METHODS FOR TEMPORAL ANALYSIS[1]

Technical Report #68

Michael T. Hannan Nancy Brandon Tuma

October 1978

Laboratory for Social Research
Stanford University

[1] Work on this paper was supported by grants from the National Institute of Education (NIE-G-76-0082) and the National Science Foundation (SOC 78-12315). Hannan was also supported by National Science Foundation grant BNS76-22943 to the Center for the Advanced Study in the Behavioral Sciences.

Contents

INTRODUCTION

Most sociological research still relies on cross-sectional analysis. Nonetheless, the field has a long history of interest in temporal analysis. Much of the traditional interest derives from the concern that causal inferences cannot be made dependably from a cross-section, because one cannot show that a variable affects change in another. This concern was frequently accompanied by exaggerated claims for the power of temporal analysis. The older literature abounds with claims that temporal designs are always superior to cross-sections. We have since realized that cross-sections give sounder results if confounding influences vary more over time than over units. As a result of this knowledge, a much more tempered view on the methodological value of temporal analysis currently pervades sociology.

Current enthusiasm for temporal analysis stems more from substantive concerns than from methodological prejudice. Macrosociology has begun to reorient to issues of structural change. Likewise the study of individual development and careers has loomed progressively larger in microanalysis. Sociologists of many stripes have come to emphasize change; temporal analysis is indispensible for the study of change, whatever its other benefits.

There are at least two literatures on temporal analysis, one dealing with discrete outcomes, the other with quantitative outcomes. Ideas and developments in one area diffuse slowly into the other. At present, progress on specifying the probabilistic mechanisms has been greater in the study of discrete outcomes; explicit stochastic models underlie many sociological studies of change in qualitative variables. Studies of changes in quantitative variables evidence an ad hoc approach to underlying stochastic mechanisms but a more systematic treatment of causal effects.

We review major perspectives on studies of change in both discrete and quantitative outcomes. We consider basic design issues as well as a variety of technical issues concerning estimation and testing. Much of the technical literature on this subject can be found outside sociology--in statistics, biometrics, econometrics, engineering, etc. We do not pretend to survey any of these fields. Rather we emphasize methods actually used in sociological research. We mention developments in allied fields when they have some obvious bearing on current research practice in sociology.

TYPES OF DESIGNS

Sociological methodology has recently favored treatments of estimation and testing rather than design. While some design issues may be sufficiently well understood that such an emphasis is appropriate, this is not the case in temporal analysis. Thus we begin by reviewing the major alternatives in the design of temporal analysis.

Qualitative Outcomes

Studies of changes in qualitative variables typically take one of four forms: panel, event-count, event-sequence, or event-history designs. Sociologists have relied mainly on panel designs which record state occupancy of a sample of units at two or more points in time. Lazarsfeld, Berelson & Gaudet's (1944) voting study is the prototype: individuals in a sample disclose their voting intentions in a sequence of surveys preceding an election. In studies of changes in cognitive and affective states, panel surveys appear to be the only alternative. However, when interest focuses on changes in state whose timing may be recalled accurately, panel data may be gathered retrospectively. The classic example is analysis of social mobility that is based on information on current occupation and on occupation at some earlier time (first job, father's job when respondent is 16 years of age, etc.)

If accuracy of recall is sufficiently high, retrospective panel designs compare favorably to designs that record outcomes contemporaneously. But they differ greatly in one respect: the sampling process. A current panel selects a sample or population and follows members forward in time; a retrospective panel selects a sample and works backwards in time. As Duncan (1966) has shown for mobility analysis, a retrospective panel systematically misrepresents earlier populations. Men from earlier generations who did not father sons or whose sons died or emigrated are not represented in a retrospective father-son mobility table. The retrospective panel yields censored samples of earlier populations. One way around the problem, as Duncan has noted, is to consider the father-son table a characterization of the status origins of those interviewed at the second "wave." But the problem is not so easily avoided if one retains an interest in the process of change.

An event-count design fills some of the gaps in the panel design: it records the number of different types of events in an interval. When a unit can be in only two states (e.g., married or not married), it records simply the number of times each state is left (e.g., the number of marriages and marital dissolutions) in a period. When there are several states (e.g., 1=employed, 2=unemployed, and 3=out-of-the-labor force), an event-count design may record the number of episodes (or spells) in each state for each unit. Still more usefully, it may give the number of transitions between pairs of states (e.g., changes from 1 to 2 may be distinguished from changes from 1 to 3). Event-count designs are comparatively rare in sociology, except for counts of a single kind of event, e.g., riots, lynchings, hospitalizations, etc. Methods specifically developed for analysis of event counts are still rarer, and our discussion below touches only briefly on methods for this design. Sociological methodology is ripe for a study of what can be learned about change processes from an event-count design as compared to either the traditional panel design or designs that supply even more information on temporal ordering.

An event-sequence design records the sequences of states occupied by each unit. It can be viewed as an elaboration of the event-count design. Suppose the possible states are 1, 2 and 3, as above. A unit's record might be (2, 1, 3, 2) for some period of time. Singer (1977) argues that an event-sequence design provides the minimal necessary information for studying careers and shows that this design improves considerably on the more common panel design. This type of design is far from new in sociology (see, e.g., Form & Miller 1949), but interest in it has only recently reawakened (see, e.g., Spilerman 1977 and Hogan 1978). We do not review literature on this design in a separate section as it is customary to analyze event sequences using techniques for panel analysis; this approach assumes that the timing of events is irrelevant.

An event-history (or sample path) design fills in the remaining gaps: it records the timing of all moves in a sequence. Many laboratory studies of small group interaction provide event-history data. Due to the opportunity to observe a group continuously, experimenters may record the timing of transitions among structural types, etc. In nonexperimental studies, event histories are necessarily retrospective. Nonetheless, they may differ markedly in the length of the recall period. The Johns

Hopkins occupational history study (Coleman et al. 1972) records dates of all job entries and exits in respondents' careers. The Seattle-Denver Income Maintenance Experiment obtains such information as well. But since families are interviewed three times a year over the study period, respondents need to recall their event histories for only four month periods (Robins & Tuma 1977).

Perhaps the most widespread application of the event-history design is in archival research. For example, C. Tilly's (see references below) pioneering study of trends in collective violence in small French political units records the dates of all events of collective violence greater than some minimal scope. The fact that Tilly typically aggregates over units (to the nation) and over time (to the year) in his analysis should not obscure the fact that the design itself records event histories to a population of small areal units. Numerous other studies of collective violence have adopted a similar design.

The four types of design are ordered in the extent of detail acquired on the process of change. Sociologists show a very strong preference for the simplest, the panel design. In some situations the panel is the only feasible temporal design. However, sociologists often forego opportunities to collect and use data on sequences and timing of events. We suspect that this tendency reflects uncertainty regarding the value of such information. Thus it is important to consider whether designs containing information on sequences and timing of events confer any important advantages.

If we are to make systematic comparisons among designs, we must be clear about the timing of measurements in panel studies. Does the measurement interval reflect some fundamental periodicity in the process under study? If so, we cannot easily compare the various designs. If, however, the timing of measurements is largely arbitrary and events may occur at any time, the appropriate mathematical specification of the process generating the data is that of a continuous-time discrete-state stochastic process. The Markov process, introduced to sociologists by Coleman (1964a), provides an important baseline stochastic process of this type.

The designs differ in their ability to discriminate among classes of continuous-time stochastic models. The classic two-wave panel design is very weak in terms of its ability to reject classes of models (Singer & Spilerman 1976a). One may test only for time-homogeneity, i.e., one can use data to accept or reject the class of models with stationary transition probabilities. A third wave of observations permits a test of the Markov property; but it does not permit, for example, distinguishing between Markov and semi-Markov processes. However, data on event-counts and event-sequences permit stronger inferences, and event-history data solve completely the so-called embedding problem (Singer 1977; Tuma, Hannan & Groeneveld 1979). That is, information on the timing of events together with event-sequences makes it possible to test for very narrow classes of models. These analytic results tell a very important lesson in design: whenever possible we should collect data on the sequences of moves and the timing of moves.

