STATISTICAL STRATEGIES FOR SMALL SAMPLE RESEARCH

RICK H. HOYLE

1999

SAGE Publications
International Educational and Professional Publisher
Thousand Oaks London New Delhi

12

Structural Equation Modeling Analysis With Small Samples Using Partial Least Squares

Wynne W. Chin & Peter R. Newsted

STRUCTURAL EQUATION MODELING (SEM) techniques as represented by software such as LISREL, EQS, AMOS, SEPath, CALIS, and RAMONA have become very popular among social scientists in the past decade. Viewed as a coupling of two traditions: an econometric perspective focusing on prediction and a psychometric emphasis that models concepts as latent (unobserved) variables that are indirectly inferred from multiple observed measures (alternately termed as indicators or manifest variables), SEM essentially offers social scientists the ability to perform path-analytic modeling with latent variables. This approach, in turn, has led some to describe it as an example of "a second generation of multivariate analysis" (Fornell, 1987).

The primary advantage that SEM-based procedures have over first-generation techniques such as principal components analysis, factor analysis, discriminant analysis, or multiple re-

gression is the greater flexibility that a researcher has for the interplay of theory and data. Overall, SEM-based approaches provide researchers with the flexibility to perform the following:

- Model relationships among multiple predictor and criterion variables.
- Construct unobservable latent variables.
- Model errors in measurements for observed variables.
- Statistically test a priori substantive/theoretical and measurement assumptions against empirical data (i.e., confirmatory analysis).

Essentially, second-generation multivariate techniques such as SEM involve generalizations and extensions of first-generation procedures. Applying certain constraints or assumptions on one or more particular second-generation techniques would result in a first-generation procedure with correspondingly less flexibility in modeling theory with data.

Yet, a number of factors, including sample size, tend to preclude the use of the predominate technique of covariance-based SEM. The objective of this chapter is to elaborate on a lesser known approach for SEM analysis called partial least squares (PLS). Rather than view PLS as a competing method, it will be shown as complementary in terms of research objectives, data conditions, and modeling. In addition to considering sample size, this chapter begins by outlining the factors that should be considered in choosing one approach over the other. The specific algorithm is then covered in a nontechnical fashion. Finally, the results of a Monte Carlo simulation are presented to see how PLS performs under varying levels of sample size and model complexity with the extreme case being 21 latent variables, 672 indicators, and a sample size of 20.

Contrasting Partial Least Squares and Covariance-Based Structural Equation Modeling

The *covariance-based approach* for SEM dates back to the original development by Jöreskog (1973), Keesling (1972), and Wiley

(1973). Its widespread popularity is due in a large part to the availability of the LISREL III program developed by Jöreskog and Sörbom in the mid-1970s and subsequent updates (see Preface, Jöreskog and Sörbom, 1989). Typically, using a maximum-likelihood (ML) function, covariance-based SEM attempts to minimize the difference between the sample covariances and those predicted by the theoretical model (i.e., $\Sigma - \Sigma(\Theta)$). Therefore, the parameter estimation process attempts to reproduce the covariance matrix of the observed measures.

Along with the benefits discussed earlier, the use of covariance-based SEM (hereafter CBSEM) involves constraints in the form of parametric assumptions, sample size, model complexity, identification, and factor indeterminacy. In order to use this approach, it is assumed that the observed variables follow a specific multivariate distribution (normality in the case of the ML function) and that observations are independent of one another. Possibly more critical is the sample size requirement, which is often beyond the range of researchers. Small samples that are not "asymptotic" in characteristics can lead to poor parameter estimates and model test statistics (Chou & Bentler, 1995; Hu & Bentler, 1995). In fact, inadmissible solutions in the form of negative variances and out-of-range covariances often occur as sample size decreases (e.g., Anderson & Gerbing, 1984; Boomsma, 1983; Dillon, Kumar, & Mulani, 1987; Gerbing & Anderson, 1987; MacCallum, 1986; van Driel, 1978). When the latent variates are dependent, fit indices tend to overreject models at sample sizes of 250 or less (Hu & Bentler, 1995, p. 95).

Equally critical with small sample sizes is the potential for Type II error, whereby a poor model can still falsely achieve adequate model fit. According to MacCallum, Browne, and Sugawara (1996), sample sizes less than 200 are inadequate to achieve the standard.80 level for a test of close fit when models have degrees of freedom at 55 or lower. Using their alternative test of "not close fit," models at samples sizes of 200 need to be even more restrictive (i.e., degrees of freedom greater than 80). Furthermore, under exploratory conditions with small-to-moderate sample sizes (i.e., 100 to 400), MacCallum (1986) demonstrated that final models derived via post hoc modifications should not be trusted.

