3

# ASSESSING RELIABILITY AND STABILITY IN PANEL MODELS

*Blair Wheaton*
YALE UNIVERSITY

*Bengt Muthén*
UNIVERSITY OF UPPSALA

*Duane F. Alwin*
INDIANA UNIVERSITY

*Gene F. Summers*
UNIVERSITY OF WISCONSIN-MADISON

* This research was supported in part by the College of Agricultural and Life Sciences, University of Wisconsin-Madison, and the National Institute of Mental Health, PHS Research Grant MI-19689. We wish to acknowledge the helpful comments of George Bohrnstedt and David Heise on an earlier draft of this paper. We, of course, assume responsibility for any errors that remain.

## *INTRODUCTION*

The proper interpretation of multiple time-point data has become a subject of increasing interest among sociologists. One concern has been with models that incorporate a single variable measured at two or more points in time. For example, Heise (1969; see also Wiley and Wiley, 1970) has dealt with a model that includes unreliability (random measurement error) in the observed variables and instability over time in the underlying "true score" variable for variables with single indicators. Heise's purpose in setting forth this type of model is to caution researchers to avoid misleading interpretations of simple test-retest correlations, and towards this purpose he provides an operational distinction between the reliability and stability of variables measured over time. Wiley and Wiley (1970) and Werts, Jöreskog, and Linn (1971) discuss the identification and estimation of reliability and stability parameters in these models under less restrictive sets of assumptions. Others (Costner, 1969; Blalock, 1970; Hauser and Goldberger, 1971) have discussed single variable multi-wave models where the variable has multiple indicators at each time point. In this case more information is available for estimating reliability and stability. Included in the treatment of both types of single variable panel models is a concern with *measurement specification*; that is, a measurement model is explicitly represented for the relationship between observed measures and their underlying constructs and for the relationships among observed measures.

A second concern with the interpretation of panel data has been in the context of the multi-wave multi-variable case (Bohrnstedt, 1969; Duncan, 1969; Heise, 1970). This second focus is important because it takes more seriously the issue of theoretical specification. By *theoretical specification* we mean the specification of the pattern of relationships among the underlying constructs. Inasmuch as this second concern is not explicitly the issue of measurement error, it lacks the special virtues of the literature cited above. Both concerns are essential in any proper interpretation of panel data, especially when observed variables are fallible. Discussions by Heise (1970), and Duncan (1972, 1975) have dealt with the issues of measurement specification and theoretical specification, but their primary concern is not with estimation issues.[1] Hannan, Rubinson, and Warren (1974) have also discussed panel models with both theoretical and measurement specifications made explicit, but their treatment of identification and estimation issues does not emphasize that given an efficient estimation procedure there are clear advantages to overidentification. A recent paper by Jöreskog and Sörbom (1976a) provides a useful and comprehensive statistical treatment of a number of types of panel models with illustrative examples. This chapter shares the orientation of Jöreskog and Sörbom to the analysis of panel data and follows the approach to estimation introduced by Jöreskog (1973; see also Jöreskog, 1976).

The purpose of the present chapter is to suggest some strategies that will improve both the measurement and the theoretical specification of models using panel data and incorporating the estimation of reliability and stability parameters. It is especially important to point out that the resulting flexibility in the specification procedure greatly increases the usefulness of panel data while providing a potential solution to some of the common problems with their use. Following the lead of Heise (1969) and the Wileys (1970), our discussion emphasizes the necessity of incorporating the issue of unreliability of measurement in any general model for the analysis of panel data. Also, following Duncan (1969) and others, we encourage the thorough specification of causal relationships among the theoretical variables included in multiple-wave data.

We begin our discussion by reconsidering the original single-variable three-wave model proposed by Heise (1969). It has become clear that for many purposes such single-variable models are inadequate at the level of theoretical specification, resulting in biased estimates of reliability and stability parameters in such models. Also involved in the issue of the theoretical specification of relationships among unobserved constructs is the reconceptualization of stability, since in the multivariate case the stability parameter is a partial regression coefficient rather than a correlation.

[1] A third literature dealing with panel analysis has addressed the issue of assessing causal priority of variables in multi-wave data using "cross-lagged" correlations (Pelz and Andrews, 1964; Rozelle and Campbell, 1969; Kenny, 1973). Others (Bohrnstedt, 1969; Duncan, 1969; Heise, 1970) have shown that this issue can be addressed in a more conventional regression framework.

We then go on to examine a set of models that are more completely specified, estimating the parameters of these models via a technique that allows us to estimate causal (structural) relations among unobserved constructs with multiple indicators (Jöreskog, 1976; Jöreskog and Sörbom, 1976b). Of special interest here is the fact that issues of specification at each level must be considered in the development of properly specified panel models.

## *THE SINGLE-VARIABLE THREE-WAVE MODEL*

The original single-variable three-wave model set forth by Heise (1969) is presented in Figure 1. First, the model states that at each point in time ($t = 1, 2, 3$) the observed variable $x_t$, is a perfect linear, additive function of two uncorrelated components: a true (reliable) component $X_t$, and a random error component $e_t$, assumed to be measurement error. Second, the model states that at

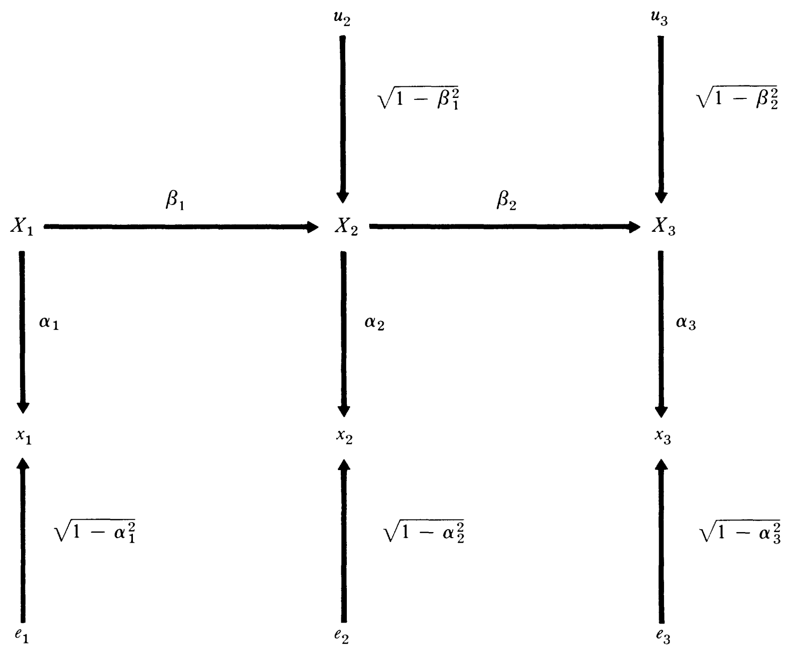


Figure 1. A Representation of the Single-Variable Three-Wave Panel Model.

each point in time the true underlying variable $X_t$ is a perfect linear additive function of two uncorrelated components: the true variable at the previous point in time, and a random disturbance representing sources of change or instability in the true variable over time. We have represented the effects of a true score on the observed score at a particular point in time as $\alpha_t$, and the effects of a true score at a particular point in time on the true score at a subsequent point in time as $\beta_j (j = 1, 2)$. In order to identify the reliability and stability parameters Heise (1969) assumes $\alpha_1 = \alpha_2 = \alpha_3$, which provides for a straightforward algebraic solution for $\alpha_t$ and the stability parameters $\beta_1$ and $\beta_2$. Wiley and Wiley (1970) point out that this solution does not make use of all the available information in the data, since it analyzes the correlation matrix instead of the variance-covariance matrix. The Wileys note that it is sufficient to assume that $\mathrm{Var}(e_1) = \mathrm{Var}(e_2) = \mathrm{Var}(e_3)$, then the variance of the unobserved true variable need not be equal across time. In this case we need not assume $\alpha_1 = \alpha_2 = \alpha_3$ in order to identify these parameters. Werts, Jöreskog, and Linn (1971) have pointed out that if the single-variable model is extended to four (or more) time points, unequal reliabilities and the stabilities for the internal measures (excluding the first and the last) can be identified without having to assume equality in the error variances.

Note that the reliability of a particular observed score at a given point in time is defined as $\alpha_t^2$ (see Lord and Novick, 1968), where for any $t$:

$$\alpha^2 = p_{xX}^2 = \frac{\mathrm{Var}(X)}{\mathrm{Var}(x)} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_E^2} \tag{1}$$

where $\sigma_X^2$ = true score variance, and $\sigma_E^2$ = measurement error variance. There is a more general definition of reliability, which we will introduce later, of which this definition is a special case.

These single-variable models do have a number of assumptions in common:

(1) The true and error components of the observed score are uncorrelated, and error components of any observed score are uncorrelated with all other true variables.

(2) Measurement errors at different points in time are uncorrelated.

(3) Disturbance terms are uncorrelated with the true variables at prior points in time.

(4) The system is assumed to be lag-1; that is, there is no direct linear effect of the true variable at time 1 on the true variable at time 3.

Problems with assumptions (1) and (2) have been discussed in a number of other papers (Blalock, 1970; Althauser and Heberlein, 1970; Costner and Schoenberg, 1973; Alwin, 1974). For our present purposes, assumption (1) expresses the measurement model used in this paper, and we will address problems resulting when the other assumptions do not hold.

An important potential problem with the single-variable model described above is one of bias in the estimates of the stability parameters. We mean by this that some external variable, say $Y$, is a determinant of the underlying true variable at two or more points in time; and, as a result, we may spuriously attribute stability to the underlying true variable. This possibility (assuming $Y$ is perfectly measured) is represented in the diagram in Figure 2.

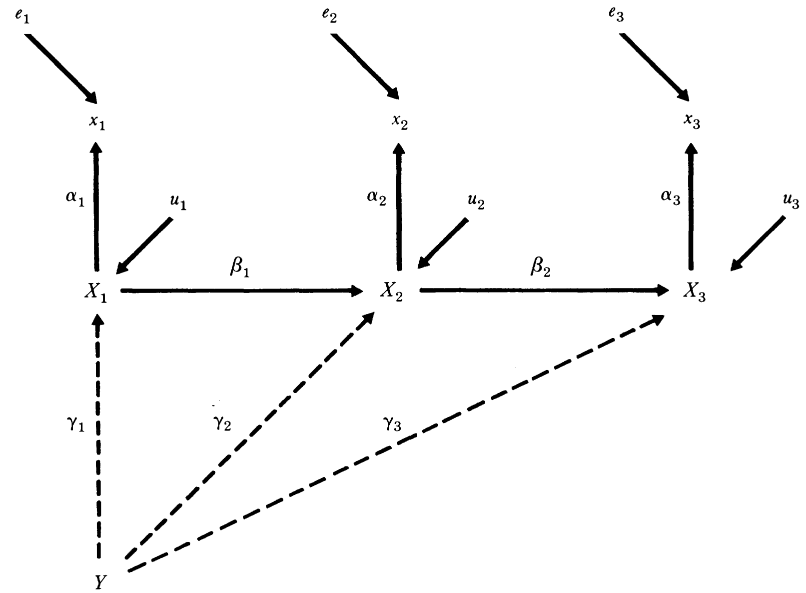


Figure 2. The Operation of an Excluded Variable in the Single-Variable Model.

Using the conventional algorithms of path analysis and assuming standardized data, we can solve for Heise's stability parameters in this model as follows (using the notation $\beta_1'$ and $\beta_2'$ for Heise's parameters):

$$\beta_1' = \frac{\rho_{13}}{\rho_{23}} = \frac{\alpha_1\alpha_3(\beta_1\beta_2 + \gamma_1\gamma_3 + \gamma_1\gamma_2\beta_2)}{\alpha_2\alpha_3(\beta_2 + \gamma_3\gamma_2 + \gamma_3\gamma_1\beta_1)}$$

$$= \frac{\alpha_1(\beta_1\beta_2 + \gamma_1\gamma_3 + \gamma_1\gamma_2\beta_2)}{\alpha_2(\beta_2 + \gamma_3\gamma_2 + \gamma_3\gamma_1\beta_1)}$$

$$\beta_2' = \frac{\rho_{13}}{\rho_{12}} = \frac{\alpha_1\alpha_3(\beta_1\beta_2 + \gamma_1\gamma_3 + \gamma_1\gamma_2\beta_2)}{\alpha_1\alpha_2(\beta_1 + \gamma_1\gamma_2)}$$

$$= \frac{\alpha_3(\beta_1\beta_2 + \gamma_1\gamma_3 + \gamma_1\gamma_2\beta_2)}{\alpha_2(\beta_1 + \gamma_1\gamma_2)}$$

Note from the above equations that $\beta_1'$ will equal $\beta_1$ when $\gamma_1\gamma_3 = \gamma_1\gamma_2 = \gamma_2\gamma_3 = 0$ and $\alpha_1 = \alpha_2$. Also note that $\beta_2'$ will equal $\beta_2$ when $\gamma_1\gamma_2 = \gamma_1\gamma_3 = 0$ and $\alpha_2 = \alpha_3$, suggesting the possibility of bias when $Y$ is related to $X$. Heise (1969) discusses the possibility, but does not present data bearing on the issue. Likewise, reliability estimates in such a model may be affected. Using Heise's assumptions ($\alpha_1 = \alpha_2 = \alpha_3 = \alpha$) for simplicity, it is clear that under most conditions

$$\alpha' = \frac{\rho_{12}\rho_{23}}{\rho_{13}} \neq \sqrt{\frac{\alpha^2(\beta_1 + \gamma_1\gamma_2)(\beta_2 + \gamma_2\gamma_3 + \beta_1\gamma_1\gamma_3)}{(\beta_1\beta_2 + \gamma_1\gamma_3 + \gamma_1\gamma_2\beta_2)}}$$

where $\alpha'$ is the estimate resulting from the Heise model. Therefore, in cases where other variables cause $X$, the Heise estimates of reliability and stability may be biased.[2]

Figure 2 introduces a new complexity in the definition and interpretation of stability. In single-variable models, the stability (standardized to a path coefficient) is automatically the total correlation between successive true variables. In the multivariate case,

[2] These principles can also be illustrated with the Wiley and Wiley (1970) model, since it is very similar to that proposed by Heise. The proofs in this case are much more cumbersome, and we will not deal with them here.

the stability parameters we will use are partial regression coefficients. This is not only a shift in operational definition, it also points to the need for a clarification of the concept of stability. Care must be taken first to specify the unit of analysis under discussion. The stability parameter is intended to represent the amount of true change in a variable across time, and in this chapter what we mean by "true change" is change of individuals' positions in the distribution relative to one another. This means that we do not address the issue of constant change over time, as reflected by changes in the intercept. Constant change of the group as a whole is another issue, which can be addressed in the context of "system" differences but is not central to the question of the relative change between individuals—which is implicit in our application of the stability concept.

