THE CAUSAL APPROACH TO MEASUREMENT ERROR IN PANEL ANALYSIS: SOME FURTHER CONTINGENCIES*

by

Michael T. Hannan

Richard Rubinson

Jean Tuttle Warren

Technical Report #48 1972

*The research reported herein was, in part, performed pursuant to a contract with the United States Department of Health, Education, and Welfare, Office of Education, and was partially supported by NSF Grant (GS-23065), "The Interrelationships of National Educational Systems with Political, Social and Economic Institutions: A Longitudinal Quantitative Study". John Boli performed many of the computations for this research.

There is a considerable stress in current methodological discussion in sociology on shifting focus from cross-sectional to longitudinal designs. Sociological research has been almost exclusively cross-sectional and our methodology is suited primarily to this case. Attempts to shift to longitudinal designs raise a number of new issues, in particular, the many discontinuities between cross-sectional thinking (where many instances of a process are observed simultaneously) and time series thinking (where a single instance is observed at many points in time). While both types of analysis are well understood, the non-experimental social scientist is typically faced with a design which falls somewhere between the two. The typical case involves observations on many instances of a process at only a few points in time. Sociologists have relied almost exclusively on the panel method to address this case.[1]

Sociologists, however, have greatly exaggerated the "power" of the panel design. While changing to a simple longitudinal design allows one to use time orderings to rule out some causal effects, it does not unambiguously resolve many questions concerning either direction or time sequencing of causal effects (Duncan 1969, 1972a, 1972b). In fact, as Heise (1970) and Duncan have shown, causal inferences in such simple longitudinal models depend on rather restrictive assumptions. It has become clear that most of the analysis problems arising in even the simplest cases are not yet well understood.

While sociologists have used the panel design primarily to resolve difficulties in causal analysis, there has been a recent emphasis in the causal models approach to measurement error on using panel observations to eliminate complications suggested by the acknowledgement of measurement imperfection. Blalock (1970) has suggested that a panel design will generally help in reducing the excess of unknowns which arises when the true sample values of substantive variables are assumed to be unknown and measured with error. Costner

(1969) earlier had demonstrated that with three indicators for each substantive variable in a recursive cross-sectional model the analyst can both test measurement models and (if appropriate) estimate causal parameters. Obtaining three indicators of each variable is often not possible, however, and Blalock (1970) has shown that a rough trade-off exists between the number of indicators in a cross-section and waves of observations in a panel under certain specified conditions.[2] In some situations, the researcher who can only obtain one or two indicators of some variables can generate tests and estimates if he can obtain repeated measurements in a panel design.

Work on measurement error in panel models has focused on only the very simplest cases. Attention has been limited to single-variable models (where a single substantive variable is measured at several points in time). This poses a serious problem for the analyst faced with a multivariate panel model (where several variables are measured at several points in time) measured with error. In attempting analysis, the researcher must deal simultaneously with the inference problems of linear panel models and with those arising from the existence of measurement error. We have begun to face such dual problems in our substantive research. This paper focuses, then, on the additional complications which arise when multivariate panel models are measured with random and non-random errors. In spirit and approach, it relies heavily on the papers cited above.

In any didactic discussion it is difficult to introduce more than one complication at a time. We are primarily interested in measurement error and are willing for the present to employ highly restrictive assumptions to rule out other complications. It is highly likely in panel models that regression disturbances (residuals) will be correlated over waves of observations due to stability in these causal variables excluded from the model (see Heise [1970]).

Since the presence of both correlated regression disturbances and measurement error will generally result in underidentification,[3] we will generally assume that the disturbances are uncorrelated *inter se* and with substantive variables in our models. We are forced to adopt this position largely because we have no *a priori* information about measurement quality. Evidence in each case is *internal* to the model. We construct models incorporating substantive arguments as likely sources of measurement error and test them with our data. We do not, however, attempt to arrive at the appropriate error model inductively. Rather, the thrust of this paper, and the literature it follows, is to emphasize the practical impossibility of solving measurement problems inductively.[4]

We follow the literature cited above in employing the technical apparatus of path analysis for testing and estimation. All of the coefficients we discuss are *standardized* by sample variances. This approach simplifies the problem of generating the large number of structural equations containing unmeasured (true values) variables arising in realistic models but has serious disadvantages for panel analysis. Wiley and Wiley (1970)have demonstrated that the assumption of stable standardized coefficients requires both the true population variances and the measurement error variances be stable over the waves of observations. This is particularly problematic in "development" models and we will discuss its implications in terms of specific models below.

The substantive research application is a cross-national study of the interrelations of national educational systems and economic, social and political development (Meyer and Hannan, 1971). In this paper we focus on one highly simplified model relating expansion of educational systems to economic development. Using data reported by the United Nations (U.N. Statistical Yearbooks) we follow a panel of 96 nations[5] through three waves of observations, 1955, 1960, and 1965. The research project is still in an early phase and we are less concerned here with substantive findings than witnessing the import of measurement complications.

We have confined our attention to lagged cross-effects rather than instantaneous effects. Since this decision may often be problematic, we will discuss our reasoning. Duncan (1969) has shown that in general one cannot take both the direction of cross-effects and the timing of causal effects as problematic. Consider the model taken from Duncan's paper drawn in Figure 1. As long as the

---

Figure 1 about here

---

analyst is unwilling to make at least one *a priori* restriction, e.g. to rule out either the lagged effects or the instantaneous effects, this model is underidentified (there are seven parameters and only six independent equations). Alternatively, one can proceed by ruling out effects (either lagged or instantaneous) in one direction. What is clear is that in the very general model drawn in Figure 1, one cannot employ only sample information to infer either lagged or instantaneous effects in one, the other, or both directions.

However, the substantive significance of the distinctions between lagged and instantaneous effects seems to vary with the process being studied and the development of the theory. In Duncan's examples, the observation points correspond to socially meaningful categories (grade in school, stage in life cycle, etc.), and the variables can conceivably increase or decrease in magnitude between waves of observations. In our research the time periods of observations are more or less arbitrary indicator points since the variables are cumulative (monotonically increasing) for almost all units. Whether the effects are lagged or instantaneous is not an important substantive issue (at least given the present state of development of the theories involved) and, ignorance of appropriate lags is less likely to produce faulty inferences when all variables are monotonic over time.[6] We will make an argument below (for didactic purposes) for lags of different lengths, i.e. an argument that the causal processes under study in a model differ in the lag with which they have an impact on the variables under study. Here Duncan's argument is compelling. The point is that for the present we take the direction and magnitude of cross-effects as particularly important and do not systematically investigate the timing of causal effects. In particular, we restrict all causal effects to be lagged effects.

A THREE-WAVE, TWO-VARIABLE, TWO-INDICATOR (3W-2V-2I)MODEL

We began our analysis with a three-wave, two-variable model. Our data enable us to select two indicators of economic development, per-capita gross national product and per-capita consumption of electricity.[7] We also select two indicators of educational expansion, the ratio of primary school students to the appropriate age-group population and the ratio of secondary school students to the age-group population.

This model, with uncorrelated residuals and purely random measurement error, is diagrammed in Figure 2. The curved arrow at the left represents the summary of the history of the operation of the postulated causal processes. We further assume that all coefficients in the model are stable over waves of observations. Throughout this paper, the underlying economic development variable

---

Figure 2 here

---

will be represented by X_i and the educational variable by Y_i, with the subscript standing for the period of observation. The indicators of economic development will be represented by X'_{ij} and the indicators of educational expansion by Y'_{ij}, with the first subscript standing for the period of observation and the second subscript for the specific indicator.

The correlation matrix for this model is presented in Table 1.

Table 1 here

The matrix includes a pattern which, on initial inspection, is very perplexing. Examine the intercorrelations of the educational ratios. The primary ratio at the first time period is most highly correlated with the secondary ratio at the last time period, next most highly correlated with the secondary ratio at the middle time period and least highly correlated with the cotemporal secondary ratio. The pattern for the primary ratio measured at the middle time period is similar in that the correlation with the secondary ratio of the last time period is the greatest. For the primary ratio measured at the last time period. the correlation is greatest at the same point in time and the correlation with secondary ratios decreases monotonically as one moves back in time.

We had originally intended to treat the two educational enrollment ratios as related only through their common relationship to the underlying variable educational expansion. We began with this hypothesis of 'common factor variance' (or "congeneric tests") because, in terms of scientific simplicity, this model represents the most parsimonious structure. Yet we find that indicators of variables measured at different points in time are more highly inter-correlated than are cotemporal measures of the variables. This result violates our conventional understanding of validity (see Campbell and Fiske (1959). And it is unlikely that such a simple causal structure could generate the observed correlation matrix. For instance, the educational portion of this model (Figure 2) fails the consistency criterion for purely random measurement error developed by Costner (1969).

We were faced with the practical problem of how to proceed with such an unforeseen result. Two lines of investigation seemed open: (1) construct more complicated two-indicator models to generate the observed pattern of inter-correlation, and (2) decompose this model to allow the indicators to be directly causally related. We will trace out the details and implications of each strategy.

To simplify the algebra we revise the model in Figure 2 to include only that portion relating to educational expansion. Since we assume that this variable is systematically affected by economic development, we allow the residuals to be correlated with the true (unobserved) values of the educational variables, where the correlations do not violate least squares restrictions. Note that this model is now more restrictive than we could like since we must postulate that U is uncorrelated with early values of educational expansion. This revised model is diagrammed in Figure 3. The new _k_ terms will be used in

Figure 3 here

the second of the models discussed below.

Using the algorithm of path analysis, we write the fifteen equations for the measurement model in Figure 3, with k=1:

$$\rho_{Y'_{11}Y'_{12}} = k^4cd \qquad (1)$$

$$\rho_{Y'_{11}Y'_{21}} = k^3c^2a \qquad (2)$$

$$\rho_{Y'_{11}Y'_{22}} = k^3cda = \rho_{Y'_{12}Y'_{21}} \qquad (3)$$

$$\rho_{Y'_{11}Y'_{31}} = k^2c^2 (ab + su) \qquad (4)$$

$$\rho_{Y'_{11}Y'_{32}} = k^2cd\ (ab + su) = \rho_{Y'_{12}Y'_{31}} \qquad (5)$$

$$\rho_{Y'_{21}Y'_{22}} = k\ cd \qquad (6)$$

$$\rho_{Y'_{21}Y'_{31}} = kc^2\ (b + srt + asu) \qquad (7)$$

$$\rho_{Y'_{21}Y'_{32}} = kcd\ (ab + su) = \rho_{Y'_{22}Y'_{31}} \qquad (8)$$

$$\rho_{Y'_{31}Y'_{32}} = cd \qquad (9)$$

$$\rho_{Y'_{12}Y'_{32}} = k^2d^2\ (ab + su) \qquad (10)$$

$$\rho_{Y'_{12}Y'_{22}} = k^3d^2a \qquad (11)$$

$$\rho_{Y'_{22}Y'_{32}} = kd^2\ (b + srt + asu) \qquad (12)$$

Consider first the possibility dismissed above that the primary and secondary ratios are 'equivalent' indicators of educational expansion and are measured with random error. In this case the $\underline{k}$ terms are assumed equal to unity. The ordering of magnitudes of the sample correlations between the primary ratio in 1955 (Y'_{11}) and the secondary ratios (Y'_{12}, Y'_{22}, Y'_{32}) gives rise to the following inequality:

$$r_{Y'_{11}Y'_{32}} > r_{Y'_{11}Y'_{22}} > r_{Y'_{11}Y'_{12}}$$

Substituting from the path equations (1), (3), and (5), we obtain corresponding values:

$$cd(ab + su) > cda > cd$$

or $(cd \neq 0)$

$$ab + su > a > 1$$

This result violates the basic model since the two terms required to exceed unity are a zero order correlation (a) and a sum of direct and indirect causal effects in standardized form (ab + su). Each term is bounded by plus and minus one. If they were not so bounded, we would be accounting for more than 100% of the variance in the dependent variables.[8] This result argues against the common factor model. Note, however, that this involves a statistical inference since it is possible to obtain sample values exceeding unity when the population parameter is less than unity.