Complex models also can be problematic relative to fit indices and computation. As the degrees of freedom increase with increasing number of indicators and latent variables, various model fit indices tend to be positively biased relative to simpler models (Mulaik et al., 1989). Pragmatically, current software packages begin to slow down or possibly fail to run as the number of indicators approaches 50 or 100 (approximately 20 to 30 in the case of distribution-free estimation).

Covariance-based SEM analysis typically requires indicators in a *reflective* mode. Under this condition, indicators are viewed as being influenced or affected by the underlying latent variable. Yet, an alternative conceptualization has the indicators in a *formative* mode. In this situation, indicators are viewed as causing rather than being caused by the latent variable. According to Bollen and Lennox (1991), such formative indicators "do not conform to the classical test theory or factor analysis models that treat indicators as effects of a construct" (p. 305). As an example, Cohen, Cohen, Teresi, Marchi, and Velez (1990) used the latent variable of socioeconomic status (SES) with education, occupational prestige, and income as indicators. In this instance, the indicators determine an individual's SES. If one of the indicators increases, the other ones need not do so. Yet, an increase in any one indicator (e.g., income) will lead to an increase in the latent variable SES. Another example of formative indicators would be job loss, divorce, recent accident, and death in the family for the latent variable life stress. Although a CBSEM analysis generally requires all latent variables to have reflective indicators, researchers may unknowingly incorporate formative ones. As Cohen et al. (1990, Table 1, p. 186) showed in a survey of 15 articles that performed CBSEM analysis, a sizable number of latent variables were indeed inappropriately modeled, treating formative indicators as reflective. Finally, as MacCallum and Browne (1993) demonstrated, any attempts to model formative indicators in a CBSEM analysis can lead to identification problems, implied covariances of zero among some indicators, and/or the existence of equivalent models. Although these problems can be managed, MacCallum and Browne (1993) argue that this would involve

"altering the original model in terms of its substantive meaning or parsimony, or both" (p. 540).

There is also an inherent indeterminacy in the CBSEM procedure. In other words, case values for the latent variables cannot be obtained in the process. Thus, it is not possible to estimate scores for the underlying latent variables in order to predict the observed indicators. In fact, an infinite set of possible scores can be created that are not only consistent with the parameter estimates, but need not be correlated. This may or may not be viewed as problematic depending on the objectives (see Maraun, 1996, and the two rounds of commentary). The CBSEM approach is ideal if the goal is to obtain population parameter estimates for explaining covariances with the assumption that the underlying model is correct. However, this procedure was not developed for predictive purposes where the researcher desires parameter estimates (i.e., weights for each individual indicator) in order to create latent variable scores that can be used to predict its own indicators or other latent variables. As will be explained in more detail later, this is the primary goal of partial least squares: predicting the variances of latent and manifest variables.

Identification problems pertaining to a unique set of estimates can also occur either algebraically or empirically under conditions where the number of indicators per construct is low, the correlation among factors is zero, or both (Rindskopf, 1984). Thus, it is generally necessary to have three or more indicators per latent variable in order to avoid identification problems in CBSEM analysis.

A final issue to consider when using a CBSEM procedure is the role theory plays in the analyses of the data. At this point, very little has been done to examine the influence of model misspecification on parameter estimates. Being a full information approach, the parameter estimates in one part of a model (e.g., loadings for one latent variable) may be unduly influenced by misspecifications in some other part of the model (e.g., missing structural paths or poor indicators for another construct). This is less of an issue for a large percentage of CBSEM studies, which follow a confirmatory mode whereby the analyses are done under strong theory and use measures that were developed from

prior studies—typically via a series of exploratory factor analyses. When the theory is still relatively tentative or the measures for each latent variable are new, however, greater emphasis may need to be placed on the data relative to the theory.