Quantitative Outcomes

Some metric outcomes change rapidly relative to our ability to measure them, e.g., size of large organizations, hours of work of individuals. Other quantitative outcomes change levels infrequently, e.g., prestige or wage rates associated with a job. For the latter, event-history designs that record both the dates of jumps and the sizes of the jumps are appropriate. In mathematical terms, the underlying stochastic process is a jump process in which one set of parameters governs holding times in states and another set controls average height of jumps (see Çinlar 1975: 90-94 for a brief discussion). Both sets of parameters may be treated as functions of exogenous variables. Though this framework appears natural for much sociological research, we are not aware of any sociological applications.

When sociologists study changes in metric variables, they typically rely on intermittant observations. This is the only feasible design for rapidly changing outcomes. We typically distinguish three such designs: A time series design records the level of the outcome at many dates for one unit. The term panel design refers to a collection of short time series (as few as two time points) on a number of units. If longer time series are available on several units, the design is called a multiple time series design.

Panel designs have been used in the study of individual social psychology (e.g., Kohn & Schooler 1978), status attainment (Kelley 1973), organizational structure and demography (Meyer 1975), and change in national social structure (Chase-Dunn 1975).

Time-series designs have been employed largely in macrosociological research. Examples include studies of levels of collective violence (Snyder & Tilly 1972), changes in voting patterns (Doreian & Hummon 1976), suicide rates (Vigderhous 1977), and studies of variations over time in labor organization and activity (Shorter & Tilly 1970). Although efforts have begun to contrast time series for different systems (e.g., Tilly, Tilly & Tilly's (1975) comparisons of rates of violent protest in France, Germany and Italy for 1830-1930), sociologists have not fully exploited multiple time-series designs.

The sociological literature contains little guidance on the choice between panel and time-series designs. If we include all the relevant causal variables and specify the proper form of the model, replications of the process over time are just as useful as replication over units. So in practice the choice between designs hinges on judgements about confounding factors. If the confounding factors are likely to vary over time but not among units at a point in time (e.g., prices in world markets), the panel design has the edge. If the confounding factors are likely to vary more across units but not over time (e.g., national culture), the time-series design has the edge.

To this point we have focused on the broadest features of designs for temporal analysis. We turn now to consideration of the details of the various strategies, discussing strengths and weaknesses of alternative approaches to modeling and estimation. We begin with issues of in the study of changes in qualitative outcomes.

EVENT-HISTORY ANALYSIS

Strategies

Three main strategies for analyzing event-history data have been used and/or discussed in sociological research. The first strategy--by far the most common--neglects some information in event histories and analyzes the data as if they were generated by some other design. Palmer's (1954) Labor Mobility in Six Cities provides a good illustration of the many outcomes that can be obtained from event histories. The data consist of work histories for the years 1940-1950 for roughly 13,000 people. Some of Palmer's findings could have been collected by a series of cross-sections (e.g., the distribution of employment status for a series of years) or by a panel (e.g., occupational status in 1950 by status in 1940). She also reports event counts (e.g., number of jobs held) in different periods. Although the range of outcomes reported is impressive, her analysis does not make clear what (if anything) was gained by the event-history design that could not have been learned by another design.

More recent analyses of event-history data have also tended to use only part of the information in event-history data. They have tended to rely on a smaller range of outcomes than Palmer, but have controlled for a larger number of variables, primarily through multivariate techniques. Ordinarily information on the dates of events is used only to compute counts of events in some period. Then these counts are analyzed as a metric variable measured either at one "time" (i.e., in one period) or at a series of "times." In short, event-history data are treated as event counts.

For example, Inverarity (1976) obtains the total number of lynchings in a period from newspaper reports on the dates of lynchings. Then he analyzes this variable through a multiple indicator, multiple cause model using a procedure developed by Jöreskog (1970). The analysis is indistinguishable from that usually performed on cross-sectional data. Similarly, Snyder and Tilly (1972) compute the count of annual collective disturbances in France from archival information on dates of violent outbreaks. Unlike Inverarity, they then use time-series analysis to investigate the relation of these counts to other time-varying characteristics of France. Similarly, Spilerman (1970) obtains the number of riots per city in different time periods from archival reports on riot dates. He not only analyzes these counts by linear regression (as in the usual cross-sectional approach) but also considers whether they could have been generated by various stochastic processes (e.g., Poisson, time-dependent Poisson, etc.). Eaton (1974) fits Poisson and negative binomial distributions to event counts taken from event histories of admissions to mental hospitals.

The second and third strategies use the information in event histories on the timing and sequence of events, as well as information on the number of events. These strategies resemble one another in assuming that a stochastic (i.e., probabilistic) process generates events and that events may occur continuously in time. (Changes that can only occur at discrete time intervals are regarded as a special case.) The two strategies differ in their additional assumptions and in the questions they ask of the data.

The exploratory strategy avoids making any additional assumptions about the process. Instead, it asks what classes of stochastic processes might have generated the data and what classes are unlikely to have generated them. Its goal is to reject types of models, i.e., to narrow the class of possible models rather than to accept any particular model. For example, after appropriate analysis, we might be able to conclude that the data are inconsistent with models in which the probability of an event per unit of time <u>increases</u> with the length of time since the last event (where an event could be, for example, a job change). We might still be unable to tell whether the probability of an event per unit of time decreases with the length of this interval, or whether it is constant over time but varies from one member of the population to another. Methods for implementing this strategy are still in a primitive state; see Singer (1977) and Singer & Spilerman (1976b) for preliminary ideas on this strategy.

The third strategy, a model-testing approach, begins by assuming some simple stochastic process, estimates its parameters, and then tests whether some of its implications fit the data. More complicated models are introduced either to test an argument or to improve fit. This strategy resembles the one used by most sociologists in analyzing cross-sectional data; it mainly differs in the kinds of models that are assumed.

A comparatively simple stochastic model often assumed to describe change in qualitative outcomes is a first-order, discrete-state, continuous-time Markov process, which includes the familiar Poisson model for the number of events in a period and the general birth-and-death model as special cases. The (simple) Markov model has been applied to a wide variety of phenomena: labor mobility (e.g., Blumen, Kogan & McCarthy 1955), changes in attitudes (e.g., Coleman 1964a), changes in friendship networks (e.g., Sørensen & Hallinan 1977), marital stability (e.g., Hannan, Tuma & Groeneveld 1977), outbreaks of collective violence (e.g., Spilerman 1970), etc.

Unfortunately the simple Markov model rarely fits sociological data well. This lack of fit has motivated various revisions and extensions of the model. It is convenient to distinguish among three types: (1) those focusing on reconceptualizing the process being studied in terms of "latent states," (2) those assuming the population studied is heterogeneous, and (3) those postulating time-dependence in the process.

Extensions

LATENT STATES In typical applications of Markov models, observed outcomes are assumed to be identical to the states of the Markov process. So, for example, if the data tell only that people hold a job or not, the states are assumed to be "holding a job" and "not holding a job." An improved conceptualization can sometimes make the application of the simple Markov model more appropriate. For example, observed states may be assumed to be related to unobserved (latent) states in some specified way. If change on the latent states is indeed Markovian but the observed and latent states are not perfectly correlated, then observed changes are generally not describable by the simple Markov model. We consider three cases.

First, suppose each observed state is composed of several unobserved states, and movement among the latent states is Markovian. Since each observed state is associated with two or more unobserved states, observed changes will not be Markovian. But an extended model may retain the stationary Markov framework and still fit the data. For example, Herbst (1963) proposed a model of interfirm mobility in which "belonging to a firm" (what the data recorded) consists of four states: undecided, temporarily committed, permanently committed and decided to leave. Mayer (1972) proposed a similar kind of model in which the data record occupational categories, but each category is composed of two latent states, one that can be left (analogous to Herbst's temporary commitment) and one that cannot (analogous to Herbst's permanent commitment).

