A COMPARATIVE ANALYSIS OF TWO STRUCTURAL

EQUATION MODELS: LISREL AND PLS

APPLIED TO MARKET DATA

Working Paper No. 276

Claes Fornell

Fred L. Bookstein

The University of Michigan

FOR DISCUSSION PURPOSES ONLY

Abstract

Marketing applications of structural equation models with unobservables have relied almost exclusively on LISREL for parameter estimation. There has apparently been little concern about the frequent inability of marketing data to meet the requirements for maximum-likelihood estimation or the common occurrence of improper solutions in LISREL modelling. In this paper we demonstrate that Partial Least Squares (PLS) can be used to overcome these two problems.

PLS is somewhat less well-grounded than LISREL in statistical theory. We show, however, that under certain model specifications the two methods produce the same results. In more general cases, they provide results which diverge in certain systematic ways. We analyze and explain these differences in terms of the underlying objectives of each method.

Introduction

Though they were introduced to marketing only recently, structural equation models with unobservables are beginning to change the conventions of marketing research methodology (Bagozzi 1980; Fornell 1982). For social science in general, the new structural equations approach is strongly identified with maximum likelihood factor analysis procedures generalized by Karl Jöreskog (1970, 1973, 1979) and the associated computer program LISREL (Jöreskog and Sörbom 1978). For marketing, in particular, nearly every application of structural modelling has used LISREL for parameter estimation. But it is not realistic to assume that all problems amenable to use of structural equation models are also suited to LISREL. There are other protocols of structural estimation which impose different assumptions about data, theory, and the ties betwen unobservables and indicators. Marketing data do not often satisfy the requirements of multinormality and interval scaling, or attain the sample size required by maximum-likelihood estimation. More fundamentally, two serious problems often interfere with meaningful modelling: improper solutions and factor indeterminacy.

Herman Wold's method of Partial Least Squares (PLS)[1] avoids many of the restrictive assumptions underlying maximum-likelihood (ML) techniques and ensures against improper solutions and factor indeterminacy. Toward a comparison of PLS with LISREL, we estimate three models for a single data set. The first model compares traditional common-factor ML estimates with PLS estimates for a case in which LISREL produces improper solutions. We show that the improper estimates do not stem from sample variance or from lack of fit but can be traced, instead, to the path-analytic fitting objective behind LISREL. It is then shown that the removal of factor

indeterminacy via PLS provides an effective cure. In a second model, wherein all factors are explicitly defined, LISREL avoids its improper solutions by giving estimates identical to those of PLS. The third model extends the second in a direction consistent with the consumer behavior theory it embodies. It too presents no improper solutions, but illustrates a systematic difference between LISREL and PLS results.

Partial Least Squares in Structural Modelling

In Partial Least Squares, the set of model parameters is divided into subsets estimated by use of ordinary multiple·regressions that involve the values of parameters in other subsets. An iterative method provides succes-sive approximations for the estimates, subset by subset, of loadings and structural parameters. Extending his theory of fixed-point estimation (Wold 1965), Herman Wold developed this method for structural models with unobservables (1974, 1975, 1980a, 1980b). As is the case with LISREL, many of the early elaborations and applications originate from Sweden (Ågren 1972; Bergström 1972; Bodin 1974; Lyttkens 1966, 1973; Noonan and Wold 1977; Areskoug et al. 1975). In the United States, Hui (1978) has extended the model to nonrecursive systems, and Bookstein (1980, 1981) has provided a geometrical restatement of its protocols.

Recent discussions of both PLS and LISREL are available in Jöreskog and Wold (1981). Applications of PLS appear in a variety of disciplines, including economics (Apel 1977), political science (Meissner and Uhle-Fassing 1981), psychology of education (Noonan 1980; Noonan and Wold 1980), chemistry (Kowalski et al. 1981), and marketing (Jagpal 1981).

Model Structure

To facilitate the comparison with LISREL, we will use Jöreskog's notation. Like LISREL, PLS puts forward two sets of equations: the structural equations ("inner relations") and the measurement equations ("outer relations"). The structural equations can be written:

$$\beta \, \eta = \Gamma \, \xi + \zeta \tag{1}$$

Here $\eta' = (\eta_1, \eta_2, \ldots, \eta_m)$ and $\xi' = (\xi_1, \xi_2, \ldots, \xi_n)$ are random vectors of unobserved criterion and explanatory variables; β (m x m) is a matrix of coefficient parameters for η, and Γ (m x n) is a matrix of coefficient parameters for ξ; and $\zeta = (\zeta_1, \zeta_2, \ldots, \zeta_m)$ is a random vector of residuals.

The measurement equations are:

$$y = \Lambda_y \eta + \varepsilon \tag{2}$$

$$x = \Lambda_x \xi + \delta \tag{3}$$

in which $y' = (y_1, y_2, \ldots, y_p)$ and $x' = (x_1, x_2, \ldots, x_q)$ are the observed criterion and explanatory variables, respectively; Λ_y (p x m) and Λ_x (q x n) are the corresponding regression matrices; and ε and δ are residual vectors.

It is assumed that $E(x) = E(y) = E(\eta) = E(\xi) = 0$; $E(\eta\zeta') = E(\zeta) = 0$; $E(\zeta\zeta') = E(\zeta\delta') = 0$; and that $E(\varepsilon\varepsilon') = \Theta_\varepsilon$, $E(\delta\delta') = \Theta_\delta$, $E(\xi\xi') = \Phi$ and $E(\zeta\zeta') = \Psi$. It is also assumed that $Var\,(\eta_i) = Var\,(\xi_j) = Var\,(x_k) = Var\,(y_r) = 1$, all i, j, k, r. Further, we compute the unobservables as exact linear combinations of their empirical indicators:

$$\eta \equiv \pi_\eta \, y \tag{4}$$

$$\xi \equiv \pi_\xi \, x \tag{5}$$

where π_η (p x m) and π_ξ (p x m) are regression matrices.

Mode Selection: The Relationship between Unobserved and Measured Variables

In estimation, the unobserved constructs (η, ξ) can be viewed as

underlying factors or as indices produced by the observables. That is, the observed indicators can be treated as reflective or formative. Reflective indicators are typical of classical test theory and factor analysis models; they are invoked in an attempt to account for observed variances or covariances. Formative indicators are more directly concerned with the delineation of abstract relationships.

The choice of indicator mode, which substantially affects estimation procedures, has hitherto received only sparse attention in the literature. Figure 1 exemplifies the choices to be made. In deciding how unobservables and data should be related, there are three major considerations: study objective, theory, and empirical contingencies.

Should the study intend to account for observed variances, reflective indicators (Figure 1, Mode A) are most suitable. If the objective is explanation of abstract or "unobserved" variance, formative indicators (Figure 1, Mode B) would give greater explanatory power. Both formative and reflective indicators can also be used within a single model. For instance, if one intends to explain variance in the observed criterion variables by way of the unobservables, the indicators of the endogenous construct should be reflective, and those of the exogenous, formative; the result is a mixed-mode estimation (Figure 1, Mode C). Modes A and B represent two separate principles--mode A minimizes the trace of the residual variances in the measurement equations, mode B minimizes the trace of the residual variances in the structural equations, both subject to certain systematic constraints which we will discuss shortly. Mode C is a compromise between the two principles.

Indicator mode is also shaped by an aspect of the substantive theory behind the model: the way in which the unobservable is conceptualized. Constructs such as "personality" or "attitude" are typically viewed as underlying factors that give rise to something that is observed. Their indicators tend to be realized, then, as reflective. On the other hand, when constructs are conceived as explanatory combinations of indicators (such as "population change"[2] or "marketing mix") which are determined by a combination of variables, their indicators should be formative.

Finally, there is an empirical element to the choice of indicator mode. In the formative mode, sample size and indicator multicollinearity affect the stability of indicator coefficients, which in this mode are based upon multiple regressions. In the reflective mode indicator coefficients, based on simple regressions, are not affected by multicollinearity. If the indicators are highly collinear but one nonetheless desires optimization of explained structural model variance, one might estimate mode B but use loadings, rather than regression weights, for interpretation. This will be illustrated in our subsequent analyses.

Should the considerations involving study objective, theory or conceptualization, and empirical contingencies be contradictory, the selection of indicator mode may be difficult. For example, one may wish to minimize residual variance in the structural portion of the model, which suggests use of formative indicators, even though the constructs are conceptualized as giving rise to the observations (which suggests use of reflective indicators). In such cases, the analyst might estimate twice, once in either mode. If the results correspond, there is no problem. If they differ, a compromise might be worked out using the factor structures of the blocks separately (as suggested by Bookstein 1981); otherwise, a decision as to the overriding concern must be made.

Fixed-Point Estimation

The PLS model is estimated by determining (1) the loadings (Λ_y, Λ_x) or weights (π_η, π_ξ) which describe how the observations relate to the unobservables, and (2) the structural relations (β, Γ), whereby values of unobservables influence values of the other unobservables in the system. Instead of optimizing a global scalar function, PLS estimates by way of a nonlinear operator for which the vector of all estimated item loadings (Λ_y, Λ_x) is a fixed point. Following its introduction by Wold in 1963, the properties of fixed point (FP) estimation have been discussed in Lyttkens (1968, 1973), Areskoug (1981), and in a collection of papers edited by Wold (1981). Several developments using some form of FP can be found in the recent psychometric literature (deLeeuw et al. 1976; Young et al. 1976; Perreault and Young 1980; Kroonenberg and deLeeuw 1980; Sands and Young 1980; Kruskal 1980; and Carroll et al. 1980).