In the case of CBSEM, it has been argued that theory is given more influence in estimating parameters as opposed to the partial least squares approach to SEM analysis. For example, if a correlated two-factor model was specified, the estimated correlation between the two abstract latent variables tends to go up as the correlations among observed indicators go down. Given a specific theoretical model, the only logical deduction for situations with low observed correlations is that there is a large amount of random error/noise affecting each indicator. According to Fornell (1989),

> ... consistent with the specification of reflective indicators, the abstract model specification plays a large role in determining the results; almost to the point that it "overrides" the data .... Thus, it is here that the researcher must make a decision about the relative weight that should be given to data vs. theory .... Fortunately, alternatives to covariance structure analysis are available when the analyst is unwilling to depart too far from the data and wants to obtain a different balance between theory and observation. One such alternative is partial least squares (PLS) developed by Herman Wold. (pp. 165–166)

As an alternative to covariance-based SEM analysis, the variance-based approach of PLS shifts the orientation from causal model/theory testing to component-based predictive modeling. Rather than focusing on building models that are meant to explain the covariances of all the observed indicators, the objective of PLS is prediction. As such, latent variables are defined as the sum of their respective indicators. The PLS algorithm attempts to obtain the best weight estimates for each block of indicators corresponding to each latent variable. The resulting component score for each latent variable based on the estimated

indicator weights maximizes variance explained for dependent variables (i.e., latent, observed, or both).

Partial least squares can be a powerful method of analysis because of the minimal demands on measurement scales (i.e., categorical to ratio level indicators can be used in the same model), sample size, and residual distributions (Wold, 1985). Although PLS can be used for theory confirmation, it can also be used to suggest where relationships might or might not exist and to suggest propositions for later testing. Being closer to the data and a limited estimation procedure, misspecifications in one part of a model have less influence on the parameter estimates in other parts of the model. Compared to the better known CBSEM, the component-based PLS avoids two serious problems: inadmissible solutions and factor indeterminacy (Fornell & Bookstein, 1982). Because the iterative algorithm performed in a PLS analysis generally consists of a series of ordinary least squares analyses, identification is not a problem for recursive models, nor does it presume any distributional form for measured variables. Furthermore, the computational efficiency of the algorithm lends itself to estimating large complex models on the order of hundreds of latent variables and thousands of indicators. The utility of the PLS method has been documented elsewhere (Falk & Miller, 1992, p. xi) as possibly more appropriate for a large percentage of the studies and data sets typically used among researchers.

In summary, if the hypothesized structural and measurement model is correct in the sense of explaining the covariation of all the indicators and the data/sample size conditions are met, the covariance-based procedure provides optimal estimates of the model parameters. It is ideal for model confirmation and estimation of the "true" underlying population parameters. However, depending on the researcher's objectives and epistemic view of data to theory, properties of the data at hand, or level of theoretical knowledge and measurement development, the PLS approach can be argued to be more suitable. Table 1 provides a summary of the key differences between PLS and CBSEM.

Table 1 Comparison of Partial Least Squares and Covariance-Based Structural Equation Modeling

Criterion	*PLS*	*CBSEM*
Objective:	Prediction oriented	Parameter oriented
Approach:	Variance based	Covariance based
Assumptions:	Predictor specification (nonparametric)	Typically multivariate normal distribution and independent observations (parametric)
Parameter estimates:	Consistent as indicators and sample size increase (i.e., consistency at large)	Consistent
Latent variable scores:	Explicitly estimated	Indeterminate
Epistemic relationship between a latent variable and its measures:	Can be modeled in either formative or reflective mode	Typically only with reflective indicators
Implications:	Optimal for prediction accuracy	Optimal for parameter accuracy
Model complexity:	Large complexity (e.g., 100 constructs and 1,000 indicators)	Small to moderate complexity (e.g., less than 100 indicators)
Sample size:	Power analysis based on the portion of the model with the largest number of predictors. Minimal recommendations range from 30 to 100 cases.	Ideally based on power analysis of specific model—minimal recommendations range from 200 to 800.

The Standard Partial Least Squares Algorithm

The basic PLS design was completed in 1977 (see Wold, 1982, p. 35) and has subsequently been extended in various ways. Lohmöller (1984, 1989) covered various inner weighting schemes. Wold (1982) discussed nonlinearities among latent variables, whereas Hui (1978, 1982) used a fixed-point PLS method to model nonrecursive (i.e., interdependent) relationships. In accord with this chapter's objective of providing a simple introduction to the PLS method, the following discussion is restricted to the basic design involving recursive models.

As discussed earlier, the objective of PLS is to help the researcher obtain determinate values of latent variables for prediction. The formal model (to be described later) explicitly defines latent variables as linear aggregates of their observed indicators. The weight estimates to create the latent variable component scores are obtained based on how the inner (i.e., structural) and outer (i.e., measurement) models are specified. As a result, the residual variances of dependent variables (both latent and observed variables) are minimized.

