Division of Research
Graduate School of Business Administration
The University of Michigan

May 1981

STRUCTURAL EQUATION MODELS WITH UNOBSERVABLE VARIABLES AND MEASUREMENT ERROR: ALGEBRA AND STATISTICS

Working Paper No. 266

Claes Fornell
The University of Michigan

David F. Larcker
Northwestern University

STRUCTURAL EQUATION MODELS WITH
UNOBSERVABLE VARIABLES AND MEASUREMENT ERROR:
ALGEBRA AND STATISTICS

ABSTRACT

Several issues relating to goodness of fit in structural equations are examined. It is shown that the convergence and differentiation criteria, as applied by Bagozzi, do not stand up under mathematical or statistical analysis.

It is argued that the choice of interpretative statistic must flow from the research objective. When this is done, it is demonstrated that the Fornell-Larcker Testing System is internally consistent and that it conforms to the rules of correspondence for relating data to abstract variables.

STRUCTURAL EQUATION MODELS WITH UNOBSERVABLE VARIABLES AND MEASUREMENT ERROR: ALGEBRA AND STATISTICS

Introduction

Although structural equation models with unobservables represent a considerable step forward for our ability to study marketing phenomena, there are several unresolved problems and many areas of confusion related to the application of these models. In his comment, Professor Bagozzi (1981) makes an important contribution to a much needed debate of some of the issues involved. He questions our results because parts of our numerical analysis do not conform with certain "logical" criteria.[1] Further, Bagozzi contends that our proposed testing system suffers from inconsistencies and violations of model specification.

This reply will show that this criticism is unfounded. Algebraic analysis will demonstrate that the criteria proposed by Bagozzi are not relevant. Statistical theory indicates that they are unreliable. It will also be shown that there are no inconsistencies or violations in our testing system.

The Algebra of Factor Analytic Structural Modeling

Our original article demonstrated the irrelevance of the overall likelihood ratio chi-square test for the evaluation of relationships betwen abstract variables (latent variables, theoretical constructs, unobservables). In view of the common practice of proclaiming support for one's theory on the basis of a good chi-square fit (sometimes without even considering that most of the degrees of freedom may be associated with the measurement model), it seemed to us that there was a need to explain the nature of the test.

Although it is perhaps "well-known" that the chi-square statistic evaluates the structure of underlying relationships and not their strengths, there is, nevertheless, a relationship between structure and strength of variable relationships. As shown in our original article, it is what we termed "structural consistency" that determines the degree of fit. It was also shown that structural consistency is easier to achieve when the observed correlations are small. This means that, given an imposed theoretical structure, high correlations will result in a higher chi-square (i.e., a worse fit) than low correlations. In other words, structural consistency can decline as both measurement and theory improve (as measured by the size of correlation coefficients) and vice versa.

According to Bagozzi, many of the "anomalous" findings of our simulations can be explained in terms of three characteristics of the correlation matrix. For example, the good fit (p=.589) of the correlation matrix in the lower right hand cell of Table 1 in Fornell and Larcker (1981) is presumably due to the fact that "each of the four cross-construct correlations is (1) in the same direction, (2) approaches statistical significance, and (3) is not grossly different in magnitude from its sisters" (Bagozzi, 1981, p. 7). We will now show that these three characteristics of the correlation matrix do not explain the goodness of fit.

First, the statistical significance of correlations is irrelevant to the chi-square statistic.[2] Second, the direction of sign is neither a necessary nor a sufficient criterion for goodness of fit. Third, uniformity among cross-correlations is sufficient, but not necessary. Instead, it is structural consistency that determines the degree of fit. From our numerical analysis it was shown that the degree of fit was accounted for by the divergence of observed correlation coefficients in the data matrix. However, the manner in which the matrices were generated (to assure matrices that were positive-definite, internally consistent, and with elements without very low values) did not distinguish between necessary and sufficient criteria. Let us, therefore, derive the definition of structural consistency algebraically.