Second, suppose true states correspond to probabilities of making an observable response, and change from one probability to another is Markovian. This is the basic idea underlying Coleman's (1964b) Models of Change and Response Uncertainty. Again, change in observed responses is not Markovian, even though the latent process is. This ingenious formulation has not been widely applied, perhaps because of its mathematical complexity. Wiggins (1973) elaborates on Coleman's (1964b) discussion.

Third, suppose change is Markovian but the true state for each episode is not always recorded accurately. If the error structure can be described, then observed changes can be expressed as a function of the true underlying Markovian process. To our knowledge this conceptualization has not yet been applied in sociological research. We mention it because it resembles the errors-in-measurement models discussed in the literature on linear models of quantitative variables.

POPULATION HOMOGENEITY Population heterogeneity has been introduced in two main ways. One approach assumes that the fundamental parameters of the Markov model have some postulated probability distribution with unknown parameters. For example, in their study of industrial mobility, Blumen, Kogan & McCarthy (1955) postulated that there are two kinds of people, movers and stayers. In effect, they assume a Bernoulli distribution on the parameters of the Markov process: a fraction, p, of the population move according to a Markov model and the rest, (1-p), do not move at all. Spilerman (1972b) and Singer & Spilerman (1974) assumed that the rate of leaving a state has a gamma probability distribution but that the conditional probability of each move is the same for everyone in the population. This way of introducing heterogeneity into Markov models has a major disadvantage. It does not permit the investigator to make inferences about the determinants of changes in qualitative outcomes.

The alternative approach assumes that the fundamental parameters of the Markov process--the instantaneous rates of change from one state to another--depend on observable variables in some specified way. Below we discuss Coleman's (1964a) approach to the study of causal effects on rates from panel data. He also proposed an extension in which rates of change are linear functions of exogenous variables, and Tuma (1976) estimated such a model. The assumption that transition rates are linear in observables can lead to a mathematically impossible situation--namely, that transition rates are negative. It seems to be both mathematically and empirically more satisfactory to assume that transition rates are log-linear functions of exogenous variables. This approach was also suggested by Coleman (1973), and it has been applied by Hannan, Tuma & Groeneveld (1977) to the study of marital stability.

TIME-STATIONARITY According to the social process being studied, authors have suggested that parameters of the Markov model depend on age (e.g., Mayer 1972), duration in a state (e.g., McGinnis 1968, Tuma 1976), experience (e.g., Sørensen 1975), and/or experimental time (e.g., Tuma, Hannan & Groeneveld 1979). The most common approach assumes that the fundamental parameters are a specific function of time, e.g., exponentially declining over time (e.g., Mayer 1972; Sørensen 1975; Sørensen & Tuma 1978). Alternatively one may divide the time axis into periods and assume that parameters are constant within periods but vary among periods (e.g., Tuma, Hannan & Groeneveld 1979). The parametric approach usually requires that fewer additional parameters be estimated. However, the nonparametric approach can be useful when little is known about the form of time-dependence.

Estimation

There are a variety of ways of using event-history data to estimate parameters in continuous-time models of change in qualitative outcomes. We consider three: moment estimation, maximum likelihood estimation and partial likelihood estimation.

MOMENT ESTIMATION Moment (M) estimation is based on equating observed sample moments (e.g., means and variances) with their expected value in the population when the postulated model is true. This approach is advocated by Coleman (1964a) for estimating parameters in the simple Markov model and by Sørensen (1977) for estimating parameters of a Poisson process from censored event-history data. Event histories are said to be censored when some events are unobserved because of some feature of the data collection procedure. For example, when retrospective life histories are collected in a survey (e.g., Coleman, Blum, Sørensen & Rossi 1972), events occurring after the interview are not recorded. The data may tell when respondents began their current jobs, but not when they will leave them. Sørensen (1977) shows that ignoring censoring gives biased estimates and proposes

M-estimators that take censoring into account.

The main advantage of M-estimators is that they can sometimes be obtained when other estimators cannot be derived or are very difficult to implement. The main disadvantage of M-estimators is that they rarely have optimal statistical properties, even in large samples. For example, Tuma & Hannan (1978) show that one of Sørensen's M-estimators that is not also maximum likelihood (ML) performs poorly compared to ML-estimators.

MAXIMUM LIKELIHOOD ESTIMATION Maximum likelihood (ML) estimators for the continuous-time, discrete-state Markov model seems to have been discussed first by biometricians (Boag 1949) and statisticians (Albert 1962). Tuma (1976) applied ML estimation to the case in which parameters depend on exogenous observables and duration in a state. Tuma & Hannan's (1978) Monte Carlo experiments show that ML-estimators based on event-history data have good properties (small bias and variance) even when sample are moderate in size and a high proportion of episodes have not yet ended (i.e., are censored). Tuma, Hannan & Groeneveld (1979) give a detailed discussion of the use of ML-estimation in event-history analysis and discuss advantages of the event-history design over panel and event-count designs.

The main advantage of ML-estimation of event histories is that it yields estimators with good properties as long as the data are generated by the postulated stochastic process. However, there is no guarantee that ML-estimators retain their good properties when the assumptions of the model are violated. That is, ML estimators <u>may</u> not be robust.

PARTIAL LIKELIHOOD ESTIMATION Partial likelihood (PL) estimation was proposed by Cox (1972) to estimate effects of exogenous variables on transition rates from event-history data when one does not know how these rates vary over time. Cox assumed that the instantaneous rate of an event (also called the hazard function), r, is:

$$r(t,x) = h(t)\exp(bx)$$

where h(t) is an unknown function of time t, x is a vector of observed exogenous variables, and b is a vector of parameters to be estimated. The likelihood function for this model is the product of three terms. Two terms depend on the unknown h(t); the last, which Cox called the partial likelihood, depends only on exp(bx) and the time ordering of events in the sample. Without specifying h(t) we cannot write the whole likelihood. Cox showed that treating the partial likelihood as though it were the whole likelihood gives consistent estimators of the b's. Efron (1977) proved that under fairly general conditions the PL-estimators of the b's are asympototically normal and maximally efficient. PL-estimation has been used to estimate effects of variables on mortality rates of heart transplant patients (Miller 1976). A sociological application has not yet been published, to the best of our knowledge. For a brief review of the statistical literature on PL-estimation, see Tuma & Hannan (1978).

The main advantage of PL-estimation is that it requires weaker assumption than ML-estimation, but still yields estimators with good statistical properties. For this reason it has generated considerable interest among statisticians. One disadvantage for investigators wishing to predict future events is that PL-estimation does not identify "the constant term." That is, though it estimates effects of variables on the rate, it does not estimate the rate. It is analogous to being able to estimate slopes but not the intercept in linear regression analysis.

PANEL ANALYSIS OF QUALITATIVE OUTCOMES

Lazarsfeld (1948) appears to have been the first sociologist to have proposed panel analysis of qualitative variables. He noted that much data studied by sociologists concerns an association between two variables X and Y. Sociologists want to know whether X induces change in Y or Y induces change in X. Observations on X and Y at a single point in time cannot tell this. Lazarsfeld suggested measuring X and Y at two times, t_0 and t_1. If X and Y are dichotomous, then at any time there are four possible response patterns. Arraying responses at time 0 by those at time 1 gives the famous 16-fold table. How should one analyze such a table (or one like it but with more waves, more variables, or more possible responses for each variable) to determine the extent to which change in one variable affects another?

Sociologists have used several approaches. One treats panel data on K qualitative variables at T points in time as a problem in analyzing a contingency table with KT variables. Another applies ordinary linear regression analysis, treating a change between successive waves as a dichotomous dependent variable. Both of these strategies implicitly assume that changes occur at discrete points in time or that the timing of changes is irrelevant to answering questions concerning the determinants of change. Another strategy assumes that changes can occur continuously in time, even though data happen to be recorded at discrete times.

