# SOME PROBLEMS OF STATISTICAL INFERENCE IN DYNAMIC SOCIOLOGICAL MODELS

Kenneth C. Land Russell Sage Foundation and Columbia University

# Abstract

The most uniquely sociological context of dynamic analysis is that of the longitudinal model of several time points of observation on multiple units. In this context, the present article emphasizes that valid statistical inferences are possible only if the model of the process under consideration is precisely and parsimoniously specified. Assuming a lag-1 stationary stochastic process with serially independent, constant variance disturbances and additive, serially independent, constant variance measurement errors to be an appropriate model, a test is given for the hypothesis that the stability parameter is constant across two or more time intervals. Given the more parsimonious model which this result provides and an appropriate a priori specification of a correlation process, a procedure is derived for estimating the degree of correlation among the measurement errors or among the disturbances. The latter result requires at least four time points of observation.

#### 1. Introduction.

Although sociologists sometimes analyze observations consisting of time series of measurements on a single social entity such as an organization or a population [7], the more uniquely sociological context of over-time statistical analysis seems to be that of a time series of observations on a number of social units which is often referred to as a panel or longitudinal analysis. Furthermore, although not a necessary condition of such designs, longitudinal studies are typically characterized by the number of time periods of observation on each social unit being smaller than the number of units observed at each point. Thus, the statistical context of dynamic sociological analysis is somewhat different from that of the time series situation which typically is treated in the statistical and economic literature. At best, the context of dynamic sociological analysis could be characterized as a situation of multiple time series [9].

Within this context of dynamic sociological analysis, questions of a statistical nature immediately arise. Several recent methodological papers have been addressed to such problems [1,6,10]. In this brief paper, I shall review and analyze some of the problems of statistical inference which have been raised by these and other authors. The emphasis will be on presenting problems and posing possible approaches to the problems rather than on the systematic development of definitive solutions.

#### 2. Specification of Models.

Consider the longitudinal model represented in Figure 1 where, for convenience, the variables are assumed to be measured in standard deviation units. Notationally, the  $x_{tn}$  are the "true" scores of a variable on a sample of n = 1, ..., Nindividuals measured at t = 1,2,3 time points, the  $x'_{tn}$  are the corresponding obtained measurements on the variable, the  $u_{tn}$  are the random disturbances or shocks affecting the true values of the variable at the time intervals, and the  $e_{tn}$  are random variables representing the errors of the measurement in the  $x_{tn}$ . This is the model for longitudinal analysis considered by Heise [6]. In brief, it represents an attempt to separate the stability and error components of observed scores on a sample of N individuals at three time points.

Heise [6, p. 98] specified the model in Figure 1 according to the following assumptions:

(i) the relationship between the true variable x and the index  $x'_{1}$  is constant over time, i.e., the reliability coefficient  $p_{xx}$  is constant over time periods;

(ii) the measurement errors  $e_{tn}$  are uncorrelated with the true scores  $x_{tn}$ ;

(iii) the random shocks u are serially uncorrelated;

(iv) the measure errors e are serially uncorrelated;

(v) the rate of change in the true scores x is approximately constant within the measurement interval.

Assumption (i) has recently been criticized by Wiley and Wiley (1970). They point out that the assumption of constant reliability requires, in general, that both true score and error variance be constant. This latter condition will not hold for any social process which results in an increase or a decrease in the variability of true scores without necessarily affecting the characteristics of the measuring instrument. Furthermore such an assumption will not allow comparisons across populations. On the other hand, many populations possess aggregate equilibrium with respect to a given variable over relatively short time intervals. That is, although the value of the variable may change for certain individuals





in the population, the aggregate result of individual changes is such that the population variances and covariances are constant. For populations governed by such processes, Heise's assumptions do not appear to be too strong.

Under the assumption outlined in the previous paragraph, Heise [6, pp. 96-97] showed that one could solve the following estimation equations

$$r_{12} = p_{xx}^{2} p_{21}$$

$$r_{23} = p_{xx}^{2} p_{32}$$

$$r_{13} = p_{xx}^{2} p_{21} p_{32}$$
(2.1)

for the following estimates of the reliability and stability coefficients

~ ~

$$\hat{\mathbf{p}}_{\mathbf{xx}}^{2} = \mathbf{r}_{12} \mathbf{r}_{23} / \mathbf{r}_{13}$$

$$\hat{\mathbf{p}}_{21} = \mathbf{r}_{12} / \mathbf{p}_{\mathbf{xx}}^{2} = \mathbf{r}_{13} / \mathbf{r}_{23} \quad (2.2)$$

$$\hat{\mathbf{p}}_{32} = \mathbf{r}_{23} / \mathbf{p}_{\mathbf{xx}}^{2} = \mathbf{r}_{13} / \mathbf{r}_{12}$$

