PACM: A Two-Stage Procedure for Analyzing Structural Models

Donald R. Lehmann
Columbia University

Sunil Gupta
University of California, Los Angeles

An alternative procedure for estimating structural equations models is described. The two-stage procedure, Path Analysis of Covariance Matrix (PACM), separately estimates the measurement and structural models using standard least-squares procedures. PACM was empirically compared to simultaneous maximum likelihood estimation of measurement and structural models using LISREL. PACM produced results similar to LISREL in many cases; it also seems to have advantages when dealing with large-scale problems, model misspecifications, collinearity among indicators, and missing data. *Index terms: causal models, confirmatory factor analysis, LISREL, path analysis, structural equations models.*

Researchers in many fields, such as social sciences, psychology, and marketing, must deal with unobservable or latent constructs (e.g., attitude, intelligence). Structural equations modeling is a research methodology that provides a way to deal with these constructs; moreover, due to its confirmatory nature, it provides a means of testing theories. It is therefore not surprising to see the growing and widespread use of this methodology in several fields (Bagozzi, 1980; Bentler, 1986; Bentler & Chou, 1987).

Structural equations modeling consists of two conceptually distinct submodels—a measurement model that specifies the relationship of observed measures to their posited underlying constructs, and a structural model that specifies the causal relationship of these constructs to each other (Anderson & Gerbing, 1982; Bentler, 1986; Jöreskog & Sörbom, 1985). With full-information estimation methods, such as those provided in the computer programs EQS (Bentler, 1985) or LISREL (Jöreskog & Sörbom, 1979, 1985), the measurement and structural submodels can be estimated simultaneously.

As Anderson and Gerbing (1988) noted, ". . . the ability to do this [estimation] in a one-step analysis approach, however, does not necessarily mean that it is the preferred way to accomplish the model-building task. . . . there is much to gain in theory testing and the assessment of construct validity from separate estimation" (p. 411). They further suggested that "separate assessment of the measurement model and the structural model preclude having good fit of one model compensate for (and potentially mask) poor fit of the other, which can occur with a one-step approach" (p. 421).

In other words, although a two-step estimation procedure may be statistically less efficient than a one-step method (Johnston, 1984), it provides substantial gain in theory testing and assessment of construct

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 13, No. 3, September 1989, pp. 301–321

0146-6216/89/030301-21$2.30

validity (Anderson & Gerbing, 1988). The purpose of this paper is to suggest such a two-step procedure for estimation of structural equations models. Not only does it deliver the advantages listed by Anderson and Gerbing (1988), but it is also conceptually simple and easy to understand. As will be demonstrated, a majority of causal models can be estimated with the proposed methodology by simply using ordinary least squares (OLS). In more complex models, two-stage least squares (2SLS) or three-stage least squares (3SLS) may be required. Although the existing procedures (e.g., LISREL) can be adapted to perform a two-stage estimation, these procedures are really designed for simultaneous estimation of the two submodels. This paper suggests a simpler two-stage estimation procedure.

The proposed procedure, Path Analysis of Covariance Matrix (PACM), essentially decomposes a covariance into its components using a least-squares procedure. In stage 1 of PACM, the measurement model is estimated by decomposing the covariance between observed measures into measure-to-construct coefficients and construct-to-construct covariances. In stage 2, PACM uses construct-to-construct covariances (estimated in stage 1) as input and estimates the structural model.

Having its roots in path analysis and econometrics, PACM is conceptually well grounded as well as easy to understand and estimate. The procedure provides several advantages over simultaneous estimation procedures such as LISREL. For example, it can handle missing covariances, it can deal with covariance matrices that are close to singular, and it provides stable estimates of the measurement model even when the structural model is changed. An added benefit of the two-stage estimation methodology is that it provides useful diagnostics about the measurement and structural models.

Background

Path analysis was developed as a way to estimate causal models. In its early applications (e.g., Blalock, 1964; Duncan, 1970; Land, 1973; Wright, 1934), each construct was measured by a single indicator and the measurement was assumed to be without error. The procedure was initially constrained to situations where the causal model was recursive. The effect of one variable on another was then deduced by "decomposing" the relation between any two variables into "paths" between connected variables so that the relation was the product of the paths linking the two variables (see Alvin & Hauser, 1975). The procedure is closely related to standard econometrics techniques (see Tukey, 1954; Wright, 1960) and has been expanded to deal with unobservable variables (see Hauser & Goldberger, 1971).

Path analysis is very restrictive due to its assumptions of no measurement error and no correlated errors, and its confinement to recursive models. Stimulated by major advances in statistics (e.g., Simon, 1954; Tukey, 1954; Wold, 1956) and biometrics (e.g., Turner & Stevens, 1959; Wright, 1960), sociological methodologists such as Blalock (1961, 1963), Boudon (1965), and Duncan (1966) demonstrated the value of combining the simplicity of path analytic representation with the rigor of specifying equations simultaneously. By the early 1970s, causal modeling was a major sociological research method (Blalock, 1971).

A breakthrough was provided by Jöreskog (1973, 1977) with the introduction of LISREL, a versatile computer program for causal modeling. Several advances have since been made in the area of causal modeling, and increasingly sophisticated estimation procedures have been introduced (e.g., Bentler, 1983, 1985; Browne, 1982, 1984). For example, Bentler (1983) and Mooijaart and Bentler (1985) have outlined an estimation procedure, called asymptotically distribution-free reweighted least squares, that is even more ambitious than any of the procedures currently implemented in such methods as EQS (Bentler, 1985) or LISREL (Jöreskog & Sörbom, 1985).

Although the increasing sophistication in estimation procedures is very useful, the desire to produce simple algorithms to replace more complex ones is widespread. For example, dummy variable regression has generally replaced MONANOVA (Kruskal & Carmone, 1969) as the method for performing conjoint measurement because of its ease of use and the closeness of the solution, even though the assumptions of regression analysis are not strictly met. Similarly, metric scaling procedures have often been used in place of nonmetric procedures even when it is not clear that the data are intervally scaled.

Bentler (1982) proposed a method for performing confirmatory factor analysis based on Hägglund's (1982) work. In summarizing the method, he stated that "the major virtue of noniterative estimates is that they can be obtained quickly and cheaply, making possible the consistent estimation of very large models that are virtually beyond the scope of current methods" (p. 421). This paper is similar in spirit to Bentler's in that it presents a simple least-squares estimation method for path analysis of unobservable constructs with multiple measures.

The PACM Approach

Causal models typically consist of two measurement models and a structural model. A simple example is shown in Figure 1. The measurement models connect the exogenous indicators ($\mathbf{X}$) and endogenous indicators ($\mathbf{Y}$) to the constructs ($\boldsymbol{\xi},\boldsymbol{\eta}$) and estimate the strength of this relationship through the loadings $\boldsymbol{\Lambda}_x$ and $\boldsymbol{\Lambda}_y$. The squared values of these loadings give the reliability of the indicators (in a standardized solution) and hence provide an assessment of how good the measurement instruments are. The structural model assesses the path coefficients ($\mathbf{B},\boldsymbol{\Gamma}$) between the exogenous constructs ($\boldsymbol{\xi}$) and endogenous constructs ($\boldsymbol{\eta}$), thus providing an empirical test of the theory.

Using the LISREL terminology, the measurement and structural models can be represented as follows:

Measurement Model 1: $\mathbf{X} = \boldsymbol{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta}$ (1)

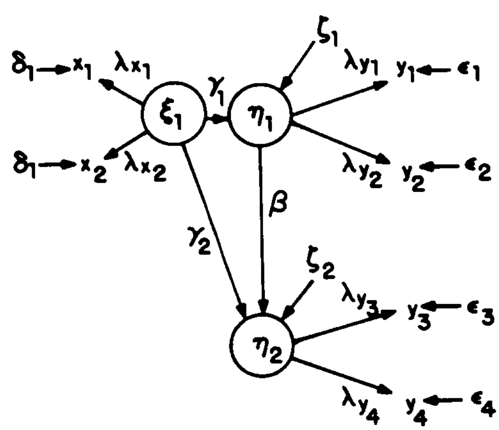
Measurement Model 2: $\mathbf{Y} = \boldsymbol{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon}$ (2)

Structural Model: $\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta}$ (3)

where $\boldsymbol{\delta}$ and $\boldsymbol{\varepsilon}$ represent measurement errors and $\boldsymbol{\zeta}$ represents structural errors.