The objective of factor analytic structural equations is to find (determine, test) an underlying model that accounts for the observed correlations. Following the assumptions given by Jöreskog (1979) and assuming standardization for all variables, the expected variance-covariance matrix, $\underset{\sim}{\Sigma}$, (i.e., the correlation matrix) for the two-indicator/two-construct model with uncorrelated measurement errors can be written:[3]

$$\underset{\sim}{\Sigma} = \begin{bmatrix} 1.0 & & & \\ \lambda_{y_2}\lambda_{y_1} & 1.0 & & \\ \lambda_{x_1}\gamma\lambda_{y_1} & \lambda_{x_1}\gamma\lambda_{y_2} & 1.0 & \\ \lambda_{x_2}\gamma\lambda_{y_1} & \lambda_{x_2}\gamma\lambda_{y_2} & \lambda_{x_2}\lambda_{x_1} & 1.0 \end{bmatrix} \qquad (1)$$

$\lambda_{y's}$ = criterion loadings
$\lambda_{x's}$ = predictor loadings
γ = structural parameter

In order for $\underset{\sim}{\Sigma}$ to equal the observed correlation matrix $\underset{\sim}{S}$, certain requirements are evident from equation (1). By multiplying corresponding cross-construct elements, we obtain the following equality:

$$(\lambda_{x_1}\gamma\lambda_{y_1}) \ (\lambda_{x_2}\gamma\lambda_{y_2}) = (\lambda_{x_2}\gamma\lambda_{y_1}) \ (\lambda_{x_1}\gamma\lambda_{y_2}) \qquad (2)$$

$$\lambda_{x_1}\lambda_{x_2}\gamma^2\lambda_{y_1}\lambda_{y_2} = \lambda_{x_1}\lambda_{x_2}\gamma^2\lambda_{y_1}\lambda_{y_2} \qquad (3)$$

For the model to be true (i.e., $\underset{\sim}{\Sigma} = \underset{\sim}{S}$) it follows that:

$$(r_{x_1y_1})\ (r_{x_2y_2}) = (r_{x_1y_2})\ (r_{x_2y_1}) \qquad (4)$$

which is a vanishing tetrad whose test was originally discussed by Spearman and Holzinger (1924) and what we refer to as structural consistency.

As will be discussed shortly, there are two estimates for every parameter in this model. If equation (4) is satisfied, the two estimates are identical. Any departure from structural consistency, and thus from the equality imposed by equation (4), implies that the pairs of correlation coefficients are in disagreement. Stated differently, the model cannot account for the observed correlations. As a result, goodness of fit will suffer. Accordingly, goodness of fit has little to do with the statistical significance of the raw correlation coefficients (as long as they are nonzero) or with the sign of these coefficients. Further, uniformity of cross-correlations is not necessary. It is the product of the pairs of cross-correlations that must be uniform.

If the model is true, the equalities of equation (4) must hold in the population. In a sample, we are likely to find divergence and two different estimates are obtained for each parameter. If these estimates are not significantly different, it is concluded that structural consistency is met. The two estimates are combined in the maximum likelihood estimation to a single estimate such that the observed correlations are reproduced as closely as possible.

To summarize, it follows from equation (4) that even if the pattern of correlations violate all three criteria suggested by Bagozzi, a perfect fit can be obtained. This is illustrated in Table 1 where cross-correlations show different signs, the correlation coefficients are low, and the magnitudes of the correlations vary. Similarly, if a correlation matrix satisfies all of Bagozzi's criteria, a poor fit can result.[4] This is illustrated in Table 2.

Convergence and Differentiation

Bagozzi suggests that there are two requirements that must be met "as a matter of logical necessity" before one can proceed with structural equation modeling. Following this reasoning, he argues that our first simulation is a poor example for demonstrating the properties of the chi-square test because several of the cells in Table 1 (Fornell and Larcker, 1981, p. 42) do not meet the requirements of convergence in measurement and differentiation in constructs. Convergence implies that all within-construct correlations are (1) high and (2) of approximately the same magnitude. Differentiation is satisfied if the cross-correlations are (1) high, (2) uniform, and (3) lower than the within-construct correlations. Should these criteria not be met, Bagozzi argues, logical grounds exist for rejecting the model.

We do not find this argument compelling, for three reasons. First, as we already have shown (equations 1-4), convergence and differentiation are not relevant to the properties of the chi-square. Thus, the criticism that parts of our empirical analysis of the chi-square do not satisfy convergence and differentiation criteria is irrelevant. This was also illustrated by our second simulation for which results equivalent to the first simulation were obtained after correlation matrices that failed to meet discriminant and convergent validity were deleted. Second, it seems "counterproductive" to evaluate structural equation models with unobservables by inspecting the raw correlation coefficients. Third, to the extent that measurement error can be removed from the structural parameters, convergence and differentiation at the observed level have little meaning. Let us now develop the rationale behind these contentions.