If we continue to accept the model diagrammed in Figure 3, we might proceed by allowing other complications in the model in order to generate the correlation matrix in Table 1. For example, we might allow the different indicators of educational expansion to be correlated at every wave because of common sources of measurement error (e.g. both pieces of information are processed by the same national bureaucracies); or, we might allow the same indicators to be correlated at different points in time because of stable sources of measurement error variance (e.g. stability for units in the bureaucratic procedures for gathering and reporting educational statistics). Working through the resulting equations (where, for instance, e_1 and e_2, e_3 and e_4, and e_5 and e_6 are correlated, or e_1 and e_3, e_2 and e_4, etc.) quickly reveals that these added complications cannot reproduce the observed pattern of inter-correlations.

Allowing inter-temporal correlation of the measurement error terms of <u>different</u> indicators (i.e. allowing Y'_{11} and Y'_{32} and Y'_{22} to share common sources of error variance) will allow us to reproduce the correlation matrix. We reject this solution, however, for it merely represents a mechanical way to generate the matrix in the absence of any substantive knowledge (i.e. it is a formalization of our ignorance).

We next develop a model which allows the random errors in variables to decrease proportionately with each wave of observation. This model is based on the assumption that with the secular trend towards national accounting, the quality of the national statistics collected increases over time. Such an assumption seems reasonable for our educational statistics since their collection and reporting is continually being supervised by the United Nations statistical office (UNESCO). To represent this secular trend in statistical quality, we allow the k term in our educational model (Figure 3) to take on values different from unity, and also require them to be between zero and one.[9] Other values will not produce decreasing random measurement error. (See Duncan, 1972b for an analogous model.)

To evaluate this model, we consider the inequalities introduced above. Since k is no longer equal to unity, we have the following:

$$k^2cd(ab + su) > k^3cd > k^4cd$$

or $(cd \neq 0)$

$$ab + su > ka > k^2$$

We can no longer reject this model on logical grounds.

We can proceed to subject the model to an additional test using the seven over-identifying restrictions (the model has eight unknowns and fifteen equations). All seven restrictions can be written as quantities equal to zero, under the hypothesis that the model is correct (Blalock, 1964). As noted previously, sampling error may produce deviations from zero even if the model is correctly specified and the analyst must make a statistical inference (where, unfortunately, the sampling distributions of the estimators are unknown). The over-identifying restrictions with the sample estimates are:

PREDICTION	ESTIMATE
$\rho_{Y'_{11}Y'_{32}} - \rho_{Y'_{12}Y'_{31}} = 0$	.219
$\rho_{Y'_{11}Y'_{22}} - \rho_{Y'_{12}Y'_{21}} = 0$	.140
$\rho_{Y'_{21}Y'_{32}} - \rho_{Y'_{22}Y'_{31}} = 0$	.057
$\rho_{Y'_{11}Y'_{12}}\rho_{Y'_{31}Y'_{32}} - \rho_{Y'_{21}Y'_{22}} = 0$	.031

(cont.) PREDICTION	ESTIMATE
$\rho_{Y'_{11}Y'_{21}}\rho_{Y'_{12}Y'_{32}} - \rho_{Y'_{12}Y'_{32}}\rho_{Y'_{11}Y'_{32}} = 0$	.074
$\rho_{Y'_{11}Y'_{21}}\rho_{Y'_{22}Y'_{32}} - \rho_{Y'_{12}Y'_{32}}\rho_{Y'_{21}Y'_{31}} = 0$	.056
$\rho_{Y'_{12}Y'_{32}}\rho_{Y'_{21}Y'_{32}} - \rho_{Y'_{11}Y'_{31}}\rho_{Y'_{21}Y'_{31}} = 0$	.018

This fit does not seem particularly close. Nonetheless, we proceed to estimate the one obviously identified coefficient,[10] k. Equations (1), (6), and (9) allow two ways of estimating k directly:

$$\hat{k} = \sqrt{\frac{\rho_{Y'_{11}}\rho_{Y'_{12}}}{\rho_{Y'_{21}}\rho_{Y'_{22}}}} = \sqrt{\frac{\rho_{Y'_{21}}\rho_{Y'_{22}}}{\rho_{Y'_{31}}\rho_{Y'_{32}}}}$$

We follow Duncan (1972a,b) and add these two expressions and insert sample estimates to arrive at an "ad hoc" estimate of k:[11]

$$\hat{k} = \frac{\sqrt{\rho_{Y'_{11}}\rho_{Y'_{12}}} + \sqrt{\rho_{Y'_{21}}\rho_{Y'_{22}}}}{\sqrt{\rho_{Y'_{21}}\rho_{Y'_{22}}} + \sqrt{\rho_{Y'_{31}}\rho_{Y'_{32}}}}$$

Inserting sample values of correlations gives an estimate of $\hat{k} = 1.04$. This estimate for k is inconsistent with the hypothesis that the sources of measurement error are stable but decreasing with each new wave of observation . This failure is not surprising given the poor fit of the entire model.

A SINGLE-INDICATOR MODEL

At this point (given our aims) we had only two choices. We could ignore one of the indicators of educational expansion and proceed with a single indicator or we could entertain the hypothesis that the two "indicators" stand in some direct causal relationship. We should mention one variation of the latter alternative which we did not pursue. The previous analysis suggests serious defects in the "common factor" approach to the educational ratios. One alternative modification would be to keep the common factor model but introduce additional direct causal links between the indicators (educational ratios). We did not puruse this approach since the number of unknown quantities becomes too large.[12]

In this section we pursue the single-indicator approach. Thus we revise the model of Figure 3 to create the three-wave, two-variable, single indicator model drawn in Figure 4. This model contains only educational variables (primary and secondary ratios measured at three points in time). For purposes of algebraic simplicity, we continue to posit uncorrelated residuals and, for the present, assume uncorrelated random measurement errors. The latter assumption will be relaxed below. This model incorporates a conceptual shift. The enrollment ratios are now taken to be abstract causal <u>variables</u> measured with random error. We continue to denote the primary ratios by Y_{i1} and the secondary ratios by Y_{i2}. Measured values are primed.

Figure 4 about here

This model allows for cross-effects in both directions. However, when we first began to examine the correlation matrix in Table 1, two of us were working on a version of the model in which the effect from secondary to primary ratios was assumed to be absent (i.e. d=0). Call this case (i). We discovered that such a model is capable of generating the correlation matrix in question given very high autocorrelation terms. In fact, the process can be represented in a simple and elegant form. To do this we alter our notation temporarily. Let P'_t and S'_t denote the primary and secondary ratios measured at time t, and P'_{t+1}, S'_{t+1} denote the measured values of the same variables at the next point in time, etc. This model is drawn at the top of Table 2 where we continue to assume purely random measurement error, stable coefficients, and uncorrelated residuals. Given this specification, the path equation for the population value of the zero-order cross-lag correlation takes on the following form:

$$\rho_{P_{t_1} S_{t+k}} = ef[\rho a^k + c\sum_{j=0}^{k-1} a^{k-j-1} c^j]$$

The behavior of these cross-lagged correlations is indeed time dependent as the difference equation representation shows. With P_t fixed, the correlation $\rho_{P_{t_1} S_{t+k}}$ will increase as k increases up to a point and then begin to decrease. Both the length of the interval over which a maximum is attained and the behavior of the correlation around that interval (e.g. rapidity of decline in magnitude) depends on the values of the coefficients of the model. The important point for our purposes is that this model, together with reasonable estimates[13] of the coefficients, produces a correlation matrix very close to that reported in Table 1. The behavior of the cross-lag correlation over ten time periods for alternative hypothesized coefficients is reported in Table 2.

Table 2 about here

Consider an alternative specification of the model relating the two educational ratios in which only the direction of the cross-effect is changed,(i.e. $d \neq 0$, $c=0$), call this case (ii). Case (i) seems preferable on substantive grounds. Increases in primary enrollments create a demand for the expansion of secondary systems as larger cohorts pass through the primary schools. However, the possibility remains that educational systems expand down from the top. Secondary expansion creates a "pull" into primary school due to the changed opportunity structure. For this reason it is useful to subject both models to test.

First we consider the "dynamics" of our second model. This model generates a correlation matrix which is the transpose of the matrix produced by the first model. Thus the correlation matrix generated in this case is not at all close to that observed in our sample. This result lends considerable support to the model where secondary ratios are taken as dependent on earlier primary ratios (case [i]).

Since this method of evaluating the competing causal models is somewhat novel in the sociological literature (and involves a number of implicit assumptions and approximations), we were interested in also conducting a more standard path analytic test of the models. To do this we write out the path equations for the two models as follows:

	Case (i)(d=0)	Case (ii)(c=0)
$\rho_{Y'_{12}Y'_{22}}$	$= e^2[a + \rho c]$	$e^2 a$
$\rho_{Y'_{12}Y'_{32}}$	$= e^2[a^2 + \rho ca + \rho bc]$	$e^2 a^2$
$\rho_{Y'_{12}Y'_{11}}$	$= e\rho f$	$e\rho f$
$\rho_{Y'_{12}Y'_{21}}$	$= e\rho bf$	$ef[d + \rho b]$
$\rho_{Y'_{12}Y'_{31}}$	$= e\rho b^2 f$	$ef[\rho b^2 + db + ad]$
$\rho_{Y'_{11}Y'_{22}}$	$= ef[c + \rho a]$	$ef\rho a$
$\rho_{Y'_{11}Y'_{32}}$	$= ef[\rho a^2 + ca + bc]$	$ef\rho a^2$
$\rho_{Y'_{11}Y'_{21}}$	$= f^2 b$	$f^2[b + \rho d]$
$\rho_{Y'_{11}Y'_{31}}$	$= f^2 b^2$	$f^2[b^2 + \rho db + \rho ad]$
$\rho_{Y'_{22}Y'_{32}}$	$= e^2[a + a\rho bc + c^2 b]$	$e^2 a$

cont.	Case (i)(d=0)	Case (ii)(c=0)
$\rho_{Y'_{22}Y'_{21}}$	$= ef[a\rho b + cb]$	$ef[ad + a\rho b]$
$\rho_{Y'_{22}Y'_{31}}$	$= ef[a\rho b^2 + cb^2]$	$ef[d + adb + a\rho b^2]$
$\rho_{Y'_{21}Y'_{32}}$	$= ef[c + acb + a^2\rho b]$	$ef[a^2d + a^2\rho b]$
$\rho_{Y'_{21}Y'_{31}}$	$= f^2b$	$f^2[b + d^2a + b\rho ad]$
$\rho_{Y'_{32}Y'_{31}}$	$= ef[cb + acb^2 + a^2\rho b^2]$	$ef[ad + a^2db + a^2\rho b^2]$

The two models do not share any prediction equations, making it difficult to choose between the two. Both cases fit better than the model tested earlier, as is shown in Table 3. If anything, case (ii) fits slightly better than case(i).