The major objective of a causal modeling technique (PACM or LISREL) is to use the covariance or correlation matrix of indicators ($\mathbf{X},\mathbf{Y}$) as input and estimate the parameters $\boldsymbol{\Lambda}_x$, $\boldsymbol{\Lambda}_y$, $\mathbf{B}$, and $\boldsymbol{\Gamma}$ (other

Figure 1
A Simple Causal Model



statistics, such as the covariances of errors, can also be estimated).

PACM works in two stages. Stage 1 estimates the measurement models while stage 2 estimates the structural model.

Stage 1

The first stage estimates the pattern coefficients (indicator-to-construct coefficients) and construct-to-construct covariances by relating them to the observed covariances in a trinary product (ignoring correlated measurement errors for the moment):

$$\begin{aligned}\mathrm{Cov}(X_i,Y_j) = {} & \text{(Pattern coefficient between } X_i \text{ and its construct)} \\ & \text{(Pattern coefficient between } Y_j \text{ and its construct)} \\ & \text{(Covariance of the two constructs)}\end{aligned} \tag{4}$$

In order to see why this is appropriate, consider the two measurement models:

$$\mathbf{X} = \mathbf{\Lambda}_x\boldsymbol{\xi} + \boldsymbol{\delta} \tag{5}$$

$$\mathbf{Y} = \mathbf{\Lambda}_y\boldsymbol{\eta} + \boldsymbol{\varepsilon} \tag{6}$$

where it is assumed that each row of $\mathbf{\Lambda}_x$ and $\mathbf{\Lambda}_y$ has a single nonzero element.

For ease of exposition, and without loss of generality, let **X** and **Y** be deviations from the mean. It therefore follows that the covariance between indicators ($\mathbf{\Sigma}_{xx}$) is

$$\mathbf{\Sigma}_{xx} = \mathrm{E}(\mathbf{XX}') = \mathrm{E}[(\mathbf{\Lambda}_x\boldsymbol{\xi}+\boldsymbol{\delta})(\mathbf{\Lambda}_x\boldsymbol{\xi}+\boldsymbol{\delta})'] = \mathbf{\Lambda}_x\mathbf{\Sigma}_{\xi\xi}\mathbf{\Lambda}_x' + \mathbf{\Theta}_\delta \quad , \tag{7}$$

where

$$\mathbf{\Sigma}_{\xi\xi} = \mathrm{E}(\boldsymbol{\xi\xi}') \tag{8}$$

$$\mathbf{\Theta}_\delta = \mathrm{E}(\boldsymbol{\delta\delta}') \quad . \tag{9}$$

Similarly

$$\mathbf{\Sigma}_{yy} = \mathbf{\Lambda}_y\mathbf{\Sigma}_{\eta\eta}\mathbf{\Lambda}_y' + \mathbf{\Theta}_\varepsilon \tag{10}$$

$$\mathbf{\Sigma}_{xy} = \mathbf{\Lambda}_x\mathbf{\Sigma}_{\xi\eta}\mathbf{\Lambda}_y' + \mathbf{\Theta}_{\delta\varepsilon} \quad . \tag{11}$$

In subscript notation, Equation 7 can be expressed as follows:

$$\mathrm{Cov}(X_i,X_j) = \lambda_{x_i}\lambda_{x_j}\mathrm{Cov}(\xi_k,\xi_l) + \theta_{\delta_i\delta_j} \quad , \tag{12}$$

where X_i and X_j are indicators connected to constructs ξ_k and ξ_l respectively. Note that Equations 7 and 12 are equivalent under the assumption that each indicator is connected to one and only one construct. Several researchers have indicated that this is a desired characteristic rather than a limitation (Anderson & Gerbing, 1982, 1988; Hunter & Gerbing, 1982). Anderson and Gerbing (1988) commented on this issue:

> A necessary condition for assigning meaning to estimated constructs is that the measures that are posited as alternative indicators to each construct must be acceptably unidimensional in nature. That is, each set of alternate indicators has only one underlying trait or construct in common (Hattie, 1985; McDonald, 1981). (p. 414)

Assignment of meaning to estimated constructs can be problematic if indicators load on more than one construct (Bagozzi, 1983; Fornell, 1983; Gerbing & Anderson, 1984). Some authors (e.g., Cattell, 1973, 1978) argue that measures tend to be factorially complex and hence may load on more than one construct. In such a case, the problem can be reconstructed so that the construct of interest is conceptualized as a higher order factor (Anderson & Gerbing, 1988; Jöreskog, 1971; Weeks, 1980).

Equation 12 can now be rewritten as

$$\mathrm{Cov}(X_i,X_j) = \lambda_{x_i}\lambda_{x_j}\mathrm{Cov}(\xi_k,\xi_l)M_{\delta_i\delta_j} \tag{13}$$

where

$$M_{\delta_i\delta_j} = 1+\left[\frac{\theta_{\delta_i\delta_j}}{\lambda_{x_i}\lambda_{x_j}\,\mathrm{Cov}(\xi_k,\xi_l)}\right] \quad . \tag{14}$$

Note that if δ_i and δ_j are uncorrelated, then $M_{\delta_i\delta_j} = 1$. Normally, correlated measurement error is expected to arise due to common method effects. A positive correlation between measurement errors will result in $M_{\delta_i\delta_j}$ greater than 1, while a negative correlation will result in $M_{\delta_i\delta_j}$ less than 1.

Equation 13 can now be expressed as

$$\log \mathrm{Cov}(X_i,X_j) = \log \lambda_{x_i}+\log \lambda_{x_j}+\log \mathrm{Cov}(\xi_k,\xi_l)+\log M_{\delta_i\delta_j}+u_{ij} \tag{15}$$

where u_{ij} is the estimation error. Equations 10 and 11 can also be written similarly to Equation 15. Because Equation 15 is in logs and the log of a negative number is undefined, absolute values of the covariances are used. The sign of the covariance can be easily recovered, as is shown in more detail below. In general form, Equation 15 can be written as

$$\log \mathrm{Cov}(\text{Indicators } i,\, j) = \sum_{\forall p} a_p D_p + \sum_{\forall k,l \text{ constructs}} a_{kl} D_{kl} + \sum_{\forall s,t \text{ errors}} a_{st} D_{st} + u_{ij} \tag{16}$$

where

$$a_p = \log \lambda_p \tag{17}$$

$$a_{kl} = \log \mathrm{Cov}(\text{Constructs } k,\, l) \tag{18}$$

$$a_{st} = \log\left\{1+\left[\frac{\mathrm{Cov}(\delta_s,\delta_t)}{\lambda_s\lambda_t\,\mathrm{Cov}(k,l)}\right]\right\} \quad , \tag{19}$$

D_p = dummy variables with a value of 1 if variable $p = i$ or j, 0 otherwise;

D_{kl} = dummy variables with a value of 1 if the k-l construct pair is involved in the relation, 0 otherwise; and

D_{st} = dummy variables with a value of 1 if $(s,t) = (i,j)$ and the errors of i,j are expected to be correlated, 0 otherwise.

Equations 17 through 19 produce λ_p, $\mathrm{Cov}(k,l)$, and $\mathrm{Cov}(\delta_s,\delta_t)$. Equation 16 can be estimated by OLS to obtain $\mathbf{\Lambda}_x$, $\mathbf{\Lambda}_y$, $\mathbf{\Sigma}_{\xi\xi}$, $\mathbf{\Sigma}_{\eta\eta}$, $\mathbf{\Sigma}_{\xi\eta}$, $\mathbf{\Theta}_\delta$, and $\mathbf{\Theta}_\varepsilon$.

Approximate standard errors of parameters can be obtained using a Taylor series expansion (Kendall & Stuart, 1977, p. 247). For example, the approximate variance of λ_p can be found as follows. Rearrangement of Equation 17 yields $\lambda_p = e^{a_p}$, hence

$$\mathrm{Var}(\lambda_p) = \left(\frac{\partial e^{a_p}}{\partial a_p}\right)^2 \mathrm{Var}(a_p) = (e^{a_p})^2\,\mathrm{Var}(a_p) = \lambda_p^2\,\mathrm{Var}(a_p) \quad . \tag{20}$$

OLS estimation gives $\mathrm{Var}(a_p)$. Hence Equation 20 can be used to find the approximate standard error of λ_p.

Note that $a_{st} = 0$ if errors δ_s and δ_t are uncorrelated. Hence, using Equation 16, a nested model test can be performed to see if there is significant improvement in fit when allowing for correlated measurement errors.

Thus the first stage of PACM consists of solving Equation 16 by OLS. Weighting schemes were tried based on the variance of the covariance or correlation [e.g., $\mathrm{Var}(r) = (1 - r^2)^2/n$, where n is the number of observations; Stuart & Ord, 1987, p. 330]. They made almost no difference in the results and are therefore not discussed further. A simple example of a causal model and the associated design matrix for the first stage of PACM analysis is given in the Appendix.