The parameter estimates obtained via PLS can be viewed as falling into three categories. The first category is the weight estimates, which are used to create the latent variable scores. The second reflects the path estimates connecting latent variables and between latent variables and their respective block of indicators (i.e., loadings). The third category pertains to the means and location parameters (i.e., regression constants) for indicators and latent variables. In order to come up with these three sets of parameter estimates, the PLS algorithm follows a three-stage process with each stage used to obtain each set of estimates, respectively. Therefore, the first stage results in obtaining the weight estimates. The second stage provides estimates for the inner model (i.e., structural relations among latent variables) and outer model (i.e., reflective or formative measurement paths). And the third stage yields the means and location estimates. In the first two stages, the indicators and latent variables are treated as deviations from their means. In the third stage,

should the researcher wish to obtain estimates based on the original data metrics, the weight and path estimates from the previous two stages are used for calculating the means and location parameters.

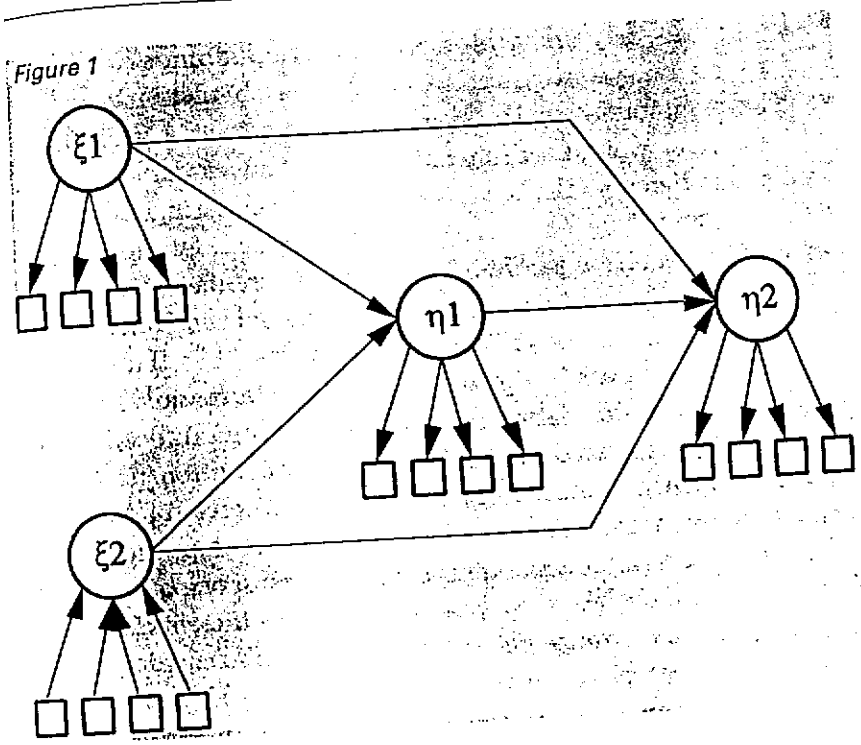
Stage 1 represents the heart of the PLS algorithm. It consists of an iterative procedure that almost always converges to a stable set of weight estimates. Essentially, component score estimates for each latent variable are obtained in two ways. The outside approximation represents a weighted aggregate of its own indicators. The inside approximation refers to a weighted aggregate of other component scores that are related to it in the theoretical model. During each iteration, the inner model estimates are used to obtain the outside approximation weights, whereas the outer model estimates are used to obtain the inside approximation weights. The procedure stops when the percentage change of each outside approximation weight relative to the previous round is less than 0.001. Thus, both the theoretical and the measurement portions of the model contribute in the estimation process. This is contrasted to a two-step approach where the measurement model is derived first before it is used at the structural level (see Fornell & Yi, 1992, for further discussion relating the one- and two-step approach toward modeling).

Multiblock Example

To illustrate the stage 1 process PLS goes through, let's look at the model provided in Figure 1. As depicted, the measures are partitioned into four blocks (two exogenous ξs and two endogenous ηs). As a starting point, the algorithm does an initial outside approximation estimation of the latent variables (LVs) by summing the indicators in each block with equal weights. The weights, in each iteration, are scaled to obtain unit variance for the latent variable scores over the N cases in the sample. Using the estimated scores for each LV as given, an inside approximation estimate of the LVs is performed.

There have been three primary "inside approximation" weighting schemes developed thus far for combining "neighboring" LVs to obtain an estimate for a specific LV: centroid, factor, and path

Figure 1



weighting. Although each weighting scheme follows a particular logic, it has been noted that the choice tends to have little influence on the results: .005 or less for structural paths and .05 or less for measurement paths (Noonan & Wold, 1982).