Table 3 about here

Since as far as we know the sampling distributions of the series of prediction equations is not known, it is difficult to choose between the two cases purely on the basis of the small difference in fits. Proceeding to the estimation stage does not reduce the uncertainty. In each case only two coefficients are obviously[14] identified, the autoregression and "epistemic correlation" for the independent variable. There are five equivalent solutions for each of the autoregression terms. Using the Duncan procedure outlined above we obtain $\hat{b} = .893$ for case (i) and $\hat{q} = 1.021$ for case (ii). Using these composite estimates, we can solve directly for $\hat{f}$ in case (i) and $\hat{e}$ in case (ii) in three equivalent ways. The combined estimates for the two quantities respectively are 1.048 and .937.

Under the model specifications for the two cases, the four terms estimated are correlations as well as path coefficients. Each case violates this assumption and neither model is satisfactory. The fact that none of the remaining parameters of either model are obviously identified greatly limits the usefulness of these models. Thus we do not continue this analysis in an effort to model the complications which might yield more acceptable estimates. Our ultimate objective is to relate these educational variables to other substantive variables (economic development in the present application). To do this we will pursue a number of the issues which arise when single-indicator models are embedded in more complex models. The most obvious extension in this substantive research is to three-variable, three-wave panel models.

Since the path analysis is indeterminate in choosing between the two models, we will take the argument based on dynamics (together with our substantive preference) as persuasive. Henceforth in the analysis we will consider case (i) as the appropriate causal model.

Before going on to more complex models, we will comment briefly on the identification problem in this simple model with one-way cross-effects. It is rather surprising that even when we assume stability in all parameters (which implies constant true and error variances), purely random measurement error, and uncorrelated disturbances, only the autoregression of the "independent" variable and the coefficient associated with the measurement term for this variable are identified.[15] What is even more surprising is the finding that the addition of new waves of observations do nothing more than provide additional tests of the model and additional estimates of the coefficients already identified even when the new waves do not add any additional unknowns. This is quite important. The bottom half of the model for case (ii) corresponds to the single indicator case discussed in the literature cited at the outset. Our result conforms to what is already known -- with single indicators measured with purely random error, uncorrelated disturbances and three waves of observations, all parameters of interest are identified. However, the addition of a second substantive variable makes clear that the more general case is considerably more complicated. This is particularly puzzling since the variable we add is exogenous (i.e. it is posited to be independent of the residuals in the regression equations for the educational ratios). We seem to have a case, then, where the addition of more information precludes the estimation of previously identified coefficients.

Inspection of the equations for case (ii) isolates the difficulty. Consider the expression for the population correlation of the primary ratios at the second and third observations (since the autoregression in this variable could be solved for before the addition of the exogenous variable):

$$r_{Y'_{22}Y'_{32}} = e^2[a + a\rho bc + c^2 b]$$

We see that the estimation complications arise because of the curved double headed arrows at the left hand side of the model (denoted by ρ) and because of the stability in all of the substantive variables. When early values of the variables are correlated due to previous operation of the causal structure under study and the variables are stable, the number of indirect paths connecting observations quickly becomes very large and expressions do not repeat themselves.[16]

This, plus the nonlinear manner in which the measurement error terms enter the equations, gives rise to systems of non-linear equations in k unknowns (where k is fairly large, e.g. 4) which are unlikely (in our experience) to yield useful solutions. As we noted above, even if the systems of equations may in principle have real roots, the actual work of solving the system (even using a computer) is enormous.

The technical problem is a failure of the <u>sufficient</u> conditions for identification. We have become accustomed to concerning ourselves only with the necessary conditions for identification. In the present context these may be stated in the form: the number of path coefficients to be estimated must not exceed the number of independent equations (Wright, 1960). Clearly we have no difficulty satisfying this condition.

Our practical problem is accentuated since we lack a readily applicable set of guidelines showing a priori for complicated cases (where a portion of the system is overidentified and another portion is underidentified) which, if any, coefficients are identified. In the usual representation of structural equations in an (unstandardized) econometric system, the application of both the necessary ("order") and sufficient ("rank") conditions to each equation is straightforward. Our difficulty has been in failing to be able to extend this approach to complicated path analytic panel models. In principle, we should be able to follow Jöreskog's (1970) representation of the covariance structure and pinpoint identification problems from an inspection of the various variance-covariance matrices. To the present we have not been able to isolate the difficulties using this approach.

This brief discussion should make plain the fact that an investigation of the usefulness of single-indicator models measured with error must focus heavily on the conditions under which in over-identified models the parameters of interest are identified (or perhaps practically estimable). Unfortunately we have proceeded on a rather ad hoc basis since we have not found any simple algorithms which allow one to make such a judgment prior to writing out the systems of equations and searching for estimates.

A THREE-VARIABLE, THREE-WAVE, SINGLE-INDICATOR MODEL (3V, 3W, 1I)

We have already argued that primary enrollment ratios affect secondary ratios (and not the reverse), and it seems a natural extension to argue that economic development affects only primary ratios directly. To further simplify our analysis, we assume for the moment no cross-effect from either educational variable to development.[17] The model drawn in Figure 5 incorporating our usual simplifying assumptions concerning disturbances and measurement error terms represents the causal structure. Again we must require all of our (standardized) coefficients to be stable or none will be identified.

Figure 5 about here

This model has an excess of 25 equations over unknown coefficients. As before each of the overidentifying restrictions (redundant equations) can be written as quantities (Spearman tetrad differences) equal to zero under the hypothesis of no specification error (i.e. the model is correct) and no sampling error. The fit of these 25 equations with our sample is extremely close. The largest deviation from the predicted value is .021 and the mean of the absolute values of the deviations is .0067. As this latter figure suggests, most of the predictions are almost exactly met.

The massive inadequacies of single indicator models (even with strong simplifying assumptions such as those made above) become inescapable when we shift attention from testing to estimation. Just as with the cases analyzed earlier, only the autoregression coefficient for economic development, the most "independent" variable and the path coefficient linking this variable with its indicator are apparently estimable. It is easy in this case to see that it is <u>only</u> the presence of measurement error which rules out the estimation of the remaining coefficients. In other words, if each of the path coefficients linking variables with indicators is a priori set equal to some value, [18] all of the remaining coefficients in the model can be solved for in a number of ways (i.e. they will not only be identified, they will be over-identified). If only partial information is available (say the value of one of the "epistemic correlations", say r) a portion of the remaining coefficients can be solved for (in this example, b, d and ρ_c).

As a further partial check on the model we estimate the available coefficients with the result that $\hat{c} = 1.01$ and $\hat{s} = .987$. In this model c is a correlation and thus should not exceed unity. However, the discrepancy is smaller than was the case previously and the estimated value of s is reasonable.

Duncan (1972b) has shown that one need not stop here in a substantive analysis. In a number of the cases he examined, estimation of the model's remaining coefficients based on "provisional" estimates of some of the unknown quantities gave rise to either logical inconsistencies (e.g., correlations greater than unity, unreal solutions, etc.) or to substantively uninterpretable results. Since we are more concerned at this point with the logic of the analysis problems, we do not pursue this strategy, but note its potential usefulness.

We have reached consistently negative conclusions about the usefulness of single-indicator models. Even under the most idealized conditions, such models do not allow us to estimate the parameters of the postulated causal structure. However, we should not totally dismiss the fact that single-indicator models are falsifiable and thus can contribute at least negative evidence to substantive problems. In addition, a focus on the dynamics of the difference equations conveys important information. Yet, any consideration of more realistic complications must inevitably lead to the study of multiple indicator models. We make this shift in the next section.

A TWO-VARIABLE, THREE-WAVE, TWO-INDICATOR MODEL (2V, 3W, 2I)

In this section, we alter the model in Figure 2, with which we began the analysis, in an effort to eliminate the causal interconnections which produced the earlier anomalous results. Specifically, we choose as our indicators of educational expansion a combined ratio of primary and secondary students to the appropriate age-specific population and the ratio of students in tertiary schools to the total population.[19] Economic development is measured by GNP/cap and KWH/cap. Unfortunately, the existing data is such that the inclusion of GNP/cap reduces our sample to the 46 (presumably) most developed nation-states.[20]

We will continue to explore the consequences of measurement error in the context of a substantive model which specifies asymmetric cross-effects from economic development to educational expansion over the time period of investigation. The revised model incorporating double indicators is diagrammed in Figure 6.

Figure 6 about here

Before analyzing this model we must digress and consider an anlysis problem raised by Blalock (1970) which seems to point to a problem with our model. To do this, we specialize the model drawn in Figure 6 in the following ways. Concentrate only on the educational "half" of this model, i.e., treat the economic development variable as unobserved and thus part of the residual. Assume (contrary to the model drawn in Figure 6) that the residuals are uncorrelated with the included variables and are not stable over waves of observations (i.e., are intertemporally uncorrelated). Assume further that the measurement error terms are correlated both simultaneously for different indicators and intertemporally for the same indicators measured at different points in time. Finally, assume that the correlations of measurement error terms are stable over waves of observations. The restrictions give us the model discussed by Blalock

(1970: Figure 5) which is diagrammed in Figure 7.

Figure 7 about here

Note that this model requires the "epistemic correlations" to be stable but allows the autoregression terms to vary between waves of observations.

This model gives rise to the following system of 15 equations:

$r_{x_{11}x_{12}} = ab + a'b'f$	(1')	$r_{x_{11}x_{22}} = abc'$	(10')
* $r_{x_{21}x_{22}} = ab + a'b'f$	(2')	$r_{x_{11}x_{32}} = abc'd'$	(11')
* $r_{x_{31}x_{32}} = ab + a'b'f$	(3')		(12')
$r_{x_{11}x_{21}} = a^2c' + (a')^2g$	(4')	* $r_{x_{12}x_{21}} = abc'$	(13')
$r_{x_{21}x_{31}} = a^2d' + (a')^2g$	(5')	* $r_{x_{12}x_{31}} = abc'd'$	(14')
$r_{x_{11}x_{31}} = a^2c'd' + (a')^2g'$	(6')	$r_{x_{21}x_{32}} = abd'$	(15')
$r_{x_{12}x_{22}} = b^2c' + (b')^2h$	(7')	* $r_{x_{22}x_{31}} = abd'$	
$r_{x_{22}x_{32}} = b^2d' + (b')^2h$	(8')		
$r_{x_{12}x_{32}} = b^2c'd' + (b')^2h'$	(9')		

(* redundant)

As Blalock notes the following estimates obviously fall out:

$$c' = \frac{r_{x_{11}x_{32}}}{r_{x_{21}x_{32}}} = \frac{abc'd'}{abd'} \; ; \quad \frac{r_{x_{11}x_{32}}}{r_{x_{11}x_{22}}} = \frac{abc'd'}{abc'}$$

However, one can obtain estimates of a and b, by for example subtracting (4') from (5') and multiplying both sides by $(ab)^2$ as long as $c' \neq d'$.

> The above procedure depends on the assumption that $c' \neq d'$. If, in fact the two stability coefficients are nearly identical in the population, then even though their sample counterparts may be slightly different, there will be very large sampling errors for the estimates of the ratio a^2/b^2 and also for all the estimates dependent on this ratio. Therefore, for all practical purposes, the procedure will be useful only if the stability coefficients c' and d' are very different. (Blalock, 1970: 109).