FP differs from ML models such as LISREL in its basic principles and assumptions. In ML estimation, the probability of the observed data given the hypothesized model is maximized. Wold's FP estimation, which is a least squares approach, minimizes residual variances. ML estimators assume a parametric model, a family of joint distributions for all observables; FP operates as a series of interdependent OLS regressions, presuming no distributional form at all. FP estimation, then, bears no resemblance at all to the search for zeroes of certain derivatives which characterizes the estimation of ML models.

The distinction between optimizing and fixed-point methods may be compared with the two main models for solving multiple regressions. If we state as our goal the constuction of a linear form $e + b_{yx_1 \cdot x_2} x_1 + b_{yx_2 \cdot x_1} x_2$

which estimates y with least error variance, we face a problem in direct

minimization. But we may instead invoke the vocabulary of path analysis,

referring to the <u>total effect</u> b_{yx_1} of variable x_1 on variable y and attempt-

ing to partition this into a <u>direct effect</u> $b_{yx_1 \cdot x_2}$ and an <u>indirect effect</u>

$b_{x_2 x_1} b_{yx_2 \cdot x_1}$ mediated by x_1's effect on x_2. The direct and indirect effects

taken together must comprise the total: we must have

$$b_{yx_1 \cdot x_2} + b_{x_2 x_1} b_{yx_2 \cdot x_1} = b_{yx_1} \tag{6}$$

and similarly for x_2. The result is a system of two equations with two

unknowns which are, of course, identical to the normal equations of the

usual approach. In this second version all explicit minimization is rele-

gated to the bivariate analyses, the coefficients b_{yx_1}, b_{yx_2}, $b_{x_1 x_2}$ being

solutions of an easier optimum problem, the simple regression. Multiple

regression may be thought of, then, as a revision by joint constraints of

simple regression coefficients independently arrived at.

A similar distinction separates PLS from LISREL in the structural

analysis of systems of latent variables. LISREL poses and solves the global

optimization problem (maximization of likelihood) explicitly. PLS limits

its explicit optimization computations to the now-familiar case of ordinary

multiple regression. The separate simple analyses are jointly adjusted by

nonlinear algebraic constraints so that "effects" can be computed meaningfully.

Analyses

We compare LISREL and PLS using a small data set, with collinear indi-

cators, in the context of a study of consumer dissatisfaction.

In his influential book of 1970, Albert O. Hirschman developed a theory of

consumer reaction to dissatisfaction. He describes two basic modes: <u>Exit</u> and <u>Voice</u>. The exiting consumer makes use of the market by switching brands, terminating usage, or by shifting patronage--all economic actions. In contrast, Voice is a political action: a verbal protest directed at the seller, and, if remedy is not obtained, sometimes via third parties. Hirschman posits that when the Exit option is blocked or when cross-elasticities are low, Voice will increase. By this reasoning, Exit should dominate in highly competitive markets, whereas the more a market resembles a monopoly, the more Voice would be expected. Since Seller Concentration is a measure of monopoly power, we hypothesize that Concentration is negatively related to Exit and positively related to Voice.

Data

Exit and Voice data were obtained from a nationwide study by Best and Andreasen (1977). The 1972 Census of Manufacturers' 4-digit concentration ratios were used as measures of market concentration. In total, seven variables were used: four measures of Concentration, two measures of Voice (aided and unaided recall), and one measure of Exit. The Voice and Exit measures were expressed as the proportion of respondents who recalled having taken action in each of those categories. From a total of 34 product and service categories in the Best and Andreasen study, 25 were retained for this analysis. Together they represent a majority of annual consumer purchases. The deleted categories either had no corresponding S. I. C. code for concentration ratios or were too general to be meaningful. The resulting correlation matrix is presented in Table 1.

Model 1

In LISREL applications, the most common way of relating unobservables

to data is by means of reflective indicators. In this mode, the model attempts to explain the <u>observed correlations</u>. Reflective indicators in a PLS model imply that the primary objective is to explain the variances of the <u>observed variables</u>. As a starting point for comparative analysis we estimated the first model using reflective indicators.

In equation form, model 1 is set forth as follows:

$$
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} = \begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix} \begin{bmatrix} \xi \end{bmatrix} + \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} \tag{7}
$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_{y2} \\ 0 & \lambda_{y3} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \tag{8}
$$

$$
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \\ \lambda x_4 \end{bmatrix} \begin{bmatrix} \xi \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \end{bmatrix} \tag{9}
$$

Results: LISREL

The results in Figure 2 illustrate a problem that most LISREL users probably have encountered more than once (cf. Jöreskog 1977; Bentler 1976; Areskoug 1981; Driel 1978): one of the variance estimates (in this case δ_3, the error of variance of x_3) is negative and the corresponding standardized loading (correlation) is greater than 1. This is an unacceptable result.

A common practice for circumventing the problem is to fix the negative variance at zero and reestimate the model, apparently on the grounds that the offending estimate is typically low and insignificant. However, this approach has both theoretical and practical flaws. The model to which it leads is based on neither the principal components nor the common-factor model (Bentler 1976). Also, forcing one offending variance to zero will quite possibly cause the problem to reappear in other variance estimates. This is illustrated in Figure 3. When δ_3 is fixed at zero, the error variance ε_3 of y_3 becomes negative.

One cause of improper solutions might be failure of the model to fit the data (Driel 1978). Given the large chi-square for the models of figures 2 and 3, this possibility cannot be ruled out without further analysis. If the model fit is to be blamed, the improper solutions should be vitiated if $\Theta_{\sim\delta}$ is specified as symmetric rather than diagonal. But the results of Figure 4 show that this modification also fails. Even though the model now fits well, it does so by virtue of several negative variances, some quite large. Thus improper solutions are not necessarily circumvented and certainly not resolved by fixing the offending parameters or improving the model's fit. This is because we are dealing with an algebraic rather than a numerical or a statistical problem: a matter not of "likelihood" or multivariate normality, but of patterns of signs and magnitudes of the correlation matrix. Recall that LISREL's objective is to reproduce as closely as possible the observed correlation matrix $\underset{\sim}{S}$ by a matrix whose entries are explicit nonlinear functions of the parameters allowed to the model. For model 1, with $\Theta_{\sim\varepsilon}$ and $\Theta_{\sim\delta}$ diagonal, $\underset{\sim}{\Sigma}$ is modelled as:

$$\underset{\sim}{\Sigma} = \begin{bmatrix} 1.0 & & & & & & \\ \gamma_1\gamma_2\lambda_{y2} & 1.0 & & & & & \\ \gamma_1\gamma_2\lambda_{y3} & \lambda_{y1}\lambda_{y2} & 1.0 & & & & \\ \lambda_{x1}\gamma_1\lambda_{y1} & \lambda_{x1}\gamma_2\lambda_{y2} & \lambda_{x1}\gamma_2\lambda_{y3} & 1.0 & & & \\ \lambda_{x2}\gamma_1\lambda_{y1} & \lambda_{x2}\gamma_2\lambda_{y2} & \lambda_{x2}\gamma_2\lambda_{y3} & \lambda_{x2}\lambda_{x1} & 1.0 & & \\ \lambda_{x3}\gamma_1\lambda_{y1} & \lambda_{x3}\gamma_2\lambda_{y2} & \lambda_{x3}\gamma_2\lambda_{y3} & \lambda_{x3}\lambda_{x1} & \lambda_{x3}\lambda_{x2} & 1.0 & \\ \lambda_{x4}\gamma_1\lambda_{y1} & \lambda_{x4}\gamma_2\lambda_{y2} & \lambda_{x4}\gamma_2\lambda_{y3} & \lambda_{x4}\lambda_{x1} & \lambda_{x4}\lambda_{x2} & \lambda_{x4}\lambda_{x3} & 1.0 \end{bmatrix} \quad (10)$$

For the model to be exactly true, it must reproduce the correlations among the indicators of concentration:

$$\lambda_{x2}\lambda_{x1} = .9594 \qquad (11)$$

$$\lambda_{x3}\lambda_{x1} = .8797 \qquad (12)$$

$$\lambda_{x3}\lambda_{x2} = .9561 \qquad (13)$$

$$\lambda_{x4}\lambda_{x1} = .7810 \qquad (14)$$

$$\lambda_{x4}\lambda_{x2} = .8664 \qquad (15)$$

$$\lambda_{x4}\lambda_{x3} = .9662 \qquad (16)$$

Notice that the correlation between these pairs of indicators is a strong negative function of the absolute difference between their subscripts—a characteristic of the familiar "Heywood" and simplex models of items in series, for which no single-factor model is adequate. For example, from equations 11-16 it is evident that $\lambda_{x1} \geq .9594$ and $\lambda_{x4} \geq .9662$. It follows that $\lambda_{x4}\lambda_{x1} \geq .9268$--but $r_{x1x4} = .7810$. Since no λ can exceed unity (as the

standardized error variance in LISREL is equal to $1 - \lambda^2$) it is not possible

for $\underset{\sim}{\Sigma}$ to fit $\underset{\sim}{S}$ without impropriety in the solutions; and the closer we come

to fitting $\underset{\sim}{S}$, the smaller will be the chi-square and the larger the magnitude

of the estimated negative variances.