Stage 2

The second stage estimates the structural equations using the first-stage estimates of $\boldsymbol{\Sigma}_{\xi\xi}$, $\boldsymbol{\Sigma}_{\eta\eta}$, and $\boldsymbol{\Sigma}_{\xi\eta}$ from Equation 18 as the covariances among the constructs. The structural equations can be written as

$$\boldsymbol{\eta} = \mathbf{B}\boldsymbol{\eta} + \boldsymbol{\Gamma}\boldsymbol{\xi} + \boldsymbol{\zeta} \quad . \tag{21}$$

This system of equations can be estimated to obtain $\mathbf{B}$, $\boldsymbol{\Gamma}$, and $\boldsymbol{\Psi} = E(\boldsymbol{\zeta}\boldsymbol{\zeta}')$ by OLS, 2SLS, 3SLS, or seemingly unrelated regression (SUR; Zellner, 1962), depending on different scenarios as indicated below.

Scenario 1. If there is a recursive set of structural equations where the error matrix $\boldsymbol{\Psi}$ is diagonal (i.e., errors of endogenous constructs are uncorrelated), and there are no restrictions on parameters across two or more structural equations, then OLS will provide unbiased and efficient estimates (Johnston, 1984, pp. 467–469). Recursive models are those where the residuals from different equations are independent of each other and the matrix of coefficients of the endogenous constructs ($\boldsymbol{\eta}$) is triangular. An example would be

$$\eta_1 = \gamma_1\xi_1 + \gamma_2\xi_2 + \zeta_1 \tag{22}$$

$$\eta_2 = \beta_1\eta_1 + \gamma_3\xi_3 + \zeta_2 \quad , \tag{23}$$

where $E(\zeta_1\zeta_2) = 0$ and there are no parameter restrictions across the two structural equations. Note that ζ_1 and ζ_2 are uncorrelated, η_1 appears on the right side of Equation 23, and η_2 does not appear on the right side of Equation 22. In other words, the matrix of coefficients of $\boldsymbol{\eta}$ is triangular. Hence this is a recursive model. In such a case, each structural equation can be estimated independently by OLS to obtain unbiased and efficient estimates.

Scenario 2. If the set of structural equations is nonrecursive but the error matrix $\boldsymbol{\Psi}$ is diagonal and there are no restrictions on parameters across two or more structural equations, then 2SLS will be appropriate (Johnston, 1984, p. 472; Judge, Griffiths, Hill, Lutkepohl, & Lee, 1985, p. 597). An example of this kind would be

$$\eta_1 = \beta_2\eta_2 + \gamma_1\xi_1 + \zeta_1 \tag{24}$$

$$\eta_2 = \beta_1\eta_1 + \gamma_2\xi_2 + \zeta_2 \quad , \tag{25}$$

where $E(\zeta_1\zeta_2) = 0$. Note that this is a nonrecursive model because η_2 appears on the right side of Equation 24 and η_1 appears on the right side of Equation 25.

Scenario 3. If the parameters across two or more structural equations have some restrictions, then a joint estimation of structural equations is essential. Joint estimation of equations is also warranted if the $\boldsymbol{\Psi}$ matrix is not diagonal, that is, if errors of two or more structural equations are correlated. The nature of estimation methodology in such cases once again depends on whether the system of equations is recursive or nonrecursive. If the system is recursive, then SUR is the appropriate method (Johnston, 1984, p. 338; Judge et al., 1985, p. 800; Zellner, 1962, p. 350). An example of such a situation is

$$\eta_1 = \gamma_1\xi_1 + \gamma_2\xi_2 + \zeta_1 \tag{26}$$

$$\eta_2 = \beta_1\eta_1 + \gamma_3\xi_3 + \zeta_2 \quad , \tag{27}$$

where $E(\zeta_1\zeta_2) \neq 0$. Here the two equations are seemingly unrelated and OLS would appear preferable. However, due to correlation between the errors ζ_1 and ζ_2, OLS estimates are inefficient. To account for this correlation between the errors, a simultaneous estimation of the two equations is warranted. The SUR method achieves this.

Scenario 4. Here the conditions of Scenario 3 hold but the system of equations is nonrecursive. In such a case 3SLS is recommended (Johnston, 1984, p. 486; Judge et al., 1985, p. 599). Conceptually the three stages can be viewed as accounting for error correlations, accounting for nonrecursiveness of the system, and finally obtaining consistent and efficient parameter estimates.

The methods of estimation to be used in different scenarios are summarized in Table 1. Estimation for stage 2 can be done using PROC SYSLIN in the SAS Institute Inc. (1984) statistical package. This program is especially attractive because it provides the flexibility of using any of the four methods of estimation listed above, it provides easy means to impose restrictions on parameters, and it can use covariances of constructs (as estimated from stage 1 of PACM) as input data. An example of stage 2 estimation is provided in the Appendix.

Features of PACM

In this section, some of the features of PACM are discussed and contrasted with the use of LISREL to perform simultaneous maximum likelihood estimation of the measurement and the structural models. PACM has advantages over the typical usage of LISREL when correlations between some variables are missing, when the correlation matrix among the measurement variables is highly collinear, when separate diagnostic information about the measurement and structural models is needed, when the structural model is changed but a change in the estimates of the measurement model is unwarranted, or when dealing with large and complex models.

A key advantage of PACM over LISREL is its ability to handle input matrices where some of the covariances (or correlations) are missing, such as might arise when data from two different sources are combined. In such cases, LISREL cannot give any estimates because it needs to invert the covariance matrix to obtain parameter estimates. PACM, on the other hand, can easily handle this situation by deleting these observations and their covariances from the design matrix in stage 1 of the estimation.

LISREL also has problems in estimation in cases where a covariance matrix has a determinant close to 0 (i.e., a highly collinear matrix). Some researchers (e.g., Jagpal, 1982) have suggested a ridge regression procedure to handle such situations. However, as Jagpal pointed out, the best ridge estimator is difficult to select and is probably beyond current knowledge for more than two constructs. PACM provides a simple solution to this problem. Stage 1 of PACM (which uses the covariance matrix as input) inverts not the covariance matrix but a far less collinear design matrix of dummy variables **D** to obtain the parameter estimates. PACM will have an estimation problem only if the covariance matrix of constructs (as estimated from stage 1 and used as input in stage 2) cannot be inverted. Inability to invert this covariance matrix indicates highly collinear constructs and hence almost no discriminant validity of constructs. Failure of estimation in stage 2 therefore provides a very useful diagnostic to the researcher about the possible lack of discriminant validity of the constructs under consideration.

Although it can be argued that PACM provides less efficient estimates due to its two-stage estimation process, this is more than compensated by the advantages of having two stages. First, PACM does not tend to inflate the path coefficients of the structural equations as is typically found in simultaneous

Table 1
Estimation Methods for Stage 2

Parameter Restrictions Across Structural Equations		Ψ	Model: Recursive	Model: Non-recursive
No	and	Diagonal	OLS	2SLS
Yes	or	Non-diagonal	SUR	3SLS

estimation with LISREL. This is because PACM estimates the measurement and structural model parameters separately and hence measurement errors are not allowed to inflate the structural parameters. That is, PACM avoids the peculiar problem of LISREL where a large measurement error in the model may in fact force LISREL to conclude that a very good structural model exists; such a result may cause a LISREL user to reach an inaccurate conclusion about the validity of a theory.

This problem of interpretational confounding occurs in the presence of misspecification (the usual situation in practice) when measurement and structural models are estimated simultaneously (Burt, 1973, 1976). Burt (1976) elaborated on this point and suggested that interpretational confounding "occurs as the assignment of empirical meaning to an unobserved variable which is other than the meaning assigned to it by an individual a priori to estimating unknown parameters" (p. 4).

Second, the two-stage estimation procedure provides separate goodness-of-fit statistics (R^2) for the measurement and structural models. This is better than a single goodness-of-fit statistic (e.g., χ^2) for the entire system. As Fornell and Larcker (1981) pointed out,

> Because the chi square is a function of the product of the variance of the measurement correlations and the variance of the theory correlations, the results show that changes in the theory model can be compensated for by changes in the measurement model. Thus, an inadequate relationship between constructs (i.e., theory) may be partially offset by measurement properties, and using the chi square statistics for theory testing would be inappropriate. . . . Therefore a testing method is needed that . . . evaluates measurement and theory both individually and in combination so as to detect compensatory effects. (p. 44)

PACM's separate fit statistics for the measurement and structural models can be very useful diagnostic information for the researcher because they provide an indication of which part needs more work—measurement or theory (or both).