While this statement is accurate as it refers to the estimation procedure Blalock employed, it is somewhat misleading if it is taken to apply more generally to the estimation of "stability" coefficients in this model.

To see this, let c' = d', and rewrite the system of equations just considered:

$$r_{x_{11}x_{12}} = ab + a'b'f = r_{x_{21}x_{22}} = r_{x_{31}x_{32}} \qquad (1'' - 3'')$$

$$r_{x_{11}x_{21}} = a^2c' + (a')^2g = r_{x_{21}x_{31}} \qquad (4'' - 5'')$$

$$r_{x_{11}x_{31}} = a^2(c')^2 + (a')^2g' \qquad (6'')$$

$$r_{x_{12}x_{22}} = b^2c' + (b')^2h = r_{x_{22}x_{32}} \qquad (7'' - 8'')$$

$$r_{x_{11}x_{22}} = abc' = r_{x_{12}x_{21}} = r_{x_{22}x_{31}} \qquad (10'' - 13'')$$

$$r_{x_{11}x_{32}} = ab(c')^2 = r_{x_{12}x_{31}} \qquad (14'' - 15'')$$

We see immediately that in this case we employ <u>different</u> equations to estimate all of the coefficients. The estimation procedes as follows:

$$\hat{c}' = \frac{r_{x_{11}x_{32}}}{r_{x_{11}x_{22}}} = \frac{ab(c')^2}{abc'} .$$

$$\hat{ab} = \sqrt{\frac{r_{x_{11}x_{22}}\, r_{x_{12}x_{21}}}{r_{x_{11}x_{32}}}} = \sqrt{\frac{a^2b^2(c')^2}{ab(c')^2}} .$$

Subtracting equation (6') from (4') and (9') from (7') and taking a ratio of these quantities gives:

$$\frac{\hat{a}^2}{\hat{b}^2} = \frac{r_{x_{11}x_{21}} - r_{x_{11}x_{31}}}{r_{x_{12}x_{22}} - r_{x_{12}x_{32}}} = \frac{a^2(c' - (c')^2)}{b^2(c' - (c')^2)}$$

Combining the above expressions yields:

$$\hat{a}^4 = \frac{r_{x_{11}x_{22}}\, r_{x_{12}x_{21}}}{r_{x_{11}x_{32}}} \cdot \frac{r_{x_{11}x_{21}} - r_{x_{11}x_{31}}}{r_{x_{12}x_{22}} - r_{x_{12}x_{32}}}$$

Given estimates of c and a, we can use equations (10' - 13') to provide estimates of b, and the disturbances.

It is obvious that this method faces a restriction similar to that discovered by Blalock. Our estimates of the ratio a^2/b^2 will have very large sample variance if c' is very close to $(c')^2$, i.e. if c' is very close to unity in the population. Thus this method is useful only if the process is such that these correlations are considerably less than unity. However, in the multivariate applications, the "stability coefficients" will be "partial" coefficients and thus may exceed unity. This is a frequent occurrence in the models with which we have worked. The implication is that the estimation method is not restricted to the case of "unstable" systems, i.e. those in which factors left out of the analysis are quite important in producing intertemporal variation in the variable under study. The method is applicable even when such variables are introduced explicitly into the analysis. The requirement is that the stability in the variable itself (as opposed to stability in other variables) not be so high as to produce almost no intertemporal variation. More concretely, if there are variables which produce systematic variation in the variable under study over time, the estimation method should not be expected to have exceedingly large sampling variance.

The important point here is that the utility of any estimation method must be evaluated relative to the substantive model under study. A method which is optimal for unstable autoregression coefficients may not be optimal for the stable case. This conclusion is reinforced by the demonstration that our method of rewriting the system of equations so as to introduce the presumed complication does not contradict Blalock's analysis of the requirement of moderately low intratemporal correlations among at least two indicators in his discussion of the three-wave, three-indicator (single variable model).

The point of this digression is to demonstrate that if our assumptions are justified we can expect to estimate the coefficients of the model drawn in Figure 6, even if the autoregression parameters are stable over waves of observations. With this assurance we can return to that model and proceed to examine the consequences of the complications thought to be most troublesome in this type of substantive application.

The first problem is familiar. With 66 equations and only eight unknowns, we face a bewildering variety of estimates for many of the coefficients. The autoregression in economic development, b, can be estimated at least 50 different ways. The problem is that the various estimates must surely differ in sampling variance. Some estimates are "direct" in the sense that they are given by ratios of two sample correlations. These "direct" estimates can then be used to solve rather more complicated systems of linear equations involving the quantity b. It is no simple matter to even identify all of the possible ways to estimate each coefficient in a model as simple as this (the complications arise through, for instance, the many possible paths connecting educational ratios between the last two time periods) nor to establish which estimates are independent. The practical problem is that the various estimators differ considerably. One would normally suspect that considerable divergence in the estimates would suggest specificaion error in the model. However, as long as the sampling distributions of the estimators is unknown, such inferences do not have any firm support. Consequently we have continued to report composite estimators which make use of some but not all of the information in the sample. We have made no attempt to exhaustively incorporate all of the logically possible independent estimates. Our estimates, then, are highly tentative as they make use primarily of the most "direct" methods of estimation. As we shall see below, this approach if carried out in substantive analyses has some serious drawbacks.

Thus we proceed to estimation of the coefficients of the model in Figure 6 assuming for the moment that j=k=l:[21]

$\hat{a} = 1.205$	$\hat{g} = .935$
$\hat{b} = 1.003$	$\hat{h} = .978$
$\hat{c} = -.210$	$\hat{p} = .772$
$\hat{e} = 1.031$	
$\hat{f} = .368$	

The only obvious difficulty is that $\hat{e}$ exceeds unity (and, as a consequence, $\hat{f}$ is small in magnitude). However, $\hat{b}$ is so close to unity that if we were limited to the estimation procedure presented as an alternative to the one suggested by Blalock, the estimates would be extremely responsive to sample error. The fact that we have three waves and six indicators results in considerably more ways to estimate the model coefficients. However, this may only serve to mask the consequences of the obvious high autoregression terms. We have not yet been able to assess the seriousness of any such problems at this time.

As before we proceed to consider the implications of several types of nonrandom measurement error. As noted earlier, among the most realistic nonrandom errors in cross-nation research is organizational "memory". All of the measures we employ are generated by national bureaucracies. It seems reasonable to argue that these bureaucracies tend to err in the same sorts of ways consistently over time. We examine this type of complication by simplifying the argument to specify a 5-year memory and allow for correlated measurement error for the same indicators over five year intervals but not longer. We have restricted the intervals to five years to simplify the problem, however, it seems reasonable to argue that organizational memory in this sense is "covariance stationary." We will continue the analysis with the last substantive model considered which is adapted for our present purposes and rediagrammed in Figure 8.

Figure 8 about here

Only 26 of the equations for this model are different from those in the previous model (Figure 7). The addition of the "memory" terms eliminates numerous simple equation systems yielding estimates of b. As a result, assuming j=k=1, the estimate of b becomes 1.06, perhaps enough greater than unity to cast doubt on this particular model. There is no apparent change in the estimation procedures for the remaining coefficients. More precisely, every valid estimate for any of the remaining terms in this model is also appropriate under the specification of the model with uncorrelated measurement error terms. Since the magnitudes of the correlations between measurement error terms are not obviously estimable, we can proceed no further without additional assumptions. The logic for proceeding with this model is not obvious since it fits the logical bounds (on magnitudes of correlation coefficients) less well than the more restrictive model discussed earlier.

We could engage in a completely parallel analysis of the proposition that the bureaucracies tend to err in the same sorts of ways in a given time period in reporting <u>different</u> national account statistics. In such a case we would allow different indicators of the same variable as well as indicators of different variables measured at the same point in time to be correlated. Since we could not estimate the systematic error components in this case, we have shifted attention to other problems.

The final class of nonrandom measurement errors we consider in the context of this model is the systematically (proportionally) decreasing random error discussed in the first section. That is we relax the restriction that j=k=1 in the model drawn in Figure 6.

By and large the estimation procedure is as above. On exception involves the proportionality terms for the decreasing error, j and k. In this case, we solve directly for these terms and then proceed as above using the estimates of j and k wherever such terms appear. Recall that our substatnive understandings require $0<j\leq 1$ and $0<k\leq 1$. Our sample estimates are

$$\hat{j} = .991$$
$$\hat{k} = 1.090$$

The result for k is much like that obtained earlier for the single indicator educational ratio model. Clearly this particular nonrandom error hypothesis is not appropriate for at least this portion of the model in Figure 7. Our examination of the data strongly suggests that failure here is due to a secular trend of decreasing cross-sectional variance in the primary-secondary ratio.[22] This trend reflects a type of "ceiling effect" which is enormously problematic in standardized models. At any rate, given the unrealistic estimate of k, there is no point in proceeding to estimate terms which depend on $\hat{k}$.

It is interesting to consider the consequences of proportionately decreasing error for the remainder of the model (the economic development portion). Allowing for this type of nonrandom error raises the estimates of path coefficients linking economic development with its indicators, $\hat{g}$ and $\hat{h}$, from .935 to .994 and from .978 to .989 respectively and lowers the estimate for the autoregression, b, from 1.003 to .879. Both types of changes are quite encouraging for substantive analysis.

We had originally become interested in this type of nonrandom error because of an interest in eventually modeling substantive processes which involve lags of <u>different</u> lengths. For example, we might argue that the lag in the causal effect of educational expansion on economic development is twice as long as the lag in the reverse effect. In cases like this we should expect that over any time period of observation the longer lagged effects will be more seriously affected by random measurement error. Unless the analyst takes the decrease in the time-dependent magnitude of random errors into account, he is likely to make incorrect inferences in comparing the magnitudes of the longer lagged and shorter lagged effects.

An example of the type of model in which this would be problematic is drawn in Figure 9. Here the lag for the effect from educational expansion is

---

Figure 9 about here

---

two waves of observations (ten years) and the lag for the effect from economic development to educational expansion is one wave (five years).

The failure of our model for decreasing random error in the educational expansion of the model in Figure 7 rules out the possibility of estimating all of the coefficients in this new model. Yet, we can see some of the consequences of this type of error using a hypothetical value of k. If we restrict j=k=1, d = -.076. However, when we assume that j=k=.991, the estimate of d is increased to -.220, a considerable increase. This exercise does suggest that our original concerns were justified and that researchers modeling processes like that under study here ought to attend to such nonrandom error.

The model drawn in Figure 9 illustrates one further difficulty with *ad hoc* estimation methods for complicated path models. The addition of a long-lagged cross-effect has only a very slight effect on ohter estimates in the model since the term appears in relatively few equations. This, given simple composit estimates, has the consequence of minimizing the difference between a model which has such an effect and one which does not. This factor will often make it very difficult to choose between two such models when both are confronted with the same sample data.