We contend that the simple change definition (operationalized as a correlation) can lead to misleading interpretations of stability. It is obvious from Figure 2 that part of this inter-temporal true score correlation is not due to the effect of the variable at $t$ on the same variable at $t + 1$. Stability, as we define it, is concerned with the amount of change or lack of change in $X_{t-1}$ and due to $X_t$ alone; that is, the degree to which one's score on $X_t$ is, in fact, the source of one's score on $X_{t+1}$. The change in $X_2$ and $X_3$ in Figure 2 caused by other agents should be interpreted separately from the stability of that variable. Such a notion of stability is implicitly recognized by Heise when discussing potential sources of distortion to stability estimates (1969, pp. 99–101).

We conclude, from the preceding discussion, that it is essential that causal models representing processes measured in panel data be correctly specified at the theoretical level—as well as the measurement level—if one is to place any confidence in estimates of the "true" stability of variables. Our orientation leads us to explore first the possibility that certain concomitant variables may be causally related to the single variable measured at three points in time in single indicator models. Also, however, by using multiple indicators for each construct we introduce the possibility of estimating correlations between error terms for the observed measures, and we create more information for the estimation of the stability parameters (Costner, 1969; Blalock, 1970). This "overidentification" in multiple indicator models is useful precisely because it results in more than one estimate of the stability parameter. Given a procedure for combining these separate estimates, the estimation of such models is enhanced. A general procedure for obtaining efficient estimation in overidentified models of the type considered here is available using a maximum-likelihood solution (Jöreskog, 1973, 1976; Jöreskog and Sörbom, 1976b), and we will make use of this method below.

## *THE SPECIFICATION AND ESTIMATION OF RECURSIVE MODELS IN PANEL DATA*

Discussions of the analysis of panel data by Duncan (1969) and Heise (1970) consider the problems inherent in making inferences from sample data in panel situations where the variables contain measurement error. Neither treatment incorporates specification of the measurement error in the variables as a part of the model to be estimated. More recent papers by Duncan (1972, 1975) develop several models for the two-wave two-variable situation that incorporate unobserved constructs and, thereby, implicitly deal with the issue of random measurement error. Also, Hannan *et al.* (1974) specify both random and nonrandom measurement errors in a series of three-wave models. We can extend these discussions using different estimation assumptions and placing them in the context of the reliability-stability issue.

We will discuss first an "unexplicated" confirmatory factor analytic model (hereafter referred to as a *CFA* model). By "unexplicated" we mean simply that we do not specify a causal system for the covariances among unobserved constructs. Later, we will distinguish this form of *CFA* from the "explicated" *CFA* model in which we postulate a causal structure among unobserved constructs, which is intended to account for the covariances and variances among these constructs. The unexplicated *CFA* model deals only with the issue of measurement specification, whereas the explicated model deals with both measurement and theoretical specification. From an econometric perspective, the explicated model could be regarded as a structural equation model extended to include multiple indicators.

Explicated confirmatory factor analytic models provide a number of advantages in estimating reliability-and-stability models with panel data. First, these methods take into account random measurement error in that the amount of this type of error is estimated directly in the model. Second, in the case of multiple indicators, certain measurement error correlations can also be estimated. Third, causal relationships between abstract constructs can be interpreted directly rather than through inference from measured variable relationships. Fourth, the postulated structure relating observed measures to unobserved constructs and unobserved constructs to each other can be tested for fit to the observed variance-covariance matrix. Fifth, the model is thus very amenable to use as a theory construction tool (see the discussion by Burt, 1973). And, sixth, the issue of reliability and stability can be addressed within the context of a general model that has increased specification flexibility and, thus, increases our chances of accurately estimating these parameters. Specifically, the models we will suggest can be used to incorporate and estimate the effects of three possible sources of distortion to reliability and stability estimates; including the effects of excluded causal variables in the Heise and Wileys' models, the effects of correlations between exogenous constructs in the model and endogenous disturbance terms (or, alternatively, violation of the lag-1 assumption), and the effects of correlated measurement error.

An example of an unexplicated *CFA* model appears in Figure 3. The well-known covariance structure implied by this model (Jöreskog, 1969) can be expressed as:

$$\Sigma = \Lambda\Phi\Lambda' + \Theta^2 \tag{2}$$

where $\Sigma$ is a $p \times p$ population variance-covariance matrix for $p = 7$ variables, composed of three types of matrices: matrix $\Lambda$ is a $p \times m$ factor pattern matrix expressing the $p$ observed variables as a linear function of $m$ unobserved constructs; matrix $\Phi$ is an $m \times m$ matrix of unobserved construct covariances and variances; and matrix $\Theta^2$ is a $p \times p$ diagonal matrix of residual error variances for the observed variables.[3] Each of these parameter matrices may contain *fixed* parameters that are assumed known *a priori*, and the other parameters in

[3] In this case, the paths for the error terms are assumed to be 1 (as in the factor analytic model), and we can then identify the variance of the error term. This is equivalent to representing the $\Theta_i$ as the path coefficients for the measurement terms with the variance of the $e_i$ equal to 1.

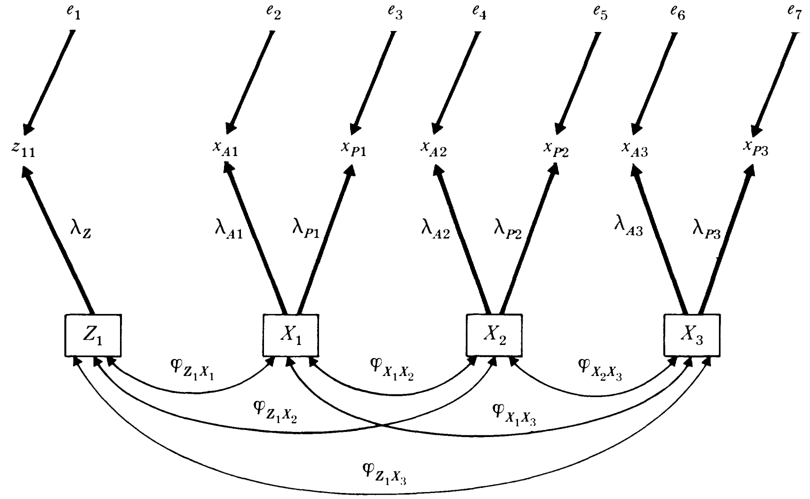


Figure 3. A Two-Variable Three-Wave *CFA* Measurement Model.

the model are then estimated. Estimated parameters can be of two types: *free* parameters that are unknown and not assumed equal to any other unknown parameters in the model, and *constrained* parameters that are set equal to other unknown parameters. The consequences of restricting certain parameters can be tested for efficacy by a $\chi^2$ goodness of fit test in the case where we restrict more parameters than are needed for identification.

In the model presented in Figure 3, there are several zero entries in matrix $\Lambda$, indicated by the absence of direct effects from certain unobserved constructs to certain observed variables (for example, $X_3$ to $x_{P2}$). The prespecification of certain zero entries in the $\Lambda$ matrix allows a covariance between factors (or "constructs," in our terminology) to be identified and to be estimated. Thus a zero effect of $X_3$ on $x_{P2}$, for instance, is not evidence of a lack of covariance; rather, there is an association between $x_{P2}$ and $X_3$ due to the fact that $x_{P2}$ is a measure of another construct, which itself *is* directly related to $X_3$.

If we feel that a particular causal structure underlies the unobserved variances and covariances in Figure 3, and we have a time-ordering established for our variables, we can postulate an

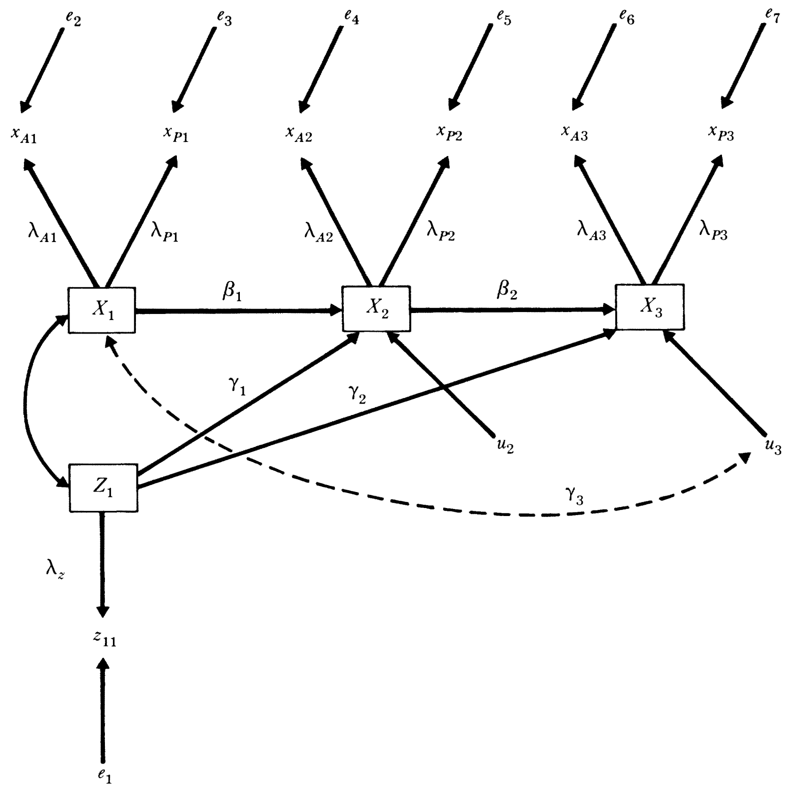


Figure 4. Explicated Causal Model for the *CFA* Model in Figure 3.

explicated *CFA* model as in Figure 4. This model will be important in our subsequent analysis. We will consider first the version of Figure 4 in which $\gamma_3 = 0$.[4] This is an explicated model containing an overidentifying restriction on the variance-covariance matrix for the constructs in addition to the restrictions on the $\Lambda$ matrix.

Following Jöreskog (1976), we can express the relationships in Figure 4 as follows, first for the relations between constructs and their measures:

[4] We specify $\gamma_3$ as the correlation between $X_1$ and $u_3$ and not the direct effect of $X_1$ on $X_3$ for reasons to be stated later.

$$\begin{bmatrix} x_{A2} \\ x_{P2} \\ x_{A3} \\ x_{P3} \end{bmatrix} = \mu + \Lambda_y \begin{bmatrix} X_2 \\ X_3 \end{bmatrix} + \begin{bmatrix} e_4 \\ e_5 \\ e_6 \\ e_7 \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} z_{11} \\ x_{A1} \\ x_{P1} \end{bmatrix} = \nu + \Lambda_x \begin{bmatrix} Z_1 \\ X_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \quad (4)$$

where $\mu$ and $\nu$ are the respective mean vectors for the measures $\Lambda_{y(p \times m)}$ = the factor loading matrix for all observed measures of endogenous unobserved constructs, $p$ = the number of observed measures and $m$ = the number of endogenous constructs, and $\Lambda_{x(q \times n)}$ = the factor loading matrix for all observed measures of exogenous unobserved constructs, $q$ = the number of observed measures, and $n$ = the number of exogenous constructs; and second, for the relations between constructs,

$$B \cdot \begin{bmatrix} X_2 \\ X_3 \end{bmatrix} = \Gamma \begin{bmatrix} Z_1 \\ X_1 \end{bmatrix} + \begin{bmatrix} u_2 \\ u_3 \end{bmatrix} \quad (5)$$

where $B_{(m \times m)}$ is a coefficient matrix for the relationships among endogenous constructs, and $\Gamma_{(m \times n)}$ is a coefficient matrix for relationships between exogenous and endogenous constructs.

In applying this model, we begin with the following assumptions: the expectation for all unobserved constructs is defined to be zero, that is, $E(Z_1) = E(X_1) = E(X_2) = E(X_3) = 0$; the expectation of the disturbances is also zero, that is, $E(u_2) = E(u_3) = 0$; the expectation of the observed variable vector is $\boldsymbol{\mu}$, $\boldsymbol{\nu}$; measurement errors are uncorrelated with each other and with all constructs in the model; exogenous constructs and disturbance terms for endogenous constructs are uncorrelated; and $B$ is nonsingular. We designate $\Phi_{(n \times n)}$ as the variance-covariance matrix of exogenous constructs, $\psi_{(m \times m)}$ as the variance-covariance matrix for disturbance terms, $\Theta_\varepsilon$ as a diagonal matrix of error standard deviations for the $p$ observed measures of endogenous constructs, and $\Theta_\delta$ as a diagonal matrix of error standard deviations for $q$ observed measures of exogenous constructs. We will then obtain the population variance-covariance matrix for the observed measures $(x_{A2}, x_{P2}, x_{A3}, x_{P3}, z_{11}, x_{A1}, x_{P1})'$ as (Jöreskog, 1976):

$$\Sigma = \tag{6}$$

$$\left[\begin{array}{c|c} \Lambda_y(B^{-1}\Gamma\Phi\Gamma' B'^{-1} + B^{-1}\Psi B'^{-1})\Lambda_y' + \Theta_\varepsilon^2 & \text{symmetric} \\ \hline \Lambda_x\Phi\Gamma' B'^{-1}\Lambda_y' & \Lambda_x\Phi\Lambda_x' + \Theta_\delta^2 \end{array}\right]$$

We should make it explicit that the variance-covariance matrix of exogenous constructs is

$$T_1 = \mathrm{Var}\begin{pmatrix} Z_1 \\ X_1 \end{pmatrix} = \Phi$$

that the variance-covariance matrix of endogenous constructs is

$$T_2 = \mathrm{Var}\begin{pmatrix} X_2 \\ X_3 \end{pmatrix} = B^{-1}\Gamma\Phi\Gamma' B'^{-1} + B^{-1}\Psi B'^{-1}$$

and that the covariance matrix between exogenous and endogenous constructs is

$$T_3 = \mathrm{Cov}\left[\begin{pmatrix} Z_1 \\ X_1 \end{pmatrix}, \begin{pmatrix} X_2 \\ X_3 \end{pmatrix}\right] = E\left[\begin{pmatrix} Z_1 \\ X_1 \end{pmatrix} \quad (X_2 X_3)\right] = \Phi\Gamma' B'^{-1}$$

Then we write $\Sigma$ in simplified form:

$$\Sigma = \left[\begin{array}{c|c} \Lambda_y(T_2)\Lambda_y' + \Theta_\varepsilon^2 & \text{symmetric} \\ \hline \Lambda_x(T_3)\Lambda_y' & \Lambda_x(T_1)\Lambda_x' + \Theta_\delta^2 \end{array}\right]$$

$$= \begin{bmatrix} \Lambda_y & 0 \\ 0 & \Lambda_x \end{bmatrix} \cdot \begin{bmatrix} T_2 & \text{sym} \\ T_3 & T_1 \end{bmatrix} \tag{7}$$

$$\cdot \begin{bmatrix} \Lambda_y & 0 \\ 0 & \Lambda_x \end{bmatrix}' + \begin{bmatrix} \Theta_\varepsilon^2 & 0 \\ 0 & \Theta_\delta^2 \end{bmatrix}$$

which is in a form analogous to our expression for $\Sigma$(equation (2)) for the unexplicated *CFA* model.