Note that PACM uses *all* available information (i.e., the full correlation matrix of **X**-**Y** variables) in stage 1. Only estimation of structural parameters is postponed to stage 2. Further, given the usual assumption of causal models that structural errors ($\boldsymbol{\zeta}$) are uncorrelated with measurement errors ($\boldsymbol{\delta},\boldsymbol{\varepsilon}$), the loss in efficiency in using a two-stage estimation method rather than a single-stage method should be minimal.

In many applications there is a need to modify the structural model without affecting the measurement model. This may be done to improve the theory or the model. In such cases the parameters of the measurement model should be the same regardless of any changes in the structural model. In other words, the reliability of the indicators should not be affected by the changes in the structural model.

PACM achieves this due to its two-stage estimation procedure, whereas LISREL changes the parameters of the measurement model (sometimes dramatically) because of its simultaneous estimation methodology. Anderson and Gerbing (1988) noted this problem: "Interpretational confounding is reflected by marked changes in the estimates of the pattern coefficients when alternate structural models are estimated" (p. 418). They also suggested that "the potential for interpretational confounding is minimized by prior separate estimation of the measurement model because no constraints are placed upon the structural parameters that relate the estimated constructs to each other" (p. 418).

The estimation approach of PACM can be especially useful when dealing with complex models with large numbers of constructs and indicators. In such cases PACM can easily provide good initial solutions by simply using OLS. Although OLS estimates may be biased due to simultaneity of the structural equations, they generally provide good initial solutions. Once the theory has been refined based on initial results, and the researcher has more confidence in the model, more appropriate methods (e.g., 3SLS) can be used to obtain final estimates. PACM should thus be useful for large and complex models where maximum likelihood estimation methods may take too much time or may result in convergence problems.

A relatively undesirable feature of PACM results from the use of logarithms. Because taking logarithms requires positive values, absolute values of covariances are required in Equation 16. The loss of the sign of covariance can be recovered by first ensuring that all indicators correlate positively with the constructs they are supposed to measure. If this is not the case, the indicator variable scale can be reversed to achieve this. This ensures that all λs are positive. Hence a negative covariance between two indicators translates into negative covariance between the constructs to which they connect. Because

$$\mathrm{Cov}(X_i,X_j) = \lambda_{x_i}\lambda_{x_j}\,\mathrm{Cov}(\xi_k,\xi_l)+\mathrm{Cov}(\delta_i,\delta_j) \quad , \tag{28}$$

a negative covariance between X_i and X_j may in fact be due to the negative measurement error covariance. In other words, if a model is misspecified by ignoring the measurement error correlation, the sign of the construct covariances could be misinterpreted, which in turn would lead to a poor model fit.

The omission of a single measurement error correlation, however, is unlikely to lead to an incorrect estimate of the sign of the construct covariance. For example, if there are two constructs, each with three measures, there would be nine correlations on which to base a decision about whether the two constructs are positively or negatively related. Assuming that a correlated measurement error affected only one of the nine pairs, the preponderance of signs (or in most cases the average) would still correctly identify the sign of the covariance between the two constructs.

When a poor fit of the model may suggest the omission of correlated measurement errors, a relatively simple method is available to search for them. By comparing the actual and predicted correlations when measurement errors are assumed to be uncorrelated, pairs with possible positive measurement error correlation (where the actual correlation is much greater than the predicted correlation) and pairs with possible negative measurement error correlation (where the actual correlation is much lower than the predicted correlation) can be identified.

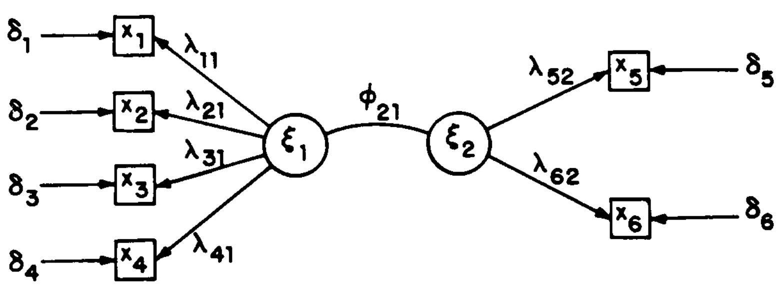
Empirical Comparison of PACM and LISREL

Example 1

This simple example taken from the LISREL VI manual (Jöreskog & Sörbom, 1985, p. III.5) examines the relationship between ability and aspiration. Figure 2 provides the underlying model.

Because no structural model is involved in this example, this problem can be analyzed using only the first stage of PACM. Results of the analysis (Table 2) give a strong indication that the proposed procedure is worth further investigation in that the results from LISREL and PACM are the same to two

Figure 2
Path Diagram for Ability and Aspiration



Adapted from Jöreskog and Sörbom (1985), LISREL VI manual, p. III.2.

Table 2
Initial and MLE LISREL Estimates and PACM OLS
Estimates for the Ability-Aspiration Model

Parameter	LISREL Estimates Initial	LISREL Estimates MLE	PACM OLS Estimates
λ_{11}	.866	.863	.876
λ_{21}	.847	.849	.850
λ_{31}	.801	.805	.796
λ_{41}	.702	.695	.697
λ_{52}	.780	.775	.780
λ_{62}	.923	.929	.923
ϕ_{21}	.664	.666	.662

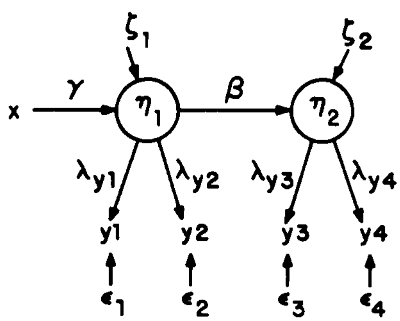
decimal places. Both methods also give the same overall fit for the model (goodness-of-fit index = .994 for LISREL, and R^2 = .998 for PACM). The initial estimates of LISREL are also very close to the other two solutions.

Example 2

The experiment analyzed by Bagozzi (1977) is used as an example that involves both the measurement and structural models. Figure 3 presents the model and Table 3 provides the results as obtained by PACM, by Bagozzi (1977, p. 216), and by LISREL VI. (Bagozzi's results are slightly different from LISREL VI results. This could be because Bagozzi used an earlier version of the LISREL program.) Both the standardized and unstandardized results are provided.

As is clear from Table 3, the results obtained from the two methods are very similar. The standardized PACM stage 1 results indicate a good fit of the measurement model (R^2 = .999). Stage 2 results suggest that the first structural equation (η_1) has a good overall fit (R^2 = .845), but the second structural equation (η_2) does not (R^2 = .368). This implies that a large portion of the variance in the second structural equation is due to the error ζ_2. This can be easily shown as follows: Figure 3 shows that $\eta_2 = \beta\eta_1 + \zeta_2$. Thus $\text{Var}(\eta_2) = \beta^2\ \text{Var}(\eta_1) + \text{Var}(\zeta_2)$. In the standardized solution, the variances of η_1 and η_2 are 1. Hence $\text{Var}(\zeta_2) = 1 - \beta^2 = .632$. Therefore R^2 for this equation is $1 - \text{Var}(\zeta_2) = .368$. This suggests that simply looking at the overall goodness-of-fit index for the model may be misleading.

Figure 3
Credibility Experiment of Bagozzi (1977)



Adapted from Bagozzi (1977), p.216

Table 3
Parameter Estimates for Bagozzi's Credibility Data

	Standardized Results			Unstandardized Results		
	Bagozzi's Estimates	LISREL VI Estimates	PACM Estimates	LISREL Estimates		PACM Estimates
Parameter	MLE	MLE	OLS	Initial	MLE	OLS
λy_1	.722	.775	.776	1.000*	1.000*	1.000*
λy_2	.893	.903	.889	1.182	1.165	1.146
λy_3	.807	.808	.812	1.000*	1.000*	1.000*
λy_4	.955	.953	.949	1.209	1.181	1.169
γ	.918	.907	.919	.000	.703	.713
β	.549	.581	.607	.631	.606	.634

*Fixed parameters.

Table 3 also reports LISREL's initial estimates for the unstandardized solution (LISREL does not give initial estimates for the standardized case). It is interesting to note that although LISREL uses a least-squares approach to obtain its initial estimates, its estimate for the γ parameter is very different. The maximum likelihood approach of LISREL improves on this estimate significantly. PACM uses OLS to obtain parameter estimates close to the maximum likelihood estimates of LISREL.