Comparing the magnitude of correlations between observed variables for the assessment of convergence and discrimination is subject to several limitations. There is no need to revert back to the ambiguous rules of thumb originally proposed by Campbell and Fiske when more objective and powerful methods are available. As also pointed out by Bagozzi, the raw correlation coefficients are affected by both random and systematic (e.g., measurement) error. Any attempt to assess convergence and differentiation by inspecting the relative size of these coefficients does not take this into account and the risk of drawing false inferences would be substantial. More powerful validity assessments can be accomplished within the context of the structural equations model, where both types of errors can explicitly be taken into account.

Let us first illustrate how the structural equation model handles systematic measurement error. The algebra for the two-indicator/two-construct model can be used to explain this theory. From equation (1) we know that:

$$\lambda_{x_1} \lambda_{x_2} = r_{x_1 x_2} \tag{5}$$

$$\lambda_{y_1} \lambda_{y_2} = r_{y_1 y_2} \tag{6}$$

Substituting (5) and (6) into equation (3) and equation (4) gives:

$$r_{x_1 x_2} \gamma^2 r_{y_1 y_2} = (r_{x_1 y_1})(r_{x_2 y_2}) \tag{7}$$

and

$$r_{x_1 x_2} \gamma^2 r_{y_1 y_2} = (r_{x_1 y_2})(r_{x_2 y_1}) \tag{8}$$

solving for γ gives

$$\gamma = \pm \left[\frac{(r_{x_1 y_1})\ (r_{x_2 y_2})}{(r_{x_1 x_2})\ (r_{y_1 y_2})} \right]^{\frac{1}{2}} = \pm \left[\frac{(r_{x_1 y_2})\ (r_{x_2 y_1})}{(r_{x_1 x_2})\ (r_{y_1 y_2})} \right]^{\frac{1}{2}}$$

If equation (9) holds, the indicators provide consistent information, systematic error is not present, and measurement error does not affect the estimate of the abstract correction, γ. As in equation (4), perfect consistency cannot be expected for the sample because of random error. If the two estimates from equation (9) are not significantly different, convergence and differentiation are not meaningful criteria for validity assessment. This is easy to understand when one realizes that the purpose of validity testing in this context is to determine the extent of measurement error. If the parameter of interest, γ, is free from measurement error, the application of convergence and differentiation as validity criteria serves no purpose.

The ability to remove measurement error from theory testing procedures is perhaps the most important methodological contribution of structural equation models. If one chooses to ignore this advantage and apply validity considerations to the raw correlations between observed variables, there is little reason to use structural equation modeling with unobservables. It is reversed logic to use the observed correlations as evidence for rejecting an untested underlying model, for it is the purpose of the model to explain the correlations; fallible as they may be. If the model fails to do this, it is rejected.

Not only can fallible observations be handled, but also observations of different quality. The requirement of uniformity for correlation coefficients is not necessary, for it would imply that all indicators reflect the unobservables equally well. Thus contrary to Bagozzi's conclusion, a valid structural equations model can be obtained even though two (observed) measures of the same thing correlate at a lower level between themselves than with some other measure.

Several things should be noted at this point. While the handling of systematic measurement error in (factor analytic) structural equation models makes the convergence and discrimination criteria at the level of observed correlations unnecessary, there is a "price" one has to pay for this advantage. This is what we attempted to show in our original article. Note from equation (9) that high and uniform measurements (in terms of high and uniform within-construct correlations) operate to decrease the estimated correlation (γ) between unobservables. Conversely, the lower the quality of observations, the higher the estimated γ. This leads to the intuitive unappealing result that measurements of high and even quality "deflate" the correlation between abstract variables, while measures of poor and uneven quality "inflate" this parameter.