Driel (1978) recommended that variables involved in improper solutions

be deleted. By eliminating variables x_2 (CR8) and x_3 (CR20) and the Exit (y_3)

variable from the data set—reducing, incidentally, the collinearity hobbling

the x-block—the model now looks like Mode A in Figure 1:

$$
\begin{bmatrix}
1.0 & & & \\
.7094 & 1.0 & & \\
.2881 & .1827 & 1.0 & \\
.2813 & .1539 & .7810 & 1.0
\end{bmatrix}
=
\begin{bmatrix}
1.0 & & & \\
\lambda_{y_1}\lambda_{y_2} & 1.0 & & \\
\lambda_{y_1}\gamma\lambda_{x_1} & \lambda_{y_2}\gamma\lambda_{x_1} & 1.0 & \\
\lambda_{y_1}\gamma\lambda_{x_2} & \lambda_{y_2}\gamma\lambda_{x_2} & \lambda_{x_1}\lambda_{x_2} & 1.0
\end{bmatrix}
\tag{17}
$$

. Again we consider the estimates necessary for an exact fit. Algebraically

(cf. Fornell and Larcker 1981b), $\lambda_{y_1}^2$ may be estimated either by $(r_{x_2y_1}r_{y_1y_2}/r_{x_2y_2})$

= 1.30 or by $(r_{x_1y_1}r_{y_1y_2}/r_{x_1y_2}) = 1.12$. The error variances for y_1 are thus

-.30 and -.12, respectively. ML estimation will conflate these two values

into a single composite estimate; their pooling thus imputes a solution which

is still improper. The impropriety is not a result of collinearity; one can

have very high correlations and still obtain interpretable estimated models if

the ratios of cross correlations $r_{x_iy_j}/r_{x_iy_k}$ are reasonably low. The improper

solution is rather a consequence of either (a) the failure of a single-factor

model to explain the correlation submatrix for a particular block (unobserv-

able), or (b) inhomogeneities of the cross-correlation submatrices between

blocks which cause them to be obviously of full rank. These are precisely the

assumptions of LISREL submodels, block by block, which are seldom verified in the course of overall modelling (which is, of course, why they cause problems).

Results: PLS

In Figure 2, the estimates via PLS for Model 1 are presented above or to the left of the corresponding LISREL estimates. Comparing them, we note lower structural parameters γ_1 and γ_2 in PLS but mostly higher loadings $\Lambda_{\sim y}$, $\Lambda_{\sim x}$. PLS does not produce improper estimates, as all residual variances are actual regression residuals; they are not inferred from the data. The PLS results are thus interpretable; they suggest a weak negative relationship between Concentration and Exit and a positive relationship between Concentration and Voice, along the lines suggested by Hirschman. The model is satisfactory insofar as the loadings, our primary focus in accounting for observed variances, are quite large. In general, PLS estimates of models with reflective indicators impute smaller measurement errors and weaker structural relationships than does LISREL. The algorithm by which the PLS estimates were obtained may be found in the Appendix to this paper.

Model 2

The cause of improper solutions is generally the attempt to account for observed correlation matrices by patterned products of model parameters that are inadequate. The problem may be circumvented by attending to variances instead of correlations, that is, by working with components rather than factors. Components, which are exact linear combinations of their indicators, "maximize variance," while factors "explain

covariance." To explore this amelioration we take advantage of the circum-
stance that certain component structures can be estimated by both PLS and
LISREL. The MIMIC model (Jöreskog and Goldberger 1975; Stapleton 1978;
Bagozzi, Fornell, and Larcker 1981) is one such case.

Let this LISREL model be:

$$\eta_1 = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \end{bmatrix} + \begin{bmatrix} \xi_1 = x_1 \\ \xi_2 = x_2 \\ \xi_3 = x_3 \\ \xi_4 = x_4 \end{bmatrix} \tag{18}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} \lambda_{y1} \\ \lambda_{y2} \\ \lambda_{y3} \end{bmatrix} \begin{bmatrix} \eta_1 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \tag{19}$$

with $\underset{\sim \varepsilon}{\Theta}$ symmetric. The LISREL estimates are presented in Figure 5. For
purposes of interpretation we make reference to the loadings for the x-block,
computed as:

$$\underset{\sim x}{\Lambda} = \underset{\sim xx}{R} \, \underset{\sim}{\Gamma} , \tag{20}$$

to yield $\lambda_{x_1} = .42$, $\lambda_{x_2} = .51$, $\lambda_{x_3} = .58$, $\lambda_{x_4} = .45$. From these loadings
we compute the elements of $\underset{\sim \delta}{\Theta}$ in Figure 5.

The PLS version of the above model is:

$$\eta = \gamma \xi + \zeta \tag{21}$$

$$\eta = \begin{bmatrix} \pi_{\eta 1} & \pi_{\eta 2} & \pi_{\eta 3} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \tag{22}$$

$$\xi = \begin{bmatrix} \pi_{\xi_1} & \pi_{\xi_2} & \pi_{\xi_3} & \pi_{\xi_4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad (23)$$

The results, according to a two-construct mode B estimation (as described in the appendix), appear in Figure 6. Clearly, the PLS and LISREL solutions have identical x-weights (γ's in LISREL, π's in PLS). Further, the loadings for the y-side (λ_y's in LISREL) are equal to the loadings of the PLS results up to a factor of $1/\gamma_{PLS}$.[4]

Thus, if LISREL is specified in a MIMIC version with Θ_ϵ symmetric, it produces results identical to those from a PLS model with formative indicators. The formative specification with $\zeta = 0$--that unobservables be exact linear combinations of their indicators--is not as restrictive as it may appear. When the errors of the y-variables are correlated, the error term ζ of the unobservable is distributed instead throughout the elements of Θ_ϵ (see Hauser and Goldberger 1971).

The PLS estimate of the structural parameter γ_{PLS} in Model 2, Figure 6, is larger than either of the two estimates γ_1, γ_2 in Model 1, Figure 2. This is because the primary objective of a formative-mode model is to minimize the trace of the residual matrix Ψ (the variance-covariance matrix of ζ), so that the measurement portion of Model 2 absorbs the largest possible part of the total residual, subject to the constraint of being a fixed point. The formative formulation, then, imputes a stronger relationship between Concen-

tration and the construct combining Exit with the two Voice measures. The

inverse relationship between Exit and Voice is reflected in the negative

weight for Exit ($\pi_{\eta 1}$ = -.80) in this construct.

Recapitulation of Models 1 and 2

We pause here to review certain important distinctions we have estab-

lished. Model 1 presumes indicators which are reflective of the constructs.

The LISREL estimation yielded negative variances and standardized loadings

(correlations) greater than one. We showed that these unacceptable estimates

ultimately derived from the LISREL objective of fitting a pattern of para-

meter products to the correlations observed. The PLS estimates, which by

construction cannot yield such improprieties, exhibited smaller measurement

variances but also lower estimated structural parameters. This accords with

the purpose of PLS mode A, which is to explain the variance of the observed

variables by minimizing mean-squared measurement residuals.

Model 2, by contrast, involved formative indicators. Estimates in this

form focus on the variance in the structural portion of the model so that

more of the net failure-of-fit is partitioned into measurement error. We

emphasize that the choice between formative and reflective indicators is not

merely a matter of empirical statistical fact. Choice of indicator mode

brings conceptual, theoretical, and empirical observations to bear together

on the objectives of the study; the partitioning of error variance can only

be manipulated insofar as it depends on this choice. In particular, a

MIMIC model specified without a disturbance term ζ but with correlated

measurement errors is equivalent to a formative PLS model with two constructs.

In this case PLS and LISREL produce identical estimated structures.

Model 3

Even though Model 2 produced consistent results from both LISREL and PLS, it is not an attractive formulation because no distinction is made, at the abstract level, between Exit and Voice. Our third model makes this distinction. Let us assume that the objective is to explain the observed y-variables (that is, in LISREL, their correlations; in PLS, their variances). In order to avoid uninterpretable results from LISREL, assume also that Concentration is formed by its indicators without any surplus variance. We arrive at a model with both formative and reflective indicators. The LISREL equations are:

$$
\begin{bmatrix} 1 & 0 \\ \beta & 1 \end{bmatrix}
\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}
=
\begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ 0 & 0 & 0 & 0 \end{bmatrix}
\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}
+
\begin{bmatrix} 0 \\ \zeta_2 \end{bmatrix}
\tag{24}
$$

$$
\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}
=
\begin{bmatrix} \lambda_{y1} & 0 \\ 0 & \lambda_{y2} \\ 0 & \lambda_{y3} \end{bmatrix}
\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}
+
\begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix}
\tag{25}
$$

where η_1 is Concentration, η_2 is Voice, and y_1 is Exit.

The PLS model is:

$$
\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}
\begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix}
=
\begin{bmatrix} \gamma_1 \\ \gamma_2 \end{bmatrix}
\begin{bmatrix} \xi \end{bmatrix}
+
\begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix}
\tag{26}
$$

$$\xi = \begin{bmatrix} \pi_{\xi_1} & \pi_{\xi_2} & \pi_{\xi_3} & \pi_{\xi_4} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \tag{27}$$

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \lambda_{y_2} \\ 0 & \lambda_{y_3} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \tag{28}$$

where η_1 is Exit (y_1), η_2 is Voice, and ξ is Concentration.

The two sets of estimates presented in Figure 7 are similar.[5] The disturbance terms in the structural portion are slightly higher in PLS, whereas the measurement residuals are higher in LISREL. These differences derive from the different fitting objectives. We may explore them in detail by use of the descriptive statistics from the testing system of Fornell and Larcker (1981a).