Example 3

This example involves a slightly more complex model taken from the LISREL VI manual (Jöreskog & Sörbom, 1985, p. III.83). Duncan, Haller, and Portes (1971) proposed this model to study the influence of peers on a person's ambition. They presented a nonrecursive model (Figure 4) and imposed the restriction of $\beta_1 = \beta_2$. Table 4 provides the results obtained from LISREL and PACM (OLS was used in stage 1 and 3SLS in stage 2 of PACM).

Once again the results demonstrate that the PACM estimates are very similar to the LISREL estimates. LISREL reports a good overall fit (goodness-of-fit index = .984). PACM provides more diagnostic information by reporting a good fit for the measurement models (R^2 for stage 1 = .997), but only moderate fit for the structural model (R^2 for stage 2 = .526). The second structural equation fits better than the first equation (R^2 for η_1 = .498, R^2 for η_2 = .597).

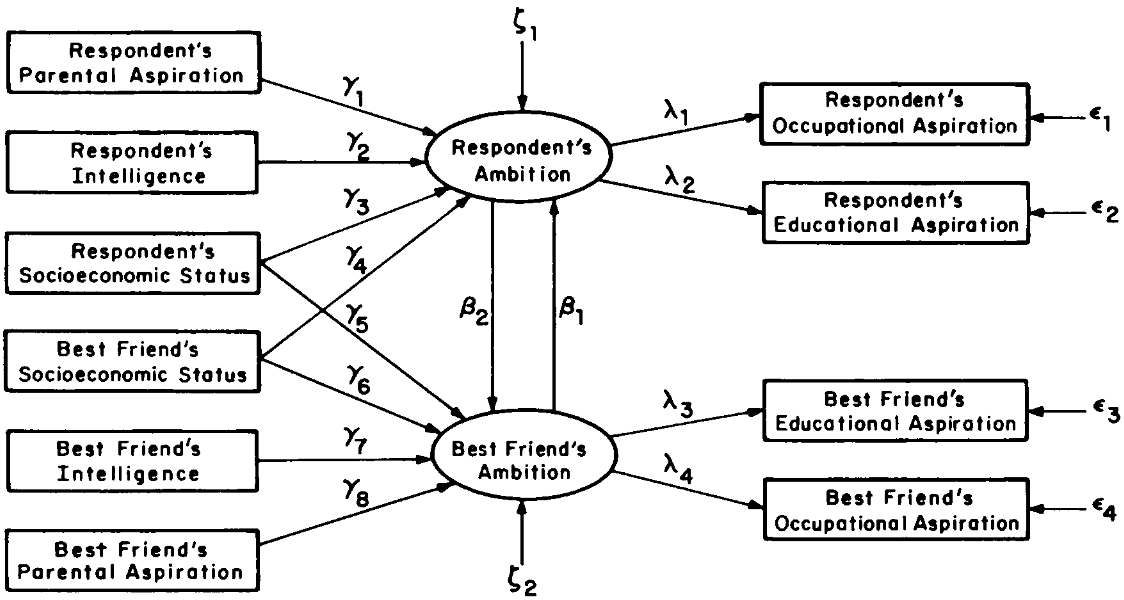
Note that LISREL imposes the restriction of $\beta_1 = \beta_2$ *only* in the unstandardized solution. In the standardized results, β_1 is approximately equal to β_2. PACM, on the other hand, can impose strict equality restrictions in both the standardized and unstandardized results. An unconstrained model was also estimated. An F test between the constrained and the unconstrained model indicated no significant difference—a result identical to that obtained by LISREL using the χ^2 test.

Example 4

As a stronger test of PACM, a fairly complex hypothetical model was taken from the LISREL VI manual (Jöreskog & Sörbom, 1985, p. III.94). This model (see Figure 5) involves several indicators including a complex indicator (X_3) that is connected to two constructs. Further, the model is nonrecursive where the structural errors ζ_1 and ζ_2 are correlated.

Jöreskog and Sörbom used this hypothetical model to demonstrate the usefulness and power of LISREL. They used data from the LISREL VI manual (p. III.97) to simulate a covariance matrix (Σ) of

Figure 4
A Model of Duncan, Haller, and Portes (1971)

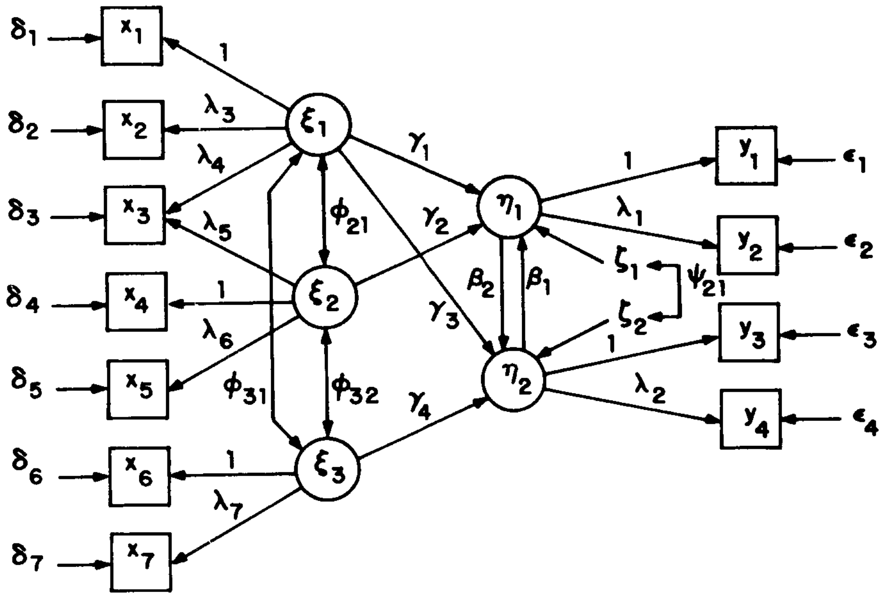


Adapted from Duncan, Heller and Portis (1971)

Table 4
Parameter Estimates for the Model in Figure 4
From a Standardized Solution With the
Restriction β_1=β_2 Imposed

	LISREL MLE		PACM	
Parameter	Estimate	t	Estimate	t
λ_1	.767	-	.790	14.57
λ_2	.813	11.92	.790	14.57
λ_3	.828	13.26	.844	14.57
λ_4	.771	-	.759	14.57
β_1	.181	4.62	.211	3.62
β_2	.179	4.62	.211	3.62
γ_1	.214	4.21	.205	5.22
γ_2	.331	6.05	.327	7.72
γ_3	.228	5.26	.272	6.42
γ_4	.101	1.83	.096	2.09
γ_5	.089	1.74	.080	1.91
γ_6	.283	5.45	.273	7.07
γ_7	.429	8.07	.420	10.51
γ_8	.197	4.22	.191	5.37
ψ_{11}	.478		.484	
ψ_{22}	.385		.391	
θ_{11}	.412		.376	
θ_{22}	.338		.376	
θ_{33}	.314		.289	
θ_{44}	.404		.424	

Figure 5
A Hypothetical Model



Adapted from Jöreskog and Sörbom (1985, p.3.95)

indicators. Then, using **Σ** as the true population covariance matrix, a random sample of 100 observations was generated, having a multivariate normal distribution with mean vector **0** and covariance matrix **Σ**. The resulting sample covariance matrix **S** is given in Table 5.

Jöreskog and Sörbom used this covariance matrix and the LISREL program to recover the true parameters. The same sample covariance matrix and PACM were also used to recover the parameters. The results obtained from the two methods are given in Table 6. The results show the LISREL and PACM estimates to be remarkably similar. Many of LISREL's initial estimates are close to its maximum likelihood estimates and PACM estimates, except for estimates of λ_4, λ_5, β_2, γ_3, and γ_4.