The centroid weighting scheme was the original procedure used by Wold. It only considers the sign of the correlations between the LV and the neighboring LVs. The strength of the correlations and the direction of the structural model are not taken into account. It is computationally simple because the resulting estimates are obtained by adding up all connected LVs with either weights of +1 or −1 depending on the sign of their correlation, which leads each LV to become similar to the centroid factor (Lohmöller, 1989). Thus, if a structurally linked LV is correlated −.30, the weight assigned to it would be −1. This approach is considered advantageous (relative to the path weighting scheme) when the LV correlation matrix is singular because the weights are based only on the bivariate correlations among component scores. A disadvantage arises when an LV correlation is close to

zero and thus may oscillate during iterations from a small positive to a small negative and back. Under this situation, these values are magnified by the corresponding +1 and −1 weights.

The factor weighting scheme, therefore, uses the correlation coefficients between the focal LV and its neighboring LVs as the weights. The LV becomes the "principal component" of its neighboring LVs. According to Lohmöller (1989), the factor weighting scheme maximizes the variance of the principal component of the LVs when the number of LVs goes to infinity.

Finally, the path weighting scheme differentially weights neighboring LVs depending on whether they are antecedents or consequents of the focal LV. This scheme, thus, attempts to produce a component that ideally can both be predicted and at the same time be a good predictor for subsequent dependent variables. To do this, all independent variables influencing the target LV are weighted by the multiple-regression coefficients, whereas all dependent LVs are weighted by the correlation coefficients. In a sense, the focal LV becomes the best mediating LV between the source and target LVs. As the only procedure among the three that takes into account the directionality of the structural model, the path weighting scheme is often used for models with hypothesized causal relations. If, on the other hand, no propositions are made regarding the associations among the LVs, the factor weighting scheme would be the logical choice.

If a factor weighting scheme were used in this example, the inside approximation estimate for ξ_1 would be the sum of the outside approximation estimates for η_1 and η_2 weighted by their respective correlation coefficients with ξ_1. ξ_2 is not included in the estimate because there is no link between it and ξ_1. η_1, on the other hand, is a weighted estimate of the other three LVs because it has structural paths with all three.

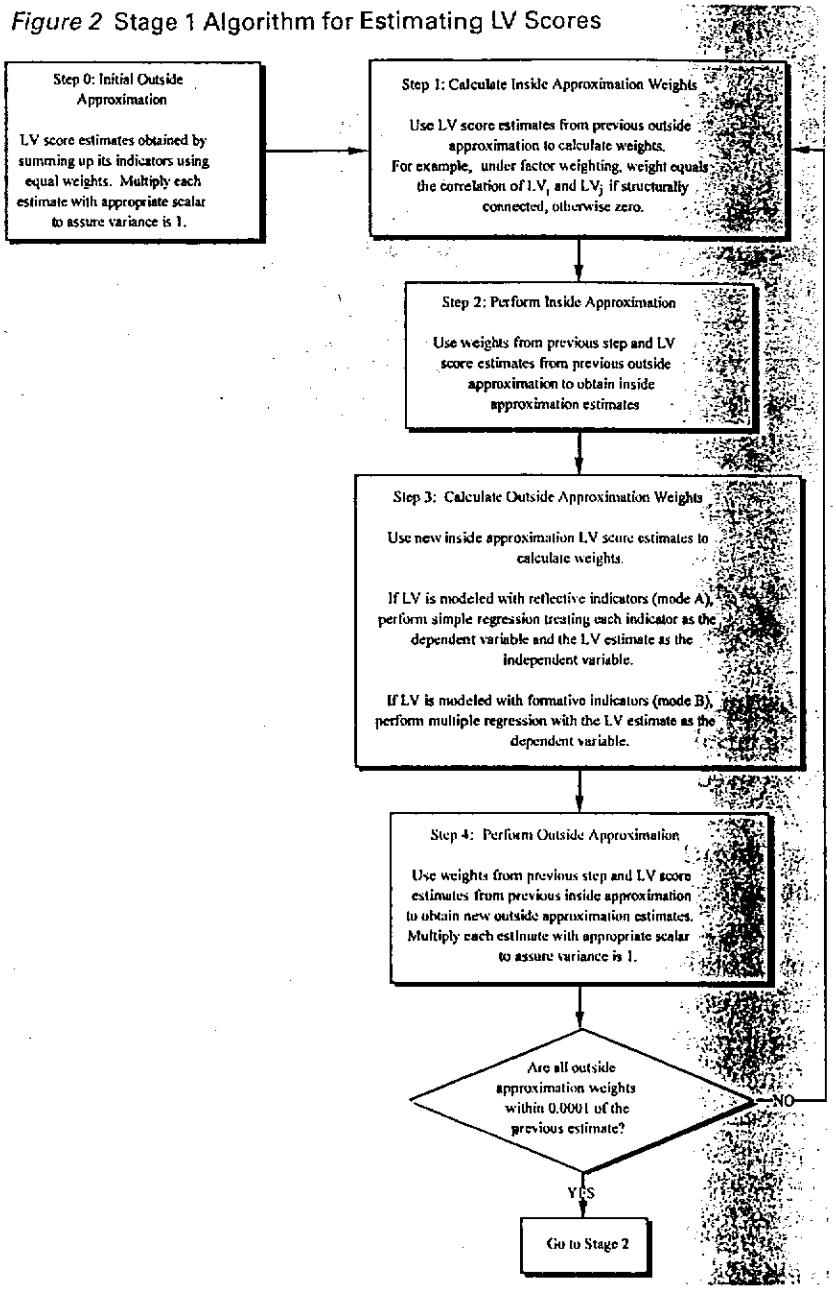
Under a path weighting scheme, there are three different inside approximation situations. As both ξ_1 and ξ_2 are exogenous constructs, they are weighted by correlation coefficients. So, for example, the estimate for ξ_1 would follow the same procedure as under the factor weighting scheme. In the case of η_2, being a pure endogenous (dependent) variable, a multiple regression is performed with the outside approximation estimates for ξ_1, ξ_2,