To guard against exaggerated interpretations, we developed a system of tests that includes analysis of average variance extracted by unobservables, the reliability of individual observations and composite unobservables, as well as examinations of the empirical associations between observed variables within the context of the model. In cases where one is not certain about the extent of measurement error removal, as when the χ^2 is large relative to the degrees of freedom in the measurement model or when relevant "third" variables are excluded, it may be useful to examine the estimated <u>model parameters</u> in terms of convergent and discriminant validity. Given equation (9), the risk of obtaining an inflated γ as a result of poor measures is offset by the fact that it becomes more difficult to satisfy validity criteria from model parameters (compared to the raw correlation coefficients). This can be shown by deriving the estimation for the loadings.

From equations (1), (9), (5), and (6) we have:

$$\lambda_{y_1} \left[\frac{(r_{x_1y_1})\ (r_{x_2y_2})}{(r_{x_1x_2})\ (r_{y_1y_2})} \right]^{\frac{1}{2}} \lambda_{x_1} = r_{y_1x_1} \tag{10}$$

$$\lambda_{y_2} \left[\frac{(r_{x_1y_1})\ (r_{x_2y_2})}{(r_{x_1x_2})\ (r_{y_1y_2})} \right]^{\frac{1}{2}} \lambda_{x_2} = r_{y_2x_2} \tag{11}$$

$$\lambda_{x_1}\lambda_{x_2} = r_{x_1x_2} \tag{12}$$

$$\lambda_{y_1}\lambda_{y_2} = r_{y_1y_2} \tag{13}$$

$$\lambda_{y_1} \left[\frac{(r_{x_1y_2})\ (r_{x_2y_1})}{(r_{x_1x_2})\ (r_{y_1y_2})} \right]^{\frac{1}{2}} \lambda_{x_2} = r_{y_1x_2} \tag{14}$$

$$\lambda_{y_2} \left[\frac{(r_{x_1y_2})\ (r_{x_2y_1})}{(r_{x_1x_2})\ (r_{y_1y_2})} \right]^{\frac{1}{2}} \lambda_{x_1} = r_{y_2x_1} \tag{15}$$

Solving for the loadings (λ_y's, λ_x's) gives:

$$\lambda_{y_1} = \pm \left[\frac{(r_{x_2y_1})}{(r_{x_2y_2})} \quad (r_{y_1y_2}) \right]^{\frac{1}{2}} = \pm \left[\frac{(r_{x_1y_1})}{(r_{x_1y_2})} \quad (r_{y_1y_2}) \right]^{\frac{1}{2}} \tag{16}$$

$$\lambda_{y_2} = \pm \left[\frac{(r_{x_2y_2})}{(r_{x_2y_1})} \quad (r_{y_1y_2}) \right]^{\frac{1}{2}} = \pm \left[\frac{(r_{x_1y_2})}{(r_{x_1y_1})} \quad (r_{y_1y_2}) \right]^{\frac{1}{2}} \tag{17}$$

$$\lambda_{x_1} = \pm \left[\frac{(r_{x_1y_2})}{(r_{x_2y_2})} \quad (r_{x_1x_2}) \right]^{\frac{1}{2}} = \pm \left[\frac{(r_{x_1y_1})}{(r_{x_2y_1})} \quad (r_{x_1x_2}) \right]^{\frac{1}{2}} \tag{18}$$

$$\lambda_{x_2} = \pm \left[\frac{(r_{x_2y_2})}{(r_{x_1y_2})} \quad (r_{x_1x_2}) \right]^{\frac{1}{2}} = \pm \left[\frac{(r_{x_2y_1})}{(r_{x_1y_1})} \quad (r_{x_1x_2}) \right]^{\frac{1}{2}} \tag{19}$$

The solutions for all parameters can be verified by applying equations (9) and (16)-(19) to the correlation matrix in Table 1. As can be seen from these equations,[5] the loadings become small and γ increases as within-construct correlations are reduced. As loadings decrease, their standard errors increase. Thus, our proposed tests for convergence and discrimination become more stringent.[6]

Research Objective and Choice of Interpretative Statistics

Another issue discussed by Bagozzi concerns the scope of analysis. Specifically, he suggests that the assessment of structural equation models should involve theoretical, methodological and statistical considerations. A more limited focus, Bagozzi warns, might lead to false and misleading conclusions. We support this argument, but it is also essential to understand the critical role of research objective in determining the nature and scope of analysis. The specification of theory and method, as well as the choice of interpretative statistics must flow from a research purpose or objective (e.g., description, prediction, explanation). Different objectives require different approaches. Our article focused on different objectives as they relate to different inferential and descriptive statistics.