The Contingency Table Strategy

Contingency table analysis has flourished within the past decade. Various authors, especially Goodman (1972a, 1972b, 1973), have developed a set of powerful methods for estimating and testing log-linear models of the entries in a contingency table. These models can be used for any number of variables and number of discrete categories per variable. We do not attempt to summarize the main features of these models because there are a variety of clear (e.g., Davis 1974) and comprehensive (e.g., Bishop, Fienberg & Holland 1975; Haberman 1974) expositions of them, and because by now they are rather well known to sociologists.

These techniques can be viewed as natural extensions of Lazarsfeld's earlier work on panel analysis of qualitative outcomes. Goodman (1973) discusses and illustrates application of these models and methods to analysis of panel data. A variety of other sociological applications to panel data have followed. One, by Hauser et al. (1975) on temporal change in occupational mobility, contains an especially clear statement of the model and a good illustration of how to interpret results based on it. For an application of this specification to parameterize age, period, and cohort effects, see Pullum (1977).

The advantages of this approach are the wide range of substantively interesting questions for which it provides an answer and the comparative ease with which it can be used. One disadvantage is that all variables included in the analysis must be changed into qualitative variables. An added disadvantage, partly arising from the total reliance on polytomous variables, is the practical problem of finding a sufficiently large sample to fill all cells of the contingency table. This is especially troublesome when a large number of variables must be considered. Another possible disadvantage concerns the value of these methods in situations in which the outcomes being studied can change continuously in time, as discussed in more detail below.

Regression Strategy

The regression strategy treats a change between two waves as a dichotomous dependent variable in a regression on a set of independent variables. Sociologists usually assume the regression is linear in the independent variables, but nonlinear approaches (see below) are often used in other fields.

Spilerman (1972a) suggests this strategy as a way to incorporate independent variables into a Markov model. Duncan & Perrucci (1976) take this approach in studying whether or not couples have migrated between two waves of a panel. Bumpass & Sweet (1972) use this method to investigate effects of causal variables on marital dissolution.

This strategy has several advantages and at least as many (if not more) disadvantages. Its main advantages are ease of application and comparatively low cost. In addition, unlike the log-linear models discussed under the contingency table strategy, a regression approach allows both quantitative and qualitative independent variables to be included in the analysis. Consequently, the "empty-cell" problem mentioned under the contingency table strategy is not likely to occur unless a great many interaction terms are included.

Some of the disadvantages of this strategy result from assuming that a dichotomous dependent variable is linear in the independent variables. These disadvantages include heteroscedasticity of disturbances, inefficiency of ordinary least squares estimates, and the possibility that predicted probabilities of a change lie outside the (0-1) range (Goldberger 1964). Various nonlinear regression methods, e.g., multivariate probit analysis and multivariate logit analysis, overcome these deficiencies of the linear model.

A potentially more disturbing disadvantage of the regression approach --one shared by the contingency table approach--arises from the fact that they ignore the timing of changes. Both approaches implicitly assume that the timing of changes is irrelevant to identification of the true underlying structure generating change. Timing is, indeed, irrelevant if changes can only occur at the times of the waves of the panel. This can happen when change occurs at discrete intervals, and the investigator knows the true lag and can arrange to collect data at this interval. But usually it is false, either because the lag is unknown or because changes can occur continuously in time.

Little is known about the consequences of applying either regression or contingency table strategies to panel analysis when the assumption mentioned above is false. Tuma (1973) has noted that the effects of independent variables vary both in magnitude and in statistical significance as the length of the time period varied in linear regression analysis of job changes. Singer & Spilerman (1976a,b) discuss a more fundamental problem. As we discuss below, identification of structural parameters in continuous-time models of change in qualitative outcomes is problematic with panel data. Moreover, these problems cannot be evaded by treating the underlying processes as occurring at discrete intervals. These disturbing conclusions give added force to suggestions that investigators collect as detailed information about change in the qualitative outcome being studied as feasible. Recognition of these problems has also promoted a renewed interest in panel analysis of qualitative outcomes using a strategy based on continuous-time models.

Continuous-Time Strategies

Coleman (1964a) is the first sociologist to have argued persuasively for basing panel analysis of qualitative outcomes on the assumption of an underlying stochastic process in which changes may occur *continuously* in time. His elaborations of this strategy are often based on the discrete-state, continuous-time Markov model discussed above. As already mentioned, the simple Markov model rarely fits data well, and various improvements have been proposed to remedy this. Coleman (1964a,b) has contributed many ideas for doing this, and his suggestions are often quite mathematically sophisticated. However, his empirical applications usually involve comparatively simple situations, e.g., two waves of observations on two endogenous dichotomous variables or on one dichotomous dependent variable and one dichotomous exogenous variable. Even models describing these rather simple interrelationships give estimation equations that are not trivial to implement. Other sociologists (e.g., Mayer, 1972) have also constructed continuous-time stochastic models with greater realism than the simple Markov model, but have not been able to estimate parameters from panel data in a satisfactory way.

In the past few years Singer & Spilerman (1974, 1976a,b) have begun to clarify what can be learned from panel data when the outcome of interest is generated by continuous-time stochastic process. These authors have not been concerned with estimating parameters in any particular model. Instead they have emphasized the development of tests for choosing among broad classes of models (compare the second strategy discussed under event-history analysis). Among their findings are the following.

First, observations on the proportion of transitions among states of the qualitative outcome being studied, which gives an estimate of the matrix of transition probabilities, cannot always be embedded in (described by) a (simple) Markov process. Moreover, sampling error can sometimes cause panel data to be unembeddable, even though they are actually generated by a Markov process. Second, even if the data are embeddable in a Markov process, there may not be a unique set of parameters that could have generated the data. Singer & Spilerman (1976a) detail a procedure for finding an exhaustive set of possibilities, but sometimes the final choice must be made on substantive grounds. Third, small changes in an observed matrix of transition probabilities (which can occur because of sampling variability) can lead to a quite different set of possible processes. A number of design features can reduce these problems, e.g., multiple waves with irregular spacing, shorter intervals between waves, etc. In short, the more closely panel data resemble event-history data, the fewer the problems in analysis.

Thus, in spite of this recent research, it is still the case that panel analysis of qualitative outcomes is a methodological mine field--if changes can occur continuously in time. While mathematical and statistical invention may clarify what we can learn from a panel design, we will not be able to answer all the questions that sociologists like to ask.

PANEL ANALYSIS OF QUANTITATIVE OUTCOMES

Strategies

The two-wave panel has also become a standard tool for the study of change in metric variables. But the problem of casting substantive arguments in operational terms within this framework is far from settled.

Researchers choose panel designs for diverse reasons; consequently there is no single methodology of panel analysis. We find three broad approaches to panel analysis in the sociological literature.

The first strategy follows Lazarsfeld (1948) in seeking an approximation to experimental design. Lazarsfeld argued that one could approximate the study of experimentally-induced changes by isolating certain classes of changes in a turnover table (such as the 16-fold table). According to this view the panel design is a special tool for detecting causal effects. The goal is to choose between two competing hypotheses: X causes Y, or Y causes X. This perspective has been taken over literally into the study of changes in quantitative variables by Campbell (1963) and Pelz & Andrews (1964). They reasoned that one might use cross-correlations (correlation of X_0 with Y_1 and Y_0 with X_1, where subscripts denote the time period of measurement) to choose between the two competing hypotheses. If $\rho_{X_0Y_1} > \rho_{Y_0X_1}$, then choose the hypothesis "X causes Y", etc.

The defects in this inference rule soon became apparent, and the procedure was recast in terms of partial cross-correlations $\rho_{X_0Y_1 \cdot Y_0}$ and $\rho_{Y_0X_1 \cdot X_0}$. Otherwise, the logic remained the same. This has become a standard procedure for choosing among rival explanations in psychological research (see, for example, Crano, Kenny & Campbell 1972).