and η_1 on the outside approximation estimate of η_2. The multiple-regression coefficients are then used as the weights for combining the outside approximation estimates of ξ_1, ξ_2, and η_1 to obtain the inside approximation estimate for η_2. Finally, the estimate for η_1 uses both multiple-regression coefficients and correlation coefficients as weights. η_1 is regressed on ξ_1 and ξ_2 to obtain weights for ξ_1 and ξ_2 and the simple correlation between η_1 on η_2 is taken for the weight of η_2.

Given the LV estimates from the inside approximation, a new set of weights from a new round of outside approximations can be obtained. Taking the inside approximation scores as fixed, either simple or multiple regression is performed depending on whether the block of indicators is in a reflective mode (termed mode A) or a formative mode (termed mode B) mode. Because ξ_1, η_2, and η_1 are modeled as mode A with arrows directed toward the indicators, each indicator in each block would be individually regressed on its respective LV estimate (i.e., inside approximation score). In the case of ξ_2, being in mode B with arrows directed toward the LV, a multiple regression of the estimate of ξ_2 on its indicators is performed. The simple- or multiple-regressions coefficients are then used as new weights for an outside approximation of each LV. Figure 2 depicts the logical flow of this iterative process.

Once the latent variable scores from stage 1 are estimated, the path relations are immediately estimated by ordinary least squares regression in stage 2. Each dependent variable in the model (either endogenous LVs or indicators in a reflective mode) is regressed on its respective independent variables (i.e., other LVs or indicators in a formative mode). When the final paths are estimated in stage 2 and it makes substantive sense (e.g., differences in the means, scale, and variances are meaningful), the means and location parameters for the indicators and LVs are estimated in stage 3. To do this, the means for each indicator are first calculated based on the original data. Then, using the weights derived in stage 1, the means for each LV are calculated. Given means for the LVs and path estimates from stage 2, the location parameter for each dependent LV is simply calculated as the difference between the just obtained mean and the systematic part accounted for by the independent LVs that influence it.

Figure 2 Stage 1 Algorithm for Estimating LV Scores



Likewise, the location parameter for a reflective indicator is simply the difference between its mean and the estimate based on its underlying LV and path loading.

Formal Specification of the Partial Least Squares Model

Having covered the PLS estimation procedure, the formal model specification that guides the process is now presented. As noted, the latent variable path models in PLS consist of three sets of relations: (1) the inner model, which specifies the relationships between LVs; (2) the outer model, which specifies the relationships between LVs and their associated observed or manifest variables (MVs); and (3) the weight relations upon which case values for the LVs can be estimated. Without loss of generality, it can be assumed that LVs and MVs are scaled to zero means and unit variances so that location parameters (i.e., constant parameter terms) can be eliminated in the following equations.