where the symbol "^" over the coefficients denotes an estimate of the corresponding population quantity based on a sample and  $r_{1,1}$  is the sample estimate of the population correlation  $\rho_{1,1}$  of the variables  $x_{1,2}$  and  $x_{2,1}$ , i, j = 1,2,3. However, Heise [6, pp. 199-100]<sup>J</sup>also found that it is not possible to estimate the parameters of the model if either the assumption of serially uncorrelated random shocks is relaxed even if one adds additional time points of observations on the sample. Specifically, for the specified model, Heise [6, p. 100] could derive only a consistency relationship among the bivariate correlation coefficients for four observation points which must hold within sampling errors if the data were drawn from the specified model:

$$r_{14} r_{23} = r_{13} r_{24}$$
 (2.3)

In this paper, we shall show that Heise's inability to relax the assumptions of his model stems from the incomplete and unparsimonious specification of the model. In particular, we shall present methods for testing the hypotheses that (1) the coefficients  $p_{21}$  and  $p_{32}$  are identical and (2) the measurement errors are uncorrelated. These results depend upon the more complete specification of the model as a stationary stochastic process.

# 3. <u>Testing for Equality of Stability Coeffi</u> <u>cients for Two or More Time Periods.</u>

We would like to specify that the <u>causal</u> <u>model</u> governing the relationship among the true scores in Figure 1 is of the form

$$x_{tn} = p_{x}x_{(t-1)n} + p_{xu}t_{tn}$$
(3.1)  
where we assume that

(i) the random shock variance is time-homogeneous for all observation points after the initial observation  $[v(u_2) = v(u_3)]$ , i.e., the process is stationary;

(ii) the random shocks  ${\tt u}_{\tt tn}$  are serially and contemporaneously independent;

(iii) the process is lag-1 (or Markov) and the parameter  $p_{\rm x}$  is constant across time intervals of equal length.

With respect to a <u>measurement model</u> for Figure 1, we would like to assume that

(iv) the measurement errors  $e_{tn}$  are independent of the true scores  $x_{tn}$ ;

(v) the measurement errors are serially independent; and

(vi) the measurement error variance is timehomogeneous

$$[v(e_1) = v(e_2) = v(e_3)].$$

Assumptions (i) and  $(v_i)$  together suffice to determine time-constant reliability. In brief, this model is the most parsimonious model for longitudinal sociological analysis which can be specified. It can be called a constant lag-1 stationary stochastic process with serially independent, constant variance disturbances and additive, serially independent, constant variance measurement errors.

The most important differences between this specification and that given by Heise are that the stationary property of the process is explicit, that the model is specified as Markovian with constant stability parameter  $p_x$ , and that the disturbances and measurement erfors are assumed to be serially independent rather than serially uncorrelated. The constant value of  $p_x$  typically has not been assumed in longitudinal sociological models of [cf., 6, 10, 1]. However, given equal time intervals between observations and the plausibility of the remaining assumptions, it would seem natural to assume that  $p_x$  is constant.

Since the remainder of the assumptions are as specified by Heise, we should like to inquire here as to how one could statistically evaluate the assumption that p is constant. One possible approach to an evaluation of the assumption of a constant stability parameter is to utilize a large-sample modification of a dummy variable method based on the analysis of covariance which was recently presented by Gujarati [5]. For N observations at three time points, the test can be represented as follows:

$$x_{tn} = a_1 x_{(t-1)n} + a_2 (Dx_{(t-1)n}) + u_{tn},$$
  
t = 2,3, n = 1,...,N (3.2)

where D = 1 if t - 1 = 2, = 0 if t - 1 = 1.

The matrix representation of (3.2) is

$$\begin{bmatrix} \mathbf{x}_{21} \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_{2N} \\ \mathbf{x}_{31} \\ \cdot \\ \cdot \\ \mathbf{x}_{31} \\ \cdot \\ \cdot \\ \mathbf{x}_{2N} \\ \mathbf{x}_{21} \\ \mathbf{x}_{21} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_{21} \\ \mathbf{x}_{21} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_{2N} \\ \mathbf{x}_{21} \\ \mathbf{x}_{21} \\ \cdot \\ \cdot \\ \cdot \\ \mathbf{x}_{2N} \\ \mathbf{x}_{$$