But the more serious problem lies with the "stepwise" method of estimation used by us, Duncan (1972a,b), and Blalock (1970). In this procedure we first estimate (inserting sample estimates) those terms which appear simply as, say, ratios of population correlations. Then those first-order estimates are used to solve more complicated expression to produce "derived" estimates for additional terms. In very complicated models like those considered in this paper, the analyst may have to go through several steps. The difficulty is the following. We can solve directly for b in the model in Figure 9 without taking into account the presence of the other cross-effect, d. Then we use this estimate of b to solve for d. However, it may be possible to solve more complicated systems of equations for b and d simultaneously. Obviously the latter procedure would be preferable since it would more faithfully represent the causal structure of the model by simultaneously taking into account the presence of both cross-effects. We should expect that estimates produced by such a procedure will ordinarily differ from those arrived at by the procedure we used. This estimation problem will loom quite important in the substantive research which motivated this analysis.

CONCLUDING REMARKS

We will not attempt to recapitulate the series of technical results scattered throughout the paper. The main point is that inferences in the multiwave, multivariable panel are much more complicated than was generally realized. Results from single-variable models with measurement error are not easily generalized to more complicated cases. Moreover the consequences of measurement error are not easily generalized but depend heavily on the specific features of the model in which it occurs. In this sense this paper reinforces the developing consensus in sociological methodology that simple formulations of the consequences of even random measurement error (e.g. attenuation) are not likely to be invariant across models. This new emphasis is beneficial since social scientists appear to have begun to rely too heavily on stock reactions to the presence of measurement error.

We have noted at numerous points difficulties of estimation. This type of work is greatly hampered by the lack of a systematic theory of identification and statistical inference for realistic panel models. The thrust of recent statements (Hauser and Goldberger, 1971, Werts, Linn and Jöreskog, 1971) is to suggest that sociologists may not have to invent such a theory but may be able to borrow formulations from econometrics and biometrics. Unfortunately, the formulations which have been applied to path models to this point are not easily generalizable to cases we have considered. It is clear that this sort of work must proceed before the causal approach to measurement error will be practically useful to sociologists employing panel models in substantive research.

FOOTNOTES

1

Considerable attention has recently been paid to the problem of merging cross-sectional and longitudinal designs in the econometric literature. See Nerlove (1971) for a clear exposition of the methodological issues involved.

2

The most important of these is that standardized coefficients remain stable over the time period of observation so that the addition of waves of observations does not add additional unknowns.

3

This is not universally true, of course. In some realistic cases, sociologists may have access to enough measurements and *a priori* restrictions on the model so that both types of complications may be dealt with. This has not been the case in our research, however.

4

This point is quite important in the substantive context, i.e. comparative research, since so much quantitiative cross-national research has attempted to deal with measurement problems inductively, e.g. by employing exploratory factor analyses. Even exploratory factor analysis requires substantive assumptions at some point. The operative issue then is whether or not the assumptions are made explicit and justified substantively.

5

When we employ two indicators of each variable, the number of observations drops to 46 due to missing data on GNP in the earliest time period. Obviously this makes the single-indicator and two-indicator models noncomparable. We have chosen this option to minimize the "ceiling effect" in primary school enrollments.

6

The difference between a cyclic process and a monotonic process in this respect is easily visualized by diagramming the process and then arbitrarily shifting the observation points along the time axis.

7

These variables have been logged to make their relationship to other variables linear.

8

Again, this argument involves implicit statistical inference since it is possible to obtain sample results that violate the restriction when it is satisfied in the population.

9

The *k* term must be less than one because the value of the total "epistemic correlation" (the decreasing component *k* times the stable random error component) *increases* over time to a maximum value of 1. In other words, the *residual paths* for the measured values (e.g. $\sqrt{1 - (k^2e)^2}$) *decrease* over time.

10
See the discussion of identification below.

11
The issues involved in estimating over-identified path models are rather complex. Hauser and Goldberger (1971) have shown that for models like ours with all recursive "arrows" allowed by the model specification to take on non-zero values, the best estimator is a maximum-likelihood procedure. Since for most of the models we use some causal connections are assumed to be absent, this method is apparently not appropriate (given the present state of our knowledge). Thus we follow Duncan's heuristic method recognizing that the properties of the resulting estimators have not been studied. This procedure seems justified only so long as we are concerned mainly with the general properties of these models as distinct from precise estimates of causal parameters. At the point where attention focuses on estimation, we would suggest following Hauser and Goldberger (1971) and Jöreskog (1969).

12
Duncan (1971b) has exhaustively treated the problems of inference in two-wave, two-variable panel models with common factors.

13
The parameter estimates are taken from regression analyses which do not allow for measurement error.

14
The estimation equations for the coefficients we estimate are obvious from the path equations. It is possible both that more complicated estimates of these terms are available and that more complicated systems of equations yield solutions for the other terms. In the cases we investigated, this did not seem to be the case. Most often the systems of nonlinear equations were not amenable to direct algebraic solution. In the cases where we were able to reduce the estimation equations to quadratic, cubic or quartic equations, we did not obtain real roots for all of the unknown terms. Since our search procedures were not entirely systematic, however, we cannot assert with confidence that no other estimates exist. In this and what follows we refer to "obvious" solutions when we refer to the estimation status of parameters.

15
In fact, with these assumptions and the assumption of no measurement error, we can allow each autoregression and cross-effect to have a unique value between each two waves of observations.

16
This is the reason that the addition of waves of observations does not eliminate the estimation problem. This only gives rise to more complicated expressions relating early "independent" variables to later "dependent" variables.

17
The literature on "returns to education" in economics suggests that this assumption does not grossly violate reality at least over the time span we are considering.

18 We mentioned at the outset that sociologists are sometimes in the position to be able to use outside information, previous research, etc. to produce reasonable estimates of measurement quality. Of course, the most common practice is setting the paths from variables to indicators equal to unity -- assuming perfect measurement.

19 This seemingly unreasonable standardization is, in fact, the one recommended by UNESCO (need citation) since the appropriate age -group is unclear. A superior measure which we will eventually incorporate is per pupil expenditures.

20 This is particularly problematic when one uses (as we do) standardized coefficients. In such a case the changing variances will create unstable (standardized) population parameters even when the slopes are invariant across time periods.

21 All of the coefficients but a and c are solved by the Duncan estimate of nine equivalent estimating expressions chosen unsystematically from the obvious possibilities. The estimation of a and c requires the solution of systems of two equations in two unknowns. There are a number of equivalent systems in the same two unknowns and the method of combining the alternative solutions is not obvious. For the lack of any better method we took the arithmetic mean of five sets of solutions.

22 In general when a model specification fails, the analyst assigns blame to specific elements with considerable risk of error. Thus in this case the inappropriateness of standardized coefficients with systematically changing variances may be at fault as we suggest, or the causal structure may be wrong, or there may be any number of additional defects.

* (From title page) The research reported herein was performed pursuant to a contract with the United States Department of Health, Education, and Welfare, Office of Education, and was partially supported by NSF Grant (GS-23065). John Boli performed many of the computations for this research.'

Table 1

Correlation Matrix of Educational Measures

	Secondary 1955	Secondary 1960	Secondary 1965
Primary 1955	.212	.303	.391
Primary 1960	.115	.149	.236
Primary 1965	.066	.089	.123

TABLE 2

SOME SAMPLE RESULTS FOR CROSS-LAG CORRELATIONS

$$
\begin{tikzcd}
S_t \arrow[r, "a"] \arrow[d, bend right, leftrightarrow, "\rho"'] & S_{t+1} \arrow[r, "a"] & S_{t+2} \\
P_t \arrow[r, "b"] \arrow[ur, "c"] & P_{t+1} \arrow[r, "b"] \arrow[ur, "c"] & P_{t+2}
\end{tikzcd}
$$

$$\rho_{P_{t_1} S_{t+k}} = \rho a^k + c \sum_{j=0}^{k-1} a^{k-j-1} b^j$$

	a=b= .9 ρ = .3 c = .1	a=b= .8 ρ = .3 c = .1	a=b= .7 ρ = .3 c = .1
k = 1	.37	.34	.31
2	.42	.35	.29
3	.46	.34	.25
4	.49	.33	.21
5	.507	.30	.17
10	.49	.173	.004

	a=b= .9 ρ = .3 c = .2	a=b= .8 ρ = .3 c = .2	a=b= .7 ρ = .3 c = .2
k = 1	.47	.44	.41
2	.60	.51	.43
3	.70	.54	.40
4	.78	.53	.34
5	.84	.51	.29
10	.86	.31	.08

Table 3

Derivations of Tetrad Differences from Zero

Case (i)	(ii)
.027	.084
.102	.001
.108	.016
.045	.052
.029	.086
.007	.065
.002	.087
.060	.014
.026	.012