Inner Model

The inner model (also referred to as the inner relations, structural model, substantive theory) depicts the relationship among latent variables based on substantive theory:

$$\eta = \beta_0 + \beta\eta + \Gamma\xi + \zeta, \tag{1}$$

where η represents the vector of endogenous (i.e., dependent) latent variables, ξ is a vector of the exogenous latent variables, and ζ is the vector of residual variables (i.e., unexplained variance).

Because the basic PLS design assumes recursive relations (i.e., one-way arrows) among LVs, each dependent latent variable η_j in this often termed "causal chain system" of LVs can be specified as follows:

$$\eta_j = \sum_i \beta_{ji}\eta_i + \sum_h \gamma_{jh}\xi_h + \zeta_j, \tag{2}$$

where β_{ji} and γ_{jh} are the path coefficients linking the predictor endogenous and exogenous latent variables ξ and η over the range specified by the indices i and h, and ζ_j is the inner residual variable.

The inner model (Equation 1) is subject to predictor specification (Wold, 1988):

$$E(\eta_j \mid \forall\eta_i, \xi_h) = \sum_i \beta_{ji}\eta_i + \sum_h \gamma_{jh}\xi_h.$$

Thus, it is assumed that each LV is a linear function of its predictors and that there are no linear relationship between the predictors and the residual,

$$E(\zeta_j \mid \forall\eta_i, \xi_h) = 0 \qquad \text{and} \qquad \text{Cov}(\zeta_j, \eta_i) = \text{Cov}(\zeta_j, \xi_h) = 0$$

for the indices i and h ranging over all predictors.

The structural form of the inner model can be also be written in reduced form by subtracting $\beta\eta$ from both sides of Equation 1 and premultiplying by $(I - \beta)^{-1}$ yielding:

$$\begin{aligned}\eta &= (I - \beta)^{-1}\Gamma\xi + (I - \beta)^{-1}\zeta \\ &= \beta^*\xi + \zeta^*,\end{aligned}$$

where β^* represents the total effect of the exogenous latent variable ξ.

To make predictor specification possible for both the structural and the reduced forms, the assumption is made that $E(\zeta_j^* \mid \forall\xi_h) = 0$ for all j endogenous LVs as they relate to the exogenous LVs impacting in the first j relations given in the structural form given in Equation 2.

Outer Model

The outer model (also referred to as outer relations or measurement model) defines how each block of indicators relates to its latent variable. The MVs are partitioned into non-overlapping

blocks. For those blocks with reflective indicators, the relationships can be defined as

$$x = \Lambda_x \xi + \varepsilon_x,$$
$$y = \Lambda_y \eta + \varepsilon_y,$$

where x and y are the MVs for the exogenous and endogenous LVs ξ and η, respectively. Λ_x and Λ_y are the loadings matrices representing simple regression coefficients connecting the LV and their measures. The residuals for the measures ε_x and ε_y, in turn, can be interpreted as measurement errors or noise.

Predictor specification, as in the case for the inner model, is assumed to hold for the outer model in reflective mode as follows:

$$E[x \mid \xi] = \Lambda_x \xi, \qquad (3)$$
$$E[y \mid \eta] = \Lambda_y \eta.$$

For those blocks in a formative mode, the relationship is defined as:

$$\xi = \Pi_\xi x + \delta_\xi,$$
$$\eta = \Pi_\eta y + \delta_\eta,$$

where ξ, η, x, and y are the same as those used in Equation 3. Π_x and Π_y are the multiple-regression coefficients for the LV on its block of indicators and δ_x and δ_y are the corresponding residuals from the regressions. Predictor specification is also in effect as

$$E[\xi \mid x] = \Pi_\xi x,$$
$$E[\eta \mid y] = \Pi_\eta y.$$

As opposed to the weight relations to be discussed next, the formative specification for outer relations refer to the MV and the true LV. This, in turn, provides the basis for the manner in which the weights are determined within the PLS estimation algorithm estimating the LV.

Weight Relations

Although the inner and outer models provide the specifications that are followed in the PLS estimation algorithm, the weight relations need to be defined for completeness. The case value for each LV is estimate in PLS as follows:

$$\hat{\xi}_h = \sum_{kh} w_{kh} x_{kh},$$

$$\hat{\eta}_i = \sum_{ki} w_{ki} y_{ki},$$

where w_{kh} and w_{ki} are the k weights used to form the LV estimates of ξ_h and η_i.

Thus, the LV estimates are linear aggregates of their observed indicators whose weights are obtained via the PLS estimation procedure as specified by the inner and outer models where η is a vector of the endogenous (i.e., dependent) latent variables, ξ is a vector of the exogenous (i.e., independent) latent variables, ζ is a vector of residuals, and B and Γ are the path coefficient matrices.