Table 5
Sample Covariance Matrix for Example 4 (N=100)

Variable	y_1	y_2	y_3	y_4	x_1	x_2	x_3	x_4	x_5	x_6	x_7
y_1	3.204										
y_2	2.722	2.629									
y_3	3.198	2.875	4.855								
y_4	3.545	3.202	5.373	6.315							
x_1	.329	.371	−.357	−.471	1.363						
x_2	.559	.592	−.316	−.335	1.271	1.960					
x_3	1.006	1.019	−.438	−.591	1.742	2.276	3.803				
x_4	.468	.456	−.438	−.539	.788	1.043	1.953	1.376			
x_5	.502	.539	−.363	−.425	.838	1.070	2.090	1.189	1.741		
x_6	1.050	.960	1.416	1.714	.474	.694	.655	.071	.104	1.422	
x_7	1.260	1.154	1.923	2.309	.686	.907	.917	.136	.162	1.688	2.684

Table 6
Parameter Estimates for the Hypothetical Model of Figure 5
(Unstandardized Solution)

Parameter	True Parameter Values	LISREL Estimates Initial	LISREL Estimates MLE	LISREL Estimates t	PACM Estimate	PACM t
λ_1	.90	.91	.92	23.0	.98	15.7
λ_2	1.10	1.14	1.14	38.0	1.18	15.7
λ_3	1.30	1.31	1.29	11.7	1.25	15.7
λ_4	.90	1.12	.92	7.7	.85	23.8
λ_5	1.20	.90	1.09	9.1	1.14	33.5
λ_6	1.10	1.05	1.08	13.5	1.06	15.7
λ_7	1.40	1.32	1.44	16.0	1.40	15.7
β_1	.49	.52	.54	0.0	.57	14.8
β_2	.60	1.19	.94	5.2	.89	13.5
γ_1	.40	.35	.21	1.4	.16	1.7
γ_2	.40	.35	.50	3.3	.51	5.4
γ_3	-1.00	-1.31	-1.22	-10.2	-1.13	-23.8
γ_4	1.20	.78	1.00	6.7	.95	14.2

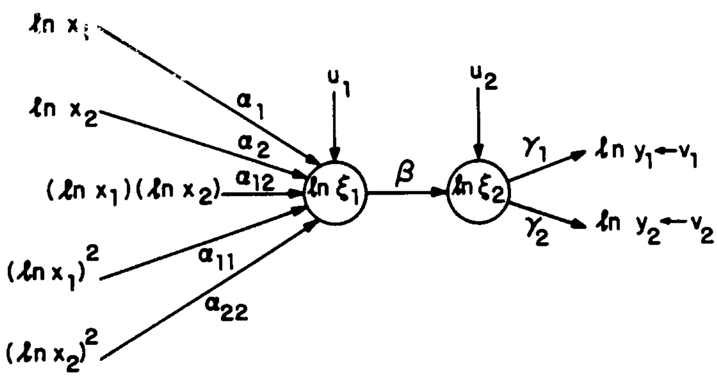
Example 5

Jagpal (1982) posited a multiplicative model relating Awareness (ξ_1) and Preference (ξ_2). The model as presented in Figure 6 is underidentified by the LISREL approach (Jagpal, 1982, p. 433). Jagpal therefore used the partial least-squares (PLS) approach (Wold, 1979).

The correlations (Table 7) among the indicators show substantial multicollinearity. This collinearity caused PLS to produce the alternating sign pattern typical of collinear predictors (even though the *sum* of the α parameters was about 1.00), as is shown in Table 8. By using ridge regression procedures, Jagpal generated estimates of the αs that were more nearly equal (and still summed to 1).

In order to use PACM (or LISREL), the indicators of the first construct (Awareness) were treated as reflexive ($X \leftarrow \xi$) rather than formative ($X \rightarrow \xi$). Thus only the parameters β, γ_1, and γ_2 are strictly

Figure 6
Jagpal's Model



Adapted from Jagpal (1982)

Table 7
Correlations for Jagpal (1982) Study

Variable	lnx_1	lnx_2	$(lnx_1)(lnx_2)$	$(lnx_1)^2$	$(lnx_2)^2$	lny_1	lny_2
lnx_1	1	.8748	.9783	.9756	.8720	.8621	.3917
lnx_2		1	.8991	.9280	.9912	.9935	.3763
$(lnx_1)(lnx_2)$			1	.9964	.9184	.8923	.3760
$(lnx_1)^2$				1	.9429	.9255	.3820
$(lnx_2)^2$					1	.9903	.3603
(lny_1)						1	.3422
(lny_2)							1

comparable. LISREL failed to produce any estimates (even initial estimates) for the correlation matrix of Table 7. It reported that the correlation matrix was not positive definite. Application of PACM to the data led to a simple conclusion (Table 8): The five X variables are all good measures of the first construct and Y_1 is a good measure of the second construct. This is exactly what inspection of the correlation matrix reveals. Because PACM assumes reflexive indicators, the α coefficients are all close to 1. Hence formative indicators would be expected to be all close to .20, essentially the ridge regression result.

The only unexpected result is that β is above 1 (a possible cause for the correlation matrix to be nonpositive definite as reported by LISREL). This results because Y_2 is more correlated with all five indicators of Awareness than it is with Y_1, and Y_1 is indistinguishable from the X variables (with the four correlations, from .8621 to .9935, essentially the same as the correlations among the X variables that range from .8720 to .9964). Hence there is no evidence that the two constructs are distinct, which a β of 1 immediately suggests. When the model was reestimated by constraining $\beta = 1$, the results for αs and γs are essentially the same (Table 8). Thus, in this case, PACM produced plausible results whereas LISREL failed to produce any results and PLS produced results only with great difficulty.

Example 6

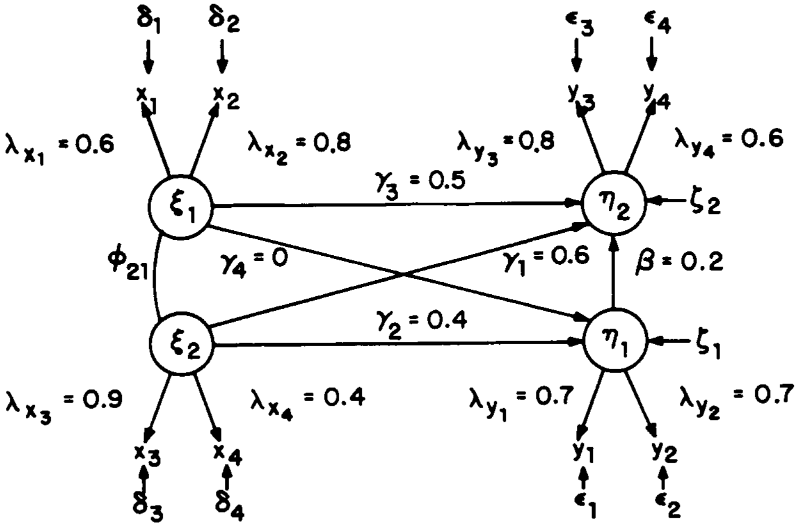
This example used simulated data generated for the model shown in Figure 7. A similar model was proposed by Bagozzi (1980) to study the effect of salespeople's self-esteem (ξ_1) and verbal intelligence

Table 8
Standardized Coefficient Estimates
for Jagpal (1982) Study

Para-	PLS		PACM	
meter	Ordinary	Ridge	Unconstrained	Constrained
α_1	.05	.16	.96	.95
α_2	.81	.27	.97	.97
α_{12}	−.62	.16	.97	.97
α_{11}	.75	.18	.99	.99
α_{22}	.01	.25	.99	.99
β	.96	.94	1.05	1.00*
γ_1	.95	.84	.92	.95
γ_2	.61	.33	.37	.38
R^2			.99	.99

*Fixed parameter.

Figure 7
Simulation Model With True Parameter Values



(ξ_2) on their performance (η_1) and job satisfaction (η_2). Based on the true (hypothetical) parameter values indicated in Figure 7, the correlation matrix of indicators was generated and analyzed by LISREL and PACM. Analysis was performed under several different scenarios.

Case 1: Correct model. In this case, the true structural model—$\eta_1 = \gamma_1\xi_1 + \gamma_2\xi_2 + \zeta_1$ and $\eta_2 = \beta\eta_1 + \gamma_3\xi_1 + \zeta_2$—was estimated. As the results from Table 9 show, both LISREL and PACM recovered the parameters fairly accurately. LISREL seems to have a slight edge over PACM in recovering pattern coefficients, whereas PACM does better at estimating the structural coefficients. Overall the results from the two methodologies are fairly similar.

Case 2: Nested model. A nested version of the structural model was estimated where γ_1 was forced to 0, that is, no link was included between self-esteem (ξ_1) and performance (η_1). Using the same input data as in case 1, LISREL and PACM were used to estimate the parameters. The results are also shown in Table 9.

