually estimating, and what assumptions are implicit when this estimator is taken to estimate other parameters; for example, the average treatment effect among the treated. One of the important assumptions examined in this work is the exclusion restriction, namely that treatment assignment (as versus treatment received) does not affect the response. (This restriction appears also in the econometric literature and can also be formulated in the literature on structural equation models in terms of direct effects.) Another important assumption, used in conjunction with the exclusion restriction, is the monotonicity assumption; under this assumption, the latent subpopulation of subjects who would not take treatment when treatment is assigned and who would take treatment when it is not assigned is nonexistent. Monotonicity is realized in many policy and prevention studies where only subjects assigned to treatment are allowed to enroll in the intervention program.

Subsequently, Imbens and Rubin (1997) extended this work, applying likelihood and Bayesian procedures to estimate LATE in both the case where the exclusion restriction holds and the case where it does not; they referred to LATE as CACE (“complier average causal effect”). Little and Yau (1998) estimated LATE by maximum likelihood in a randomized study of unemployed workers, where the treatment consists of several training sessions and the response is the reduction in depression; in this study, almost half of the subjects assigned to treatment did not actually take the treatment. Whereas Little and Yau assumed that the exclusion restriction holds, Jo (1999), arguing that there should be a direct effect of treatment assignment on the response among noncompliers, extended this work to that case.

In observational studies, the assumption that treatment assignment is random, given observed covariates (strong ignorability), allows consistent estimation of effects. Matching or stratifying on these covariates is desirable, if possible, facilitating estimation with minimal additional assumptions. If this is not possible, Rosenbaum and Rubin (1983a) showed that it suffices to match or stratify on the propen-

sity score. (In a study with treatment or no treatment, this is the probability of treatment, given the covariates.) Alternatively, model-based adjustments may be made. However, rarely is the state of knowledge in the social sciences adequate for a researcher to feel confident that he or she has measured all of the relevant covariates. The possibility of hidden bias (relevant unobserved covariates) is great. Thus it is important to develop methods for assessing the sensitivity of an analysis to this possibility, as in Rosenbaum and Rubin (1983b). (For further material, see Rosenbaum 1995, who has done extensive work on this topic.) Economists, who sometimes refer to strong ignorability as selection on observables (Heckman and Hotz 1989), have also pointed to the importance of unobserved variables (selection on unobservables), proposing various methods for using longitudinal data, in conjunction with modeling assumptions, to estimate treatment effects that correct for this type of selection (Heckman and Robb 1985, 1988). (For a more extensive review of this literature, see Winship and Morgan 1999.)

Another useful approach develops upper and lower bounds on the values of effects of interest (assuming that the response is bounded). Although these effects are not identified in observational studies, components of the effects are identified. Assumptions about the upper and lower bounds on the nonidentified components are then used in conjunction with the identified components to obtain the bounds. When the nonidentified components take on the maximum and the minimum values allowable, the bounds can be quite wide. Nonetheless, these can be narrowed considerably under quite reasonable assumptions compatible with the state of existing knowledge. Robins (1989) first developed these bounds. Manski (1995) also used these bounds, and also developed bounds for several other contexts of special relevance to social scientists.

Although the interplay between statistics and social sciences has proven very useful, many issues need further work. In Holland’s (1988) work on path analysis, subjects are randomized to a group encouraged to study and a group not encouraged to study. The amount of time studied and the score on a subsequent test are recorded. Assuming that the “direct effect” of encouragement on test scores is 0 for all subjects (an exclusion restriction) and that the effect of time studied on the response is identical for all subjects (the assumption of constant effect), the value of this effect is the instrumental variable estimand. It would be useful to extend this work to psychological prevention studies, where the treatment targets several mediators, which in turn affect the response, and researchers inappropriately use structural equation models to estimate the effects of the mediators on the response. It would also be useful to do this without the strong assumption of constant effect typically made in conjunction with linear models. New study designs for estimating the mediated effects could then be developed. Second, in psychology and sociology, researchers often use latent variables derived from factor analyses as causes, covariates, and effects. Little attention has been paid to the issues that arise in attempting to make appropriate causal inferences in such studies (Sobel 1994, 1997). Finally, so-

cial scientists are often concerned with interdependence among units. One form occurs when the response of a subject under a given condition depends on the conditions to which other subjects are assigned; the assumption of no interference between units discussed by Cox (1958) or, in Rubin's terms, SUTVA or stability does not hold. Violations of stability certainly have been alluded to by social scientists (Garfinkel, Manski, and Michalopoulos 1992; Sobel 1995), and some work on this subject has been done in epidemiology (Halloran and Struchiner 1995), but much remains to be done in both experimental and nonexperimental contexts.