Referring to Table 2, the Average Variance Accounted for (AVA) is the mean-squared structural parameter. This statistic is slightly higher in LISREL. Because of the similarity of model specifications, the difference in AVA is small. Average Variance Extracted by the unobservables (AVE) is the mean-squared loading for blocks of indicators separately. This statistic is consistently higher in PLS than in LISREL, due to the smaller imputed measurement errors. For each endogenous construct, Redundancy (which is the product of the squared structural parameter by AVE) measures the power of the exogenous constructs for predicting the y-variables. As Exit (y_1) has no measurement error by construction, its AVE is 1. Thus, Redundancy for y_1 is γ^2 in

as well as frequency of convergence. Unlike ML techniques, PLS makes minimal demands about measurement scales, sample size, or the distribution of residuals. Small sample sizes—sometimes fewer than the number of variables (Wold 1980c)—can be sufficient for meaningful PLS analyses; and, in contrast to ML, PLS estimation does not involve a statistical model, thereby avoiding the need for assumptions regarding scales of measurement. Nominal, ordinal, and interval-scaled variables are permissible in PLS in the same ways as in ordinary regression.

A primary difference between PLS and LISREL concerns the structure of unobservables. LISREL specifies the residual structures, while PLS specifies the unobservables. This difference bears important implications which have long been debated in the psychometric literature (see the review by Steiger, 1979). The main defense of the factor model is that it allows for imperfect measurement by assigning surplus variance to the unobservables. But such measurement error implies certain disturbing consequences. An infinite number of unobservables may bear the same pattern of correlations with observed variables and yet be only weakly or even negatively correlated with each other (Mulaik and McDonald 1978). For exploratory analyses, such indeterminacy can be very problematic. In confirmatory structural equations, modelling indeterminacy has been thought to be less of a problem by reason of presumed existence of "prior knowledge" ruling out conflicting explanations. Because the chi-square statistic of fit in LISREL is identical for all possible unobservables satisfying the same structure of loadings, a priori knowledge is necessary. However, indeterminacy can create difficulties for confirmatory studies as well. There have been cases where several hypothesized models account for the same data equally well

PLS and $\lambda^2_{y_1}$ in LISREL. Redundancy for predicting the observed Voice indicators is calculated as the product of γ^2_2 (PLS) or β^2 (LISREL) with the corresponding AVE. Because the measurement errors are smaller, Redundancy is higher for PLS.

The sharpest difference between the PLS and LISREL estimates is seen in Operational Variance (OV), which is equal to Redundancy times the squared multiple correlation of the exogenous indicators with their unobservable construct. This statistic expresses the degree to which the variances of the y-observables are accounted for, via the model, by the observed x-variables. In the PLS solution, because all unobservables are exactly defined, it is identical to Redundancy; in LISREL it is attenuated and therefore lower. For instance, when the Redundancy of Voice is multiplied by the squared multiple correlation of Concentration with its indicators,[6] there results an OV of .072 in LISREL; the corresponding value for PLS is .105.

Discussion

Under the classifical assumptions of independence and normally distributed residuals, ML and OLS estimates in regression analysis are identical. In structural equation modelling, this is not the case. Except as applied to certain MIMIC models, PLS and LISREL have different objectives and present systematically different results. However, as illustrated in Models 2 and 3, the more similar the model specification, the more similar the results will be.

As we have argued above, LISREL attempts to account for observed correlations, whereas PLS aims to account for variances at the observed or abstract level (depending upon indicator mode). Other major differences between the models include assumptions about factor structure, mechanisms of statistical inference, matters of identification,[7] interpretation of measurement error,

(see Mulaik 1976). Thus, confirmatory studies are not necessarily free from the problem of having several interpretations. In this paper we suggest an equally serious problem: only indeterminate factors can have improper loadings. Since improper loadings lead to negative variances, such results do not have several interpretations, but, rather, none at all.

For the discipline of marketing, which is often concerned with prediction and control, other drawbacks flow from indeterminacy. As factor scores cannot be calculated, specific case predictions are not possible without prior estimation of the scores themselves. Similarly, testing for outliers and modelling of factor scores cannot be done either. PLS avoids factor indeterminacy by explicitly defining the unobservables. Factor scores for prediction or further modelling are then readily available.

Yet PLS estimators lack the precision of ML estimators. Given multivariate normality, LISREL estimates are efficient in large samples, and support analytic estimates of asymptotic standard errors. In exchanging far greater a priori assertion for statistical inference, LISREL is a model for theory testing. PLS is applicable over a wider range of problems, in particular, when prior information is wanting and theory is less developed. However, the theoretical underpinnings of PLS estimation are also less well developed. Although it appears to converge more quickly than LISREL (Areskoug 1981), an analytic proof of PLS convergence for general models has yet to be worked out.

The example used in this paper is not amenable to statistical inference because the sample is small and nonrandom. But PLS modelling is not necessarily devoid of statistical inference. Although statistical tests are not integrated within the protocol of estimation, one may invoke the classic

standard-error formulae of multiple regression or (perhaps more sensibly) use a Stone-Geisser (Stone 1974; Geisser 1974) test for predictive significance. This provides jackknifed standard errors for the individual parameters, whereas LISREL calculates standard errors from the inverse of the information matrix. Both methods have limitations. In PLS, there may be a problem in selecting the appropriate subgroup size for estimating pseudo-values in jackknifing. However, the larger the subsample size, the more reliable the statistic. The standard errors estimated by LISREL are also fallible. Since each unobservable requires a scale identification restriction affecting the information matrix, standard errors vary by choice of restriction, and can yield rather unpleasant paradoxes of interpretation (Pijper and Saris 1981). By contrast, in recursive models, the fixed-point estimation scheme of PLS is free of identification problems.

The tests for predictive significance, Redundancy, and Operational Variance can be used to assess the overall "fit" of both LISREL and PLS models. A corresponding test in LISREL is the likelihood ratio chi-square. However, this test has some undesirable properties which limit its ability to support any conclusions beyond those dealing narrowly with accounting for observed correlations. Fornell and Larcker (1981a, 1981b) showed that the strength of variable relationships is _negatively_ associated with goodness of fit in models otherwise isomorphic.

Summary

For the marketing analyst, the choice between LISREL and PLS is neither arbitrary nor straightforward. Both apply to the same class of models--structural equations with unobservables and measurement error--but they have different structures and objectives.

LISREL attempts to account for observed correlations, while PLS aims at explaining variances (of variables observed or unobserved).

LISREL offers statistical precision in the context of stringent as-asumptions; PLS trades efficiency for simplicity and fewer assumptions.

The factor model underlying LISREL allows more errors in measurement than the components model invoked by PLS. In LISREL, unobservables are truly unobservable; PLS foresakes the consequent enhancements of theoretical explanation in order to avoid the ambiguities and improprieties which often ensue.

In sum, nothing less than the general research setting can determine the appropriate modelling approach. It is within this context that the LISREL benefits of statistical efficiency and higher estimated relationships among unobservables should be weighed against the problems associated with indeterminacy. The analyses of the paper suggest that when relevant correlation submatrices (block by block) are not of the appropriate reduced rank, factor indeterminacy is more serious than generally acknowledged. The frequent occurrences of improper and uninterpretable solutions advise against the use of LISREL unless its assumptions are verifiably true; and, when they are not, we submit that PLS is more likely to provide meaningful analysis.

TABLE 1

CORRELATION MATRIX

Exit	(y_1)	1.0					
Voice 1	(y_2)	-.0079	1.0				
Voice 2	(y_3)	-.0797	.7094	1.0			
CR 4	(x_1)	-.1024	.2881	.1827	1.0		
CR 8	(x_2)	-.1757	.2759	.1820	.9594	1.0	
CR 20	(x_3)	-.2000	.3057	.1959	.8797	.9561	1.0
CR 50	(x_4)	-.1311	.2813	.1539	.7810	.8664	.9662	1.0

TABLE 2

Model 3: Descriptive Statistics*

|	PLS	LISREL
<u>Average Variance Accounted (AVA) for</u> <u>The Structural Model</u>	.165	.177
<u>Average Variance Extracted (AVE) by</u> <u>The Unobservables</u>		
ρ_{vc} (market concentration)	.256	.228
ρ_{vc} (voice)	.856	.706
ρ_{vc} (exit	1.0	1.0
<u>Redundancy</u>		
$\overline{R}^2_{y_1}$ /market concentration	.212	.230
$\overline{R}^2_{y_2, \, y_3}$ /market concentration	.105	.087
<u>Operational Variance (OV)</u>		
$\hat{\overline{R}}^2_{y_1}/x_1 \; \cdots \; x_4$	.212	.230
$\hat{\overline{R}}^2_{y_2, y_3}/x_1 \; \cdots \; x_4$	.105	.072

*See Fornell and Larcker (1981a) for formulas and a detailed description of these statistics.
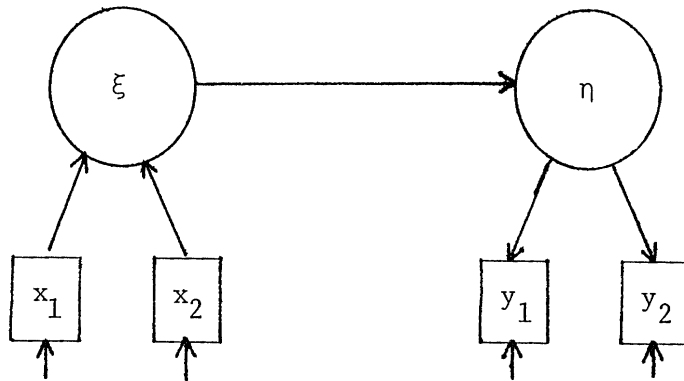
FIGURE 1