Comparing the results of the two cases leads to several observations:

1. According to LISREL, the overall fit of the nested model does not suffer dramatically (goodness of fit = .982 for the nested model and 1.0 for the true model). PACM, on the other hand, indicates that although the fit of the measurement model and the second structural equation (η_2) are the same in case 1 and case 2, R^2 for the first structural equation (η_1) drops from .712 to .41. This is precisely what would be expected because in case 2 γ_1 has been forced to 0. This example also demonstrates the usefulness of separate fit statistics compared to an overall goodness-of-fit criterion.
2. Whereas PACM does not change the parameters of the measurement model (λ_x, λ_y, ϕ) across the two cases, LISREL produces some dramatic changes. For example, in case 2, LISREL estimated $\lambda_{x_3} = .581$ instead of .9, $\lambda_{x4} = .307$ instead of .4, and $\phi_{21} = .718$ instead of .4. These results therefore imply that even when the same indicators are used in two cases, their reliability can change dramatically as the structural model changes. For example, the reliability of indicator X_3 is $(.9)^2 = .81$ in case 1 and $(.581)^2 = .34$ in case 2. Clearly this is undesirable. This interpretational confounding in LISREL

Table 9
Standardized Solution for a Simulated Example

Parameter	True Value	True Model LISREL	True Model PACM	Nested Model LISREL	Nested Model PACM	Misspecified Model LISREL	Misspecified Model PACM	Missing Correlation LISREL[d]	Missing Correlation PACM
λ_{x_1}	.6	.600	.600	.600	.600	.553	.600	-	.600
λ_{x_2}	.8	.800	.800	.800	.800	.698	.800	-	.800
λ_{x_3}	.9	.900	.900	.581	.900	.900	.900	-	.897
λ_{x_4}	.4	.400	.400	.307	.400	.400	.400	-	.401
λ_{y_1}	.7	.699	.683	.700	.683	.700	.683	-	.683
λ_{y_2}	.7	.699	.717	.700	.717	.700	.717	-	.717
λ_{y_3}	.8	.800	.781	.800	.781	.799	.781	-	.778
λ_{y_4}	.6	.600	.614	.600	.614	.599	.614	-	.617
ϕ_{21}	.4	.405	.400	.718	.400	.567	.400	-	−.400
γ_1	.6	.593	.600	.000[a]	.000[a]	.91	.760	-	.600
γ_2	.4	.402	.400	1.009	.640	.000[a]	.000[b]	-	.400
γ_3	.5	.562	.534	.584	.534	1.436	.540	-	.534
γ_4	.0	.000[a]	.000[a]	.000[a]	.000[a]	−.084[b]	.026[b]	-	.000[a]
β	.2	.113[b]	.155[c]	.094[b]	.155[c]	−.735[b]	.134[b]	-	.155[c]
GFI		1.000		.982		.982			
R^2 for Stage 1			.999		.999		.999		.999
R^2 for η_1			.712		.410		.578		.712
R^2 for η_2			.435		.435		.435		.435

[a]Fixed parameter.
[b]Not significant at α=.10.
[c]Significant at α=.10.
[d]No estimates produced by LISREL.

estimates is a result of simultaneous estimation of the measurement and structural models. Due to its two-stage estimation procedure, PACM does not suffer from this confounding.

3. Forcing $\gamma_1 = 0$ in the first structural equation has the effect of increasing the γ_2 value. However, note that this inflation of the γ_2 parameter is much more pronounced in LISREL than in PACM. In fact, LISREL provides such an inflated estimate of γ_2 (1.009) that it forces the variance of error ζ_1 to be negative.

Case 3: Misspecified model. In practice, model misspecification is not uncommon. This scenario is represented by positing no link between verbal intelligence (ξ_2) and performance (η_1) (i.e., $\gamma_2 = 0$) and instead allowing a link between verbal intelligence (ξ_2) and job satisfaction (η_2). That is, the true model specifies that $\gamma_2 \neq 0$ and $\gamma_4 = 0$, whereas in the misspecified model $\gamma_2 = 0$ and $\gamma_4 \neq 0$. LISREL and PACM results for this case are presented in Table 9. Comparing these results with the true parameter values and the results of case 1 leads to the following observations:

1. The separate fit statistics of PACM provide more diagnostic information than the overall goodness-of-fit index.
2. Once again LISREL changes the parameter estimates of the measurement model (for example, λ_{x_3} = .553 in case 3 and .6 in case 1). PACM results for the measurement model remain the same.
3. Both models correctly estimated γ_4 to be not significantly different from 0.
4. LISREL tends to inflate the structural parameters far more than PACM. For example, γ_1 (true value =

.6) was estimated as .912 by LISREL and .76 by PACM. Similarly, γ_3 (true value = .5) was estimated as 1.436 by LISREL and .54 by PACM.

Results of cases 2 and 3 indicate that, compared to LISREL, PACM shows a tendency to be more robust to model misspecification.

Case 4: Missing correlation. As previously indicated, PACM can handle missing correlations very easily. LISREL, on the other hand, provides no estimates in such a case. As an example, in the simulated data the correlation of X_3 and Y_3 was assumed to be missing (the true correlation was .236). As the results in Table 9 indicate, PACM estimates show virtually no impact of this missing correlation.

These examples demonstrate that PACM can handle simple and complex models very effectively. In certain cases PACM shows clear advantages over LISREL. Finally, it is very easy to use; in most cases both stages of PACM can be estimated by simply using OLS.

Conclusions

PACM is specifically designed to implement a two-stage estimation procedure in a simpler format than is possible in two-stage adaptations of LISREL. Not only does it produce results comparable to LISREL in most cases, but it also seems to have a number of advantages. Specifically, it seems to be more robust with respect to collinearity among measurement variables, missing data, and model misspecification. Also, it provides separate information on the fit of the measurement model and each equation in the structural model. It also seems better suited to large-scale problems.

PACM can estimate a majority of structural models encountered in research by simply using OLS. Thus, although PACM will not necessarily replace LISREL, it seems to have enough promise to be worth further investigation. Finally, PACM may expand the usage of structural equations to new users by providing them a methodology that is easier to understand and simpler to use.

Appendix

Design Matrix for Stage 1 Analysis of PACM

Construction of a design matrix of dummy variables **D** for stage 1 analysis of PACM can best be illustrated by a simple example. Consider the model in Figure 3. As indicated in Equation 13, correlation between two indicators can be expressed as a trinary product of the path from indicator 1 to construct 1, the path from indicator 2 to construct 2, and the covariance between construct 1 and construct 2. Each correlation provides one observation in the design matrix. The dependent variable is the log of the absolute value of the correlation. The independent variables are derived by assigning a value of 1 to the dummy variables of the three components of the trinary product. All other dummy variables are 0 for that observation. For example, the correlation between y_1 and y_3 will be coded as $D_{y_1} = 1$, $D_{y_3} = 1$, $D_{\eta_1\eta_2} = 1$, and all other dummies 0. Further, the model specifies $X = \xi$ (i.e., $\lambda_x = 1$). Because log λ_x is the coefficient of D_x and $\log(1) = 0$, this restriction can be easily imposed by not including D_x in the design matrix. Thus the correlation between X and Y_1 will be coded as $D_{y_1} = 1$ and $D_{\xi\eta_1} = 1$. The complete design matrix for the above model is given in Table 10.

Note that the dummies $D_{\xi\xi}$, $D_{\eta_1\eta_1}$, and $D_{\eta_2\eta_2}$ are not included in the design matrix. This implies that their coefficients are forced to 0 in the regression equation. Because the coefficient of $D_{\xi\xi}$ is log Cov(ξ,ξ) [i.e., log Var(ξ)], this implies forcing Var(ξ) = 1. Similarly, Var(η_1) and Var(η_2) are forced equal to 1. This produces a standardized solution. It is easy to recode the design matrix (by including $D_{\xi\xi}$, etc.) if an unstandardized solution is desired. Similarly, if some λ_x or λ_y are to be forced to 1, the corresponding dummy D_x or D_y can be eliminated from the design matrix.

Table 10
Design Matrix for Stage 1 Analysis of Figure 3

Dependent Variable: Log of Correlation	Independent Variable						
	D_{y_1}	D_{y_2}	D_{y_3}	D_{y_4}	$D_{\xi\eta_1}$	$D_{\xi\eta_2}$	$D_{\eta_1\eta_2}$
x, y_1	1	0	0	0	1	0	0
x, y_2	0	1	0	0	1	0	0
x, y_3	0	0	1	0	0	1	0
x, y_4	0	0	1	0	0	1	0
y_1, y_2	1	1	0	0	0	0	0
y_1, y_3	1	0	1	0	0	0	1
y_1, y_4	1	0	0	1	0	0	1
y_2, y_3	0	1	1	0	0	0	1
y_2, y_4	0	1	0	1	0	0	1
y_3, y_4	0	0	1	1	0	0	0

This design matrix allows for all the measurement parameters to be unequal. However, the design matrix can be easily changed to constrain parameters to be equal. For example, if it is desired that λ_{y_1} and λ_{y_2} be equal in Figure 3, a new dummy variable $D_{y_1y_2} = D_{y_1} + D_{y_2}$ can be created to replace the dummies D_{y_1} and D_{y_2}. Note that $D_{y_1y_2} = 0$ if neither Y_1 nor Y_2 is involved, $D_{y_1y_2} = 1$ if either Y_1 or Y_2 is involved, and $D_{y_1y_2} = 2$ if both Y_1 and Y_2 are involved. Because the constrained model is nested in the unconstrained model, a standard nested-model test can be used to assess the appropriateness of this constraint.