We proposed three different tests associated with different research objectives. If the objective is theory testing at the abstract level, an F-test for the γ estimate was suggested. If the objective is of a more pragmatic nature, such as predicting (via the model) the variance of the observed criterion variables from the unobserved explanatory constructs, focus should center on Redundancy. Finally, if one is interested in accounting for the variance of the y-variables from the x-variables, the Redundancy measure was adjusted by removing factor indeterminacy. We termed the resulting statistic "Operational Variance."

Bagozzi suggests that there is a contradiction involved in these three tests because the F-test can indicate a significant structural model parameter while the tests for Redundancy and Operational Variance may imply that the overall relationships are insignificant. This reflects a misunderstanding of the relationship between the tests. No contradiction exists because we are dealing with progressively stronger standards. It is likely that one may find a significant relationship between unobservables even though the shared variance with their respective indicators is small (see equations 9-19). The Redundancy measure takes the lack of shared variance between the construct and its indicators into account. This is analogous to the situation in canonical analysis where a high canonical correlation coefficient can be obtained even though redundancy is low (Fornell, 1979). Operational Variance adjusts for both lack of shared variance between constructs and indicators and for factor indeterminacy. Consequently, we do not find it disconcerting that the test of C_{14} (100% measurement, 15% theory) in Table 9 (Fornell and Larcker, 1981, p. 48) rejects the model based on Operational Variance. The reason is that the shared variance between constructs (γ^2) is only about 7%. After adjusting for factor indeterminacy and for loadings that are less than unity (note that 100% measurement refers to the starting point of a reduction in measurement; it does not imply perfect measurements), the resulting shared variance (i.e., Operational Variance) is too low to be significant. We see no reason for suggesting that the chi-square statistic provides a "more reasonable" interpretation here.

Is "Causal" Analysis Necessary?

In addition to a comprehensive evaluation of structural equation results, our proposed testing system also offers information about the extent to which the abstract analysis is removed from the level of observations. Bagozzi criticizes this aspect of our system because it "shifts the conceptual focus of research away from causal hypotheses and toward empirical associations". However, there are no strong arguments for limiting structural equations to causal modeling. There are many situations in which it is useful to examine relationships at the empirical level relative to those at the corresponding abstract level. As we have shown, there can be a substantial difference between the two. While the marketing theorist is interested in the testing of hypotheses which relate abstract variables, the applied marketer is perhaps more confined to that which can be observed.

Moreover, it is not possible to infer causal ordering from the estimated parameter values per se. Causal models and structural equations are sometimes used synonymously. We have deliberately avoided the former label in our discussion. The model used in our example (two-indicator/two-construct) can be shown to be formally identical to a factor analysis with oblique rotation and has little to do with causality between constructs. In order to infer causality in the structural equation, one typically needs not only larger models (with more restrictions) but also <u>a priori</u> knowledge about the values of structural parameters. In addition, several external criteria must be met (see for example, Duncan, 1975; Blalock, 1961; Bagozzi, 1980). Forcing the researcher to make causal inferences when the justification for causal order is questionable would only serve to artificially limit the usefulness and applicability of structural equation models.

Rules of Correspondence

A related issue concerns the rules for relating observed data to abstract variables. The treatment of an unobservable as a function of its indicators violates the theoretical model we discuss. As pointed out by Bagozzi, the model under consideration here is consistent with classical test theory in the sense that the indicators are assumed to be "reflective" (effects) of their respective unobservables. We have no quarrel here. However, we take issue with the idea that (1) "formative" indicators (i.e., producers of or contributors to abstract variables) are contradictory to the "rules of correspondence as formulated by philosophers of science" (Bagozzi, 1981, p.) and (2) that it is necessary to infer unidirectionality in the analysis of shared variance. There are many examples in the literature where indicators are formative (see, Blalock, 1969; Heise, 1972; Land, 1970; Schönemann and Steigler, 1976) and where both reflective and formative indicators have been used within the same model (Wold, 1980; Hauser and Goldberger, 1971; Fornell, 1979; Fornell and Bookstein, 1981; Jöreskog and Goldberger, 1975). One can also find support for the notion of formative indicators from general systems theory (Bertalanffy, 1968), as well as from philosophy of science (Kaplan, 1964)[7].