which is an aid to visualizing the test. The test of equality of the regression coefficients for the two time intervals is as follows: one computes the least squares estimates of a, and a<sub>2</sub>. Under the additional assumption that the dísturbances are drawn from a normal distribution, the least squares estimators of a, and a, will be the maximum likelihood estimators. This insures that they will be the best asymptotically normal (BAN) estimators. Furthermore, the estimated approximate covariance matrix of the vector [a,  $a_2$ ]' is  $\hat{\sigma}_u^2$  (M<sub>XX</sub>)<sup>1</sup> where  $\hat{\sigma}_u^2$  is the estimated variance of the disturbance term and M is the moment matrix of the x variables in (3.2) [cf. 2, p. 386]. Since the estimators of  $a_1$  and  $a_2$  are asymptotically normal, it follows that one can form the ratio  $\hat{a}_2/s(\hat{a}_2)$  and use it to test the hypothesis H :  $a_2 = 0$  under the normal distribution where s  $(\hat{a}_2)$  is the estimated standard deviation of  $\hat{a}_2$ . Of course this test is only approximately appropriate for small samples. If  $\hat{a}_2$  is statistically significant, then  $\hat{p}_{32}=\hat{a}_1+\hat{a}_2$  while  $\hat{p}_{21}=\hat{a}_1$  so that one must reject the null hypothesis that  $p_x$  is constant.

On the other hand, if  $\hat{a}_2$  is statistically insignificant, then  $\hat{p}_x = a_1$  for both time periods.

This test can also be utilized to test the equality of constant terms in case the variables are not measured in deviation or standard deviation units. For such a test, one merely enters the dummy variable into the equation in an additive fashion. In the present context, the method can also be generalized to test for equality of the coefficients for an arbitrary number of time intervals. Thus, given that the other specifications of the model are maintained, one can statistically test the hypothesis that the coefficient  $p_x$  is constant across time intervals. Hopefully, sociologists will be encouraged to utilize this test, because of the more parsimonious representation which it facilitates.

Finally, it should be noted that the assumption of constant p has very little effect on the estimators of the reliability and the stability coefficients as computed by Heise. In particular, the formulae corresponding to those in (2.2) are

$$\hat{p}_{xx}^2 = r_{12}^2 / r_{13}$$
 (3.4)  
 $\hat{p}_x = r_{13} / r_{12}$ .

Another implication of this model is that  $r_{12} = r_{23}$ . In other words, the correlations between adjacent time periods should be equal. In fact, for variables measured in standard deviation units, a test of this equality is equivalent to the above test of the equality of the stability coefficients.

# 4. Serial Correlation.

As we noted above, under his assumptions, Heise [6] was unable to derive estimates of the stability and reliability coefficients if either of the assumptions of serially uncorrelated shocks or of serially uncorrelated measurement errors was modified. Yet, Heise [6, p. 98] himself offers a strong argument for seeking a model in which serial correlation of measurement errors can be accounted for: "The assumption may be violated when respondents recall earlier answers and try to be consistent in their responses. In such a case, distortions occurring in early measurements will tend to be reproduced over time." Since the test of a constant stability coefficient among time intervals in the preceding section allows a simpler representation of the longitudinal model, we should like to inquire here as to whether this simplification can be utilized to allow the model to account for serial correlation of the disturbances or of measurement errors.

Before considering the estimation of serial correlation of measurement errors, we should possess some method of ascertaining whether or not serial correlation is a critical problem for a given set of observations. One possibility is

to apply the Durbin-Watson d statistic [3, 4]. However, the validity of that statistic rests on the assumption that all of the explanatory variables in a regression equation are exogenous variables (as distinguished from lagged values of the dependent variable). Since this assumption is not valid in the present case, the statistic would be biased. Another possible approach is to regenerate the correlations between the variables on the basis of the estimated reliability and stability coefficients and to compare these with the correlation coefficients computed from the observations. The difference between the observed and the derived correlation coefficients is the maximumlikelihood estimator of the correlation between the residuals (consisting of both random shocks and measurement errors) as Land [8] has shown. Under suitable assumptions about the normality of the distributions of the residuals, this estimator can be subjected to the standard test of significance for correlation coefficients. However, an indication of positive correlation among the computed residuals by this test does not resolve the issue of whether the correlation is due to correlated measurement errors or to correlated disturbances. To resolve this issue, substantive theoretical considerations must be brought to bear on the interpretation. If the measurements consist of some social psychological test administered in relatively short time intervals and for which one would theoretically expect a strain for consistency, then one may interpret a positive correlation of residuals as a consequence of correlated measurement errors. On the other hand, if one suspects that there is some unmeasured variable that has stability over time and has a substantial effect on the true scores x<sub>tn</sub>, then one may prefer to interpret a posi-tive correlation of residuals as due to correlated disturbances.

Given that one has satisfied himself that the correlation among measurement errors is substantial enough so that one should take it into account, then the relevant question now becomes whether or not we can estimate the value of the correlation. The answer is that one can estimate the correlation among the measurement errors if one obtains at least a fourth time point of observations on the sample and if one places suitable a priori restrictions on the process which is assumed to have generated the correlation among the measurement errors. To illustrate, suppose that we have a fourth set of observations on the variable x. Then we possess effectively only one additional piece of independent statistical information -the observed correlation r<sub>14</sub>. This severely restricts the class of processes which one can assume to have generated the correlation among the measurement errors. Specifically, one can consider only one-parameter processes for this set of observations.