FIGURE 1*

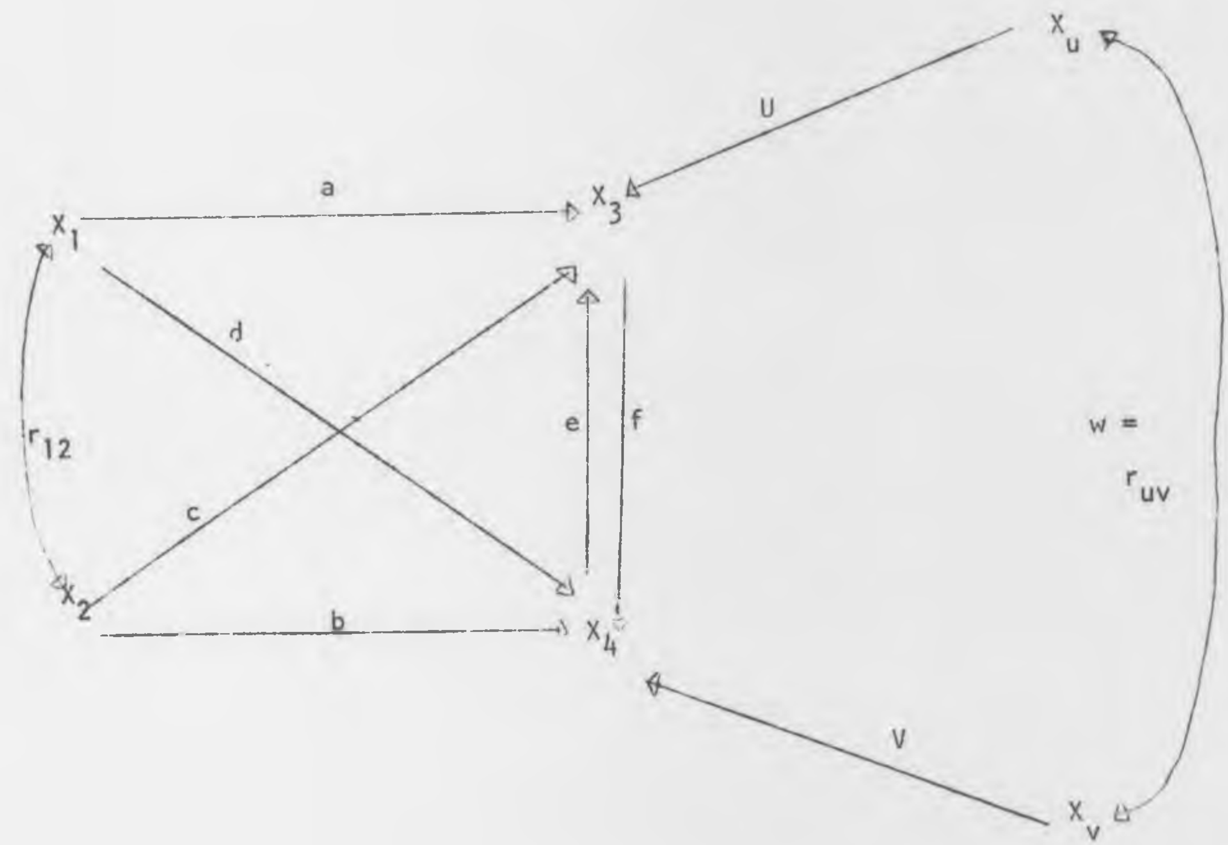
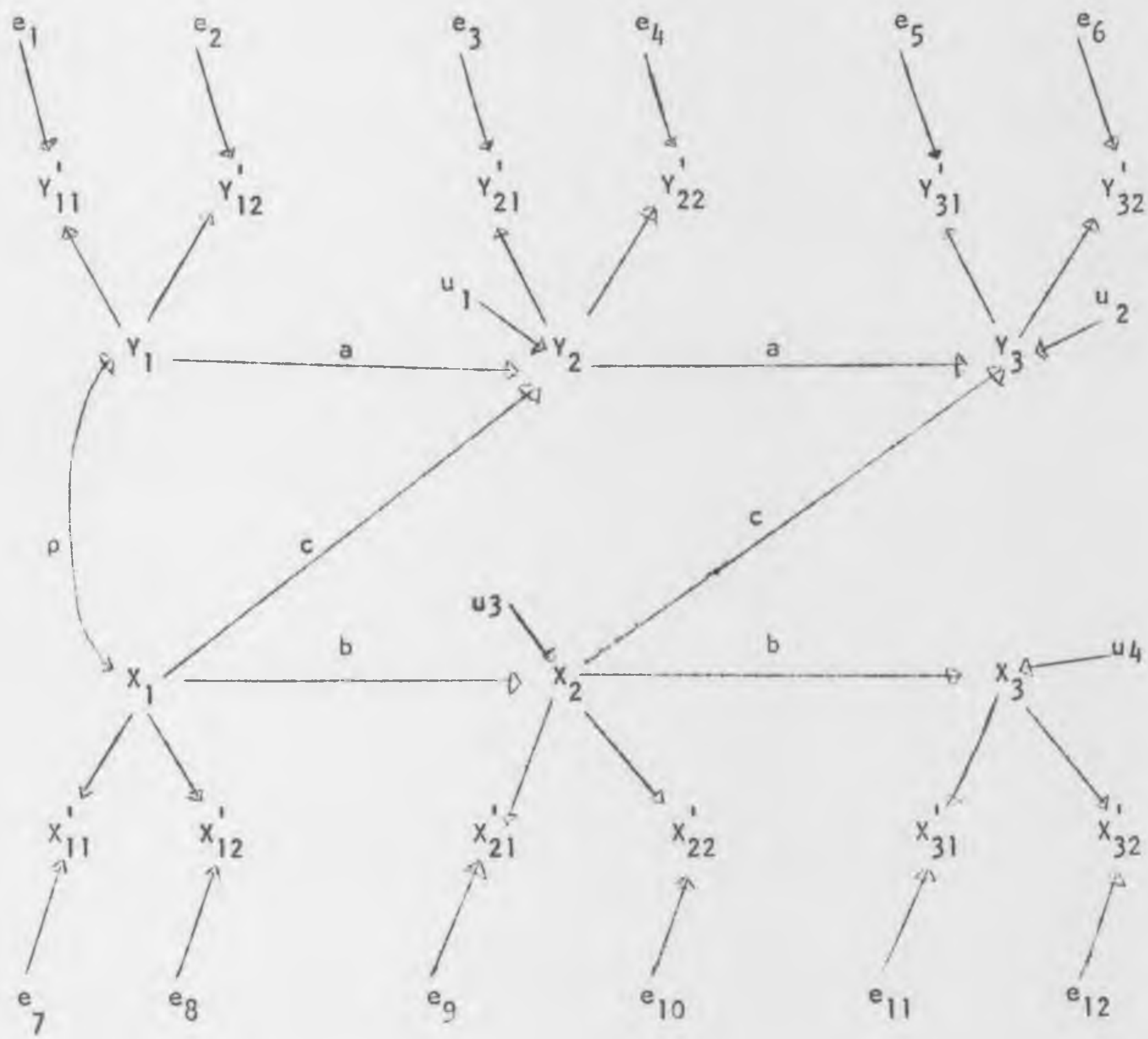


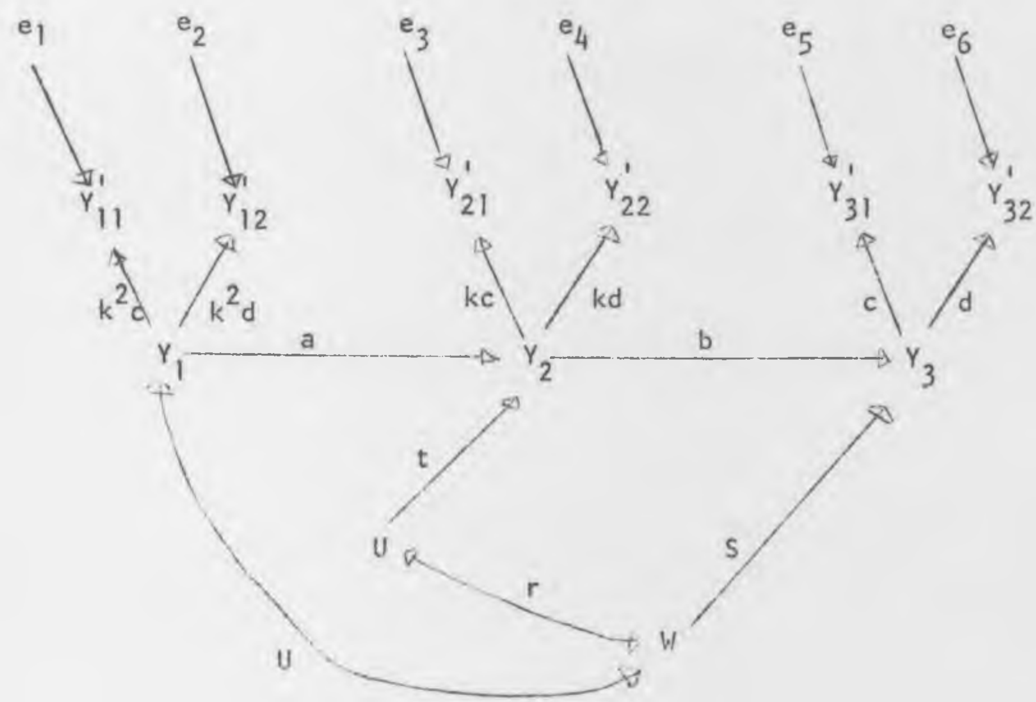
*Figure taken from Duncan (1969: Figure 1)

FIGURE 2



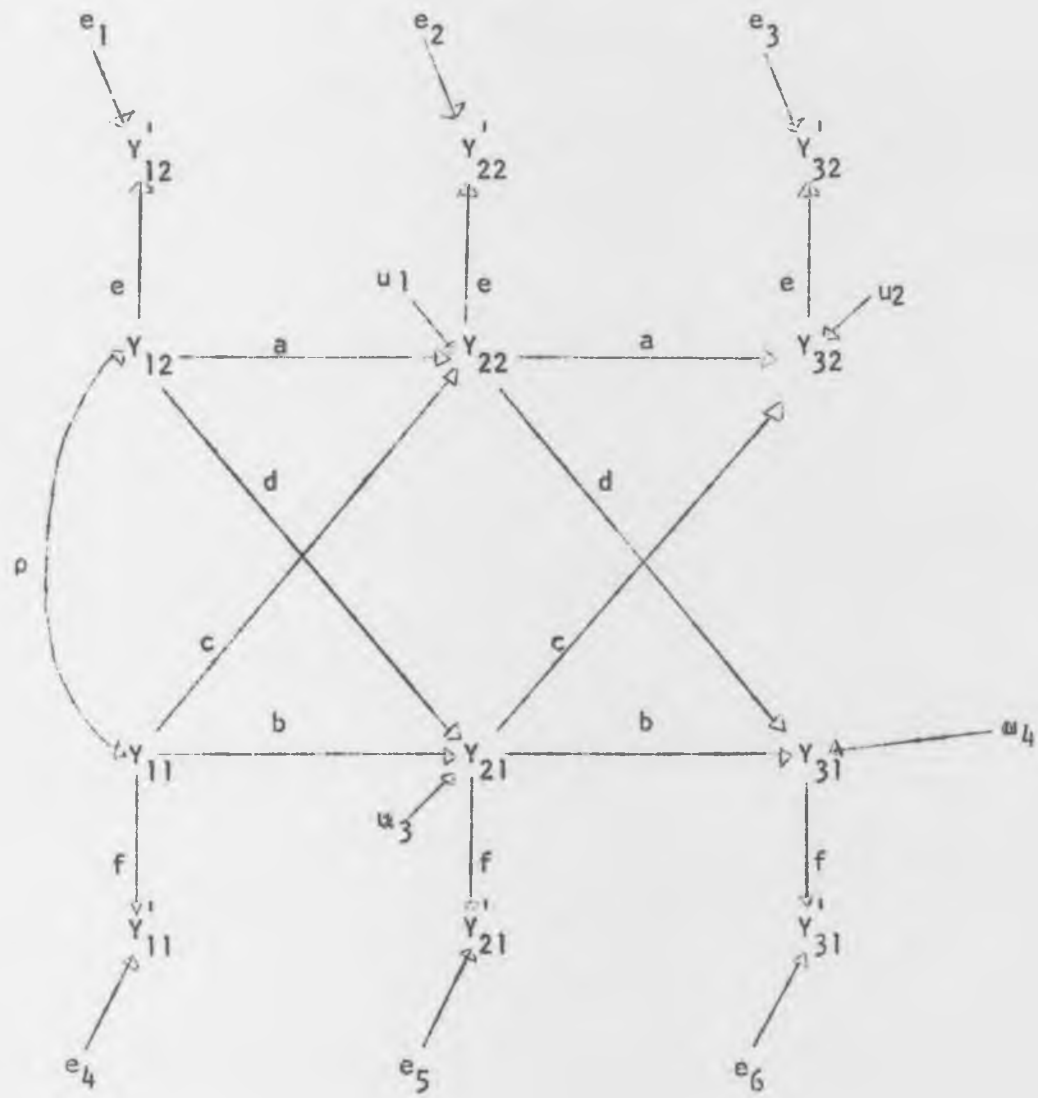
e1
e2
e3
e4
e5
e6
Y'11
Y'12
Y'21
Y'22
Y'31
Y'32
u1
u2
Y1
a
Y2
a
Y3
ρ
c
c
u3
b
b
u4
X1
X2
X3
X'11
X'12
X'21
X'22
X'31
X'32
e7
e8
e9
e10
e11
e12