Regardless of whether the model is estimated with reflective or formative indicators, we do not find that there is a conceptual problem in determining the amount of shared variance between the abstract variables and their indicators. Nor is there a mathematical problem. Formally, the shared variance between standardized indicators and unobservables is equal to one minus measurement error variance (Jöreskog, 1979; Wold, 1980). In essence, the measure of shared variance is the squared correlation coefficient; it presumes neither unidirectionality nor causality between the variables involved.

Generalizations

Bagozzi argues that our proposed system is not generalizable to larger and more complex models. Although there are no fundamental problems involved in generalizing our methods for measurement evaluation, there are several difficulties associated with providing a single measure of overall explanatory power in complex multivariate relationships. If the inferential properties of Redundancy and Operational Variance are to remain intact through generalizations, several orthogonality restrictions would have to be imposed. We agree with Bagozzi that these restrictions severely limit the possibility of statistical inference in larger models. In fact, we questioned the feasibility of developing a single measure that is useful in representing all the relationships in a more complicated model. Nevertheless, separate assessments of the measurement model and the structural model are still possible. For example, the redundancy between constructs[8] (Fornell and Larcker, 1981, p. 20) provides the average variance in $\underset{\sim}{\eta}$ that is explained by $\underset{\sim}{\xi}$. As such, it has a straightforward interpretation. It is also similar to what was recommended by Cramer and Nicewander (1979) in their analysis of six different measures of multivariate association. This statistic does not require the structural equation model to be specified as a canonical model.

Measures of explanatory power are of vital importance in full information maximum likelihood estimation of structural equations. For example, if redundancy between constructs is low, the endogenous unobservables are poorly accounted for, which suggests the possibility of misspecification as a result of omitted variables. In full information estimation, all parameters of the model are affected by omitted variables. Regardless of structural consistency, the parameter estimates are likely to deviate substantially from their true value if relevant exogenous variables that are correlated with included variables are omitted. Under these circumstances, substantive inference, let alone causal interpretation, becomes little more than an exercise in speculation.

Another useful measure which presents no problems in terms of generalization is the Upper Bound of explanatory power (Fornell and Larcker, 1981, p. 48). Both Redundancy between constructs and the Upper Bound are inferential beyond simple models and can be tested by the method provided by Miller (1975). As for Redundancy (for observed variables) and Operational Variance, statistical inference is not possible unless the assumptions mentioned previously are made. At best, these measures should be regarded as descriptive indices of explanatory power that can be averaged over two-construct systems within larger models.

Summary

Several aspects of the goodness of fit test in structural equation models were examined. Algebraic analysis demonstrated that the criteria presented by Bagozzi are not related to goodness of fit. It was also shown that convergence and differentiation at the level of observed correlations suffer from several limitations and that these criteria are not applicable to structural equation modeling. Further, the purported violations and logical inconsistencies of our proposed testing system were shown to be unfounded.

Given the many advantages of structural equation models for theory testing and theory building, it is possible that they may fundamentally alter the way in which empirical research in marketing will be conducted. Professor Bagozzi has illustrated this in many pioneering applications. To be sure, there are several statistical, philosophical, and methodological issues that are unresolved and in need of further analysis and discussion. The difficulties posed by these issues should not be overlooked. However, recent research has been directed to several of these problems. Regarding the problem of sample size as pointed out by Bagozzi, important insights have been gained into the sensitivity of the chi-square statistic (Geweke and Singleton, 1980) and maximum likelihood estimation (Boomsa, 1981; Sprott, 1980). These problems, as well as several others, (e.g., indeterminacy of factors, improper solutions, distributional requirements, variable scale requirements, R^2 analogue) are more or less specific to factor analytic structural modeling (e.g., LISREL). It has been shown (Fornell and Bookstein, 1981; Fornell and Denison, 1981) that many of them can be overcome by other models of structural equations.