One possible set of a priori assumptions

regarding measurement errors which can be fitted in the present context is to assume that the errors e follow a first-order (Markov) auto-regressive scheme

$$e_{tn} = p_{e} e_{(t-1)n} + v_{tn}$$
 (4.1)

where  $p_e < 1$  and  $v_{tn}$  satisfies the assumptions for all t

$$E(v_{tn}) = 0$$
  

$$E(v_{tn}v_{(t+s)n}) = \sigma_v^2 \text{ is } s = 0 \qquad (4.2)$$
  

$$= 0 \text{ if } s \neq 0$$

where E denotes the expectation operator. In other words, we assume that the v are serially uncorrelated. Substantively, this set of assumptions implies that any effect on the current response of memory of earlier responses is carried through the most recent response and is a scalar function of that response. For this set of assumptions, the estimation equations which must be solved for the estimates of the stability and reliability coefficients are

$$r_{12} = p_{xx}^2 p_x + p_{xe}^2 p_e$$
  

$$r_{13} = p_{xx}^2 p_x^2 + p_{xe}^2 p_e$$
  

$$r_{14} = p_{xx}^2 p_x^3 + p_{xe}^2 p_e^3.$$
(4.3)

Recalling that  $p_x = 1 - p_x$ , we see that this is a system of three equations in three unknowns. Although the equations are nonlinear in the unknown, they can be solved by modern numerical procedures. For example, provided that all of the observed correlations are positive, one simple approach to a solution of (4.3) is to take logarithms of the equations which then become linear and can be solved by standard linear procedures.

The method of dealing with correlated measurement errors outlined here can be generalized. The crucial assumption is that  $p_x$  is constant. This gives one more degree of freedom which can be utilized to study correlated measurement errors. For observations at four time points, one must specify a priori a one-parameter process as governing the correlation among the errors. The Markov scheme given above is only one possibility. Another likely candidate could be derived from an exponential decay model of memory. Of course, additional time points of observation would provide more degrees of freedom which could be utilized to fit more complex processes of correlation among measurement errors or among disturbances. Hopefully, sociologists will soon possess longitudinal surveys consisting of many time points of observation.

#### 5. Concluding Comments.

This paper has been written on the assumption that models for longitudinal sociological analysis should consist of precise and parsimonious a priori specifications of the processes which are assumed to generate the observations. From this perspective, we have shown that one may test the hypothesis that the stability coefficients relating the true scores on a variable are constant over time. Given the more parsimonious model which results therefrom, we have indicated how certain kinds of correlations among measurement errors can be estimated under appropriate assumptions. Although we have dealt only with measurements of a single variable over time, the procedures presented here can be generalized to situations of measurements on several variables over time if certain conditions of identifiability of the parameters are maintained.

### References

- Blalock, H. M. Jr.
   1970 "Estimating measurement error using multiple indicators and several points in time." <u>American Sociological Review</u> 35: 101-111.
- [2] Christ, C. F. 1966 <u>Econometric Models and Methods</u>, New York: Wiley.
- [3] Durbin, J. and G. S. Watson.
   1950 "Testing for serial correlation in least squares regression. I." <u>Biometrika</u> 37: 409-428.
- [4] 1951 "Testing for serial correlation in least squares regression. II." <u>Biometrika</u> 38: 159-178.

- [5] Gujarati, D. 1970 "Use of dummy variables in testing for equality between sets of coefficients in two linear regressions: a note." <u>The</u> <u>American Statistician</u> 24: 50-52.
- [6] Heise, D. R. 1969 "Separating reliability and stability in test-retest correlation." <u>American</u> <u>Sociological Review</u> 34: 93-101.
- [7] Land, K. C.
   1970a "A mathematical formalization of Durkheim's theory of the causes of the division of labor." in E. F. Borgatta (ed.), <u>Sociological Methodology: 1970.</u> San Francisco: Jossey-Bass.
- [8] 1971 "Identification, parameter estimation, and hypothesis testing in linear staticstochastic simultaneous-equation sociological models." Presented at a Conference on Structural Equation Models at the University of Wisconsin at Madison, November 12-16, 1970 and to be published in a volume of the of the Conference papers. Mimeo. Russell Sage Foundation: New York.
- [9] Quenouille, M. H. 1957 <u>The Analysis of Multiple Time Series.</u> New York: Hafner.
- [10] Wiley, D. E. and J. A. Wiley. 1970 "The estimation of measurement error in panel data." <u>American Sociological</u> <u>Review:</u> 35: 112-117.