Kenny (1973, 1975) has explicated the logic of this procedure as a "test for spuriousness." He actually specifies a particular covariance structure among unmeasured X's and Y's and their measured values and argues that cross-lag correlation tests correspond to certain meaningful restrictions on the covariance structure. In particular, if the covariance structure does not contain "causal effects" relating X and Y, and if a number of other strong conditions hold (such as constant variances of latent and measured variables over time), the cross-lag partial correlations will be zero on average.

The "test for spuriousness" depends on a particular specification of the covariance structure--in short, on a model. Moreover, some of Kenny's conditions appear not to hold in many situations, e.g., X and Y often have very different stabilities over time. In many reasonable situations, cross-lag correlation tests give exactly the wrong answer, i.e. suggest that X causes Y when the reverse is true (Rogosa 1978a).

Many difficulties that beset cross-lag correlation analysis can be traced to the main question: does X cause Y <u>or</u> Y cause X? Though the question admits the possibility that neither effect exists, it does not anticipate that both effects may hold.

The structural equation approach to panel analysis permits systematic treatment of more general questions. Instead of viewing panel designs as a special tool for testing, it focuses on estimating parameters of the joint distribution of variables measured at two or more points in time. The sociological literature shows that one may form simple models that embody the various alternative causal structures relating X and Y (Duncan 1969; Heise 1970). The panel design may thus be treated as a special case of the usual nonexperimental cross-sectional design. Then, as Goldberger (1971) argued, there is no need for any special estimation and testing theory for panel analysis. Standard and widely available methods for structural-equation analysis apply.

The view that panel analysis has been subsumed as a special case of structural-equation methods seems to be widely held in sociology. However, a third view contends this claim. This perspective, advocated by

Coleman (1964a, 1968), follows Lazarsfeld in emphasizing change. But it agrees with the structural-equation perspective that inferences concerning change cannot be model-free. It argues that explicit dynamic models are needed if panel analysis is to yield meaningful substantive results. In one sense, the usual structural-equation models for panel analysis fit these criteria, since the equations may be considered stochastic difference equations. But, if as we argued earlier, most social processes do not have fixed lag structures and may change at any instant, the proper specification is a continuous-time process. The structural relations are expressed as time-differential equations. The usual panel regressions can then be viewed as particular forms of the solution of the equations of the process, i.e., as integral equations. The relation between integral equations and panel regressions permits use of data with discrete spacing to estimate the parameters of a process changing continuously in time. We argue below that this perspective has considerable advantages. However, to date this approach has been used only sparingly in sociological research (for example, see Freeman & Hannan 1975; Hummon, Doreian & Teuter 1975; Doreian & Hummon 1976; Sørensen & Hallinan 1977 and Hannan & Freeman 1978).

Estimation

The recent sociological literature contains treatments of special complications that arise in the various approaches to panel analysis. In some cases, these developments tell cautionary tales, in others they suggest alternative estimation strategies.

Duncan (1969) raised a fundamental objection to the then widely held view that panel analysis offers a "free lunch", namely that it obviates the need to use a model in making inferences. He considers a two-wave, two variable (2W2V) panel design and supposes that the analyst assumes that relations are linear-additive but wishes to remain agnostic concerning the direction of causation. The most general linear-additive model then applies by default:

$$Y_1 = \alpha_0 + \alpha_1 Y_0 + \alpha_2 X_0 + \alpha_3 X_1 + u \tag{1a}$$

$$X_1 = \beta_0 + \beta_1 X_0 + \beta_2 Y_0 + \beta_3 Y_1 + v \tag{1b}$$

Note that this model contains both lagged and instantaneous effects. It is easy to show that the number of parameters to be estimated exceeds the number of covariances available with which to estimate them in a 2W2V design; none of the parameters are identified. Since the parameters may not be estimated uniquely from data, no numerical calculations tell us anything about the causal structure.

Sociological researchers rarely estimate models like (1). Instead they typically use models with <u>only</u> lagged effects such as:

$$Y_1 = \alpha_0 + \alpha_1 Y_0 + \alpha_2 X_0 + u' \tag{2a}$$

$$X_1 = \beta_0 + \beta_1 X_0 + \beta_2 Y_0 + v' \tag{2b}$$

As long as the disturbances are uncorrelated with the regressors (as can happen if there is no instantaneous reciprocal causation), all parameters of (2) may be identified in a 2W2V design. Of course, the identifying restrictions may be wrong; there may be causal effects with lags shorter than the lag built into the design. If so, we will not have improved matters by using the restricted model with only lagged effects.

So identification, the fundamental issue in panel analysis, turns on the problem of using the "right" lag structure. Heise (1970) discusses some consequences of using the wrong lag. The problem of course is that we rarely if ever have enough information about the detailed structure of a process to specify the true lag exactly (Davis 1978). As long as we focus on discrete-time processes, lack of such knowledge is a massive obstacle to analysis.

A major advantage of the continuous-time specification is that it makes the timing between waves irrelevant (Coleman 1968). Thus, for at least the class of linear differential equation models, the identification problem that concerns Heise (1970) and Davis (1978) does not arise. Consider the following simple case. Let the rate of change in both X and Y depend linearly on X and Y:

$$dY(t)/dt = a_0 + a_1Y(t) + a_2X(t) \tag{3a}$$

$$dX(t)/dt = b_0 + b_1X(t) + b_2Y(t) \tag{3b}$$

The integral equations corresponding to this system, subject to initial conditions $X(0) = X_0$ and $Y(0) = Y_0$, have the form:

$$Y(t) = \gamma_0 + \gamma_1Y_0 + \gamma_2X_0 \tag{4a}$$

$$X(t) = \delta_0 + \delta_1X_0 + \delta_2Y_0 \tag{4b}$$

where the γ's and δ's are complex functions of the parameters of the system (3) and of elapsed time between t_0 and t. Inspection of these functions shows that the spacing of observations is taken into account in a perfectly natural way. Moreover, this feature permits systematic comparison of estimates from studies with different lags. Thus the continuous-time perspective solves two of the major practical difficulties in conventional quantitative panel analysis: choosing a lag and comparing findings from analyses with different lags.

Identification issues aside, the most troublesome feature of quantitative panel analyses concerns the specification of the omitted factors, whose effects are summarized in a disturbance term. The usual practice of applying ordinary least squares (OLS) estimators to models such as (2) implies that errors are uncorrelated over time. But if these factors are stable over time, i.e., autocorrelated, the disturbance term cannot be uncorrelated with the right-hand side variables in the conventional model, (2). Consequently, OLS estimators of the parameters of the conventional two-wave panel model are biased whenever the disturbance is autocorrelated (Johnston 1972). Evidence that autocorrelation bias is large in the designs and research situations favored by sociologists has accumulated rapidly. Thus progress in analysis of sociological panels depends critically on solutions to the problem of autocorrelation.

The main obstacle to such progress has been the heavy reliance on the two-wave panel with single measurements of each variable. Recent work shows that reasonably satisfactory solutions to the problem can be achieved by either increasing the number of waves of observations or by using multiple measures of each variable. In each case, one obtains information sufficient both to estimate structural parameters and to adjust for some types of autocorrelation. Each development requires moving beyond ordinary least squares estimators, as we discuss below.

The use of multiple measures of latent variables in panel designs first attracted attention in sociology as a framework within which to cope with measurement error (Blalock 1970; Duncan 1972; Hannan, Rubinson & Warren 1974). This early literature recognized that structural parameters could still be identified in some such models even when measurement errors are autocorrelated. More recent work has shown that disturbances associated with the latent variables may also be autocorrelated without destroying identification if one places sufficient restrictions on the model.

Current work in this tradition focuses on efficient estimation and model testing. The key innovation is Jöreskog's (1970) development of "full information" maximum likelihood procedures for linear structural equation systems. The advantages of this approach are discussed by Jöreskog & Sörbom (1976) and Wheaton, Alwin & Summers (1977). This procedure has been implemented in empirical research by Bielby, Hauser & Featherman (1977), Kohn & Schooler (1978) and Esmer (1979).