THREE DIFFERENT MODES OF RELATING UNOBSERVABLES TO EMPIRICAL INDICATORS
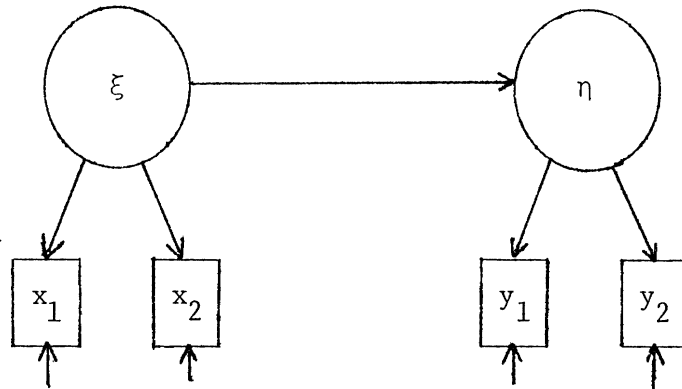
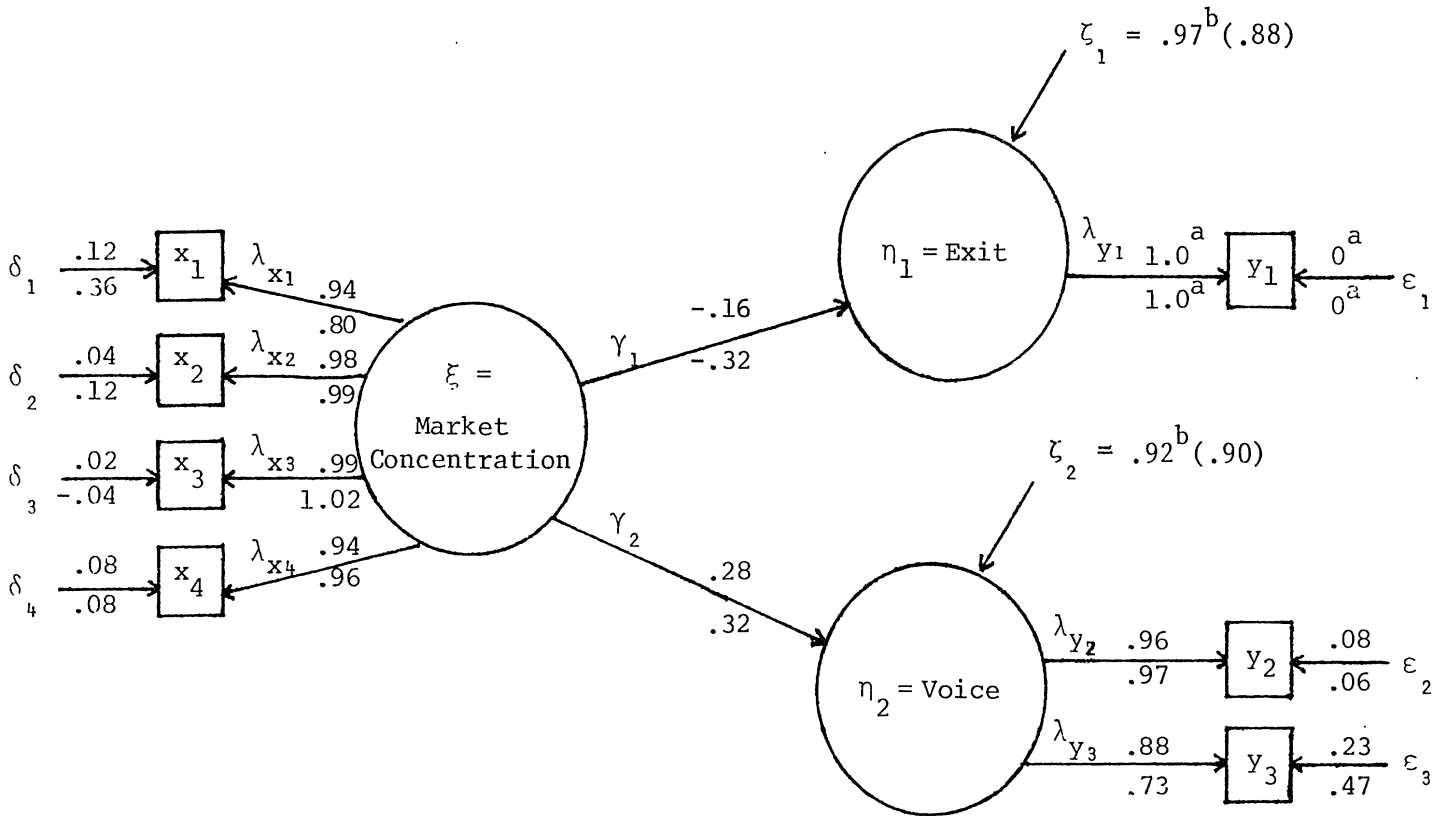Mode A: Reflective indicators
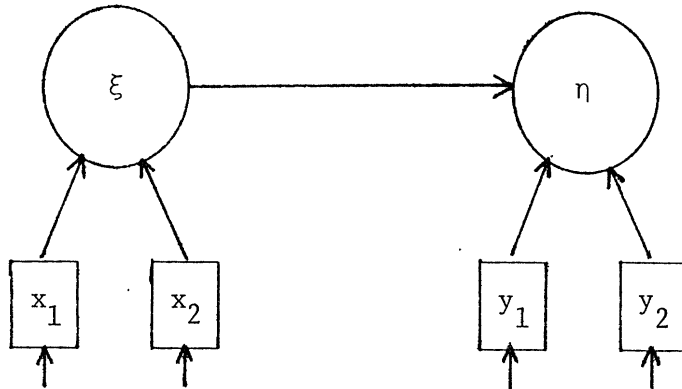
Mode B: Formative indicators

Mode C: Formative indicators for the exogenous construct,
         reflective indicators for the endogenous construct

FIGURE 2

MODEL 1: PLS AND LISREL ESTIMATES WITH REFLECTIVE INDICATORS*

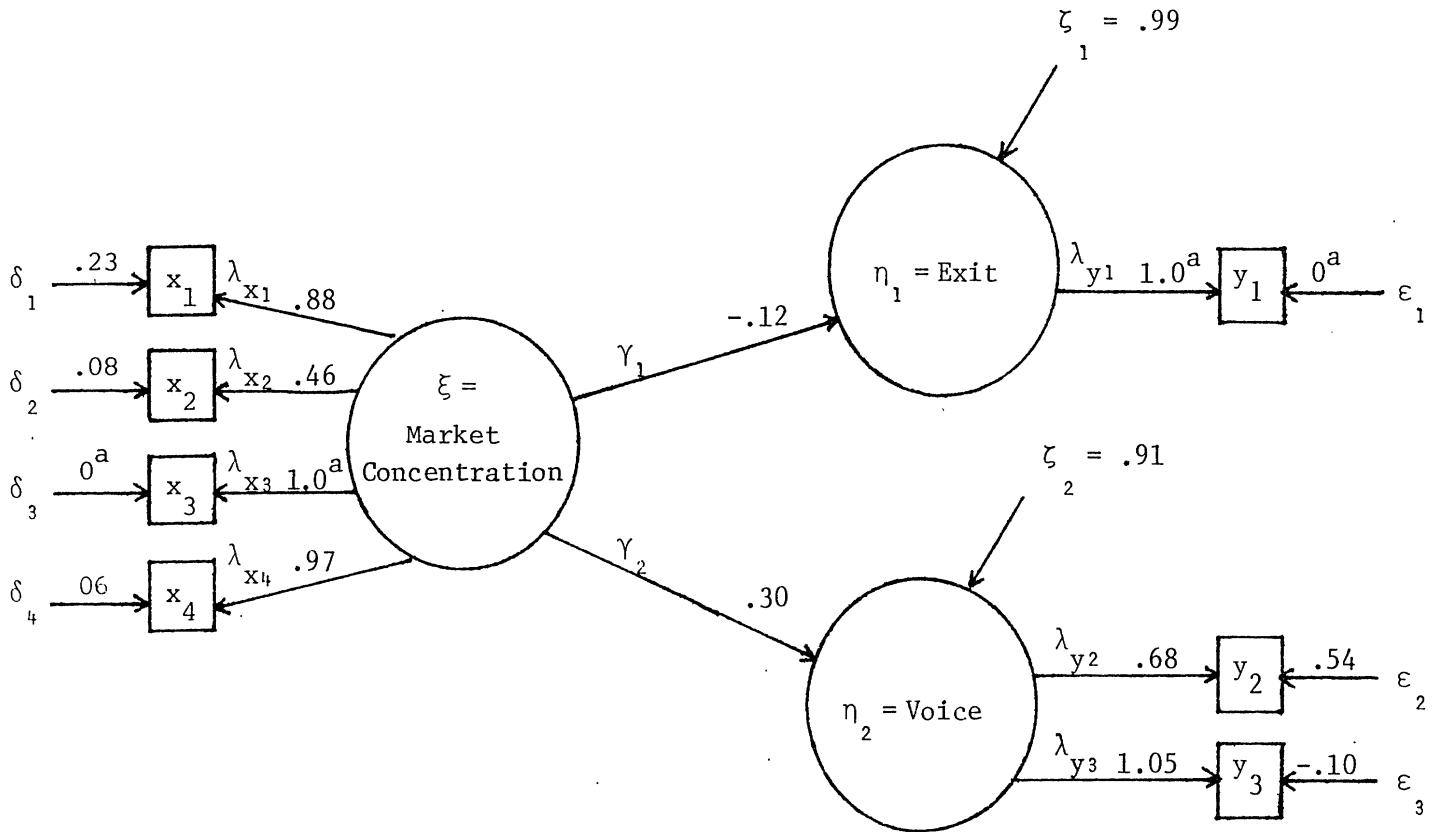*PLS Standardized Estimates above arrow, LISREL Standardized Estimates below arrow