In most disciplines within the social sciences, empirical work has yet to be much informed by the aforementioned developments. As this changes, a dramatic transformation should be seen in the way social research is conducted and the purposes that this research serves. To take but one example, much of quantitative political science and sociology may be characterized as a highly stylized search for new causes of effects. Researchers typically begin with one or more outcomes and a list of causes identified by previous workers. Potentially new causes are then listed; if these account ("explain") for additional variability in the response, then the new causes are held to affect the outcome, and the significant coefficients in the new model are endowed with a causal interpretation. The process is repeated by subsequent workers, resulting in further "progress." When researchers realize that they are merely adding more and more variables to a predictive conditioning set, one wonders what will take the place of the thousands of purported (causal) effects that currently fill the journals.

REFERENCES

- Angrist, J. D., and Imbens, G. W. (1995), "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity," *Journal of the American Statistical Association*, 90, 431–442.
- Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996), "Identification of Causal Effects Using Instrumental Variables" (with discussion), *Journal of the American Statistical Association*, 91, 444–472.
- Bernert, C. (1983), "The Career of Causal Analysis in American Sociology," *British Journal of Sociology*, 34, 230–254.
- Cox, D. R. (1958), *The Planning of Experiments*, New York: Wiley.
- Cox, D. R., and Wermuth, N. (1993), "Linear Dependencies Represented by Chain Graphs" (with discussion), *Statistical Science*, 9, 204–218, 247–283.
- Duncan, O. D. (1966), "Path Analysis: Sociological Examples," *American Journal of Sociology*, 72, 1–16.
- Garfinkel, I., Manski, C., and Michalopoulos, C. (1992), "Micro-Experiments and Macro Effects," in *Evaluating Welfare and Training Programs*, eds. C. Manski and I. Garfinkel, Cambridge, MA: Harvard University Press, pp. 253–273.
- Geweke, J. (1984), "Inference and Causality in Economic Time Series Models," in *Handbook of Econometrics*, Vol. 2, eds. Z. Griliches and M. D. Intriligator, Amsterdam: North-Holland, pp. 1101–1144.
- Goldberger, A. S. (1964), *Econometric Theory*, New York: Wiley.
- Granger, C. W. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods," *Econometrica*, 37, 424–438.
- Halloran, M. E., and Struchiner, C. J. (1995), "Causal Inference in Infectious Diseases," *Epidemiology*, 6, 142–151.
- Heckman, J. J., and Hotz, V. J. (1989), "Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training" (with discussion), *Journal of the American Statistical Association*, 84, 862–880.
- Heckman, J. J., and Robb, R. (1985), "Alternative Methods for Evaluating the Impact of Interventions," in *Longitudinal Analysis of Labor Market Data*, eds. J. J. Heckman and B. Singer, Cambridge, U.K.: Cambridge University Press, pp. 156–245.
- (1988), "The Value of Longitudinal Data for Solving the Problem of Selection Bias in Evaluating the Impact of Treatment on Outcomes," in *Panel Surveys*, eds. G. Duncan and G. Kalton, New York: Wiley, pp. 512–538.
- Holland, P. W. (1988), "Causal Inference, Path Analysis, and Recursive Structural Equation Models" (with discussion), in *Sociological Methodology*, 1988, ed. C. C. Clogg, Washington, DC: American Sociological Association, pp. 449–493.