Our object is to fit this matrix to the corresponding sample variance-covariance matrix $S$ by an appropriate choice of parameter estimates. We use the following fitting function (Jöreskog, 1976):

$$F = \tfrac{1}{2}[\log|\Sigma| - \log|S| + \text{tr.}\,(S\Sigma^{-1}) - (p + q)] \tag{8}$$

which will be minimized with respect to the elements in

$$\Lambda_y, \Lambda_x, B, \Gamma, \Phi, \Psi, \Theta_\varepsilon, \Theta_\delta$$

With the additional assumption of a multivariate normal distribution for the observed variable vector, the assumptions stated earlier ensure that minimizing $F$ is equivalent to maximizing the likelihood of the sample values; that is, our parameter estimates will be *ML*-estimates with optimal large-sample properties. This additional assumption is not an unreasonable approximation for the data we will use, and our sample size seems large enough ($N = 932$) for large-sample properties to be relevant. The model as represented in Figure 4 may be conveniently estimated using the fitting function above via the Jöreskog and Sörbom computer program LISREL (1976b). For the case in which we impose restrictions on the relationships among constructs, the *ML*-estimates have the property of optimal efficiency compared to the ordinary least squares estimates. In some cases in which this part of the model is just-identified, the estimates are in fact the same. LISREL can be applied as easily in both situations. However, the method of analysis used by LISREL is especially advantageous in the case that the researcher has theoretical reasons for imposing restrictions *a priori* on relationships among constructs.

Using LISREL, we can also obtain a $\chi^2$ goodness of fit test (see Jöreskog, 1973, p. 90) for any overidentified model, with degrees of freedom equal to the number of distinct elements in $S[(p + q) \cdot (p + q + 1)/2]$ minus the number of different parameters to be estimated. The $\chi^2$ value for a model is determined by both the sample size and the minimum value of our fitting function $F_0$ as follows:

$$\chi^2_{df} = 2(N - 1) \cdot F_O \tag{9}$$

This means that we are testing the null hypothesis ($H_0$) that the $\Sigma$ matrix for a particular constrained model is correct as specified against the alternative ($H_1$) that an unconstrained $\Sigma$ matrix (which is, of course, still positive definite and symmetric) is the correct model. Obviously, the $\Sigma$ matrix under $H_1$ will reproduce the $S$ matrix perfectly, and $F_1 = 0$. In practice, this hypothesis testing procedure implies that larger values of $\chi^2$ reflect poorer fit for the model under $H_0$.

The $\chi^2$ test statistic is very often significant in samples of large size, suggesting rejection of the proposed model. Thus, it is used frequently in a descriptive fashion to decide between models on the basis of *relative* fit to the data. In applying this $\chi^2$, we will want to assess varying $\chi^2$/d.f. ratios across models in order to get a rough indication of fit per degree of freedom. For our sample size, we judge a ratio of around 5 or less as beginning to be reasonable, based on our experience in inspecting the sizes of residuals which accompany varying $\chi^2$ values. The residuals matrix is the difference between the $\Sigma$ matrix of the estimated variances and covariances implied by the model and the observed variance-covariance matrix ($S$). In order to give an indication of the kinds of residuals that accompany different $\chi^2$/d.f. ratios we will present such information at various points in the discussion.

## *EMPIRICAL ESTIMATION OF RECURSIVE MODELS FOR PANEL DATA*

### The Data

The empirical examples that we present to illustrate the estimation techniques suggested in the foregoing section make use of data from a longitudinal study of the effects of industrial development in a rural region in Illinois. The study was designed such that the data were collected on the same variables at three points in time (1966, 1967, and 1971) in each of two regions—an experimental region where a Jones & Laughlin Steel Corporation cold rolling mill was being constructed, and a control region where there was no such development taking place. For purposes of the present analyses the two samples are combined, producing a total of 932 cases.[5]

[5] See Summers *et al.* (1969) for a further description of the research setting and O'Meara (1966) for a detailed description of the sampling design.

Beginning with a traditional "social structure and personality" perspective, our original interest was in the reliability and stability of measured attitudes and the possible effects of selected social status factors on these attitudes over time. The initial interest in the stability of attitudes stemmed from a concern with whether social psychological variables (such as alienation and expressed social distance toward minority groups) are either highly volatile and, therefore, amenable to change, or are relatively stable over time. The following attitude scales were administered at all three points in time: an *alienation scale,* composed of items from Srole (1956) and items developed for this study by Summers; an *anomia subscale* of the above scale, from Srole (1956); a *powerlessness subscale* from the alienation scale, comprised of the items developed by Summers; a *religious importance scale,* measuring the degree to which the individual attaches importance to religion in his own life and composed of items from the scale developed by Snell and Middleton (1961); a *Latin American social distance scale* from Bogardus (1925), composed of a summated 6-item scale, each item was coded with a dichotomous acceptance-rejection designation; and a *Negro social distance scale,* a modification of items in Bogardus (1925), also a summed 6-item scale wherein each item was coded with a dichotomous acceptance-rejection designation.

In addition, two measures of the respondent's current socioeconomic position are also included in the present analysis: the respondent's educational attainment is expressed as years of schooling completed, and the status of the respondent's occupation is expressed in terms of the Duncan Socioeconomic Index (*SEI*), using the average *SEI* of the major census occupational grouping in which his occupation falls. For those not currently in the labor force, the *SEI* score for their last full-time job was used.

## Single-Indicator Models

In Table 1 we present the estimates of reliability and stability of each of the six attitude measures in a single-variable model using the computational formulae of both Heise (1969) and Wiley and Wiley (1970). It is remarkable that, given the time period involved (from one to four years), the stability of these social psychological variables appears to be rather high. It appears also that three different reliability coefficients estimated from the Wileys' model are

TABLE 1

Reliabilities and Stabilities Estimated from Both Single-Indicator Models for Six Selected Attitudes.

| | Heise Model | | | | Wiley and Wiley Model | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Reliabilities | Stabilities | | | Reliabilities | | | Stabilities | | |
| | $\alpha^{2}$* | $\beta_1$ | $\beta_2$ | $\beta_1 \cdot \beta_2$ | $\alpha_1^2$ | $\alpha_2^2$ | $\alpha_3^2$ | $\beta_1$ | $\beta_2$ | $\beta_1 \cdot \beta_2$ |
| Alienation | 0.75 | 0.97 | 0.81 | 0.79 | 0.78 | 0.75 | 0.76 | 0.95 | 0.80 | 0.76 |
| Anomia | 0.71 | 0.94 | 0.79 | 0.74 | 0.74 | 0.71 | 0.73 | 0.92 | 0.78 | 0.72 |
| Powerlessness | 0.64 | 0.98 | 0.81 | 0.79 | 0.69 | 0.64 | 0.67 | 0.94 | 0.80 | 0.75 |
| Religious importance | 0.68 | 0.82 | 0.77 | 0.63 | 0.63 | 0.68 | 0.50 | 0.85 | 0.91 | 0.77 |
| Latin American social distance | 0.43 | 1.15 | 0.96 | 1.11 | | | | | | |
| Negro social distance | 0.52 | 1.08 | 0.96 | 1.03 | | | | | | |

*In this table, estimated parameters will not include "^" notation, although it is to be understood that they are sample estimates.

usually similar over time points. Heise's assumption of constant reliability may indeed be reasonable for these variables. As Wiley and Wiley (1970) note, their time 2 reliability coefficient is equal to Heise's overall reliability—an assertion which is empirically demonstrated by the estimates presented in Table 1.

However, there are some problems with these results. In computing the parameter estimates in the two versions of the single-variable panel model for "Latin American social distance" and "Negro social distance" some unexpected results were obtained. For these variables some of Heise's stabilities exceed unity, and for corresponding coefficients in the Wiley models negative variances were obtained in the disturbances for the true score variables at times 2 and 3. This makes the application of the Wiley formulas problematic. Of course, given some sampling error, these results are possible and would then lead to the interpretation that the variable in question is, in fact, perfectly stable. However, such an interpretation is not convincing without first testing alternative explanations.

The above discussion suggests the possibility that, at least for certain variables, single-variable models may be inappropriate in the sense that some of the assumptions are not met and that they are, therefore, misspecified in important ways. In particular, we are concerned first with the possibility that important causal variables are excluded from these single-variable models (that is, assumption (3), noted earlier, is violated).[6] This circumstance is not, however, limited to the cases in which the stability estimates exceed unity in Heise's model. Even those stability estimates that are within reasonable limits may be biased upward, and this is a possibility we wish to explore. For instance, the estimated stabilities for alienation in Table 1 are 0.95 to 0.97 for $\beta_1$ and 0.80 to 0.81 for $\beta_2$ in the two models, respectively, and we feel these results demand closer inspection as well.

Before we proceed, however, it is important to illustrate the limitations imposed on alternative specifications by three-wave single-indicator models. Using LISREL, rather than the computational formulae, we can find reliability and stability estimates for the just-identified Wileys' model. In the case of the "Latin American

[6] We shall retain the lag-1 assumption in its original form, and choose to relax assumption (3) instead, where it is possible.

social distance" variable, for instance, we estimate the following parameters: $\alpha_1^2 = (0.654)^2$, $\alpha_2^2 = (0.639)^2$, $\alpha_3^2 = (0.727)^2$, $\beta_1 = 1.031$, and $\beta_2 = 0.981$. Both of the stability paths are close to one. The question is whether we can assess with other models the plausibility of these estimates. The first step might be to estimate a model as in Figure 2, adding a variable with proven theoretical and empirical relevance to the phenomenon of expressed social distance. Our inclination, then, is to assess the effects of "years of education" in this model, assuming that social distance towards minority groups (or "out groups") is itself a proxy for generalized prejudice. The correlations underlying this model are contained in Table 2. Without presenting the results of the whole model, we need only mention that effects of education as measured in 1966 on later states of social distance are essentially zero (where we apply the criterion for statistical significance that a coefficient should exceed twice its standard error).[7] Thus, changes in estimates of stability are minor.

In cases where the added variable's effects are unsuccessful, or only partially successful—and if we want to asses the potential influence of other causal variables which could distort the stability estimates via indirect paths between $X_1$ and $X_3$—we can specify the model so that $p_{X_1u_3} \neq 0$.[8] However, without imposing further restrictions, the model is under-identified in the single indicator case. It should be pointed out that there is more than one way to render the model just-identified, and we have no statistical criteria for deciding between these models. The resulting estimates are affected by the particular restrictions we choose, and the fit to the observed data is perfect for each model. It may be useful to impose some contrasting over-identifying restrictions on the model in order to test their efficacy with the $\chi^2$ test.

If we want to estimate $p_{X_1u_3}$ (a path which stands for other possible sources of further distortion to the stability estimates), we

[7]The LISREL program produces estimates of the asymptotic standard errors, and in a large sample such as ours the above ratio has an approximate normal distribution.

[8]This relaxes the third assumption we began with in applying LISREL. This can be done by specifying $u_3$ as an extra exogenous construct correlated with $X_1$ and causing $X_3$ with the unstandardized path fixed to one. It should be noted that this parameter does not constitute a special case for the *ML* method of estimation. *ML* produces a *consistent* and *asymptotically efficient* estimate of the parameter.

must impose restrictions on other parameters in the model. Researchers may have *a priori* reasons for choosing certain kinds of restrictions. In this case, we arbitrarily fix $\gamma_2 = \gamma_3 = 0$ for illustrative purposes, and this allows us to estimate the measured-to-unobserved regressions, given that they are constrained equal, and $p_{X_1 u_3}$. For ease in interpretation, we can discuss the standardized solution for this model for selected important parameters given below:

$$\hat{\beta}_1 = 0.99 \qquad \hat{\alpha}_3 = 0.66$$
$$\hat{\beta}_2 = 1.13 \qquad \hat{p}_{X_1 u_3} = -0.52$$
$$\hat{\alpha}_1 = 0.74 \qquad \hat{\gamma}_1 = -0.31$$
$$\hat{\alpha}_2 = 0.66$$

The estimate of $p_{X_1 u_3}$ does not, in fact, represent a significant path despite its size, due to the relatively large standard error in comparison to the corresponding unstandardized parameter ($Z$ value $= -1$). However, the $\chi^2$ for this model is 4.58 with 1 degree of freedom, denoting a relatively good fit. Note that the stabilities in this case are again close to one.

For a final model, we removed the education variable from the model—returning to the single-variable case—and assumed $p_{X_1 u_3} = 0$. It should be noted that we have no way in single-indicator models of attempting to estimate the possible influence of correlated measurement error. We suggest, therefore that our *best* estimate of the stability is, in fact, one. We add the restriction in our final model for social distance that $V(u_2) = V(u_3) = 0$. In effect, this assumes that the correlations among the constructs are each one. Forcing the model to lag-1 form (as in Figure 2) produces estimates of $\hat{\beta}_1 = 1$ and $\hat{\beta}_2 = 1$. The fit of this model is, of course, better than the last, with a $\chi^2$ value of 0.59 with 2 degrees of freedom (probability level for the model $= 0.74$). Given the alternatives, our conclusion is that the best estimate is that "Latin American social distance" is perfectly stable over this five-year period.

In sum, we feel that estimates of reliabilities and stabilities in single-indicator models are difficult to interpret in view of the limited alternatives available to these models. Our *ad hoc* procedures for exploring the reasonableness of Wileys' specification are, at best, risky. The inherent lack of flexibility is the most important limitation in applying these models, and we are led to suggest that the increased flexibility in multiple indicator models makes them more useful for estimating both reliability and stability parameters.

### Multivariate Panel Models with Multiple Indicators

In this section we will estimate several multivariate models for panel data in which we incorporate, as above, parameters representing the magnitude of measurement error and the stability of the variables over time. Now, however, we wish to include a variety of forms and extensions of the general model presented in Figure 4 and discussed previously. We will focus on one variable with multiple indicators (in this case, alienation) in successive specifications for the purpose of facilitating comparisons across models. The models which we present begin to specify more completely the relationships between indicators of alienation and other variables. In placing alienation in a substantive context, we wish to assume a "social structure and personality" perspective. Our interest is in the possible consequences that position in the social structure can have for the individual's psychological functioning. This is, of course, a classic sociological issue, one which is pertinent in a search for causal variables which can be added to single-variable models involving attitudes.