[a]Fixed Parameter

[b]PLS Estimate; Corresponding LISREL Estimate in Parentheses

$\chi^2 = 49.72$ (13 d.f.) $p \approx 0$

FIGURE 3

MODEL 1: LISREL ESTIMATES WITH OFFENDING VARIANCE ESTIMATE FIXED TO ZERO
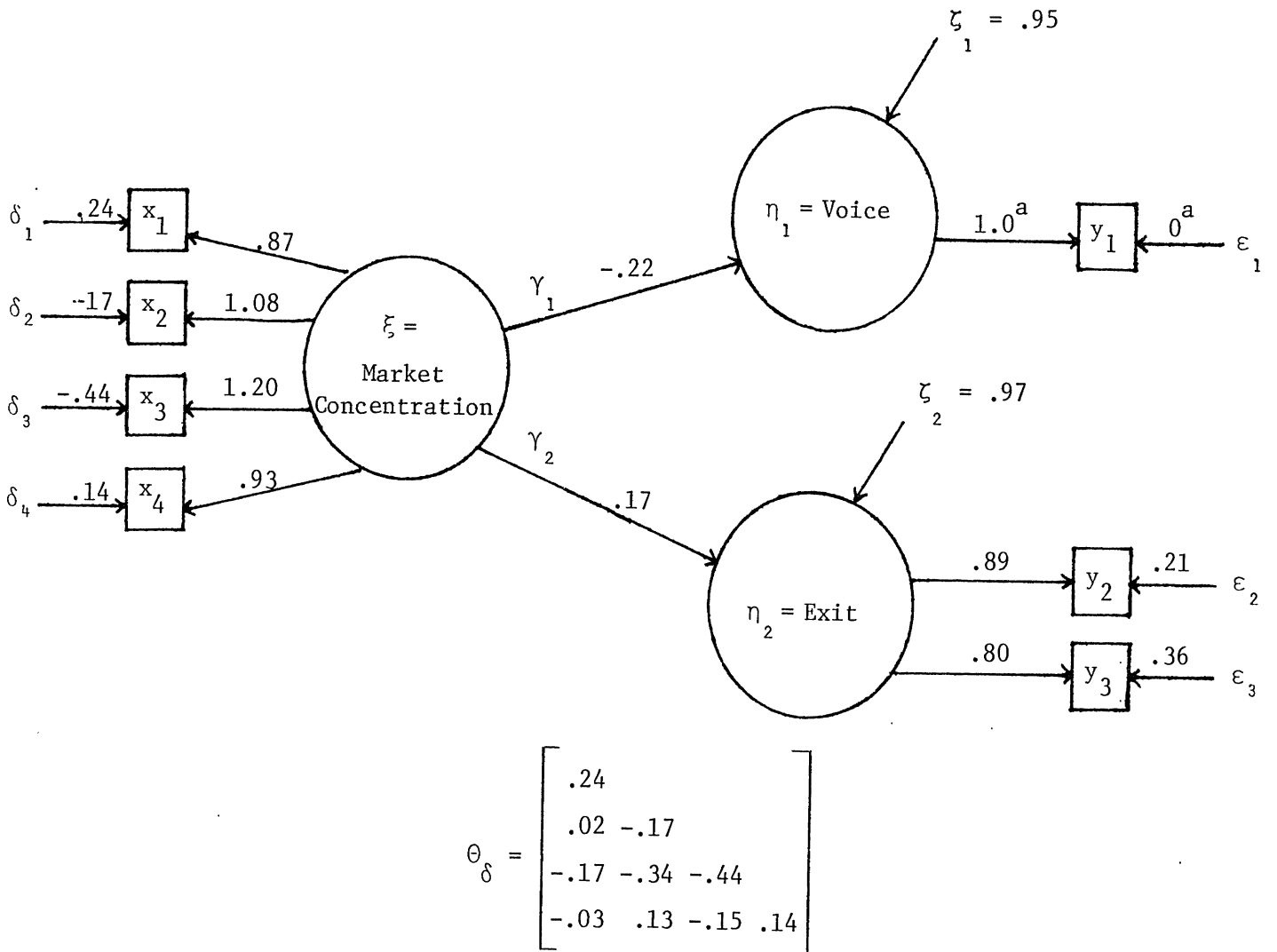

[a]Fixed Parameter

$\chi^2 = 121.17$ (14 d.f.) $p \approx 0$

FIGURE 4

MODEL 1:   LISREL ESTIMATES WITH CORRELATED MEASUREMENT ERRORS FOR
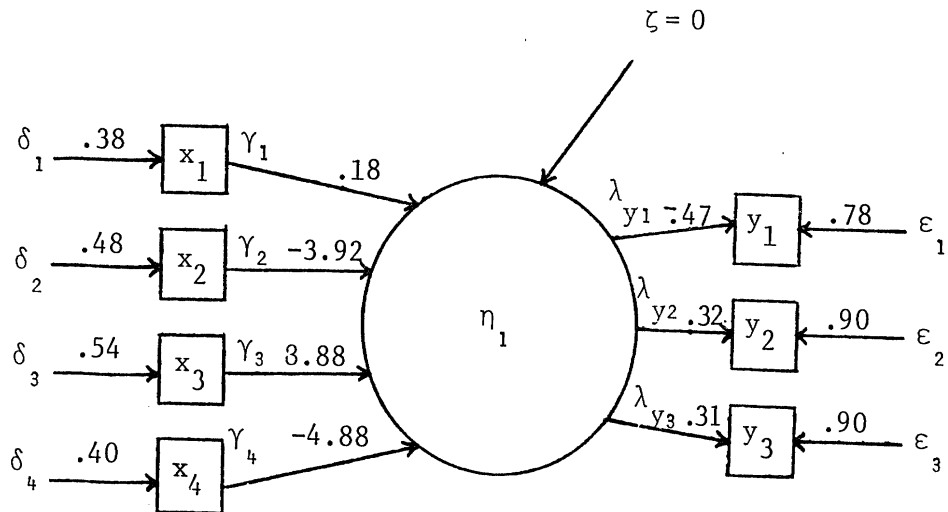
THE EXOGENEOUS CONSTRUCT

$$\Theta_\delta = \begin{bmatrix} .24 & & & \\ .02 & -.17 & & \\ -.17 & -.34 & -.44 & \\ -.03 & .13 & -.15 & .14 \end{bmatrix}$$

$\chi^2 = 2.41$ (d.f. = 7) $p \doteq .89$

[a]Fixed Parameter

FIGURE 5

MODEL 2: LISREL ESTIMATES

$$\theta_\epsilon = \begin{bmatrix} .78 & & \\ .14 & .90 & \\ .06 & .61 & .90 \end{bmatrix}$$

$\chi^2 = 2.66$ (d.f. = 6) $p \approx .85$
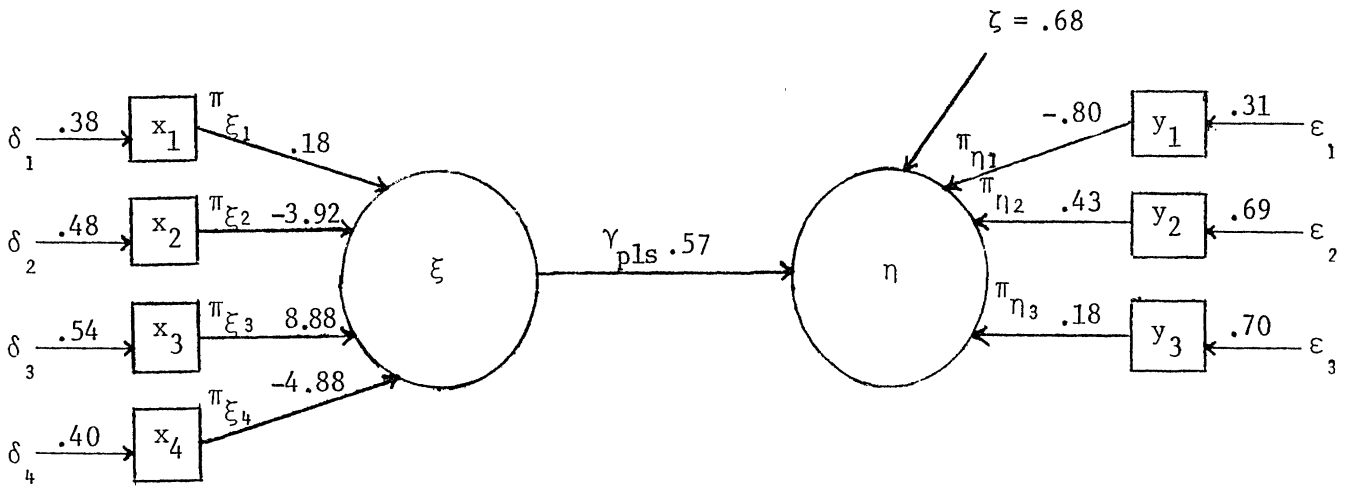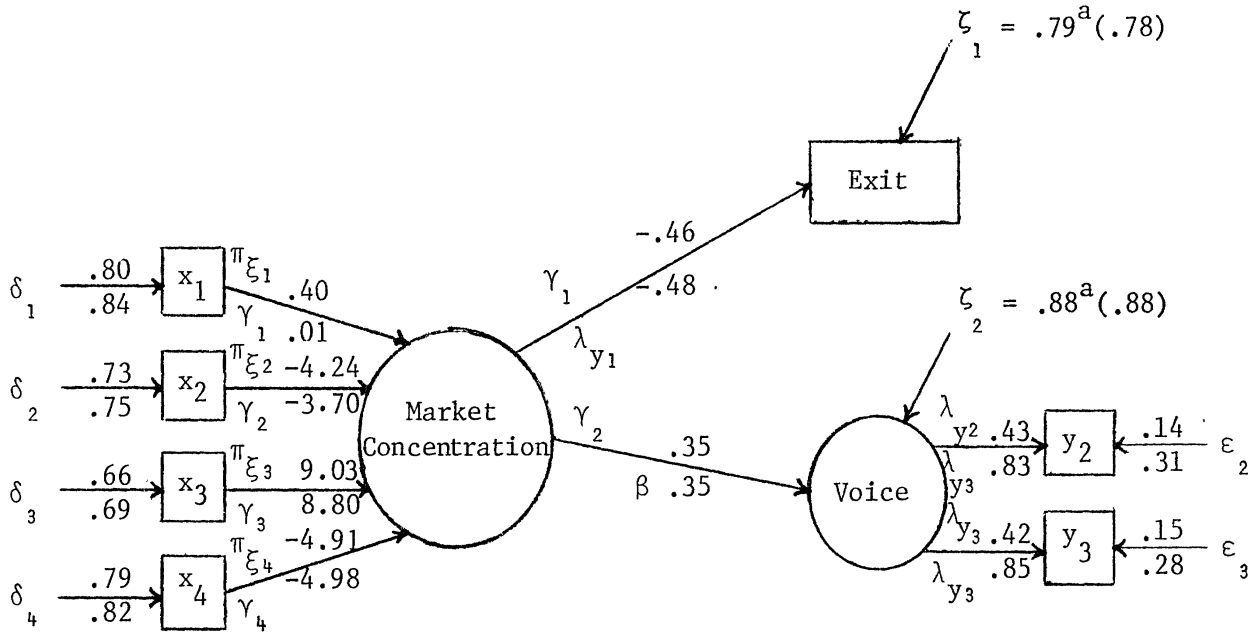
FIGURE 6