- Holland, P. W., and Rubin, D. B. (1983), "On Lord's Paradox," in *Principals of Modern Psychological Measurement*, eds. H. Wainer and S. Messnick, Hillsdale, NJ: Lawrence Erlbaum, pp. 3–35.
- (1988), "Causal Inference in Retrospective Studies," *Evaluation Review*, 12, 203–231.
- Imbens, G. W., and Rubin, D. B. (1997), "Bayesian Inference for Causal Effects in Randomized Experiments With Non-Compliance," *The Annals of Statistics*, 25, 305–327.
- Jo, B. (1999), "Estimation of Intervention Effects With Noncompliance: Model Misspecification Sensitivity Analysis," unpublished manuscript.
- Jöreskog, K. G. (1977), "Structural Equation Models in the Social Sciences: Specification, Estimation and Testing," in *Applications of Statistics*, ed. P. R. Krishnaiah, Amsterdam: North-Holland, pp. 265–287.
- Little, R. J., and Yau, L. H. Y. (1998), "Statistical Techniques for Analyzing Data From Prevention Trials: Treatment of No-Shows Using Rubin's Causal Model," *Psychological Methods*, 3, 147–159.
- Manski, C. F. (1995), *Identification Problems in the Social Sciences*, Cambridge, MA: Harvard University Press.
- Neyman, J. (1923), "On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9," translated into English and edited by D. M. Dabrowska and T. P. Speed (1990), *Statistical Science*, 5, 463–472.
- Pearl, J. (1998), "Graphs, Causality, and Structural Equation Models," *Sociological Methods and Research*, 27, 226–284.
- Robins, J. M. (1989), "The Analysis of Randomized and Nonrandomized AIDS Treatment Trials Using a New Approach to Causal Inference in Longitudinal Studies," in *Health Services Research Methodology: A Focus on AIDS*, eds. L. Sechrest, H. Freedman, and A. Mulley, Rockville, MD: U.S. Department of Health and Human Services, pp. 113–159.
- Robins, J. M., and Greenland, S. (1992), "Identifiability and Exchangeability for Direct and Indirect Effects," *Epidemiology*, 3, 143–155.
- Rosenbaum, P. R. (1995), *Observational Studies*, New York: Springer-Verlag.
- Rosenbaum, P. R., and Rubin, D. B. (1983a), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, 70, 41–55.
- (1983b), "Assessing Sensitivity to an Unobserved Binary Covariate in an Observational Study With Binary Outcome," *Journal of the Royal Statistical Society, Ser. B*, 212–218.
- Rubin, D. B. (1974), "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies," *Journal of Educational Psychology*, 66, 688–701.
- (1977), "Assignment to Treatment Groups on the Basis of a Covariate," *Journal of Educational Statistics*, 2, 1–26.
- (1978), "Bayesian Inference for Causal Effects: The Role of Randomization," *The Annals of Statistics*, 6, 34–58.
- (1980), Comment on "Randomization Analysis of Experimental Data: The Fisher Randomization Test," by D. Basu, *Journal of the American Statistical Association*, 75, 591–593.
- (1991), "Practical Implications of Modes of Statistical Inference for Causal Effects and the Critical Role of the Assignment Mechanism," *Biometrics*, 47, 1213–1234.
- Simon, H. A. (1954), "Spurious Correlation: A Causal Interpretation," *Journal of the American Statistical Association*, 49, 467–492.
- Sobel, M. E. (1990), "Effect Analysis and Causation in Linear Structural Equation Models," *Psychometrika*, 55, 495–515.
- (1994), "Causal Inference in Latent Variable Models," in *Latent Variables Analysis: Applications for Developmental Research*, eds. A. von Eye and C. C. Clogg, Thousand Oaks, CA: Sage, pp. 3–35.