We have included two aspects of a person's position in the social structure—educational achievement and occupational status—as a way of placing alienation in a more completely specified, albeit rudimentary, theoretical context. Although there may be important mediating processes which more thoroughly explain how such aspects of social structure affect these variables, our main interest here is in estimating their total effects.

The correlation matrix for all observed variables that appear in our earlier analyses of single-indicator models, and all subsequent analyses, is presented in Table 2. It is apparent from a perusal of this matrix that there are consistent patterns of relatively strong negative correlations between the selected attitude variables, on the one hand, and the two measures of social position, on the other. These results are consistent with the findings of a number of other studies (Thompson and Horton, 1960; Middleton, 1963; Photiadis and Schweiker 1971; Simpson, 1970; and Anderson, 1973).

TABLE 2

Correlation Matrix for all Attitude and Status Variables Used.

| | | $X_1$ | $X_2$ | $X_3$ | $X_4$ | $X_5$ | $X_6$ | $X_7$ | $X_8$ | $X_9$ | $X_{10}$ | $X_{11}$ | $X_{12}$ | $X_{13}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Education (1966) | $X_1$ | | | | | | | | | | | | | |
| *SEI* (1966) | $X_2$ | 0.54 | | | | | | | | | | | | |
| Anomia (1966) | $X_3$ | −0.34 | −0.25 | | | | | | | | | | | |
| Powerlessness (1966) | $X_4$ | −0.41 | −0.30 | 0.70 | | | | | | | | | | |
| Latin American social distance (1966) | $X_5$ | −0.25 | −0.21 | 0.27 | 0.31 | | | | | | | | | |
| *SEI* (1967) | $X_6$ | 0.53 | 0.82 | −0.26 | −0.31 | −0.14 | | | | | | | | |
| Anomia (1967) | $X_7$ | −0.36 | −0.30 | 0.67 | 0.55 | 0.26 | −0.30 | | | | | | | |
| Powerlessness (1967) | $X_8$ | −0.41 | −0.29 | 0.55 | 0.63 | 0.28 | −0.33 | 0.66 | | | | | | |
| Latin American social distance (1967) | $X_9$ | −0.16 | −0.19 | 0.27 | 0.26 | 0.49 | −0.11 | 0.25 | 0.28 | | | | | |
| *SEI* (1971) | $X_{10}$ | 0.54 | 0.81 | −0.25 | −0.25 | −0.21 | 0.78 | −0.28 | −0.26 | −0.18 | | | | |
| Anomia (1971) | $X_{11}$ | −0.35 | −0.29 | 0.53 | 0.50 | 0.27 | −0.32 | 0.56 | 0.47 | 0.29 | −0.31 | | | |
| Powerlessness (1971) | $X_{12}$ | −0.37 | −0.28 | 0.42 | 0.51 | 0.24 | −0.26 | 0.44 | 0.52 | 0.23 | −0.28 | 0.67 | | |
| Latin American social distance (1971) | $X_{13}$ | −0.22 | −0.16 | 0.26 | 0.26 | 0.47 | −0.14 | 0.25 | 0.27 | 0.41 | −0.19 | 0.29 | 0.28 | |
| | Mean | 10.90 | 37.49 | 13.70 | 14.85 | 1.11 | 36.72 | 13.61 | 14.76 | 0.89 | 37.47 | 14.13 | 14.90 | 0.78 |
| | S. D. | 3.10 | 21.22 | 3.60 | 3.31 | 1.40 | 21.00 | 3.44 | 3.06 | 1.27 | 20.98 | 3.54 | 3.16 | 1.25 |

We can begin by estimating a baseline model for alienation,[9] which can be used as a starting point for comparison. This model is represented as a sub-model of Figure 4, including all paths *except* those involving $Z_1$, $z_{11}$, and $e_1$. Also, we assume that $p_{X_1 u_3} = \gamma_3 = 0$. The underlying model involving $X_1$, $X_2$, $X_3$, and their indicators is a familiar one and in various forms has been discussed elsewhere (Costner, 1969; Blalock, 1970). We need only remind the reader that it is overidentified. Each alienation construct at each point in time ($X_1$ = alienation (1966); $X_2$ = alienation (1967); $X_3$ = alienation (1971)) has two indicators, an anomia sub-scale ($x_{A1}$, $x_{A2}$, $x_{A3}$) and a powerlessness sub-scale ($x_{P1}$, $x_{P2}$, $x_{P3}$).

We should lay out our general strategy in estimating this and subsequent models. Of course, the first issue is one of identification. By identification we specifically mean that no two sets of distinct parameter values should be able to produce the same $\Sigma$ matrix.[10] Before attempting to estimate any model, one should check for identification by ensuring analytically that a solution exists for every distinct parameter in terms of the variance-covariance elements of the population $\Sigma$ matrix. We will not describe this process for every model (see Jöreskog and Sörbom, 1976a, for a discussion of identification in selected panel models). However, the specification consequences which follow from consideration of identification issues will be made explicit where necessary.[11]

In all the following models we analyze, we will incorporate the assumption that the unstandardized regressions of observed measures on constructs (the so-called "factor loadings" — see Werts, Linn, and Jöreskog, 1974, p. 273) are equal at each point in time for

[9] These models implicitly take a stand on the conceptualization of alienation that must be mentioned. There has been controversy as to the generality versus specificity of the alienation concept. Our position is that anomia and powerlessness can be interpreted as indicators of the more general and abstract construct of alienation, even though anomia and powerlessness can be distinguished as separate dimensions by some factor analytic criteria.

[10] Equivalently, in the case of *exact* multinormality of the observed variables, the model is not identified if for a given sample the same likelihood value can result from different sets of values for our parameters.

[11] We discourage using the LISREL program itself to decide issues of identification. It is true that if the program suceeds in computing standard errors of the estimates this is a good indication that the model is identified. However, this may be a costly trial-and-error procedure. Also, the program does not tell the user what the implications of different restrictions are in terms of identifying the model.

each indicator of alienation. In Figure 4 this means constraining $\lambda_{A1} = \lambda_{A2} = \lambda_{A3}$ and $\lambda_{P1} = \lambda_{P2} = \lambda_{P3}$. This regression weight $\lambda$ states the relationship between the unit of measurement of the observed variable and that of the construct (Werts,Linn, and Jöreskog, 1974, p. 273). In the case of the anomia measure, for instance, the equations are:

$$x_{A1} = \mu_{A1} + \lambda_{A1} X_1 + e_2$$

$$x_{A2} = \mu_{A2} + \lambda_{A2} X_2 + e_4$$

$$x_{A3} = \mu_{A3} + \lambda_{A3} X_3 + e_6$$

The $\lambda$s are unstandardized regression coefficients (and when standardized they correspond to validity coefficients). The equality assumption stated above has a specific empirical interpretation: for a given alienation true score $X$, a unit change in alienation will produce the same amount of change in $x_{A1}$, $x_{A2}$, and $x_{A3}$. This amounts to specifying that each of the measures of alienation are the "same" measures across time. We include this set of restrictions for reasons of parsimony, and in order to emphasize that we feel successive applications of the same measure are tapping the same construct. These restrictions are not necessary for identification and are testable by the $\chi^2$ test in the case of multiple indicators.

The earlier definition of reliability given by equation (1) for the single-indicator model should be modified for the more general multiple indicator case. Here we define the reliability as:

$$\alpha^2 = \frac{\lambda^2 \sigma_X^2}{\lambda^2 \sigma_X^2 + \sigma_E^2} \tag{10}$$

This definition is the same as the previous expression (where $\lambda$ is implicitly assumed to be 1) in the sense that the ratio is the portion of the observed variance "explained" by the construct and it is, in fact, the square of the correlation between the observed measure and its construct. Thus, the square root of this expression is the validity (defined operationally as the correlation between an observed measure and its construct). The definition we use points out the fact that $\lambda$ and $\sigma_X^2$ cannot be identified separately in the classical test theory model underlying our earlier definition for the single-indicator case. Note that, although we have restricted these loadings to be equal, the true score variances, the measurement errors—and thus the reliabilities—can vary over time.

Throughout our analyses we will fix the scale of the alienation construct by setting the variance of $X_1$ to one. The scale for $X_2$ and $X_3$ is then determined by setting construct loadings equal at each point in time. The variance of the endogenous constructs $X_2$ and $X_3$ may, thereby, vary in relation to the variance of $X_1$. The endogenous variances are, in fact, determined by the model, since

$$V(X_2) = \beta_1^2 V(X_1) + \sigma_{u_2}^2$$
$$V(X_3) = \beta_2^2 V(X_2) + \sigma_{u_3}^2$$

in the Figure 4 sub-model involving alienation measures only. Thus we should not assume these variances are known (for example, equal to 1), but rather let the model generate estimates of these variances.

To summarize, the consequences of these features of our application of the explicated *CFA* model are important in comparing multiple-indicator panel models with the single-indicator versions. The Heise model necessitates the assumption of equal reliability at three points in time, and this assumption implies that the ratio of true score to observed variance remains invariant; either due to no change in the true score variance and error variance, or to a change in both which happens not to effect the ratio (Wiley and Wiley, 1970, pp. 112–113). This rather restrictive assumption was relaxed in the Wileys' case to allow for changes in reliability, but it was necessary in their case to assume constant error variance even though the true score variance could vary. With models as in Figure 4, however, we can allow both error variance and true score variance to vary for the variables with multiple indicators; and, thus, reliability will vary as a function of changes in both. This is especially useful in estimating reliabilities of the same measure in different populations and across different models in the same population.

The proper data matrix to use as input in these analyses is the sample variance-covariance matrix *S*. It will be true that for many panel models this will be equivalent to using the correlation matrix *R* as input. In this case the model is said to be scale-free; that is, "a change in the unit of measurement in one or more of the observed variables can be appropriately absorbed by a corresponding change in the parameters" (Werts, Linn, and Jöreskog, 1971, p. 404). The models we estimate cannot be scale-free because of the equality restrictions on the loadings for similar measures. This means that if we had used the $R$ matrix instead of the $S$ matrix, the results of our models would have been different in terms of fit and in terms of the specific parameter estimates—possibly leading to quite different interpretations.

This does not mean that we are uninterested in the standardized estimates of the parameters of our panel models. In order to produce reliability and stability estimates we re-scale the unstandardized estimates produced by an analysis of the $S$ matrix. The standardized solution we report sets the construct variances at 1, so that the regression coefficients for relations between constructs are path coefficients. For the regressions of observed measures on constructs, we report the validities, that is, the path coefficients where the measured variances are also set to 1. Note that this standardized solution does not reproduce the same $\Sigma$ matrix as the unstandardized version. It is scaled so as to reproduce the input correlation matrix $R$ with the same overall fit, but it is *not* the same standardized solution that would have been produced from a direct analysis of the $R$ matrix. Our standardized solution is the accurate description of the relationships in the model when the model is not scale-free.

The estimates for the three-wave two-indicator model for alienation are presented in the first two columns of Table 3, under the heading of sub-model IA. Table 3 includes estimates from a succession of models which are all varieties of Figure 4, and the parameters labeled in Figure 4 are listed in Table 3. Models are numbered according to the following scheme: roman numerals are changed due to respecification of Figure 4 by either adding or changing a variable, and letters A and B stand for whether the assumption is made that $p_{X_1u_3} = \gamma_3 = 0$ or it is free in the model.

Recalling our definition of reliability, we can illustrate the relationship between the unstandardized loadings $\lambda$ and the standardized loadings $\lambda^*$. Given, for example, $\hat{\lambda}_{P2} = 2.66$, $\hat{V}(X_2) = 0.823$, and the estimated observed variance of $x_{P2} = 9.3636$, we can calculate the reliability as follows, using equation (10):

$$\hat{\alpha}^2 = \frac{\hat{\lambda}^2 \hat{\sigma}_X^2}{\hat{\lambda}^2 \hat{\sigma}_X^2 + \hat{\sigma}_E^2} = \frac{\hat{\lambda}^2 \cdot \hat{V}(X_3)}{\hat{V}(x_{P2})}$$

$$= \frac{(2.66)^2 \cdot 0.823}{9.3636} = 0.622$$

We note that $\hat{\alpha} = \sqrt{0.622} = 0.787$, which is the reported $\hat{\lambda}^*_{P2}$, the validity coefficient.

The reported validities range from $\hat{\lambda}^*_{A1} = 0.859$ to $\hat{\lambda}^*_{P2} = 0.787$, and in general the validities for the anomia measure are slightly higher. The two stabilities estimated by this model are $\hat{\beta}^*_1 = 0.897$ and $\hat{\beta}^*_2 = 0.768$. These are, of course, the total estimated correlations $r_{X_1X_2}$ and $r_{X_2X_3}$ as well, with $r_{X_1X_3}$ estimated as 0.689. Though these estimates are the standard we will use in comparisons with subsequent models, we should point out first the difference between the single-indicator stabilities for alienation, anomia, and powerlessness and these stabilities in the multiple-indicator case. The estimates of $\beta_1$ and $\beta_2$ are, in fact, slightly lower here—although not markedly so (see Table 1). This model produces a $\chi^2 = 184.08$ with 9 d.f., which results in a descriptive fit ratio of 20.45. According to most criteria, this is an unsatisfactory fit.

The upper diagonal of Table 4 gives the residuals for model IA($\hat{\Sigma} - S$) expressed as a signed proportion of the corresponding sample variances and covariances $S$. We have avoided reporting the residuals as correlations since our $\hat{\Sigma}$ matrix is attempting to reproduce the observed variances as well. Also, however, the raw residuals given by the differences between corresponding elements in $\hat{\Sigma}$ and $S$ have little intuitive value for purposes of interpretation. Therefore, we have divided each residual by the absolute value of the corresponding element in $S$ in order to present residuals as a proportion of input variance or covariance unexplained by the model. It should be made clear at this point that we do not advocate using the residuals as evidence for modifications in specification (see Costner and Schoenberg (1973) and Sörbom (1975) for discussions of the problems inherent in such a procedure), and we do not use the residuals for these purposes in this chapter. Residuals are presented in order to asses the *overall* magnitude of unexplained covariances and variances in the model. We should emphasize that our fitting function, Equation 8, is not aiming at the simple minimization of the $(\Sigma_{ij} - S_{ij})/|S_{ij}|$ elements, but at the minimization of $F_0$. Of course, the two are related, and we compare residuals to varying $\chi^2$/d.f. ratios (where $\chi^2$ for a given $N$ is determined by $F_0$) in order to assess the level of correspondence between them, and in order to show what types of residuals accompany models whose fit we deem as unsatisfactory or satisfactory.