MODEL 2: PLS ESTIMATES

FIGURE 7

MODEL 3: PLS AND LISREL ESTIMATES *

$\chi^2 = 3.347$ (d.f. = 8) $p \approx .91$

*PLS Standardized Estimates above arrow, LISREL Standardized Estimates below arrow

[a] PLS Estimate, corresponding LISREL Estimate in parentheses

FOOTNOTES

[1] Previously known as NIPALS (Nonlinear Iterative Partial Least Squares) or NILES (Nonlinear Iterative Least Squares).

[2] "Population change" is presumed to be determined by natality, mortality, and migration. This example is from Hauser's (1973) criticism of sociologists' overreliance on reflective indication.

[3] No attempt was made to pool concentration ratios for several categories to obtain a better fit between the Best and Andreasen data and the S. I. C. codes. Therefore, substantive conclusions based on the analysis reported here must be interpreted with caution. Also, given the nature of the data, the sample size, and the sampling procedure, statistical inference will play a very minor role in our analyses.

[4] The y-loadings in PLS are:

$$\underset{\sim}{\Lambda}_y = \frac{1}{\gamma_{PLS}} \begin{bmatrix} \lambda_{y1} & \lambda_{y2} & \lambda_{y3} \end{bmatrix}$$

where γ_{PLS} is the estimated structural parameter from PLS, and $\lambda_{y1} \ldots \lambda_{y3}$ are the estimated loadings from LISREL. Numerically, we have

$$\underset{\sim}{\Lambda}_y = \frac{1}{.57} \begin{bmatrix} -.47 & .31 & .31 \end{bmatrix} = \begin{bmatrix} -.83 & .56 & .55 \end{bmatrix}$$

which are the y-loadings. This can easily be verified by computing the corresponding loadings for the PLS solution as $\underset{\sim}{\Lambda}_y = \underset{\sim}{R}_{yy} \underset{\sim}{\pi}_\eta$.

The estimate γ_{PLS} is not available from LISREL but can be computed as:

$$\left(\underset{\sim}{R}_{\eta_1 y} \underset{\sim}{R}_{yy}^{-1} \underset{\sim}{R}_{\eta_1 y} \right)^{\frac{1}{2}} .$$

Since γ_{PLS} is the largest eigenvalue of $\underset{\sim}{R}_{yy}^{-1} \underset{\sim}{R}_{yx} \underset{\sim}{R}_{xx}^{-1} \underset{\sim}{R}_{xy}$, both PLS and LISREL provide general models for canonical correlation.

[5] In other comparisons we have found significant structural relations in a LISREL model while corresponding PLS estimates have been insignificant, even when the latter was estimated via Mode-B to minimize the structural error. As shown by Bagozzi, Fornell, and Larcker (1981), this paradox is traceable to the factor score indeterminacy in LISREL.

[6] This adjustment is identical to the index of factor score indeterminacy proposed by Green (1976).

[7] For recursive models, PLS basically has no problem on this score -- cf. Bookstein, 1981.

REFERENCES

Ågren, A. (1972), Extensions of the Fixed-Point Method, published doctoral
    dissertation, Department of Statistics, University of Uppsala, Sweden.

Apel, H. (1977), Simulation sozio-ökonomischer Zusammanhange-Kritik and
    Modification von Systems Analysis, doctoral dissertation, J. W. von
    Goethe University, Frankfurt and Main, Germany.

Areskoug, B. (1981), "Some Asymptotic Properties of PLS Estimators and a
    Simulation Study for Comparisons between LISREL and PLS," in Systems
    under Indirect Observation: Causality, Structure, Prediction, K. G.
    Jöreskog and H. Wold, eds. Amsterdam: North Holland, in press.

_____, H. Wold, and E. Lyttkens (1975), "Six Models with Two
    Blocks of Observables as Indicators for One or Two Latent Variables,"
    Research Report No. 6, Department of Statistics, University of
    Gothenburg, Sweden.

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

_____, Claes Fornell, and David F. Larcker (1981), "Canonical
    Correlation Analysis as a Special Case of a Structural Relations Model,"
    Multivariate Behavioral Research, 16, in press.

Bentler, P. M. (1976), "Multistructure Statistical Model Applied to Factor
    Analysis," Multivariate Behavioral Research, 11, 3-25.

Bergström, R. (1972), "An Investigation of the Reduced Fixed-Point Method,"
    seminar paper, Department of Statistics, University of Uppsala, Sweden.

Best, Arthur and Alan R. Andreasen (1977), "Consumer Response to Unsatisfac-
    tory Purchases: A Survey of Perceiving Defects, Voicing Complaints and
    Obtaining Redress," Law & Society, 11, 701-42.

Bodin, L. (1974), Recursive Fixed-Point Estimation: Theory and Application,
    published doctoral dissertation, Department of Statistics, University
    of Uppsala, Sweden.

Bookstein, Fred L. (1980), "Data Analysis by Partial Least Squares," in Evalu-
    ation of Econometric Models, J. Kmenta and J. B. Ramsey, eds. New York:
    Academic Press, 75-90.

_____ (1981), "The Geometric Meaning of Soft Modeling with Some
    Generalizations," in Systems under Indirect Observation: Causality,
    Structure, Prediction, K. G. Jöreskog and H. Wold, eds. Amsterdam:
    North Holland, in press.

Carroll, J. Douglas, Sandra Pruzansky, and Joseph Kruskal (1980), "CANDELINC:
    A General Approach to Multidimensional Analysis of Many-Way Arrays with
    Linear Constraints on Parameters," Psychometrika, 45, 3-24.

deLeeuw, J., F. W. Young, and Y. Takane (1976), "Additive Structure in Qualitative Data: An Alternating Least Squares Method with Optimal Scaling Features," Psychometrika, 41, 471-503.

Driel, Otto van (1978), "On Various Causes of Improper Solutions in Maximum Likelihood Factor Analysis," Psychometrika, 43, 225-43.

Fornell, Claes (1982), ed. A Second Generation of Multivariate Analysis in Marketing, New York: Praeger.

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

_____ and David F. Larcker (1981b), "Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics," Journal of Marketing Research, 18 (August), 382-88.

Geisser, S. (1974), "A Predictive Approach to the Random Effect Model," Biometrika, 61, 101-7.

Green, Bert F. (1976), "On the Factor Score Controversy," Psychometrika, 41, 263-66.

Hauser, Robert M. (1973), "Disaggregating a Social-Psychological Model of Educational Attainment," in Structural Equation Models in the Social Sciences, A. S. Goldberger and O. D. Duncan, eds. New York: Seminar Press, 255-84.

_____ and Arthur S. Goldberger (1971), "The Treatment of Unobservable Variables in Path Analysis," in Sociological Methodology, H. L. Costner, ed. San Francisco: Jossey-Bass, 81-117.

Hirschman, Albert O. (1970), Exit, Voice, and Loyalty--Responses to Decline in Firms, Organizations, and States. Cambridge: Harvard University Press.

Hui, B. S. (1978), "The Partial Least Squares Approach to Path Models of Indirectly Observed Variables with Multiple Indicators," unpublished doctoral thesis, University of Pennsylvania.

Jagpal, Harsharanjeet S. (1981), "Measuring Joint Advertising Effects in Multiproduct Firms," Journal of Advertising Research, 21, No. 1, 65-69.

Jöreskog, Karl G. (1970), "A General Method for Analysis of Covariance Structures," Biometrika, 57, 239-51.

_____ (1973), "A General Method for Estimating a Linear Structural Equation System," in Structural Equation Models in the Social Sciences, A. S. Goldberger and O. D. Duncan, eds. New York: Seminar Press, 85-112.

_____ (1979), "Structural Equation Models in the Social Sciences: Specification, Estimation and Testing," in Advances in Factor Analysis and Structural Equation Models, Karl G. Jöreskog and Dag Sörbom, eds. Cambridge, Mass.: ABT Books, 105-27.