Table 1

A MODEL BASED ON A CORRELATION MATRIX
WHERE CORRELATION COEFFICIENTS ARE INSIGNIFICANT
AND OF VARYING SIGN AND MAGNITUDE

	Y_1	Y_2	X_1	X_2
Y_1	1			
Y_2	-.260	1		
X_1	-.126	.190	1	
X_2	.114	-.171	-.109	1

	Standardized Maximum Likelihood Estimates
λ_{y_1}	.416
λ_{y_2}	-.625
λ_{x_1}	.348
λ_{x_2}	-.313
γ	-.873
Ψ	.238
$\underset{\sim}{\Theta}_{\varepsilon}$	$\begin{bmatrix} .827 & \\ & .609 \end{bmatrix}$
$\underset{\sim}{\theta}_{\delta}$	$\begin{bmatrix} .879 & \\ & .902 \end{bmatrix}$

Chi-Square Statistic = 0
d.f. = 1
P ≃ 1 (n = 100)

Table 2

A MODEL BASED ON A CORRELATION MATRIX WHERE CORRELATION COEFFICIENTS ARE SIGNIFICANT, OF THE SAME SIGN, AND FAIRLY UNIFORM IN MAGNITUDE

	Y_1	Y_2	X_1	X_2
Y_1	1			
Y_2	.625	1		
X_1	.327	.367	1	
X_2	.422	.327	.640	1

	Standardized Maximum Likelihood Estimates
λ_{y_1}	.843
λ_{y_2}	.741
λ_{x_1}	.751
λ_{x_2}	.853
γ	.564
Ψ	.682
$\underset{\sim}{\Theta}_{\varepsilon}$	$\begin{bmatrix} .289 & \\ & .451 \end{bmatrix}$
$\underset{\sim}{\theta}_{\delta}$	$\begin{bmatrix} .437 & \\ & .273 \end{bmatrix}$

Chi-Square Statistic = 6.4745
d.f. = 1
$P \simeq .01$ ($n = 100$)

FOOTNOTES

1 According to Bagozzi, our analyses are accurate in a purely mathematical sense but violate certain logical desiderata (P.1). To us, it is a logical impossibility for mathematics to violate logic. It may be that we are talking about different modes of logic (Bagozzi, 1981, footnote 2). Nevertheless, since Bagozzi draws upon the logic of philosophy of science in support of his arguments and because we will use algebraic analysis to repudiate them, it is essential that the meaning of logic be in agreement with mathematics.

2 Except that when correlations are low and insignificant, it is easier to satisfy structural consistency. This was shown in our numerical analysis and will be evident from the subsequent algebraic analysis.

3 For the reader familiar with path analysis, the expected correlation matrix Σ can also be derived from the tracing rule (Jacobson and Lalu, 1974; Blalock, 1964; Werts, Linn and Jöreskog, 1974; Costner, 1969; Duncan, 1975, Jöreskog, 1970; Goldberger, 1971).

4 While the cross-correlations are not identical, they are not grossly different from each other.

5 It is also evident from equations (9) and (16) - (19) that one can obtain λ's > 1 and $\gamma > 1$ which would result in negative measurement error and negative structural equation error, respectively. Such results are termed "improper solutions" and constitute a common problem in factor analytic maximum likelihood structural equations (Fornell and Bookstein, 1981).

6 In the testing of these two criteria, sampling error can then be taken into account by using Fisher's z transformation of r, which evaluates the significance of the difference in the estimated parameter values.

7 See Fornell and Bookstein (1981) for a discussion of how to choose between reflective and formative indicator modes.

8 Note that redundancy between constructs is different from the overall redundancy index. Redundancy between constructs is applicable to larger models with multiple η's and ξ's and, in contrast to overall redundancy, does not account for construct-indicator relationships.

REFERENCES

Bagozzi, Richard P. (1980). Causal Models in Marketing. New York: Wiley and Sons.

______________ (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error: A Comment." Journal of Marketing Research 18 (August).

Bertalanffy, L. Von (1968). General System Theory. New York: George Braziller.

Blalock, Hubert M., Jr. (1961). Causal Inferences in Nonexperimental Research. Chapel Hill, North Carolina: The University of North Carolina Press.

______________ (1969). "Multiple Indicators and the Causal Approach to Measurement Error." American Journal of Sociology 75 (September): 264-272.

Boomsa, A. (1981). "The Robustness of LISREL Against Small Sample Sizes in Factor Analysis Models." in K. G. Jöreskog and H. Wold (eds.), Systems Under Indirect Observation: Causality, Structure, Prediction. Amsterdam: North Holland, in press.