- (1995), "Causal Inference in the Social and Behavioral Sciences," in *Handbook of Statistical Modeling for the Social and Behavioral Sciences*, eds. G. Arminger, C. C. Clogg, and M. E. Sobel, New York: Plenum, pp. 1–38.
- (1997), "Measurement, Causation, and Local Independence," in *Latent Variable Modeling and Applications to Causality*, ed. M. Berkane, New York: Springer-Verlag, pp. 11–28.
- Spirites, P., Glymour, C., and Scheines, R. (1993), *Causation, Prediction*

- and Search, New York: Springer-Verlag.
- Whittaker, J. (1990), *Graphical Models in Applied Multivariate Statistics*, New York: Wiley.
- Winship, C., and Morgan, S. L. (1999), "The Estimation of Causal Effects From Observational Data," forthcoming in *Annual Review of Sociology*.
- Wright, S. (1928), "Appendix," in *The Tariff on Animal and Vegetable Oil*, by P. G. Wright, New York: MacMillan.

Political Methodology: A Welcoming Discipline

Nathaniel L. BECK

1. INTRODUCTION

Although empirical political science can be dated as far back as Aristotle, and some of the earliest work in statistics was on political methodology (Petty 1690), the self-definition of political science as a science goes back only a century or so (with a convenient date being the founding of a "scientific" department at Columbia University in 1880). Political methodology is a young subfield within political science. The coming of age of that subfield is best evidenced by the arrival of the journal *Political Methodology*, now named *Political Analysis*. By that date, the subfield of political methodology is about 25 years old.

Although serious quantitative political analyses were conducted in the nineteenth and early twentieth centuries (Gow 1985), the increased use of quantitative analysis went hand-in-hand with the post-World War II "behavioral revolution" in political science (Dahl 1961). The behavioralists had begun using more and more sophisticated methods, though few, if any, would have labelled themselves as "methodologists." The most sophisticated behavioralists acquired their methodological training from other disciplines (first sociology, later economics) and usually did not teach courses in methodology to pass on their sophistication to their graduate students. But starting some time in the late 1960s, we did begin to develop a small cadre of people who did at least a portion of their scholarship as methodologists.

Chris Achen, in his final editorial for *Political Methodology*, summed up the situation:

When [*Political Methodology*] began in the mid-1970s, methodology was more often an avocation than a vocation. No political science journal welcomed methodological articles, and many journals rejected them out of hand. Certainly no Political Methodology Society existed to give shape and organization to the needs of political methodologists. In the face of these difficulties, John Sullivan and George Marcus created *Political Methodology*... By [the 1980s], the field had come of age, and the rapid development of the last few years were possible (Achen 1985).

In this essay, I look at, from my own perspective, a few major features of this burgeoning of political methodology

and how they might play out in the early part of the new millennium. In particular, I look at the relationship of political methodology to other social science disciplines and to statistics. I then look at the changing nature of data collection.

2. POLITICAL METHODOLOGY AND OTHER SOCIAL SCIENCES

Political science as a discipline is substantively, not methodologically, defined. Political scientists use a variety of methods to attack questions related to political institutions and behavior. Although the methodological issues are defined by our political questions, we freely use whatever methodological solutions are available. Thus political methodology has freely drawn on insights from econometrics, psychometrics, sociology, and statistics. One clear trend, however, is our increased reliance on our own methodological expertise.

Quantitative political analysis in the 1950s and early 1960s was largely the analysis of contingency tables. By the 1960s, path analysis, imported from sociology, was becoming the state of the art. By the 1970s, multiple regression had become the dominant tool, as it remains, and political methodology turned more and more to econometrics for its basic tools (and to econometricians for basic training). (Overviews of the political methodology subfield have been given in Achen 1983 and Bartels and Brady 1993.)

Whereas political methodologists have always been a diverse lot, the modal political methodologist looked a lot like an applied econometrician (and had received advanced training by taking the econometrics sequence at top departments). The ease of use of modern software has brought a decreasing lag between the introduction of a new method in econometrics and its use in political science. Thus political science articles now routinely have citations to state-of-the-art econometrics papers in such areas as *time series analysis*, *analysis of choice* (and other limited dependent-variable methods) and *analysis of duration data*.

Although it might be tempting to belittle political methodology in the 1970s and 1980s as a "borrowing" disci-

Nathaniel L. Beck is Professor, Department of Political Science, University of California, San Diego, La Jolla, CA 92093 (E-mail: beck@ucsd.edu) and Editor of *Political Analysis*. The author thanks Henry Brady and Gary King for helpful comments.