An alternative strategy involves pooling waves of a multi-wave panel. The resulting design, called a pooled cross-section and time series design, tacitly assumes that the same structure operates in each pair of adjacent waves. If so, the information in excess of that generated by a two-wave panel can be used to estimate parameters of a postulated autocorrelation process. One promising specification of the autocorrelation process uses the classical variance-components model. It assumes that the disturbance consists of two (or more--see below) unrelated components: one component is truly random; the other is a constant that characterizes the unit of observation (e.g., genetic composition, enduring features of personality, features of constitutional systems, etc.). Under this specification, the disturbances are autocorrelated only because of the unit-specific components. If the latter are considered to be fixed effects, pooled within-unit regressions eliminate autoregression bias (Maddala 1971). If the unit-specific effects are considered random variables drawn from some distribution, one may use generalized least squares estimators that have good large sample properties and reasonably good small sample properties as well (Nerlove 1971; Hannan & Young 1977).

The pooled cross-section and time-series estimators have been extended to deal with further practical complications. Lillard & Willis (1976) have estimated models with fixed individual effects and random disturbances that are themselves autocorrelated (with a first-order autoregressive scheme). Nielsen & Hannan (1977) have used an estimator that accommodates for individual-specific effects and heteroscedasticity of the random component.

It is also straightforward to add period-specific effects as well (Kuh 1959; Balestra & Nerlove 1966). The period effect summarizes the environmental factors that are unique to the measurement period and affect all units alike. These effects may also be considered as fixed factors or as realizations of some stochastic process generating environmental variability. Simple extensions of the fixed effects and generalized least squares estimators apply to these specifications.

The pooled cross-section and time-series design seems a natural framework within which to study age, period, and cohort effects (see

Ryder 1965 for a discussion of the importance of distinguishing these components). It is well known that the three effects cannot be identified in cross-sections. However, as long as one assumes an additive structure, two of the three may be identified in such designs (Mason et al. 1973). In a pooled model, period effects may be estimated without difficulty; however, age and cohort (viewed as an individual-specific effect) may not be distinguished without further restrictions on the model.

One last estimation issue deserves mention. The sociological and economic literatures have pursued different tracks in estimating systems of linear differential equations. The integral equations corresponding to systems contain matrix functions of the form $\exp(\underset{\sim}{B}t)$ where $\underset{\sim}{B}$ is a k by k matrix when the system contains k equations. Sociologists, following Coleman (1968)--but see Kaufman (1976)--use what is known as a spectral decomposition of this matrix function to relate regression estimates to dynamic parameters. But this strategy does not permit use of constraints on elements of $\underset{\sim}{B}$ in estimation. Consequently, estimation is not fully efficient. Econometricians, seeking efficient estimators, have focused on discrete approximations to the differential equation systems that permit the use of constraints on parameters (Bergstrom 1976). It is not yet known whether the approximation errors introduced by this approach compensate for the ability to utilize constraints.

TIME SERIES ANALYSIS

We will only briefly indicate the main lines of development of time series analysis in sociological research. Many of the issues of strategy and estimation parallel those already discussed. Morever, the statistical theory of time series estimation is far more codified than is the case for panel analysis.

Recent time-series literature, especially in economics, has often focused on questions similar to those posed by Lazarsfeld. In an influential paper, Granger (1969) defined direction of causality in terms of predictability in multiple time series. He proposed that one time series, (X_t), causes another, (Y_t), if current values of Y can be predicted from past values of X, partialling for the effects of past values of Y. This conception resembles that underlying cross-lag correlation analysis--with the important exception that Granger explicitly includes the possibility of joint causation. Nonetheless, much has been made of Sims' (1972) use of distributed-lag estimators to determine whether the stock of money causes income variations _or_ vice versa.

It turns out that translating Granger's criteria for causation into two-wave panel format does not give a cross-lag correlation test. Instead, it implies that X causes Y if the structural cross-lag parameter labeled α_2 in equation (2a) is nonzero and that Y causes X if β_2 in (2b) is nonzero (Rogosa 1978b).

Time-series analysis is the standard procedure for estimating continuous-time dynamic models. For examples, see Doreian & Hummon (1976) and Pitcher, Hamblin & Miller (1978). However, the structural-equation perspective, with discrete lags, is more commonly applied to sociological time series. Then the standard econometric literature on time series with its focus on autocorrelation of disturbances applies (see Hibbs 1974 for a review). The econometric literature stress two forms of autocorrelation, autoregressive and moving average processes. Much recent work follows Box & Jenkins (1976) in specifying a very general mixture of the two processes as a model of the noise process. This strategy has swept the field of applied time-series analysis but has barely penetrated sociological research. Hibbs (1977) discusses the potential value of the Box-Jenkins approach to the study of policy interventions when long time series are available, and Vigderhous (1977) has illustrated its value in forecasting social trends. Finally, much theoretical work on time series uses a spectral representation of the series that transforms from a time domain to a frequency domain. The goal is to decompose a long series into components of different frequency just as sound may be so decomposed. One may then wish to smooth high-frequency (or short-period) waves so as to achieve a clearer representation of the longer cycles of the process. Possible sociological applications of this strategy have been discussed by Mayer & Arney (1974).

CONCLUSIONS

The notion that temporal analysis automatically yields conclusive inferences dies hard. However, the thrust of most recent methodological developments has been to argue cogently against this view. We have emphasized that the stock tools of temporal analysis in sociology, the two-wave panel for qualitative and quantitative outcomes, admits multiple interpretations. In the qualitative case, when changes may occur at any time, one cannot identify structural parameters from only two waves of panel data. Event counts, event sequences and event histories permit much finer model testing and should be used more often in sociological research. The identification problem plagues the quantitative case as well. If the model assumes a discrete-time process, one must know the timing of the causal lags. Overall these recent methodological developments reemphasize the importance of substantive theory and models for making good use of temporal data.

The situation is not wholly bleak, however. Sociologists have begun to devote more attention to modeling change processes. We propose that such developments, particularly the use of continuous-time stochastic models of change, will permit a much richer use of temporal data than in past sociological research. Not only will such models enrich sociological analysis, they also focus attention squarely on change processes. They emphasize that temporal data is not just like cross-sectional data, but that it contains information on the manner in which change comes about.

Finally we have commented separately on analysis of qualitative and quantitative outcomes. But many of the most interesting issues in sociological theory concern linked changes in quality and quantity. Sociologists have not even begun systematic study of coupled changes in qualitative and quantitative outcomes. One major obstacle to the development of explicit process models for quality and quantity is that we use different mathematical structures in the qualitative and quantitative cases. For the former we use stochastic models; for the latter we use deterministic models (see the discussion on Coleman 1964a: 526-8). Clearly there is a need to develop stochastic models for changes in quantitative variables. Unfortunately this leads to considerable mathematical complexity (see Jazwinski 1970 for a discussion). Nonetheless, this seems a necessary next step if we are to use temporal data to address many fundamental issues.

Literature Cited

Albert, A. 1962. Estimating the infinitesimal generator of a continuous time, finite state Markov process. Ann. of Math. Stat. 33: 727-53.

Balestra, P., Nerlove, M. 1966. Pooling cross-section and time series data in the estimation of a dynamic model: the demand for natural gas. Econometrica 34: 585-612.

Bergstrom, A. R. ed. 1976. Statistical Inference in Continuous Time Economic Models. Amsterdam: North-Holland. 333 pp.

Bielby, W. T., Hauser, R. M., Featherman, D. L. 1977. Response errors of nonblack males in models of intergenerational tranmission of socioeconomic status. Am. J. Sociol. 82: 1242-88.

Bishop, Y. M. M., Fienberg, S. E., Holland, P. W. 1975. Discrete Multivariate Analysis: Theory and Practice. Cambridge: The MIT Press. 557 pp.

Blalock, H. M., Jr. 1970. Estimating measurement error using multiple indicators and several points in time. Am. Sociol. Rev. 35: 101-117.

Blau, P. M., Duncan, O. D. 1967. The American Occupational Structure. New York: Wiley. 520 pp.