_____ and Arthur S. Goldberger (1975), "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable," Journal of the American Statistical Association, 70, 631-39.

Jöreskog, Karl G. and Dag Sörbom (1978), LISREL IV: Analysis of Linear Structural Relationships by the Method of Maximum Likelihood. Chicago: National Educational Resources.

Jöreskog, Karl G. and Herman Wold (1981), eds. Systems under Indirect Observation: Causality, Structure, Prediction. Amsterdam: North Holland, in press.

Kowalski, B. R., R. W. Gergerlach, and H. Wold (1981), "Chemical Systems under Indirect Observation," in Systems under Indirect Observation: Causality, Structure, Prediction, K. G. Jöreskog and H. Wold, eds. Amsterdam: North Holland, in press.

Kroonenberg, Pieter M. and Jan deLeeuw (1980), "Principal Component Analysis of Three-Mode Data by Means of Alternating Least Squares Algorithms," Psychometrika, 45, 69-97.

Kruskal, Joseph (1980), "An Elegant New/Old Approach to Estimating Path Models (Structural Equation Models) with Unobserved Variables," technical report. Murray Hill, N. J.: Bell Laboratories.

Lyttkens, E. (1973), "The Fixed-Point Method for Estimating Interdependent Systems with the Underlying Model Specification," Journal of the Royal Statistical Society, A136, 353-94.

_____ (1968), "On the Fixed-Point Property of Wold's Iterative Estimation Method for Principal Components," in Multivariate Analysis, P. R. Krishnaiah, ed. New York: Academic Press, 335-50.

Meissner, W. and M. Uhle-Fassing (1981), "PLS--Modeling and Estimation of Politimetric Models," in Systems under Indirect Observation: Causality, Structure, Prediction, K. G. Jöreskog and H. Wold, eds. Amsterdam: North Holland, in press.

Mulaik, Stanley A. (1976), "Comments on the Measurement of Factorial Indeterminacy," Psychometrika, 41, 249-62.

_____ and Roderick P. McDonald (1978), "The Effect of Additional Variables on Factor Indeterminacy in Models with a Single Common Factor," Psychometrika, 43, 177-92.

Noonan, Richard (1980), "School Environments and School Outcomes: An Empirical Comparative Study Using IEA Data," working paper series No. 26, Institute of International Education, University of Stockholm, Sweden.

_____ and Herman Wold (1977), "NIPALS Path Modelling with Latent Variables: Analyzing School Survey Data Using Nonlinear Iterative Partial Least Squares," Scandinavian Journal of Educational Research, 21, 33-61.

_____ (1980), "PLS Path Modelling with Latent Variables: Analyzing School Survey Data Using Partial Least Squares—Part II," Scandinavian Journal of Educational Research, 24, 1-24.

Perreault, William D. and Forrest W. Young (1980), "Alternating Least Squares Optimal Scaling: Analysis of Nonmetric Data in Marketing Research," Journal of Marketing Research, 17 (February), 1-13.

Pijper, de W. M. and W. E. Saris (1981), "The Effect of Identification Restriction on the Test Statistic in Latent Variable Models," in Systems under Indirect Observation: Causality, Structure, Prediction, K. G. Jöreskog and H. Wold, eds. Amsterdam: North Holland, in press.

Sands, R. and F. W. Young (1980), "Component Models for Three-Way Data: An Alternating Least Squares Algorithm with Optimal Scaling Features," Psychometrika, 45, 39-87.

Stapelton, D. C. (1978), "Analyzing Political Participation Data with a MIMIC Model," in Sociological Methodology, San Francisco: Jossey-Bass, 52-74.

Steiger, James H. (1979), "Factor Indeterminacy in the 1930's and the 1970's—Some Interesting Parallels," Psychometrika, 44, 157-67.

Stone, M. (1974), "Cross-Validatory Choice and Assessment of Statistical Predictions," Journal of the Royal Statistical Society, B36, 111-33.

Wold, Herman (1963), "Toward a Verdict on Macroeconomic Simultaneous Equations," in Semaine d'étude sur le rôle de l'analyse économétrique dans la formulation des plans de développement, P. Salviucci, ed., Scripta Varia 28 (Pontifical Academy of Science, Vatican City). Cited in Wold, H. (1981) ed. The Fixed Point Approach in Interdependent Systems. Amsterdam: North Holland.

_____ (1965), "A Fixed-Point Theorem with Econometric Background, I-II," Arkiv för Matematik, 6, 209-40.

_____ (1974), "Causal Flows with Latent Variables: Partings of the Ways in the Light of NIPALS Modeling," European Economic Review, 5, 67-86.

_____ (1975), "Path Models with Latent Variables: The NIPALS Approach," in Quantitative Sociology: International Perspectives on Mathematical and Statistical Model Building, H. M. Blalock et al., eds. New York: Academic Press, 307-57.

_____ (1980a), "Model Construction and Evaluation When Theoretical Knowledge Is Scarce—Theory and Application of Partial Least Squares," in Evaluation of Econometric Models, J. Kmenta and J. G. Ramsey, eds. New York: Academic Press, 47-74.

_____ (1980b), "Soft Modelling: Intermediate between Traditional Model Building and Data Analysis,". Mathematical Statistics, 6, 333-46.

_____ (1980c), "Factors Influencing the Outcome of Economic Sanctions: An Application of Soft Modeling," paper presented at the Fourth World Congress of the Econometric Society, AIX-en-Provence, France.

_____ (1981) ed. The Fixed Point Approach to Interdependent Systems. Amsterdam: North Holland.

Appendix

PLS Algorithms for Models 1, 2, and 3 of the text

The computations of PLS estimation are performed by iterations of explicit simple and multiple regressions. This can easily be accomplished within such computer packages as MIDAS, SAS, and TROLL. Specialized PLS programs are also available. Interested readers may contact the first author at the Graduate School of Business Administration, The University of Michigan, Ann Arbor, Michigan, 48109, concerning these programs.

Model 1 (Figure 2).

Initialize. Set $\eta_1 = y_1 = $ Exit, $\eta_2 = y_2 + y_3 = $ Voice, $\xi = x_1 + x_2 + x_3 + x_4 = $ Concentration.

Loop. Normalize η_1, η_2, and ξ to variance unity.

Regress η_1, η_2 on x_1, x_2, x_3, x_4 separately:

$$\eta_1 = \lambda_{1i} x_i + \varepsilon_{1i} ,$$

$$\eta_2 = \lambda_{2i} x_i + \varepsilon_{2i} , \quad i = 1, 2, 3, 4.$$

Construct

$$\hat{\xi} = \sum_{i=1}^{4} (\lambda_{1i} - \lambda_{2i}) x_i ,$$

the minus sign because η_1 and η_2 have covariances of opposite sign with ξ.

Regress $\hat{\xi}$ on y_2 and y_3 separately:

$$\hat{\xi} = \lambda_{3i} y_i + \varepsilon_{3i} , \quad i = 2, 3.$$

Compute $\eta_2 = \sum_{i=2}^{3} \lambda_{3i} y_i .$

There is no η_1, since this block has only one indicator.

Test. If $\hat{\xi}$ is not equal to ξ or $\hat{\eta}_2$ to η_2, Loop again. Otherwise,

Finish. Regress η_1 and η_2 separately on ξ for the structural parameters γ_1, γ_2.

Model 2   (Figure 6).

Initialize. Set $\eta = y_1 + y_2 + y_3$, $\xi = x_1 + x_2 + x_3 + x_4$.

Loop. Normalize η and ξ to variance unity.

Regress η on x_1, x_2, x_3, x_4 jointly:

$$\eta = \sum_{i=1}^{4} \pi_{\eta i} x_i + \epsilon_\eta .$$

Construct

$$\hat{\xi} = \sum_{i=1}^{4} \pi_{\eta i} x_i .$$

Regress $\hat{\xi}$ on y_1, y_2, y_3 jointly:

$$\hat{\xi} = \sum_{i=1}^{3} \pi_{\xi i} y_i + \epsilon_\xi .$$

Construct $\hat{\eta} = \sum_{i=1}^{3} \pi_{\xi i} y_i .$

Test. If $\hat{\xi}$ is not equal to ξ or $\hat{\eta}$ to η , Loop again.

Otherwise,

Finish. Regress η on ξ for the structural parameter γ.

Model 3   (Figure 7).

Initialize. Set $\eta_1 = y_1 = $ Exit, $\eta_2 = y_2 + y_3 = $ Voice,
$\xi = x_1 + x_2 x_3 + x_4 = $ Concentration.

Loop. Normalize η_1, η_2, and ξ to variance unity.

Regress η_1, η_2 on x_1, x_2, x_3, x_4 jointly:

$$\eta_1 = \sum_{j=1}^{4} \pi_{1j} x_j + \epsilon_1 ,$$

$$\eta_2 = \sum_{j=1}^{4} \pi_{2j} x_j + \epsilon_2 .$$

Compute $\hat{\xi} = \displaystyle\sum_{j=1}^{4} (\pi_{1j} - \pi_{2j}) x_j$.

Regress $\hat{\xi}$ on y_2, y_3 separately:

$$\hat{\xi} = \lambda_{y_2} y_2 + \varepsilon_{y_2} \; ,$$

$$\hat{\xi} = \lambda_{y_2} y_2 + \varepsilon_{y_2} \; .$$

Compute $\hat{\eta}_2 = \displaystyle\sum_{i=2}^{3} \lambda_{y_i} y_i$.

Test. If $\hat{\xi}$ is not equal to ξ or $\hat{\eta}_2$ to η_2, Loop again. Otherwise,

Finish. Regress η_1 and η_2 separately on ξ for the structural parameters γ_1, γ_2.