FIGURE 3



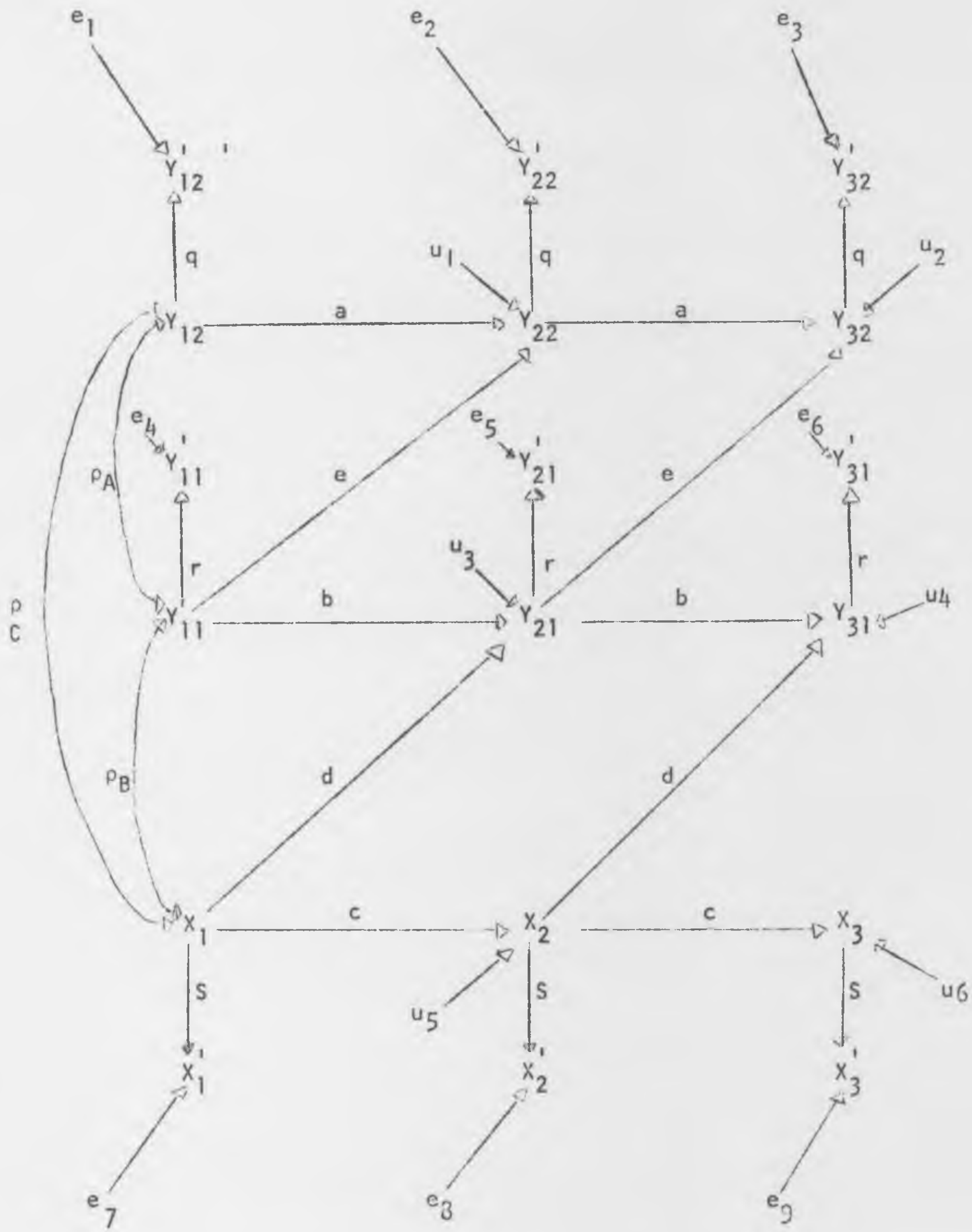
e1
e2
e3
e4
e5
e6
Y'11
Y'12
Y'21
Y'22
Y'31
Y'32
k2c
k2d
kc
kd
c
d
Y1
Y2
Y3
a
b
t
U
S
r
W
U

FIGURE 4



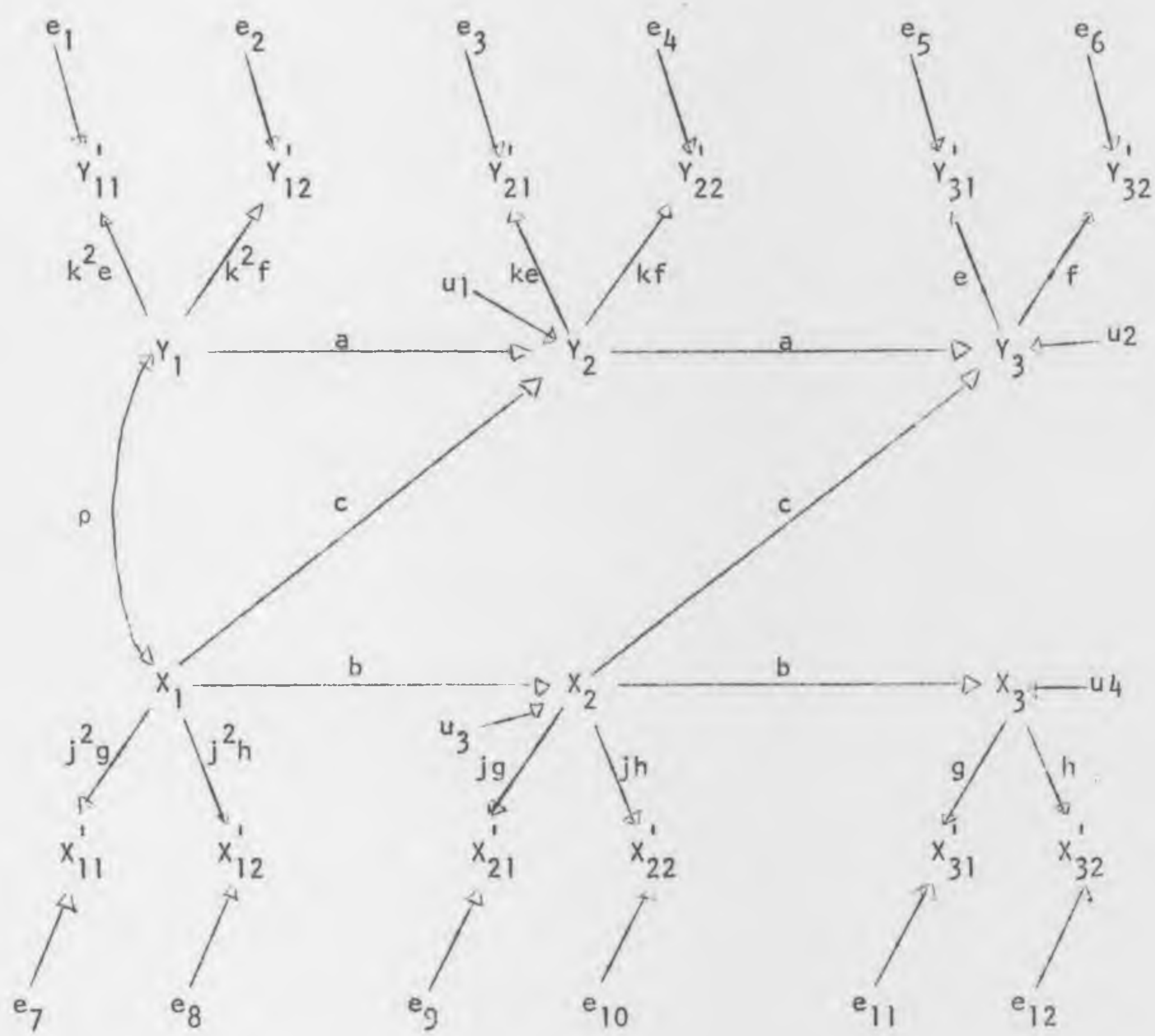
e1
e2
e3
Y'12
Y'22
Y'32
e
e
e
u1
u2
Y12
a
Y22
a
Y32
d
d
p
c
c
b
b
Y11
Y21
Y31
u4
u3
f
f
f
Y'11
Y'21
Y'31
e4
e5
e6

FIGURE 5



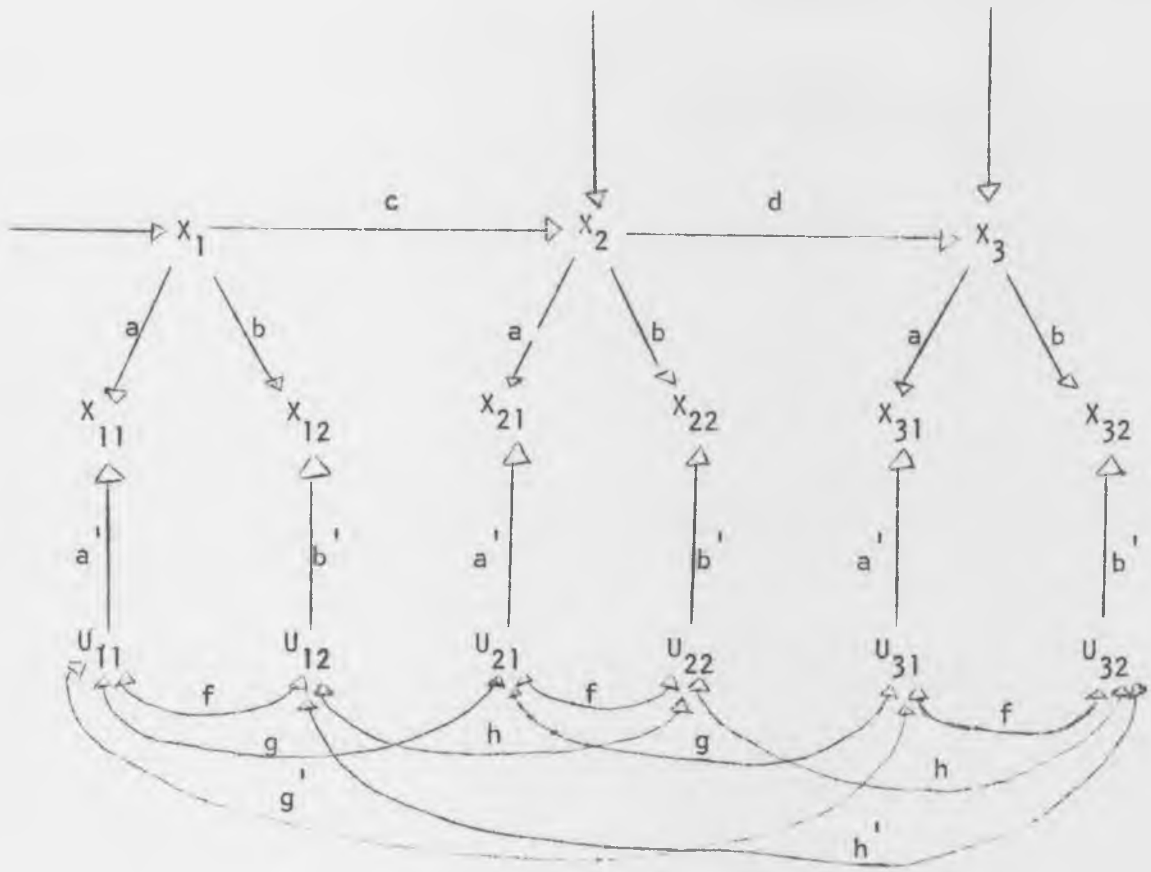
e1
e2
e3
Y'12
Y'22
Y'32
q
q
q
u1
u2
Y12
a
Y22
a
Y32
e4
e5
e6
Y'11
Y'21
Y'31
ρA
e
e
r
r
r
u3
ρ
C
Y11
b
Y21
b
Y31
u4
ρB
d
d
X1
c
X2
c
X3
S
S
S
u5
u6
X'1
X'2
X'3
e7
e8
e9