The residuals for model IA are often near ten percent and range up to fifteen percent of the original $S$ covariance elements. Although these levels are not extremely high, the magnitude of these residuals do suggest that important relationships among these variables are presently left unspecified. With a large sample size such as ours we can explore some of these possibilities without inordinately capitalizing on chance.

We choose to specify "real" alternatives to this model before including "hypothetical" alternatives. Thus, we include first the causal influence of variables which we have measured and which have substantive relevance to the phenomenon of alienation before including paths involving unmeasured alternatives, $\gamma_3 \neq 0$ for example. We estimate now a model for Figure 4 in which $Z_1$ is educational attainment (unobserved) and $z_{11}$ is the measured score for this variable. According to this model, the education construct is assumed to be perfectly correlated over this time period—thus, the 1967 and 1971 measures are not included in the model. In order to render the model identified, some assumptions must be made about the relationship between $Z_1$ and $z_{11}$. For this purpose we assume that the reliability of the education measure is known *a priori,* based on the Siegel and Hodge (1968) reported reliability of educational attainment of 0.9332.[12] In the present data our assumption of a perfect cross-temporal education correlation given this reliability is not inappropriate: the validity is $\sqrt{0.9332} = 0.966$, and the cross-temporal correlations between measures are 0.956, 0.915, and 0.910; which would result in stability estimates near unity in the Heise model. In order to estimate $\sigma_{e1}$ in Figure 4, we fix $V(Z_1) = 1$ and constrain $\lambda_z$ so that it corresponds to a validity of 0.966 when transformed to $\lambda_z^*$. Here we do this by solving for $\lambda$ in the formula:

$$\alpha = \frac{\lambda \sigma_X}{\sigma_0}$$

[12] Siegel and Hodge (1968) report a test-retest correlation of $(0.966)^2 = 0.9332$ for educational attainment based on matches from the 1960 census and the March, 1962 Current Population Survey.

derivable from equation (10), and where $\sigma_0$ is the population standard deviation of the observed measure. Since we know that for an exogenous indicator in this situation the $ML$-estimate for $\sigma_0$ is equal to the corresponding sample standard deviation:

$$0.966 = \hat{\lambda}(1)/3.10$$

$$\hat{\lambda} = (0.966) \cdot (3.10) = 2.99$$

In this and the following models we should note that, as usual, comparisons of absolute sizes of parameters should be between unstandardized coefficients (see Schoenberg, 1972). However, the standardized estimates are more interpretable in many situations and are the source of our validity and stability coefficients. We shall discuss them as well.

In model IIA, we assess the causal influence of education on alienation in 1967 and 1971 and thereby hope to account for a portion of the unanalyzed stability paths in model IA via indirect paths betweeen $X_1$ and $X_2$, and $X_1$ and $X_3$. The result is a corresponding drop in the stabilities of about 0.04 (see Table 3). This we assess to be important enough to justify the inclusion of education in the model—since the education effects denoted by $\gamma_1$ and $\gamma_2$ are significant in the negative direction, and education and alienation in 1966 are rather strongly correlated, $r_{X_1 Z_1} = -0.46$. In other words, the model as it stands lends some support to a social structure and personality hypothesis about attitude influence, although it does not address the possibility of reciprocal influence—and there are numerous possibilities for spuriousness in $\gamma_1$ and $\gamma_2$. The descriptive fit ratio here is $200.08/12 = 16.67$, which reflects an overall improvement in fit compared to model IA.

The residuals for this model are presented in the lower diagonal of Table 4 (without parentheses). For the relationships among alienation measures, this model is no more effective than IA: some residuals are slightly lower and others are slightly higher for model IIA. Also, residuals for the relationship between education and alienation measures are, on the average, not much better. There are clearly not large differences in residuals corresponding to the change in the descriptive fit ratio. In short, we feel the fit is still, on the whole, unsatisfactory—despite the fact that certain residuals are fairly low. We should point out that both of these first two models

TABLE 3

Parameter Estimates for a Variety of Specifications of Figure 4 for the Alienation Construct, $N = 932$.[1, 2]

| | Sub-Model IA | | Model IIA | | Model IIB | | Model IIIA | | **Model IIIB** | |
|---|---|---|---|---|---|---|---|---|---|---|
| | U | S | U | S | U | S | U | S | **U** | **S** |
| $\lambda_z$ | | | 2.990[f] | 0.967 | 2.990[f] | 0.968 | 19.820[f] | 0.934 | **19.82[f]** | **0.934** |
| $\lambda_{A1}$ | 3.120[a] | 0.859 | 3.090[a] | 0.850 | 3.090[a] | 0.849 | 3.110[a] | 0.857 | **3.110[a]** | **0.856** |
| $\lambda_{P1}$ | 2.660[b] | 0.813 | 2.690[b] | 0.822 | 2.690[b] | 0.823 | 2.670[b] | 0.815 | **2.670[b]** | **0.815** |
| $\lambda_{A2}$ | 3.120[a] | 0.824 | 3.090[a] | 0.815 | 3.090[a] | 0.822 | 3.110[a] | 0.822 | **3.110[a]** | **0.830** |
| $\lambda_{P2}$ | 2.660[b] | 0.787 | 2.690[b] | 0.797 | 2.690[b] | 0.803 | 2.670[b] | 0.789 | **2.670[b]** | **0.795** |
| $\lambda_{A3}$ | 3.120[a] | 0.846 | 3.090[a] | 0.836 | 3.090[a] | 0.836 | 3.110[a] | 0.844 | **3.110[a]** | **0.845** |
| $\lambda_{P3}$ | 2.660[b] | 0.794 | 2.690[b] | 0.803 | 2.690[b] | 0.803 | 2.670[b] | 0.796 | **2.670[b]** | **0.795** |
| $\beta_1$ | 0.814 | 0.897 | 0.772 | 0.852 | 0.764 | 0.836 | 0.786 | 0.866 | **0.778** | **0.850** |
| $\beta_2$ | 0.806 | 0.768 | 0.756 | 0.721 | 0.513 | 0.493 | 0.772 | 0.736 | **0.518** | **0.497** |
| $\gamma_1$ | | | −0.091 | −0.100 | −0.095 | −0.103 | −0.083 | −0.091 | **−0.085** | **−0.094** |
| $\gamma_2$ | | | −0.093 | −0.098 | −0.201 | −0.211 | −0.080 | −0.084 | **−0.170** | **−0.178** |
| $\gamma_3$ | 0.000[f] | 0.000 | 0.000[f] | 0.000 | 0.174 | 0.270 | 0.000[f] | 0.000 | **0.202** | **0.309** |
| $\mathrm{Cov}(X_1, Z_1)$ | | | −0.460 | −0.460 | −0.459 | −0.459 | −0.347 | −0.347 | **−0.347** | **−0.347** |
| $V(X_1)$ | 1.000[f] | 1.000 | 1.000[f] | 1.000 | 1.000[f] | 1.000 | 1.000[f] | 1.000 | **1.000[f]** | **1.000** |
| $V(Z_1)$ | | | 1.000[f] | 1.000 | 1.000[f] | 1.000 | 1.000[f] | 1.000 | **1.000[f]** | **1.000** |
| $\sigma_{u2}$ | 0.400 | 0.442 | 0.391 | 0.429 | 0.420 | 0.459 | 0.392 | 0.432 | **0.423** | **0.463** |
| $\sigma_{u3}$ | 0.611 | 0.641 | 0.603 | 0.694 | 0.645 | 0.678 | 0.605 | 0.636 | **0.655** | **0.687** |
| $\sigma_{e1}$ | | | 0.784 | 0.253 | 0.781 | 0.252 | 7.580 | 0.357 | **7.580** | **0.357** |
| $\sigma_{e2}$ | 1.860 | 0.512 | 1.920 | 0.527 | 1.920 | 0.529 | 1.870 | 0.516 | **1.880** | **0.517** |
| $\sigma_{e3}$ | 1.910 | 0.583 | 1.860 | 0.569 | 1.860 | 0.568 | 1.900 | 0.580 | **1.890** | **0.579** |
| $\sigma_{e4}$ | 1.950 | 0.567 | 1.990 | 0.579 | 1.960 | 0.570 | 1.950 | 0.569 | **1.910** | **0.557** |
| $\sigma_{e5}$ | 1.890 | 0.617 | 1.850 | 0.605 | 1.830 | 0.596 | 1.890 | 0.614 | **1.860** | **0.606** |
| $\sigma_{e6}$ | 1.870 | 0.533 | 1.930 | 0.549 | 1.930 | 0.549 | 1.880 | 0.536 | **1.880** | **0.535** |
| $\sigma_{e7}$ | 1.940 | 0.608 | 1.900 | 0.596 | 1.900 | 0.596 | 1.930 | 0.606 | **1.940** | **0.606** |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $V(X_2)$ | 0.823 | 1.000 | 0.822 | 1.000 | 0.836 | 1.000 | 0.823 | 1.000 | 0.838 | 1.000 |
| $V(X_3)$ | 0.908 | 1.000 | 0.904 | 1.000 | 0.904 | 1.000 | 0.907 | 1.000 | 0.908 | 1.000 |
| $\chi^2$ | 184.080 | | 200.080 | | 195.930 | | 190.200 | | 185.670 | |
| d.f. | 9.000 | | 12.000 | | 11.000 | | 12.000 | | 11.000 | |
| Probability level | 0.000 | | 0.000 | | 0.000 | | 0.000 | | 0.000 | |

f = fixed parameters
a = parameters constrained equal
b = parameters constrained equal
U = unstandardized solution
S = standardized solution

[1] All coefficients in this table exceed twice their standard error.

Variable designations common to all models are: $X_1$ = Alienation (1966); $X_2$ = Alienation (1967); $X_3$ = Alienation (1971); $x_{A1}$ = Anomia (1966); $x_{P1}$ = Powerlessness (1966); $x_{A2}$ = Anomia (1967); $x_{P2}$ = Powerlessness (1967); $x_{A3}$ = Anomia (1971); $x_{P3}$ = Powerlessness (1971).

For models IIA and IIB: $Z_1$ = Education (unobserved) 1966 and $z_{11}$ = years of education in 1966 (measured).

For models IIIA and IIIB: $Z_1$ = *SEI* (unobserved) 1966 and $z_{11}$ = *SEI* score in 1966.

[2] The "^" notation is not included for estimated parameters, but it is to be understood that all parameters are sample estimates.

TABLE 4

Relative Evaluation of Fit of Models IA, IIA, and IIIA to Observed Variances and Covariances ($S$ Matrix), Residuals expressed as a Proportion of $S$ Values

$$\hat{\Sigma}_{ij} - S_{ij}/|S_{ij}|$$

Model IA above diagonal; Model IIA below diagonal without parentheses; Model IIIA below diagonal in parentheses.

| | $z_{11}$ | $x_{A1}$ | $x_{P1}$ | $x_{A2}$ | $x_{P2}$ | $x_{A3}$ | $x_{P3}$ |
|---|---|---|---|---|---|---|---|
| Covariances | | | | | | | |
| $z_{11}$ | | | | | | | |
| $x_{A1}$ | −0.1217 (−0.1213) | | −0.0051 | −0.0455 | 0.1149 | −0.0547 | 0.1398 |
| $x_{P1}$ | 0.1173 (0.1291) | −0.0029 (−0.0048) | | 0.0786 | −0.0970 | −0.0704 | −0.1292 |
| $x_{A2}$ | −0.0752 (−0.0007) | −0.0648 (−0.0494) | 0.0812 (0.0794) | | −0.0170 | −0.0534 | 0.1512 |
| $x_{P2}$ | 0.0739 (0.0026) | 0.1175 (0.1156) | −0.0742 (−0.0922) | −0.0158 (−0.0167) | | 0.0815 | −0.0659 |
| $x_{A3}$ | −0.0365 (−0.0021) | −0.0714 (−0.0586) | −0.0659 (−0.0699) | −0.0736 (−0.0570) | 0.0827 (0.0823) | | 0.0053 |
| $x_{P3}$ | 0.0416 (0.0036) | 0.1453 (0.1405) | −0.1049 (−0.1246) | 0.1525 (0.1521) | −0.0435 (−0.0608) | 0.0033 (0.0048) | |
| Variances | | | | | | | |
| IA | | 0.0173 | −0.0215 | −0.0030 | 0.0035 | −0.0155 | 0.0210 |
| IIA | 0.0004 | 0.0183 | −0.0204 | −0.0038 | 0.0043 | −0.0157 | 0.0193 |
| IIIA | −0.0001 | 0.0182 | −0.0221 | −0.0039 | 0.0046 | −0.0155 | 0.0205 |

reproduce the variances in $S$ fairly well, with approximately 2 percent the highest proportion of unexplained variance.

Model IIB adds the possibility that $X_1$ is correlated with the disturbance term for $X_3$. With multiple indicators, we can identify this path without adding further restrictions to the model. There are at least two possible interpretations of a covariance between $X_1$ and $u_3$: another exogenous variable or set of variables presently excluded from the model could covary with $X_1$ and have a causal effect on $X_3$, or $X_1$ could affect $X_3$ directly. We do not have the space to explore these alternatives here, so we will leave $\gamma_3$ specified as in Figure 4. We see that $\hat{\gamma}_3^*$ for model IIB in Table 3 denotes a significant correlation between $X_1$ and $u_3$ ($= 0.27$); and the influence on our stabilities is important $-\hat{\beta}_1$ is only reduced to 0.764 from 0.772, but $\hat{\beta}_2$ drops from 0.756 to 0.513, amounting to a $-0.243$ change in this stability estimate. The effect of education on alienation at time 3 appears to be larger than its effect on alienation at time 2 in this model, but there is more than one possible interpretation of this result: either the one year lag underestimates the true causal lag for this relationship or there are factors excluded from the model prior to time 1 which covary with $X_1$ and $Z_1$ and affect $X_3$ directly (resulting in an inflated $\gamma_2$).

The resulting stability coefficients are now 0.836 and 0.493 respectively. The question is: how do we assess the improvement in fit of this model? In this case we compare the change in $\chi^2$ from model IIA to IIB, which *itself* is a $\chi^2$ variable with degrees of freedom equal to the difference in degrees of freedom of the two models, and not the descriptive fit ratio across the two models. This is true in comparing all "nested" models; that is, models where no new variables are specified, and we impose or relax constraints while holding specification in the rest of the model constant. Comparing model IIB to IIA, we see that the $\chi^2$ difference is:

$$\chi_1^2 = \chi_{12}^2 - \chi_{11}^2 = 200.08 - 195.93 = 4.05$$

meaning that the decrease in $\chi^2$ due to respecification is significant beyond the 0.05 level ($\chi^2_{1\text{d.f.}, p=0.05} = 3.84$), and that model IIB stands for an improvement in fit.[13]

[13] Of course, this significant $\chi^2$ value is merely a re-statement of the significant $Z$-value for the correlation between $X_1$ and $u_3$. This is true in our analyses when nested models differ only in one parameter that is fixed to zero in the first model.