Blumen, I., Kogan, M., McCarthy, P. J. 1955. The industrial mobility of labor as a probability process. Cornell Studies in Industrial and Labor Relations 6. Ithaca: Cornell University Press.

Boag, J. W. 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. J. Roy. Stat. Soc. Series B 11: 15-53.

Box, G. E. P., Jenkins, G. M. 1976. Time Series Analysis: Forecasting and Control. San Francisco: Holden-Day. 575 pp. 2nd ed.

Bumpass, L. L., Sweet, J. A. 1972. Differentials in marital instability: 1970. Am. Sociol. Rev. 37: 754-66.

Campbell, D. T. 1963. From description to experimentation: interpreting trends as quasi-experiments. In Problems in Measuring Change, ed. C. W. Harris, pp. 212-42. Madison: University of Wisconsin Press.

Chase-Dunn, C. 1975. The effects of international economic dependence on development and inequality: a cross-national study. Am. Sociol. Rev. 40: 720-38.

Chirot, D., Ragin, C. 1975. The market, tradition and peasant rebellion: the case of Romania in 1907. Am. Sociol. Rev. 40: 428-44.

Çinlar, E. 1975. Introduction to Stochastic Processes. Englewood Cliffs, NJ: Prentice-Hall. 402 pp.

Coleman, J. S. 1964a. Introduction to Mathematical Sociology. New York: The Free Press. 554 pp.

Coleman, J. S. 1964b. Models of Change and Response Uncertainty. Englewood Cliffs, NJ: Prentice-Hall. 130 pp.

Coleman, J. S. 1968. The mathematical study of change. In Methodology in Social Research, ed. H. M. Blalock, Jr., A. Blalock, pp. 428-78. New York: McGraw-Hill.

Coleman, J. S. 1973. The Mathematics of Collective Action. Chicago: Aldine. 191 pp.

Coleman, J. S., Blum, Z. D., Sørensen, A. B., Rossi, P. 1972. White and black careers during the first decade of labor force experience. Soc. Sci. Res. 1: 243-70.

Cox, D. R. 1972. Regression models and life tables. J. Roy. Stat. Soc. Series B 34: 187-220.

Crano, W. D. 1977. What do infancy mental tests test? A cross-lagged panel analysis of selected data from the Berkeley Growth Study. Child Dev. 48: 144-51.

Crano, W. D., Kenny, D. A., Campbell, D. T. 1972. Does intelligence cause achievement? A cross-lagged panel analysis. J. Educ. Psychol. 63: 258-75.

Davis, J. A. 1974. Hierarchical models for significance tests in multivariate contingency tables: an exegesis of Goodman's recent papers. In Sociological Methodology 1973-1974, ed. Herbert L. Costner, pp. 189-231. San Francisco: Jossey-Bass.

Davis, J. A. 1978. Studying categorical data over time. Soc. Sci. Res. 7: 151-79.

Doreian, P., Hummon, N. P. 1976. Modeling Social Processes. Amsterdam: Elsevier. 172 pp.

Duncan, O. D. 1966. Methodological issues in the analysis of social mobility. In Social Structure and Social Mobility in Economic Development, ed. N. Smelser, S. M. Lipset, pp. 51-97. Aldine.

Duncan, O. D. 1969. Some linear models for two-variable panel analysis. Psychol. Bull. 72: 177-82.

Duncan, O. D. 1972. Unmeasured variables in linear models for panel analysis. In Sociological Methodology 1972, ed. H. Costner, pp. 36-82. San Francisco: Jossey-Bass.

Duncan, O. D. 1975. Introduction to Structural Equation Models. New York: Academic Press. 180 pp.

Duncan, R. P., Perrucci, C. C. 1976. Dual occupation families and migration. Am. Sociol. Rev. 41: 252-61.

Eaton, W. W., Jr. 1974. Mental hospitalization as a reinforcement process. Am. Sociol. Rev. 39: 252-60.

Efron, B. 1977. The efficiency of Cox's likelihood function for censored data. J. Amer. Stat. Assoc. 72: 557-65.

Esmer, Y. 1979. Political mobilization and economic development: a confirmatory factor analysis. In National Development and the World System: Educational, Economic, and Political Change, 1950-1970, ed. J. W. Meyer & M. T. Hannan. Chicago: University of Chicago Press. In press.

Form, W. H., Miller, D. C. 1949. Occupational career patterns as a sociological instrument. Am. J. Sociol. 54: 317-29.

Freeman, J., Hannan, M. T. 1975. Growth and decline processes in organizations. Am. Sociol. Rev. 40: 215-28.

Goldberger, A. S. 1964. Econometric Theory. New York: Wiley. 399 pp.

Goldberger, A. S. 1971. Econometrics and psychometrics: a survey of communalities. Psychometrika 36: 83-107.

Goodman, L. A. 1972a. A modified multiple regression approach to the analysis of dichotomous variables. Am. Sociol. Rev. 37: 28-46.

Goodman, L. A. 1972b. A general model for the analysis of surveys. Am. J. Sociol. 77: 1035-86.

Goodman, L. A. 1973. Causal analysis of data from panel studies and other kinds of surveys. Am. J. Sociol. 78: 1135-91.

Granger, C. W. J. 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37: 424-38.

Haberman, S. J. 1974. The Analysis of Frequency Data. Chicago: University of Chicago Press. 419 pp.

Hannan, M. T., Freeman, J. 1978. The internal politics of growth and decline. In Environments and Organizations, ed. M. Meyer, pp. 177-99. San Francisco: Jossey-Bass.

Hannan, M. T., Rubinson, R., Warren, J. T. 1974. The causal approach to measurement error in panel analysis: some further contingencies. In Measurement in the Social Sciences, ed. H. M. Blalock, Jr., pp. 293-324. Chicago: Aldine.

Hannan, M. T., Tuma, N. B., Groeneveld, L. P. 1977. Income and marital events: evidence from an income maintenance experiment. Am. J. Sociol. 82: 1186-1211.

Hannan, M. T., Young, A. A. 1977. Estimation in panel models: results on pooling cross-sections and time series. In Sociological Methodology 1977, ed. D. R. Heise, pp. 52-83. San Francisco: Jossey-Bass.

Hauser, R. M., Koffel, J. N., Travis, H. P., Dickinson, P. J. 1975. Temporal change in occupational mobility: evidence from men in the United States. Am. Sociol. Rev. 40: 279-97.

Heise, D. R. 1969. Separating reliability and stability in test-retest correlations. Am. Sociol. Rev. 34: 93-101.

Heise, D. R. 1970. Causal inference from panel data. In Sociological Methodology 1970, ed. F. Borgotta, G. Bohrensfedt, pp. 3-27. San Francisco: Jossey-Bass.

Herbst, P. G. 1963. Organizational commitment: a decision process model. Acta Sociol. 7: 34-46.

Hibbs, D. A., Jr. 1974. Problems of statistical estimation and causal inference in time-series regression models. See Davis 1974, pp. 252-308.

Hibbs, D. A., Jr. 1977. On analyzing the effects of policy interventions: Box-Jenkins and Box-Tiao versus structural equation models. See Hannan & Young 1977, pp. 137-79.

Hogan, D. P. 1978. The variable order of events in the life course. Am. Sociol. Rev. 43: 573-86.

Hummon, N. P., Doreian, P., Teuter, K. 1975. A structural control model of organizational change. Am. Sociol. Rev. 40: 813-24.

Inverarity, J. 1976. Populism and lynching in Louisiana, 1889-1896: a test of Erickson's theory of the relationship of boundary crises and repressive justice. Am. Sociol. Rev. 41: 262-79.

Jazwinski, A. H. 1970. Stochastic Processes and Filtering Theory. New York: Academic Press. 376 pp.

Johnston, J. 1972. Econometric Methods. New York: McGraw-Hill. 437 pp. 2nd ed.

Jöreskog, K. G. 1970. A general method for the analysis of covariance structures. Biometrika 57: 239-51.