FIGURE 6



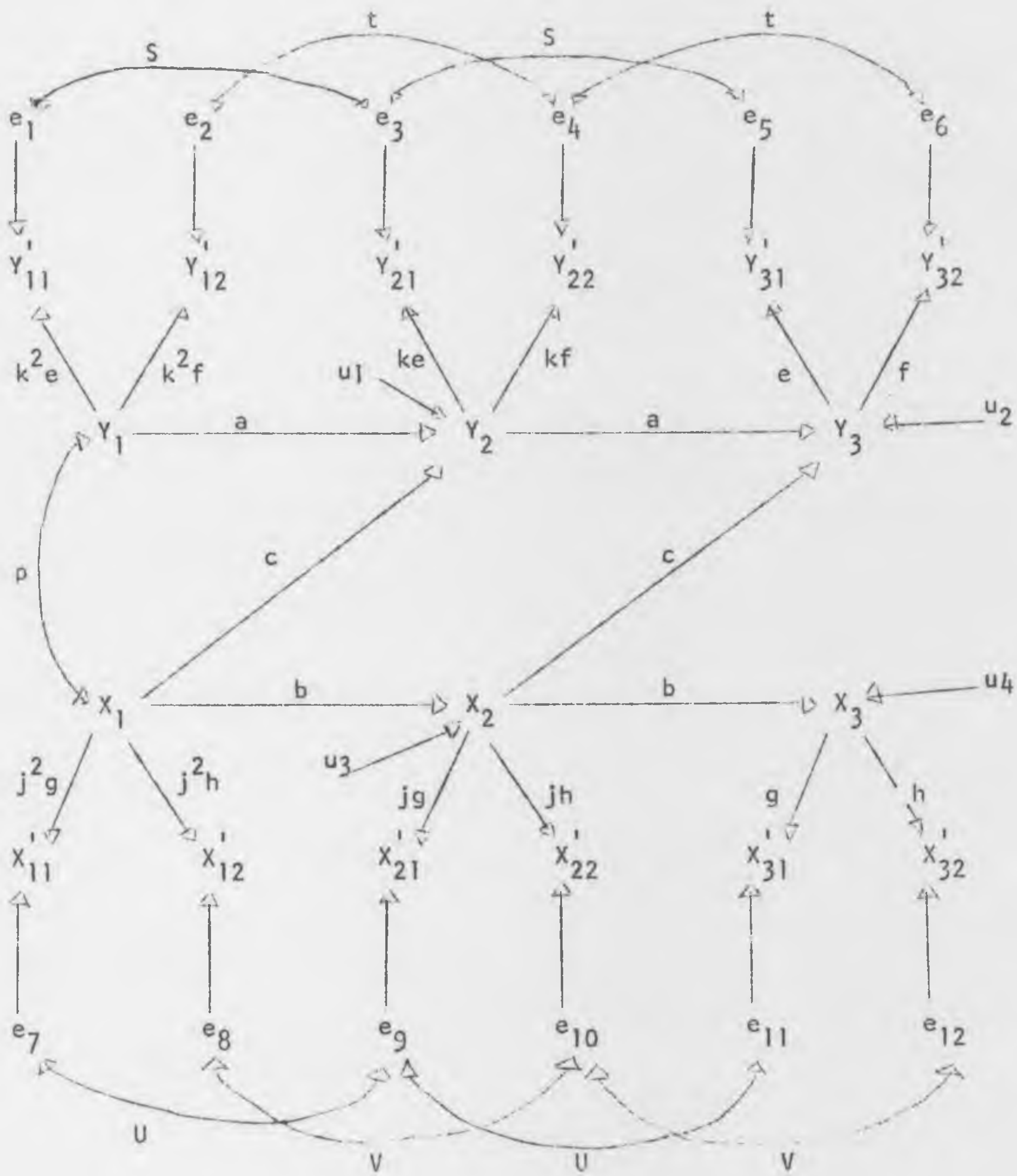
e1
e2
e3
e4
e5
e6
Y'11
Y'12
Y'21
Y'22
Y'31
Y'32
k2e
k2f
u1
ke
kf
e
f
Y1
a
Y2
a
Y3
u2
p
c
c
X1
b
X2
b
X3
u4
j2g
j2h
u3
jg
jh
g
h
X'11
X'12
X'21
X'22
X'31
X'32
e7
e8
e9
e10
e11
e12

FIGURE 7*



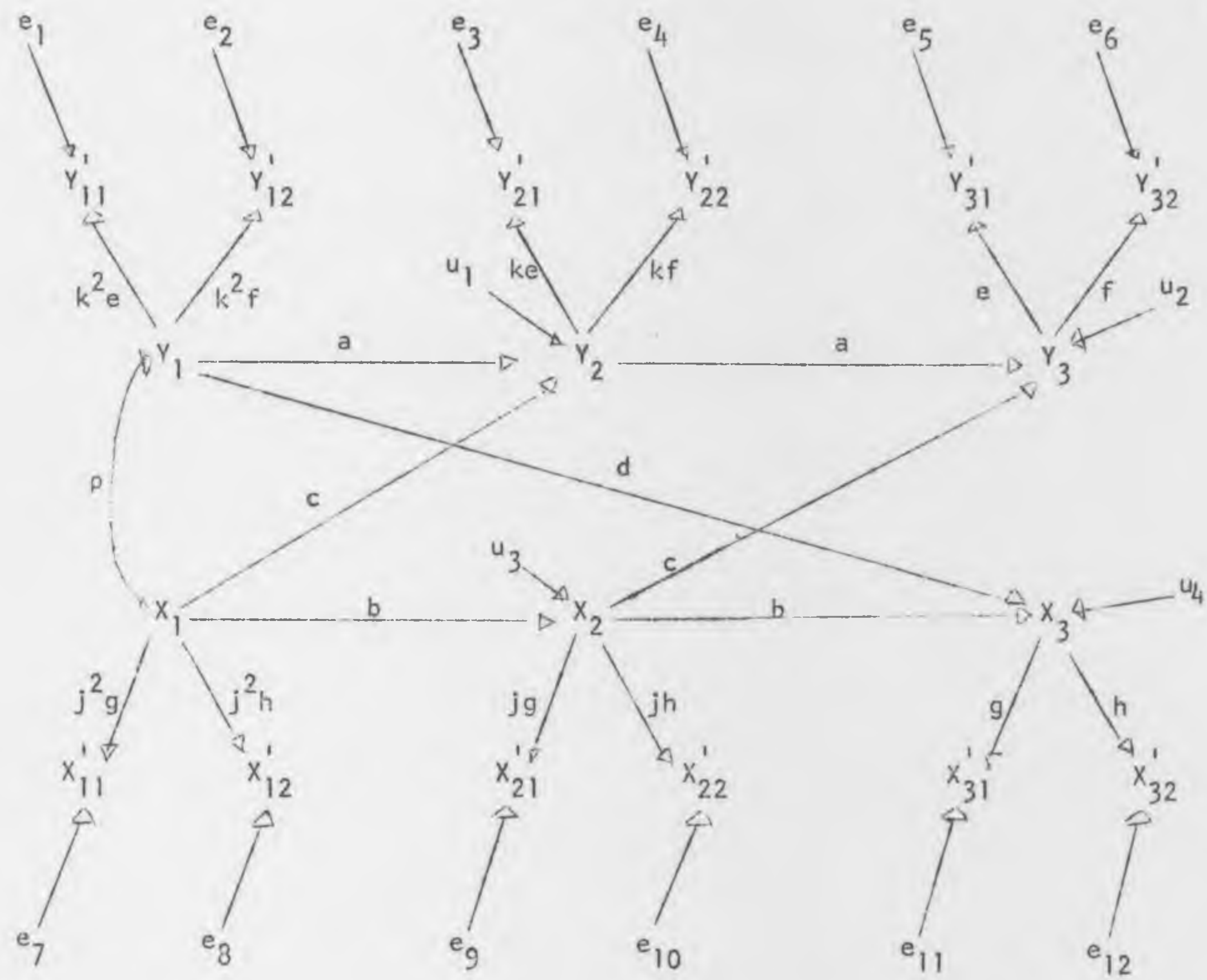
*Blalock (1970: Figure 5)

FIGURE 8



t
S
S
t
e1
e2
e3
e4
e5
e6
Y'11
Y'12
Y'21
Y'22
Y'31
Y'32
k2e
k2f
u1
ke
kf
e
f
Y1
a
Y2
a
Y3
u2
p
c
c
X1
b
X2
b
X3
u4
u3
j2g
j2h
jg
jh
g
h
X'11
X'12
X'21
X'22
X'31
X'32
e7
e8
e9
e10
e11
e12
U
V
U
V

FIGURE 9



e1
e2
e3
e4
e5
e6
Y'11
Y'12
Y'21
Y'22
Y'31
Y'32
k2e
k2f
u1
ke
kf
e
f
u2
Y1
a
Y2
a
Y3
ρ
c
d
u3
c
X1
b
X2
b
X3
u4
j2g
j2h
jg
jh
g
h
X'11
X'12
X'21
X'22
X'31
X'32
e7
e8
e9
e10
e11
e12

REFERENCES

Blalock, Hubert M.

1964 Causal Inferences in Nonexperimental Research. Chapel Hill: University of North Carolina Press.

1970 "Estimating measurement error using multiple indicators and several points in time," American Sociological Review 35(February):101-111.

Campbell, Donald T., and Donald W. Fiske

1959 "Convergent and discriminant validation by the multitrait-multimethod matrix," Psychological Bulletin 56(March):81-105.

Costner, Herbert L.

1969 "Theory, deduction and rules of correspondence," American Journal of Sociology 75(September):245-263.

Duncan, Otis Dudley

1969 "Some linear models for two-wave, two-variable panel analysis," Psychological Bulletin 72(September):177-182.

1972a "Some linear models for two-wave, two-variable panel analysis, with one way causation and measurement error," in H. M. Blalock et al., (eds.) Mathematics and Sociology (forthcoming).

1972b "Unmeasured variables in linear models for panel analysis: some didactic examples," in H. Costner (ed.) Sociological Methodology (forthcoming).

Hauser, Robert M. and Arthur S. Goldberger

1971 "Treatment of unobservable variables in path analysis," in Herbert L. Costner (ed.) Sociological Methodology 1971. San Francisco: Jossey-Bass.

Heise, David R.

1970 "Causal inference from panel data," in Edgar F. Borgotta and George W. Bohrenstedt (eds.) Sociological Methodology 1970. San Francisco: Jossey-Bass.

Jöreskog, K.

1970 "A general method for the analysis of covariance structures," Biometrika 57( ):239-251.

Meyer, John and Michael T. Hannan

1971 "The interrelationships of national educational systems with political and economic institutions," Unpublished research proposal submitted to the National Science Foundation.

Nerlove, Marc

1971 "Further evidence on the estimation of dynamic economic relations from time series," Econometrics 39(March):359-382.

Werts, Charles, E., Robert L. Linn and Karl G. Jöreskog

1971 "Estimating the parameters of path models involving unmeasured variables," in Hubert M. Blalock (ed.) Causal Models in the Social Sciences. Chicago: Aldine.

Wiley, David E. and James A. Wiley

1970 "The estimation of measurement error in panel data," American Sociological Review 35(February):112-116.

Wright, Sewell

1960 "Path coefficients and path regressions: alternative or complementary concepts," Biometrics 16 ( ):189-202.