Before proceeding to our next model, we must discuss other possible specifications of model IIB. There are other kinds of correlated error between disturbance terms and exogenous variables ($p_{X_1 u_2}$)and between disturbance terms($p_{u_2 u_3}$)which could affect stability estimates. However, $p_{X_1 u_2} \neq 0$ cannot be identified. The same possibility for indirect relationships through other variables between $X_2$ and $X_3$ would be reflected by $p_{u_2 u_3} \neq 0$, and this path can be identified. We cannot identify both $p_{u_2 u_3}$ and $p_{X_1 u_3}$ since the disturbance terms are free parameters in this model. We did estimate a variant of model IIB in which $p_{u_2 u_3}$ was not constrained to zero and $p_{X_1 u_3} = \gamma_3$ was constrained to zero. The results of this model are not reported in Table 3, but they are important to the ensuing discussion. The estimated correlation $\hat{p}_{u_2 u_3}$ is $-0.193$, which fails to be significant according to our criterion. Our expectation is that this "test" is not seriously affected by constraining $\gamma_3$ to zero and forcing this covariance through other paths in the model. The fit of the model reflected by the $\chi^2_{11\mathrm{df}}$ is 195.93, exactly the same as model IIB. This reflects the fact that the construct part of the model is just-identified in both cases. Given that we can specify either $p_{X_1 u_3} \neq 0$ or $p_{u_2 u_3} \neq 0$, and it turns out in the above case that $p_{u_2 u_3} = 0$, we choose to specify subsequent models with $p_{X_1 u_3}$ unconstrained as the alternative, and we will not discuss the possibilities connoted by $p_{u_2 u_3}$ further.

Model IIIA replaces education and its measure with *SEI* (unobserved) and its measure in 1966. For this model $Z_1$ in Figure 4 is now occupational status, as measured by Duncan's *SEI* Index ($z_{11}$). Again, our specification assumes that $Z_1$ is stable over the five-year period in the sample. Using the Siegel and Hodge (1968) reliability for occupational reporting of 0.8726, we estimate the autocorrelations of *SEI* over time to be near unity and, thus, we use the specification as in Figure 4. Again, we must constrain the construct loading for $z_{11}$, but in this case we want to set $\lambda_z$ such that $\lambda_z^* = 0.9341 = \sqrt{0.8726}$. By fixing $V(Z_1) = 1$, we calculate $\lambda_z$ to be 19.82. Loadings for anomia and powerlessness are again set equal.

The model estimates the stabilities for alienation to be higher than for Model IIA, although they are very close. Our hope in specifying occupational status as $Z_1$ was that this variable would improve on previous estimates of stability by accounting, in part, for $\gamma_3$ in Model IIB. In fact, *SEI* is slightly less correlated with alienation in 1966 ($-0.347$ compared to $-0.46$ for education), and its lagged effects are also slightly smaller. However, these effects are still significant, although we do not yet know how confounded they are with the education effects. The fit for this model as evaluated by our fit ratio is $190.12/12 = 15.85$, the best overall fit to this point.

We add the specification that $\gamma_3 \neq 0$ for model IIIB, and the results reflect the changes in estimates between models IIA and IIB. Most importantly, there is a significant reduction in the estimate of the stability of alienation between 1967 and 1971 ($\beta_2^*$ drops from 0.736 to 0.497). $\gamma_3$ is still significant in this model and is higher than for model IIB, reflecting the fact that *SEI* accounts for indirect effects less effectively. The $\chi^2$ change from model IIIA to IIIB of $-4.53$ denotes a significant improvement in fit for a model for the relationship between occupational status and alienation.

Finally in Table 3, we should note the similarity in validity estimates across all models estimated to this point. Changes in validities are slight, amounting to $\pm 0.01$. Model IA does as well as any in estimating these parameters. This invariance should be expected, however, since we have not changed either the measurement specification or the type of concomitant theoretical construct in the model. Estimates of these validities suggest that the reliabilities of similar measures across time may change slightly; and, of course, longer time periods could produce larger differences. Since differences in estimates of reliability for similar measures across time range up to 0.06 here, the ability to allow validity to vary is of some importance.

In Table 4 we compare the residuals for models IIA and IIIA in the lower diagonal. Proportional residuals for model IIIA are in parentheses, while entries for model IIA are not. In line with the similarity of the fit ratios for these two models, the residuals are also similar; the major difference being the residuals between $z_{11}$ and the alienation measures at time 2 and time 3. Column one shows the superiority of model IIIA in reproducing these covariances, which may account for the slightly better fit ratio for this model.

The logical next step is to estimate a model in which the effects of both education and occupational status are included simultaneously in order to assess the confounding in previous models. The model in Figure 5 can be used for this purpose. This model allows for some instability in *SEI* (designated as $S_1$ through

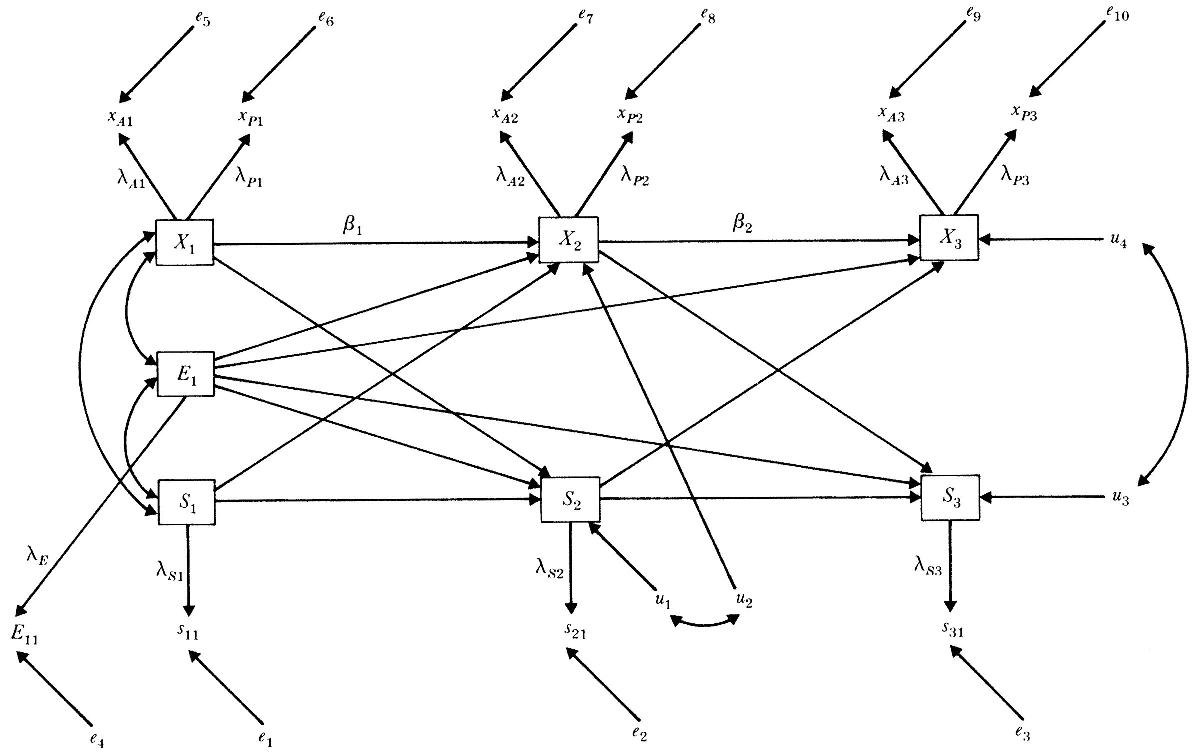


Figure 5. Explicated Causal Model for Alienation with Both Education and *SEI* Included as Separate Constructs, *SEI* at Three Time Points.

$S_3$ in the model). Thus, we include measures in 1966 ($s_{11}$), 1967 ($s_{21}$), and in 1971 ($s_{31}$). We must allow for the possibility here that education ($E_1$) will account for a portion of the previous stability estimates for *SEI.* This model will be useful for social researchers who might want to assume some instability in more than one variable in their panel models. Figure 5 represents just one type of cross-lagged model that can be used with panel data, and it is based on the Heise (1970) model extended to three time points.

We saw in Figure 4 that education and occupational status exhibited very similar effects on alienation. Figure 5 is mainly concerned with the size of the correlation between $E$ and $S$ variables, and the distribution of their individual effects. Without reporting the complete solution of this model, we do wish to summarize the estimates we derive. First, the correlation between education and occupational status is substantial, $r_{E_1 S_1} = 0.642$; and the consequences for the rest of the model are far-reaching. Lagged effects for education and *SEI* separately on alienation are all now nonsignificant, and in this sense their individual effects are similar. The estimates of $\beta_1$ and $\beta_2$ in this model are thus close to those for model IA. The model is unsatisfactory in that it does not describe parsimoniously the relationships between these variables. Evidence from this model and models II and III suggests that education and *SEI* do not affect alienation differently in this sample. Moreover, the argument could be advanced that the *SEI* and education measures are tapping the common construct of socioeconomic status, and that this construct explains the covariance between these measures.

This leads us to estimate a final set of models which are all variants of the model in Figure 6. This model argues that occupational status in 1966, $s_s$, and education in 1966, $s_E$, measure in common a construct $S_1$ we call socioeconomic status (*SES*). For certain substantive purposes this specification may be questionable; however, introducing *SES* as a summary concept here may, in fact, increase the interpretive value of the model since now we do not have to assume reliabilities for education and *SEI.* Also, this measurement specification is a possibility clearly suggested by previous models. The *SES* part of the model can be identified by fixing $V(S_1) = 1$, and allowing both $\lambda_S$ and $\lambda_E$ to be free. Also, we need not restrict $\sigma_{e1}$ and $\sigma_{e2}$.

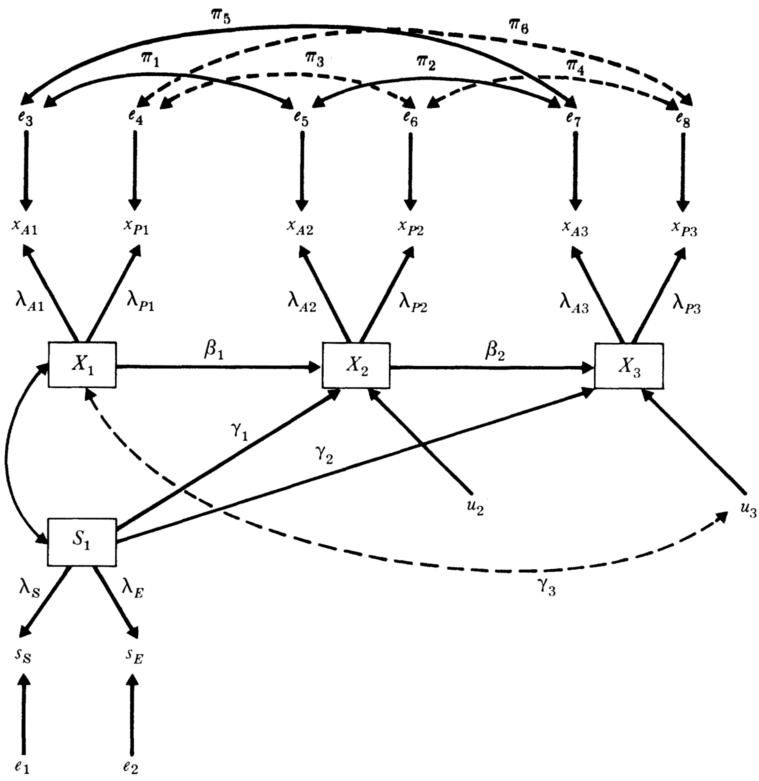


Figure 6. Final Explicated Model for the Relationship between *SES* and Alienation, with a Two-Indicator Specification for an *SES* Construct Assumed Perfectly Stable, 1966–1971.

## Panel Models with Multiple Indicators and Correlated Error between Measures

The other major change in specification in Figure 6 concerns the possibility of correlated measurement errors. Thus, we now wish to relax assumption (2) for the single-indicator models. These correlated errors stand for sources of nonrandom measurement error in the measures, such as a yeasaying or naysaying bias, "memory effects," etc. (Costner, 1969; Alwin, 1974). The first source is a real possibility in statements containing judgments, and the second becomes important when the time lag is not sufficient to create independence of successive measures. If this second type of problem is, in fact, the source of correlated error, we have no *a priori* reason for expecting correlated errors to be equal across successive lags in this model. Our error specification allows us to assess the influence of both sources of nonrandom error, although we do not separate the total covariance into different "methods bias" components. To represent these possibilities, we include all error correlations for similar measures across time. Thus, we are not detecting correlated errors by the more systematic stepwise procedure suggested by Sörbom (1975), which is intended to produce the best-fitting error model; rather, we impose the *a priori* restriction that correlated error should occur between the same measures only, and then we estimate the most flexible error model possible given this restriction. Changing the measurement specification with these error correlations can affect *both* validity and stability estimates and, indirectly, our estimates of other parameters.

We wish to emphasize that the correlated errors in Figure 6 can be identified without imposing any equality restrictions, which allows the differential time lag to have the effect it should. The identification issues which arise in this model are worthy of discussion since equality restrictions on correlated errors have often been assumed in identifying these errors over time, for example, $\pi_1 = \pi_2$ and $\pi_3 = \pi_4$. Also, separate parameters for $\pi_5$ and $\pi_6$ are sometimes assumed not to be identifiable in two-indicator models. First, consider a version of Figure 6 in which $S_1$ and its accompanying measures are excluded from the model. In making a preliminary inspection of the identifiability of the model, we introduce the possibility that a new $\lambda'_{A1} = \lambda_{A1} \cdot a$, where $a$ is any nonzero constant not equal to one. Our definition of identification demands that no two sets of parameters can reproduce the same population variance-covariance matrix $\Sigma$. However, we *can* compensate for the change in $\lambda_{A1}$ by a change in $\lambda_{P1}$ equal to $1/a$, thereby reproducing the same covariance between $x_{A1}$ and $x_{P1}$, and the same holds as a result of equivalent changes in $\lambda_{Ai}$ at later points in time. In order to reproduce the same cross-temporal covariance for the same indicators, we must compensate in the error covariances $\pi_i$, which is possible since they are all free parameters. Finally, in order to reproduce the same $\Sigma$ variances, we can compensate by the proper changes in the error variances since they are also all free. Since we can produce two sets of parameters for the model that both reproduce $\Sigma$, the model is not identified—even where we have equality restrictions for the construct loadings of the same indicators across time.