Jöreskog, K.G., Sörbom, D. 1976. Statistical models and methods for analysis of longitudinal data. In Latent Variables in Socioeconomic Models, ed. D. J. Aigner, A. S. Goldberger, pp. Amsterdam: North-Holland.

Kaufman, R. L. 1976. The solution and interpretation of differential equation models. Am. Sociol. Rev. 41: 746-47.

Kelley, J. 1973. Causal chain models for the socioeconomic career. Am. Sociol. Rev. 38: 481-93.

Kenny, D. A. 1973. Cross-lagged and synchronous common factors in panel data. In Structural Equation Models in the Social Sciences, ed. A. S. Goldberger, O. D. Duncan, pp. 153-67. New York: Seminar Press.

Kenny, D. A. 1975. Cross-lagged panel correlation: a test for spuriousness. Psychol. Bull. 82: 345-62.

Kohn, M. L., Schooler, C. 1978. The reciprocal effects of the substantive complexity of work and intellectual flexibility: a longitudinal assessment. Am. J. Sociol. 84: 24-52.

Kuh, E. 1959. The validity of cross-sectionally estimated behavior equations in time series application. Econometrica 39: 341-58.

Lazarsfeld, P. F. 1948. The use of panels in social research. Am. Philos. Soc. 92: 405-10.

Lazarsfeld, P. F. 1972. Mutual effects of statistical variables. In Continuities in the Language of Social Research, ed. Paul F. Lazarsfeld, Ann K. Pasnella & Morris Rosenberg, pp. 388-98. New York: The Free Press.

Lazarsfeld, P. F., Berelson, B. R., Gaudet, H. 1944. The People's Choice. New York: Duell, Sloan, & Pearce. 178 pp.

Lillard, L. A., Willis, R. J. 1976. Dynamic aspects of earnings mobility. Working Paper No. 150. National Bureau of Economic Research, Inc., Stanford, Calif.

Maddala, G. S. 1971. The use of variance components models in pooling cross-section and time series data. Econometrica 39: 341-58.

Mason, K. O., Mason, W. M., Winsborough, H. H., Poole, W. K. 1973. Some methodological problems in cohort analysis of archival data. Am. Sociol. Rev. 38: 242-57.

Mayer, T. F. 1972. Models of intragenerational mobility. In Sociological Theories in Progress, ed. J. Berger, M. Zelditch, Jr., B. Anderson, Vol. 2, pp. 308-57. Boston: Houghton-Mifflin.

Mayer, T. F., Arney, W. R. 1974. Spectral analysis and the study of social change. See Davis 1974, pp. 305-55.

McGinnis, R. 1968. A stochastic model of social mobility. Am. Sociol. Rev. 33: 712-22.

Meyer, M. 1975. Leadership and organizational structure. Am. J. Sociol. 81: 514-42.

Miller, R. G. 1976. Least squares regression with censored data. Biometrika 63: 449-464.

Nerlove, M. 1971. Further evidence on the estimation of dynamic economic relations from a time series of cross-sections. Econometrica 39: 359-82.

Nielsen, F., Hannan, M. T. 1977. The expansion of national educational systems: tests of a population ecology model. Am. Sociol. Rev. 42: 479-90.

Palmer, G. L. 1954. Labor Mobility in Six Cities. New York: Social Science Research Council. 177 pp.

Pelz, D. C., Andrews, F. M. 1964. Causal priorities in panel study data. Am. Sociol. Rev. 29: 836-48.

Pitcher, B. L., Hamblin, R. L., Miller, J. L. L. 1978. The diffusion of collective violence. Am. Sociol. Rev. 43: 23-35.

Pullum, T. W. 1977. Parameterizing age, period, and cohort effects: an application to U.S. delinquency rates, 1964-1973. In Sociological Methodology 1978, ed. K. Schuessler, pp. 116-140. San Francisco: Jossey-Bass.

Robins, P., Tuma, N. B. 1977. Change in rates of entering and leaving employment under a negative income tax program: evidence from the Seattle and Denver income maintenance experiments. Research Memorandum #48, SRI International, Menlo Park, CA.

Rogosa, D. 1978a. Causal models in longitudinal research: rationale, formulation, and interpretation. In Longitudinal Research in Human Development: Design and Analysis, ed. J. R. Nesselroode, P. B. Baltes. New York: Academic Press. In press.

Rogosa, D. 1978b. Time and time again: some analysis problems in longitudinal research. Chicago: Educational Finance and Productivity Center, University of Chicago.

Ryder, N. B. 1965. The cohort as a concept in the study of social change. Am. Sociol. Rev. 30: 843-61.

Shorter, E., Tilly, C. 1970. The shape of strikes in France. Comparative Stud. Society Hist. 13: 60-86.

Sims, C. A. 1972. Money, income, and causality. Am. Econ. Rev. 62: 540-52.

Singer, B. 1977. Individual histories as the focus of analysis in longitudinal surveys. SSRC Conference on the National Longitudinal Surveys, Washington, D. C.

Singer, B., Spilerman, S. 1974. Social mobility models for heterogeneous populations. See Davis 1974, pp. 256-401.

Singer, B., Spilerman, S. 1976a. The representation of social processes by Markov models. Am. J. Sociol. 82: 1-54.

Singer, B., Spilerman, S. 1976b. Some methodological issues in the analysis of longitudinal surveys. Ann. of Econ. and Soc. Measurement 5: 447-74.

Snyder, D., Tilly, C. 1972. Hardship and collective violence in France, 1830-1960. Am. Sociol. Rev. 37: 520-32.

Sørensen, A. B. 1975. The structure of intragenerational mobility. Am. Sociol. Rev. 40: 456-71.

Sørensen, A. B. 1977. Estimating rates from retrospective questions. See Hannan & Young 1977, pp. 209-23.

Sørensen, A. B., Hallinan, M. T. 1976. A stochastic model for change in group structure. Soc. Sci. Res. 5: 43-61.

Sørensen, A. B., Hallinan, M. T. 1977. A reconceptualization of school effects. Sociol. Educ. 50: 273-89.

Sørensen, A. B., Tuma, N. B. 1978. Labor market structures and job mobility. Working Paper #505-78. University of Wisconsin, Madison.

Spilerman, S. 1970. The causes of racial disturbances: a comparison of alternative explanations. Am. Sociol. Rev. 35: 627-49.

Spilerman, S. 1972a. The analysis of mobility processes by the introduction of independent variables into a Markov chain. Am. Sociol. Rev. 37: 277-94.

Spilerman, S. 1972b. Extensions of the mover-stayer model. Am. J. Sociol. 78: 599-626.

Spilerman, S. 1977. Careers, labor market structure, and socioeconomic achievement. Am. J. Sociol. 83: 551-93.

Tilly, C., Tilly, L., Tilly, R. 1975. The Rebellious Century 1830-1930. Cambridge: Harvard University Press. 354 pp.

Tuma, N. B. 1973. Incorporating independent variables into stochastic models of social mobility. Presented at Ann. Meet. Am. Sociol. Assoc. NY.

Tuma, N. B. 1976. Rewards, resources and rate of mobility: a nonstationary multivariate stochastic model. Am. Sociol. Rev. 41: 338-60.

Tuma, N. B., Hannan, M. T., Groeneveld, L. P. 1979. Dynamic analysis of event histories. Am. J. Sociol. 84. In press.

Tuma, N. B., Hannan, M. T. 1978. Approaches to the censoring problem in analysis of event histories. In Sociological Methodology 1979, ed. K. Schuessler. In press.

Vigderhous, G. 1977. Forecasting sociological phenomena: applications of Box-Jenkins methodology to suicide rates. See Pullum 1977, pp. 20-51.

Wheaton, B., Muthen, B., Alwin, D. F., Summers, G. F. 1977. Assessing reliability and stability in panel models. See Hannan & Young 1977, pp. 84-136.

Wiggins, L. M. 1973. Panel Analysis. Latent Probability Models for Attitude and Behavior Processes. San Francisco: Jossey-Bass. 255 pp.