Stage 2 Analysis

Stage 1 provides the regression parameters of the dummy variables $D_{\xi\eta_1}$, $D_{\xi\eta_2}$, and $D_{\eta_1\eta_2}$. These parameters are the log of the covariance between $\xi\eta_1$, $\xi\eta_2$, and $\eta_1\eta_2$ respectively (see Equation 18). In stage 2 these covariances are read as input (e.g., using SAS) and a least-squares analysis is performed for each structural equation. For Figure 3 this means running two regressions:

$$\eta_1 = \gamma_1\xi_1 + \zeta_1 \tag{A1}$$

$$\eta_2 = \beta\eta_1 + \zeta_2 \quad . \tag{A2}$$

Note that this is a recursive set of equations where the errors ζ_1 and ζ_2 are uncorrelated. Hence separate estimation of these two equations by OLS will provide unbiased and efficient estimates of the parameters.

References

Alvin, D. F., & Hauser, R. M. (1975). The decomposition of effects in path analysis. *American Sociological Review, 40,* 37–47.

Anderson, J. C., & Gerbing, D. W. (1982). Some methods for respecifying measurement models to obtain unidimensional construct measurement. *Journal of Marketing Research, 19,* 453–460.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103,* 411–423.

Bagozzi, R. P. (1977). Structural equation models in experimental research. *Journal of Marketing Research, 14,* 209–226.

Bagozzi, R. P. (1980). *Causal models in marketing.* New York: Wiley.

Bagozzi, R. P. (1983). Issues in the application of covariance structure analysis: A further comment. *Journal of Consumer Research, 9,* 449–450.

Bentler, P. M. (1982). Confirmatory factor analysis via noniterative estimation: A fast, inexpensive method. *Journal of Marketing Research, 19,* 417–424.

Bentler, P. M. (1983). Some contributions to efficient statistics in structural models: Specification and estimation of moment structures. *Psychometrika, 48,* 493–517.

Bentler, P. M. (1985). *Theory and implementation of EQS, a structural equations program.* Los Angeles: BMDP Statistical Software.

Bentler, P. M. (1986). Structural modeling and Psychometrika: An historical perspective on growth and achievements. *Psychometrika, 51,* 35–51.

Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods and Research, 16,* 78–117.

Blalock, H. M. (1961). Correlation and causality: The multivariate case. *Sociological Forces, 39,* 246–251.

Blalock, H. M. (1963). Making causal inferences for unmeasured variables from correlations among indications. *American Journal of Sociology, 69,* 53–62.

Blalock, H. M. (1964). *Causal inferences in nonexperimental research.* Chapel Hill NC: University of North Carolina Press.

Blalock, H. M. (1971). *Causal models in the social sciences.* Chicago: Aldine-Atherton.

Boudon, R. (1965). A method of linear causal analysis: Dependence analysis. *American Sociological Review, 30,* 365–374.

Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Ed.), *Topics in applied multivariate analysis* (pp. 72–141). Cambridge: Cambridge University Press.

Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology, 37,* 62–83.

Burt, R. S. (1973). Confirmatory factor-analytic structures and the theory construction process. *Sociological Methods and Research, 2,* 131–187.

Burt, R. S. (1976). Interpretational confounding of unobserved variables in structural equation models. *Sociological Methods and Research, 5,* 3–52.

Cattell, R. B. (1973). *Personality and mood by questionnaire.* San Francisco: Jossey-Bass.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences.* New York: Plenum.

Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology, 72,* 1–16.

Duncan, O. D. (1970). Partial, partitions, and paths. In E. F. Borgatta & G. W. Bohrnstedt (Eds.), *Sociological methodology 1970* (pp. 38–47). San Francisco: Jossey-Bass.

Duncan, O. D., Haller, A. O., & Portes, A. (1971). Peer influence on aspiration: A reinterpretation. *American Journal of Sociology, 74,* 119–137.

Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. *Journal of Consumer Research, 9,* 443–448.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement errors. *Journal of Marketing Research, 18,* 39–50.

Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research, 11,* 572–580.

Hägglund, G. (1982). Factor analysis by instrumental variables methods. *Psychometrika, 47,* 209–222.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9,* 139–164.

Hauser, R. M., & Goldberger, A. S. (1971). The treatment of unobservable variables in path analysis. In H. L. Costner (Ed.), *Sociological methodology 1971* (pp. 81–117). San Francisco: Jossey-Bass.

Hunter, J. E., & Gerbing, D. W. (1982). Unidimensional measurement, second order factor analysis, and causal models. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior, 4* (pp. 267–299). Greenwich CT: JAI Press.

Jagpal, H. S. (1982). Multicollinearity in structural equation models with unobservable variables. *Journal of Marketing Research, 19,* 431–439.

Johnston, J. (1984). *Econometric methods* (3rd ed.). New York: McGraw-Hill.

Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests, *Psychometrika, 36,* 109–133.

Jöreskog, K. G. (1973). Analysis of covariance structures. In P. R. Krishnaiah (Ed.), *Multivariate analysis-III* (pp. 263–285). New York: Academic Press.

Jöreskog, K. G. (1977). Structural equation models in the social sciences: Specification, estimation and testing. In P. R. Krishnaiah (Ed.), *Applications of statistics* (pp. 265–287). Amsterdam: North-Holland.

Jöreskog, K. G., & Sörbom, D. (1979). *Advances in factor analysis and structural equation models.* Cambridge MA: Abe Books.

Jöreskog, K. G., & Sörbom, D. (1985). *LISREL VI.* Uppsala, Sweden: University of Uppsala.

Judge, G. C., Griffiths, W. E., Hill, R. C., Lutkepohl, H., & Lee, T.-C. (1985). *The theory and practice of econometrics* (2nd ed.). New York: Wiley.

Kendall, M., & Stuart, A. (1977). *The advanced theory of statistics, Vol. 1.* New York: Macmillan.

Kruskal, J. B., & Carmone, F., Jr. (1969). MONANOVA: A FORTRAN-IV program for monotone analysis of variance. *Behavioral Science, 14,* 165–166.

Land, K. C. (1973). Principles of path analysis. In E. Borgatta (Ed.), *Sociological methodology 1973* (pp. 3–37). San Francisco: Jossey-Bass.

McDonald, P. E. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34,* 100–117.

Mooijaart, A., & Bentler, P. M. (1985). The weight matrix in asymptotic distribution-free methods. *British*

Journal of Mathematical and Statistical Psychology, 38, 190–196.

SAS Institute Inc. (1984). *SAS/ETS user's guide, version 5 edition.* Cary NC: Author.

Simon, H. A. (1954). Spurious correlation: A causal interpretation. *Journal of the American Statistical Association, 49,* 467–479.

Stuart, A., & Ord, J. K. (1987). *Kendall's advanced theory of statistics, Vol. 1* (5th ed.). New York: Oxford University Press.

Tukey, J. W. (1954). Causation, regression, and path analysis. In O. K. Kempthorne, T. A. Bancroft, J. W. Gowen, & J. L. Lush (Eds.), *Statistics and mathematics in biology* (pp. 35–66). Ames IA: Iowa State University Press.

Turner, M. E., & Stevens, C. D. (1959). The regression analysis of causal paths. *Biometrics, 15,* 236–258.

Weeks, D. G. (1980). A second-order longitudinal model of ability structure. *Multivariate Behavioral Research, 15,* 353–365.

Wold, H. (1956). Causal inference from observational data: A review of ends and means. *Journal of the Royal Statistical Society, A, 119,* 28–61.

Wold, H. (1979). Estimation and evaluation of models where theoretical knowledge is scarce: An example of partial least squares. In J. Ramsey & J. Kmenta (Eds.), *Evaluation of econometric models.* New York: Academic Press.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics, 5,* 161–215.

Wright, S. (1960). Path coefficients and path regressions: Alternatives or complementary concepts. *Biometrics, 16,* 189–202.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association, 57,* 348–368.

Acknowledgments

The authors thank many colleagues who provided helpful comments on this paper, particularly Greg Carpenter, Claes Fornell, Kamal Gupta, Srinivas Reddy, and Wilfried Vanhonacker. Computer assistance of Alfonso Ang is also acknowledged. Both authors contributed equally to the project.

Author's Address

Send requests for reprints or further information to Donald R. Lehmann, Graduate School of Business, Columbia University, New York NY 10027, U.S.A.