Now we can consider the consequences of adding $S_1$ to the model. For example, in order to leave the covariance between $x_{A1}$ and $s_S$ in $\Sigma$ intact, and assuming for the moment that we leave the variances and covariances for constructs the same, we would have to divide $\lambda_S$ by $a$. However, in order to avoid changing the covariance between $x_{P1}$ and $s_S$ we would have to *multiply* $\lambda_S$ by $a$. Therefore, we have a contradiction, implying that we cannot obtain the same $\Sigma$ with these two sets of parameters. We could, without changing $\lambda_S$, change the parameters for relationships between constructs in order to reproduce $\Sigma$, but this would change the reproduced covariances for alienation indicators. For instance, we could attempt to alleviate the above contradiction by compensating in $\mathrm{Cov}(X_1, S_1)$, but this would reproduce either the covariance between $s_S$ and $x_{A1}$ or $s_S$ and $x_{P1}$, but not both. For example, if we have divided $\lambda_S$ and $\lambda_{P1}$ by $a$ as above, then in order to reproduce $\mathrm{Cov}(s_S, x_{P1})$ we would have to multiply $\mathrm{Cov}(X_1, S_1)$ by $a^2$. However, this would again change the reproduced covariance between $s_S$ and $x_{A1}$. Now we know that none of the parameters at the first time point can be altered. Changes in the $\lambda$s at later points in time for the alienation measures result in similar contradictions in attempting to reproduce covariances with $s_S$. Using this approach, it seems unlikely that the same matrix could be reproduced by two different sets of parameters. A rigorous analysis of identifiability will prove this to be correct in this case. Therefore, the model becomes identifiable as we specify it due to the inclusion of the concomitant variable $S_1$ (of which only one indicator is necessary for identification pruposes). There are two indirect paths that the observed cross-temporal covariances between alienation indicators can take: through the constructs, and through the error correlations. $S_1$ helps to determine the $\lambda$s for these indicators, and we can thereby separate the indirect path through the constructs from the indirect path through the error terms—two components which are otherwise inseparable.

Our first specification of Figure 6, however, assumes there is no correlated error, as in previous models (that is, $\pi_1 = \pi_2 = \pi_3 =$

$\pi_4 = \pi_5 = \pi_6 = 0$). We include $\gamma_3 \neq 0$ because we know that it is significant in all versions of Figure 4. Results for this model are under model VA in Table 5. First we see that our validity for *SEI* in the *SES* construct is somewhat lower than what was assumed for the single-indicator construct in Figure 4. As a consequence of this $S_1$ specification, relationships between *SES* and alienation are affected: the correlation at time 1 is now $-0.525$, and the lagged effects of *SES* on alienation are also correspondingly larger than in Figure 4 versions, $\hat{\gamma}_1^* = -0.137$ and $\hat{\gamma}_2^* = -0.263$. As a result, the estimate of $p_{X_1 u_3}$ is a bit smaller, but still significant. The effect on the stabilities is also predictable.

The standardized estimates are slightly lower still, than for previous models, with $\hat{\beta}_1^* = 0.812$ and $\hat{\beta}_2^* = 0.467$. This is a reasonable model for assessing the influence of position in the social structure on a generalized attitude such as alienation, and we are beginning to produce stability estimates which are closer to expectation. We should expect, however, that for attitudes which are not object-specific there should be some stability over a five year period. The descriptive fit ratio is $200.64/16 = 12.54$, and we judge this specification to be somewhat superior to models IIB and IIIB in which education and *SEI* are considered separately. Despite the general interpretative use of this model, it is still unsatisfactory in terms of fit; and there are other specifications we must now consider. By changing the theoretical specification in various ways, we have reduced the fit ratio from 20.54 to 12.54. Considering the fact that changes at this level of specification have not been incorporated in the reliability/stability literature, this is an important improvement in fit. Of course, what we have illustrated with one concomitant variable states our point only to a degree, and other variables could by added to this model which would specify the context within which alienation changes over time more completely.

We turn finally to changes in measurement specification with model VB. In this model, $\pi_1$ through $\pi_6$ are free (as outlined above), but in all other respects the model is the same as model VA.[14] The results of this model are interesting in terms of estimating reliability and stability. First we inspect the results for $\pi_1$ through

[14] In Jöreskog and Sörbom (1976a) it is shown how correlated errors are specified in LISREL.

TABLE 5
Parameter Estimates for Various Specifications of Figure 6.[1,2]

| | Model VA | | Model VB | | Model VC | |
|---|---|---|---|---|---|---|
| | U | S | U | S | U | S |
| $\lambda_S$ | 13.670 | 0.645 | 13.500 | 0.636 | 13.480 | 0.635 |
| $\lambda_E$ | 2.600 | 0.838 | 2.630 | 0.848 | 2.640 | 0.850 |
| $\lambda_{A1}$ | 3.090[a] | 0.849 | 2.880[a] | 0.792 | 2.860[a] | 0.788 |
| $\lambda_{P1}$ | 2.690[b] | 0.823 | 2.900[b] | 0.886 | 2.920[b] | 0.889 |
| $\lambda_{A2}$ | 3.090[a] | 0.823 | 2.880[a] | 0.764 | 2.860[a] | 0.759 |
| $\lambda_{P2}$ | 2.690[b] | 0.803 | 2.900[b] | 0.863 | 2.920[b] | 0.871 |
| $\lambda_{A3}$ | 3.090[a] | 0.837 | 2.880[a] | 0.780 | 2.860[a] | 0.775 |
| $\lambda_{P3}$ | 2.690[b] | 0.802 | 2.900[b] | 0.859 | 2.920[b] | 0.866 |
| $\beta_1$ | 0.742 | 0.812 | 0.648 | 0.711 | 0.651 | 0.712 |
| $\beta_2$ | 0.486 | 0.467 | 0.359 | 0.346 | 0.383 | 0.370 |
| $\gamma_1$ | −0.125 | −0.137 | −0.170 | −0.187 | −0.167 | −0.183 |
| $\gamma_2$ | −0.250 | −0.263 | −0.310 | −0.327 | −0.295 | −0.312 |
| $\gamma_3$ | 0.157 | 0.246 | 0.210 | 0.302 | 0.198 | 0.287 |
| $\mathrm{Cov}(X_1, S_1)$ | −0.525 | −0.525 | −0.527 | −0.527 | −0.526 | −0.526 |
| $V(X_1)$ | 1.000[f] | 1.000 | 1.000[f] | 1.000 | 1.000[f] | 1.000 |
| $V(S_1)$ | 1.000[f] | 1.000 | 1.000[f] | 1.000 | 1.000[f] | 1.000 |
| $\sigma_{u2}$ | 0.415 | 0.454 | 0.516 | 0.566 | 0.520 | 0.569 |
| $\sigma_{u3}$ | 0.638 | 0.671 | 0.695 | 0.733 | 0.689 | 0.727 |
| $\sigma_{e1}$ | 16.220 | 0.765 | 16.370 | 0.771 | 16.390 | 0.772 |
| $\sigma_{e2}$ | 1.690 | 0.546 | 0.164 | 0.529 | 1.630 | 0.526 |
| $\sigma_{e3}$ | 1.920 | 0.528 | 2.220 | 0.611 | 2.240 | 0.616 |
| $\sigma_{e4}$ | 1.860 | 0.569 | 1.520 | 0.464 | 1.500 | 0.458 |
| $\sigma_{e5}$ | 1.950 | 0.569 | 2.220 | 0.645 | 2.240 | 0.651 |
| $\sigma_{e6}$ | 1.830 | 0.597 | 1.550 | 0.505 | 1.500 | 0.491 |
| $\sigma_{e7}$ | 1.920 | 0.548 | 2.190 | 0.626 | 2.220 | 0.633 |
| $\sigma_{e8}$ | 1.910 | 0.597 | 1.640 | 0.512 | 1.600 | 0.501 |
| $\pi_1$ | 0.000[f] | 0.000 | 2.300 | 0.467 | 2.380 | 0.474 |
| $\pi_2$ | 0.000[f] | 0.000 | 1.740 | 0.357 | 1.830 | 0.368 |
| $\pi_3$ | 0.000[f] | 0.000 | 0.077* | 0.033 | 0.000[f] | 0.000 |
| $\pi_4$ | 0.000[f] | 0.000 | 0.202* | 0.080 | 0.000[f] | 0.000 |
| $\pi_5$ | 0.000[f] | 0.000 | 1.450 | 0.297 | 1.500 | 0.302 |
| $\pi_6$ | 0.000[f] | 0.000 | −0.012* | −0.005 | 0.000[f] | 0.000 |
| $V(X_2)$ | 0.836 | 1.000 | 0.831 | 1.000 | 0.836 | 1.000 |
| $V(X_3)$ | 0.905 | 1.000 | 0.898 | 1.000 | 0.900 | 1.000 |
| $\chi^2$ | 200.640 | | 16.900 | | 18.180 | |
| d.f. | 16.000 | | 10.000 | | 13.000 | |
| Probability level | 0.000 | | 0.0767 | | 0.1507 | |

f = fixed parameters
a = parameters constrained equal
b = parameters constrained equal
* = coefficients less than twice their standard error (designated nonsignificant)
U = unstandardized solution
S = standardized solution

[1]Variables are the same as in previous models except $S_1$ = socioeconomic status (1966), $s_S$ = *SEI* (1966), $s_E$ = education (1966).

[2]The "^" notation is not included for estimated parameters, but it is to be understood that all parameters are sample estimates.

$\pi_6$. The correlated errors for the anomia measure are significant and surprisingly strong: $\hat{\pi}_1^* = 0.467$, $\hat{\pi}_2^* = 0.357$ and $\hat{\pi}_5^* = 0.297$. We can see that relaxing equality restrictions in this case is an important feature of the model. On the other hand, the evidence suggests that there is no correlated error for the powerlessness measure, with $\hat{\pi}_3$, $\hat{\pi}_4$, and $\hat{\pi}_6$ near zero.

The consequences for our validity estimates are as one would expect: for the first time, estimates of validity for anomia are now lower than for powerlessness. Estimates of validity for anomia drop about 0.06, while estimates for powerlessness increase about 0.06. The distortion due to assuming there is no correlated measurement error in previous models is considerable. In fact, powerlessness is the more reliable measure of alienation, and its reliability is higher when we specify correlated errors over time for the anomia measure. This specification also has important effects on the stability parameters. Note that $\hat{\beta}_1$ is affected more in this model than any estimated previously. The change in the unstandardized estimate is $-0.094$, resulting in a stability $\hat{\beta}_1^* = 0.711$. We now see that alienation in 1966 accounts for only about fifty percent of the variance in alienation in 1967. $\hat{\beta}_2$ is also drastically affected in this model, dropping $-0.127$ from 0.486 to 0.359 in the unstandardized case and resulting in a stability of just 0.346. This estimate is still significant, but rather low.

The correlated error across anomia measures also indirectly affects estimates of the substantive relationships in the model. The effects of *SES* on alienation are somewhat larger than in model VA. This suggests that the correlated error in the model was previously manifested as both a component of the stabilities (causing inflated estimates) and a counteractive component of indirect effects on $X_2$ and $X_3$ through $S_1$, causing inhibition of the $S_1$ effects. Also, $\hat{\gamma}_3^*$ is larger again, now an estimated correlation of 0.302. The fact that $\gamma_3$ remains significant in these models can be interpreted in at least two ways, as noted earlier.

The single most succinct reflection of the change in this model is our measure of overall fit. $\chi^2$ is now 16.90 with 10d.f., producing a fit ratio of 1.69. For the first time we see that the probability level for the model differs from zero. This we judge to be a satisfactory model in terms of fit, except that we should re-estimate the model with $\pi_3 = \pi_4 = \pi_6 = 0$, since this is indicated in model VB. Given the number of different specification possibilities we have now considered, we can start to believe the reliability/stability estimates this model produces. Also, we have shown that changes in both theoretical and measurement specification are necessary in producing more reasonable reliability and stability estimates.

Thus we go on to specify model VC with $\pi_3 = \pi_4 = \pi_6 = 0$, producing our final model for the alienation construct. Estimates for model VC are quite similar to VB, except that here there is even more of a difference between the anomia and powerlessness validities, there is a slight increase in $\hat{\beta}_2$ and in $\hat{\pi}_1$, $\hat{\pi}_2$, and $\hat{\pi}_5$, and there are slight decreases in *SES* effects and $\hat{\gamma}_3$.All effects are significant in this model. The correlated errors are substantial, with $\hat{\pi}_1^* = 0.474$, $\hat{\pi}_2^* = 0.368$ and $\hat{\pi}_5^* = 0.302$. This is a potentially disturbing result for many researchers interested in panel models, but it does reflect a real possibility which one can address only with multivariate multiple indicator models. Our final estimates of the validities for anomia are below 0.8, and for powerlessness they range from 0.87 to 0.89. These validities reflect reliability estimates $\hat{\lambda}_i^{*2}$ ranging from 0.58 for anomia in 1967 to about 0.79 for powerlessness in 1966. This is a considerable range and reflects the importance of allowing both true score variance and error variance to vary in these models. Our final stability estimates are 0.712 and 0.370, and we consider these as approaching a lower limit. Owing indirectly to the effect of correlated measurement error in the model, the total correlations between alienation constructs estimated by the model are somewhat lower for model IIA: $r_{X_1X_2} = 0.804$, $r_{X_2X_3} = 0.689$, and $r_{X_1X_3} = 0.684$. These are still in marked contrast with our stability estimates.

The descriptive fit ratio for this model is the best we have achieved, equal to 1.40 (with an accompanying improvement in the probability level). Residuals for model VC are compared with those for model VA in Table 6 (entries for VC are in parentheses). Differences in fit of the two models are clearly reflected by these residuals, especially for covariances among alienation indicators. Model VC is a clear improvement here and in reproducing relationships of these indicators with $s_E$, but it seems less effective in reproducing relationships between $s_S$ and the alienation indicators at times 2 and 3. This is the only area in which the residuals increase across the two models, and the problem is especially noticeable in the case of the latter two anomia indicators. We see that the negative covariances between $s_S$ and $x_{A2}$ and $s_S$ and $x_{A3}$ are overestimated by the model (taking into account the sign of the residual). Despite this, the overall fit of model VC suggests that this model is adequately specified for interpretative purposes. It is true that this fit can be improved, but we do not feel that further modification will seriously affect our estimates of reliabilities and stabilities.