Costner, Herbert L. (1969). "Theory, Deduction, and Rules of Correspondence." American Journal of Sociology 75 (September): 254-263.

Cramer, Elliot M., and W. Alan Nicewander (1979). "Some Symmetric, Invariant Measures of Multivariate Association." Psychometrika 44, No. 1 (March): 43-54.

Duncan, Otis Dudley (1975). Introduction to Structural Equation Models. New York: Academic Press.

Fornell, Claes (1979). "External Single-Set Components Analysis of Multiple Criterion/Multiple Predictor Variables." Multivariate Behavioral Research 14: 323-338.

Fornell, Claes, and Fred L. Bookstein (1981). "A Comparative Analysis of Two Structural Equation Models: LISREL and PLS Applied to Market Data." Working Paper, The Graduate School of Business Administration, The University of Michigan.

Fornell, Claes, and Daniel R. Denison (1981). "Evaluating Convergent, Discriminant, and Nomological Validity Via Confirmatory Multidimensional Scaling." Working Paper, The Graduate School of Business Administration, The University of Michigan.

Fornell, Claes, and David F. Larcker (1981). "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." Journal of Marketing Research 18 (February): 39-50.

Geweke, John F., and Kenneth J. Singleton (1980). "Interpreting the Likelihood Ratio Statistic in Factor Models When Sample Size is Small." Journal of the American Statistical Association 75 (March): 133-137.

Goldberger, Arthur S. (1971). "Econometrics and Psychometrics: A Survey of Communalities." Psychometrika 36, No. 2 (June): 83-107.

Hauser, R. M., and A. S. Goldberger (1971). "The Treatment of Unobservable Variables in Path Analysis." In Sociological Methodology, edited by H. L. Costner, pp. 81-177. San Francisco: Jossey-Bass.

Heise, David R. (1972). "Employing Nominal Variables, Induced Variables, and Block Variables in Path Analyses." Sociological Methods and Research 1 (November): 147-174.

Jacobson, Alvin L., and N. M. Lalu (1974). "An Empirical and Algebraic Analysis of Alternative Techniques for Measuring Unobserved Variables." In Measurement in the Social Sciences: Theories and Strategies, edited by Hubert M. Blalock, Jr., pp. 215-242. Chicago: Aldine Publishing Company.

Jöreskog, Karl G. (1970). "A General Method for Analysis of Covariance Structures." Biometrika 57: 234-251.

_______________ (1979). "Structural Equation Models in the Social Sciences: Specification, Estimation and Testing." In Advances in Factor Analysis and Structural Equation Models, edited by Karl G. Jöreskog and Dag Sörbom, pp. 105-127. Cambridge, Massachusetts: ABT Books.

Jöreskog, Karl G., and A. S. Goldberger (1975). "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable." Journal of the American Statistical Association 10: 631-639.

Kaplan, Abraham (1964). The Conduct of Inquiry New York: Chandler Publishing Company.

Land, Kenneth C. (1970). "Path Coefficients for Unmeasured Variables." Social Forces 48 (June): 506-511.

Miller, John K. (1975). "The Sampling Distribution and a Test for the Significance of the Bimultivariable Redundancy Statistic: A Monte Carlo Study." Multivariate Behavioral Research 6 (April): 233-244.

Schönemann, P. H., and J. H. Steiger (1976). "Regression Component Analysis." British Journal of Mathematical Statistical Psychology 29: 175-189.

Spearman, C., and K. Holzinger (1924). "The Sampling Error in the Theory of Two Factors." British Journal of Psychology, 15, 17-19.

Sprott, D. A. (1980). "Maximum Likelihood in Small Samples: Estimation in the Presence of Nuisance Parameters." Biometrika 67, 515-523.

Werts, C. W.; R. L. Linn; and K. G. Jöreskog (1974). "Quantifying Unmeasured Variables." In Measurement in the Social Sciences: Theories and Strategies, edited by Hubert Blalock, Jr., pp. 270-292. Chicago: Aldine Publishing Company.

Wold, Herman (1980). "Model Construction and Evaluation When Theoretical Knowledge is Scarce." In Evaluation of Econometric Models, edited by Jan Kmenta and James B. Ramsey, pp. 47-74. New York: Academic Press.