We began with the constraint that the loadings for the same alienation indicator over time would be assumed equal. This is an important assumption in reliability/stability models, and it merits further attention. First, we feel that this equality is implied in any model that purports to interpret a variable's stability: it is the specification of equal loadings when using the same measures over time, which means that we are expecting to measure the same construct. Strictly speaking, the concept of stability demands this specification, but we *can* test for the equality assumption by comparing $\chi^2$ differences between a constrained "same construct" model and an unconstrained model in which loadings for the same indicators are not set equal.[15] We first tested model IA for the efficacy of this equality assumption by specifying an alternative model in which the loadings for anomia and powerlessness are both free at time 1 (with $V(X_1) = 1$), and the loadings for anomia at times 2 and 3 are set at 1 (in order to give the construct a metric),

[15] A word of caution is in order here: successive application of this $\chi^2$ difference test to a series of nested models affects the validity of its probability level and forces an analysis of specification problems to become exploratory and speculative without further data.

TABLE 6

Relative Evaluation of Fit of Models VA and VC to Observed Variances and Covariances ($S$ Matrix), Residuals Expressed as a Portion of $S$ Values

$$\hat{\Sigma}_{ij} - S_{ij}/|S_{ij}|$$

Model VA without parentheses; Model VC in parentheses.

| | $s_S$ | $s_E$ | $x_{A1}$ | $x_{P1}$ | $x_{A2}$ | $x_{P2}$ | $x_{A3}$ | $x_{P3}$ |
|---|---|---|---|---|---|---|---|---|
| Covariances | | | | | | | | |
| $s_S$ | | | | | | | | |
| $s_E$ | −0.0001 (−0.0000) | | | | | | | |
| $x_{A1}$ | −0.1607 (−0.0613) | −0.1095 (−0.0450) | | | | | | |
| $x_{P1}$ | 0.0828 (0.0197) | 0.1275 (−0.0000) | −0.0032 (0.0004) | | | | | |
| $x_{A2}$ | 0.0072 (0.1037) | −0.0756 (−0.0001) | −0.0714 (0.0157) | 0.0727 (−0.0158) | | | | |
| $x_{P2}$ | −0.0067 (−0.0622) | 0.0743 (−0.0061) | 0.1088 (0.0173) | −0.0821 (−0.0156) | 0.0006 (0.0039) | | | |
| $x_{A3}$ | 0.0306 (0.1312) | −0.0442 (0.0361) | −0.0392 (−0.0065) | −0.0342 (−0.0938) | −0.0893 (−0.0123) | 0.0635 (−0.0173) | | |
| $x_{P3}$ | 0.0193 (−0.0273) | 0.0352 (−0.0410) | 0.1843 (0.1111) | −0.0752 (0.0142) | 0.1320 (0.0460) | −0.0611 (0.0140) | 0.0036 (0.0018) | |
| Variances | | | | | | | | |
| VA | 0.0001 | −0.0000 | 0.0200 | −0.0222 | −0.0045 | 0.0048 | −0.0168 | 0.02 07 |
| VC | 0.0000 | 0.0000 | −0.0188 | −0.0178 | 0.0030 | −0.0005 | −0.0200 | 0.02 27 |

and the loadings for powerlessness are free. The test assesses whether the *set* of loadings at times 2 and 3 differs from the *set* of loadings at time 1. The $\chi^2$ test for improvement in fit is

$$\chi^2_{9\text{df}} - \chi^2_{7\text{df}} = 184.08 - 181.47$$

$$\chi^2_{2\text{df}} = 2.60$$

which is not significant at the 0.05 level, $p = 0.27$. Thus, the equality assumption holds for this model. We can apply the same test to model VC, since this model includes a concomitant variable, correlated error, and a construct-to-endogenous disturbance correlation not included in model IA. Here, respecification of the constrained version with the free loading version reduces the $\chi^2$ value 5.3, which with 2d.f. just fails to be a significant difference at the 0.05 level, $p = 0.08$. Thus, the assumption of loading equality is more in question in our final model, but it is still not unreasonable. This is especially true when we consider the residuals for model VC: excluded correlations between $s_S$ and the errors for the anomia indicators, for instance, would presently operate through paths in the model which include the construct loadings—and this could distort these estimates slightly. We conclude that the equality assumption is still tenable and that we are discussing the same construct over time in this model.

The conclusions one can infer from this model can now be summarized. First, powerlessness is a more reliable measure of alienation than anomia. Second, this anomia measure exhibits some nonrandom measurement error which is correlated over time. Third, alienation shows a moderate amount of stability over a one year period, but a considerable amount of instability over a four year period. Fourth, in terms of an analysis of trends over a five year period in this sample, higher *SES* in 1966 leads to lower levels of alienation in 1967 and 1971. Fifth, evidence indicates that either other causal variables are related to alienation in 1966 and cause alienation in 1971, or that alienation in 1966 is directly related to alienation in 1971. This second explanation is possible, for instance, if alienation in 1971 depends on original position in the alienation distribution in 1966.

## *SUMMARY AND CONCLUSIONS*

It should be obvious at this point that specification flexibility is a mandatory requirement in the adequate estimation of reliability and stability parameters with panel data. Our presentation has emphasized that the issue of stability should be addressed within a thoroughly specified theoretical and measurement context. There is a statistical basis for this assertion in that stability coefficients are subject to bias if obtained from single-indicator single-variable models. Also, we have seen that only with multiple-indicator multivariate models can we begin to test our reliability and stability estimates.

We should point out that the stability of a variable over time should be an issue in the theory needed to explain that variable. Basically, we could speculate that stable variables should be less sensitive to situationally based influences, and unstable variables should be more sensitive to these sources of influence. This is the sense in which the stability issue is inseparable from the substantive context and, thus, the causal system within which the variable operates. We should be sensitive to the point at which estimates of stability begin to "bottom out" in the light of successive respecifications of a model, since evidence for completeness of a causal system partially depends on this lower-limit stability value.

We began with stability estimates for alienation in model IA of 0.897 and 0.768, and finally concluded with model VC that the stabilities were probably closer to 0.712 and 0.370, respectively. Obviously, there is some danger in taking a "barefaced" approach to single-indicator or single-construct models. Additionally, we can best address theoretical issues only when measurement issues are also incorporated into our models.

The superiority of multiple-indicator models is not only on methodological grounds. In fact, these models are in harmony with meta-theoretical arguments which suggest the need for multiple-indicators of all theoretical concepts. This argument is usually made in terms of the increased content validity of the unobserved construct with increasing numbers of indicators. Thus, there are convenient rationales available at both the empirical and meta-theoretical levels for choosing multiple-indicator panel models where possible.

In general, our conclusion is that there is a need to be explicit about both measurement and theoretical specification in stating relationships between variables in panel models in order to maximize the accuracy of estimation of parameters in these models. We have shown that concerns at one level of specification can affect the estimation of parameters at another. And since we believe the stability coefficient has implications for our definition and interpretation of a concept in the process of theory construction, these specification issues are important for theorists to consider as well. Where we depend on parameter estimates to make decisions with theoretical implications, adequacy at both levels of specification is necessary if our decisions are to be taken seriously.

## *REFERENCES*

ALTHAUSER, R. B., AND HEBERLEIN, T. A.

1970 "Validity and the Multitrait-multimethod Matrix." In E. F. Borgatta and G. W. Bohrnstedt (Eds.), *Sociological Methodology 1970.* San Francisco: Jossey-Bass, pp. 151–169.

ALWIN, D. F.

1974 "Approaches to the Interpretation of Relationships in the Multitrait-multimethod Matrix." In H. L. Costner (Ed.), *Sociological Methodology 1973-1974.* San Francisco: Jossey-Bass, pp. 79–105.

ANDERSON, B. D.

1973 "School Bureaucratization and Alienation from High School." *Sociology of Education* 46:315–334.

BLALOCK, H. M., JR.

1970 "Estimating Measurement Error Using Multiple Indicators and Several Points in Time." *American Sociological Review* 35:101–111.

BOGARDUS, E. S.

1925 "Measuring Social Distances." *Journal of Applied Sociology* 9:299–308.

BOHRNSTEDT, G. W.

1969 "Observations on the Measurement of Change." In E. F. Borgatta (Ed.), *Sociological Methodology 1969.* San Francisco: Jossey-Bass, pp. 113–133.

BURT, R. S.

1973 "Confirmatory Factor-Analytic Structures and the Theory Construction Process." *Sociological Methods and Research* 2:131–190.

COSTNER, H. L.

1969 "Theory, Deduction, and Rules of Correspondence." *American Journal of Sociology* 75:245–263.

COSTNER, H. L., AND SCHOENBERG, R.

1973 "Diagnosing Indicator Ills in Multiple Indicator Models." In A. S. Goldberger and O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press, pp. 167–199.

DUNCAN, O. D.

1969 "Some Linear Models for Two-Wave, Two-Variable Panel Analysis." *Psychological Bulletin* 72:177–182.

1972 "Unmeasured Variables in Linear Models for Panel Analysis." In H. L. Costner (Ed.), *Sociological Methodology 1972.* San Francisco: Jossey-Bass, pp. 36–82.

1975 "Some Linear Models for Two-Wave, Two-Variable Panel Analysis, with One-Way Causation and Measurement Error." In H. M. Blalock, Jr., A. Agankegian, F. N. Borodkin, R. Boudon, and V. Capecchi (Eds.), *Quantitative Sociology: International Perspectives on Mathematical and Statistical Modeling.* New York: Academic Press, pp. 285–306.

HANNAN, M. T., RUBINSON, R., AND WARREN, J. T.

1974 "The Causal Approach to Measurement Error in Panel Analysis: Some Further Contingencies." In H. M. Blalock, Jr. (Ed.), *Measurement in the Social Sciences: Theories and Strategies.* Chicago: Aldine, pp. 293–323.

HAUSER, R. M., AND GOLDBERGER, A. S

1971 "The Treatment of Unobservable Variables in Path Analysis." In H. L. Costner (Ed.), *Sociological Methodology 1971.* San Francisco: Jossey-Bass, pp. 81–117.

HEISE, D. R.

1969 "Separating Reliability and Stability in Test-Retest Correlation." *American Sociological Review* 34:93–101.

1970 "Causal Inference from Panel Data." In E. F. Borgatta and G. W. Bohrnstedt (Eds.), *Sociological Methodology 1970.* San Francisco: Jossey-Bass, pp. 3–27.

JÖRESKOG, K. G.

1969 "A General Approach to Confirmatory Maximum Likelihood Factor Analysis." *Psychometrika* 34:183–202.

1973 "A General Method for Estimating a Linear Structural Equation System." In A. S. Goldberger and O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press, pp. 83–112.

1976 "Structural Equation Models in the Social Sciences: Specification, Estimation, and Testing." Department of Statistics, University of Uppsala. Invited paper for the Symposium on Applications of Statistics, Ohio, June 14–18, 1976.

JÖRESKOG, K. G., AND SÖRBOM, D.

1976a "Statistical Models and Methods for Analysis of Longitudinal Data." In D. J. Aigner and A. S. Goldberger (Eds.), *Latent Variables in Socioeconomic Models.* Amsterdam: North-Holland.

1976b "LISREL – Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods." A computer program manual. Chicago: International Educational Services.

KENNY, D. A.

1973 "Cross-Lagged and Synchronous Common Factors in Panel Data." In A. S. Goldberger and O. D. Duncan (Eds.), *Structural Equation Models in the Social Sciences.* New York: Seminar Press, pp. 153–165.

LORD, F. M., AND NOVICK, M. R.

1968 *Statistical Theories of Mental Test Scores.* Reading, Mass.: Addison-Wesley.

MERTON, R. K.

1957 *Social Theory and Social Structure.* Glencoe: Free Press.

MIDDLETON, R.

1963 "Alienation, Race, and Education." *American Sociological Review.* 28:973–977.

O'MEARA, J.

1966 *Sample Design: Bench-Mark Survey.* Memoranda No. 1 to Project No. 007. Urbana, Ill.: Survey Research Laboratory, University of Illinois.

PELZ, D. C., AND ANDREWS, F. M.

1964 "Detecting Causal Priorities in Panel Data." *American Sociological Review* 29:836–848.

PHOTIADIS, J. D., AND SCHWEIKER, W. F.

1971 "Correlates of Alienation: the Marginal Businessman." *Rural Sociology* 36:20–30.

ROZELLE, R. M., AND CAMPBELL, D. T.

1969 "More Plausible Rival Hypotheses in the Cross-Lagged Panel Correlation Technique." *Psychological Bulletin* 71:74–80.

SCHOENBERG, R.

1972 "Strategies for Meaningful Comparison." In H. L. Costner (Ed.), *Sociological Methodology 1972.* San Francisco: Jossey-Bass, pp. 1–35.

SIEGEL, P. M., AND HODGE, R. W.
1968 "A Causal Approach to the Study of Measurement Error." In H. M. Blalock, Jr. and A. B. Blalock (Eds.), *Methodology in Social Research.* New York: McGraw-Hill, pp. 21–59.

SIMPSON, M. E.
1970 "Social Mobility, Normlessness, and Powerlessness in Two Cultural Contexts." *American Sociological Review* 35:1002–1013.

SNELL, P. AND MIDDLETON, R.
1961 "Dimensions and Correlates of Religious Idiologies." *Social Forces* 39:285–290.

SÖRBOM, D.
1975 "Detection of Correlated Errors in Longitudinal Data." *British Journal of Mathematical and Statistical Psychology* 28:138–151.

SROLE, L.
1956 "Social Integration and Certain Corollaries: An Exploratory Study." *American Sociological Review* 21:709–716.

SUMMERS, G. F., HOUGH, R. L., SCOTT, J. T., AND FOLSE, C. L.
1969 *Before Industrialization: A Rural Social System Base Study.* Bulletin No. 736. Urbana, Ill.: Illinois Agricultural Experiment Station, University of Illinois.

THOMPSON, W. E., AND HORTON, J. E.
1960 "Political Alienation as a Force in Political Action." *Social Forces* 38:190–195.

WERTS, C. E., JÖRESKOG, K. G., AND LINN, R. L.
1971 "Comment on 'the Estimation of Measurement Error in Panel Data'." *American Sociological Review* 36:110–113.

WERTS, C. E., LINN, R. L., AND JÖRESKOG, K. G.
1971 "Estimating the Parameters of Path Models Involving Unmeasured Variables." In H. M. Blalock, Jr. (Ed.), *Causal Models in the Social Sciences.* Chicago: Aldine, pp. 400–409.
1974 "Quantifying Unmeasured Variables." In H. M. Blalock, Jr. (Ed.), *Measurement in the Social Sciences: Theories and Strategies.* Chicago: Aldine, pp. 270–292.

WILEY, D. E., AND WILEY, J. A.
1970 "The Estimation of Measurement Error in Panel Data." *American Sociological Review* 35:112–117.

*Received April 15, 1975.*