Predictor Specification

Predictor specification (Presp), therefore, forms the basis for PLS modeling. Whereas the covariance-based ML estimation rests on the assumptions of a specific joint multivariate distribution and independence of observations, the PLS approach does not make these hard assumptions. Instead, the PLS technique of model building uses very general, soft distributional assumptions, which often lead to this approach being termed "soft modeling." Thus, as Lohmöller (1989, p. 64) noted, "it is not the concepts nor the models nor the estimation techniques which are 'soft,' only the distributional assumptions."

Presp "is imposed on relations that the investigator wants to use for prediction, be it in theoretical or estimated form, and Presp provides the ensuing predictions" (Wold, 1988, p. 589).

Lohmöller (1989) further states that Presp "starts with the purpose of prediction (not primarily a structural explanation) [and] sets up a system of relations preferably linear, where the structure of the relations must be founded in the substance of the matter, and the predictive purpose should not jeopardize a structural-causal interpretation of the relation (causal-predictive relation)" (p. 72). Presp adopts the statistical assumptions for a linear conditional expectation relationship between dependent and independent variables, which can be summarized as

$$\begin{aligned} y = \alpha + Bx + \nu, \qquad & \hat{y} \equiv E[y \mid x] = \alpha + Bx \\ & \Rightarrow E[\nu] = 0 \\ & \Rightarrow \mathrm{Cov}[x, \nu] = \mathrm{Cov}[\hat{y}, \nu] = 0 \\ & \Rightarrow \mathrm{Cov}[x, y] = \mathrm{Cov}[x, \hat{y}] = B\,\mathrm{var}[x], \end{aligned} \tag{4}$$

where y and x are $m \times 1$ and $n \times 1$ matrices of dependent and independent variables, ν is an $m \times 1$ matrix of residuals, and B the $m \times n$ matrix of coefficient relations between y and x. The implications are that, for a given x and y:

1. x is a predictor (cause or stimulus) of y, and not the other way around (i.e., nonreversability).
2. $\hat{y}$ is the systematic part of y, with respect to x.
3. The systematic part, $\hat{y}$, is a linear function of x.

The observational or empirical representation of Equation 4 would follow simply by including the index n for observations $1, \ldots, N$:

$$\begin{aligned} y = \alpha + \beta x_n + \nu_n, \qquad & \hat{y}_n \equiv E[y_n \mid x_n] = \alpha + \beta x_n \\ & \Rightarrow E[\nu_n] = 0 \\ & \Rightarrow \mathrm{Cov}[x_n, \nu_n] = \mathrm{Cov}[\hat{y}_n, \nu_n] = 0 \\ & \Rightarrow \mathrm{Cov}[x_n, y_n] = \mathrm{Cov}[x_n, \hat{y}_n] = \beta\,\mathrm{var}[x_n]. \end{aligned}$$

Therefore, it should be noted that identical distributions are not assumed. For any two cases, say n and $n+1$, no assumption is made that the residuals ν_n and ν_{n+1} have the same distribution. Nor is independence of cases required because no specification was made regarding the correlation between two different cases

(i.e., Cov[ν_n, ν_{n+1}]). In general, a sufficient condition for consistency of least squares estimates is that, as the number of observations approaches infinity, the sum of the correlations between cases must stay below infinity (i.e., $\sum_i |\text{cor}(\nu_n, \nu_{n+1})| < \infty$; Wold, 1988).

Thus, predictor specification can be viewed as a least squares counterpart to the distributional assumptions of ML modeling. It avoids the assumptions that observations follow a specific distributional pattern and that they are independently distributed. Therefore, no restriction is made on the structure of the residual covariances and, under least squares modeling, the residual variance terms are minimized. In summary, Wold (1988) states that Presp "provides a general rationale for (i) LS [least squares] specification and (ii) LS estimation, and thereby also for the application of (iii) the cross-validation test for predictive relevance ... and (iv) the assessment of SEs by Tukey's jackknife" (p. 587), which are used for model evaluation.

Sample Size Requirements Based on the Inside and Outside Approximations

With the formal model specification and the basic PLS estimation process described, the requirements for sample size become reasonably clear for all three stages. As our previous example demonstrates, either simple or multiple regressions are performed, depending on the mode for each block of indicators and the inner weighting scheme. Due to the partial nature of the estimation procedure, where only a portion of the model is involved at any one time, only that part that requires the largest multiple regression need be found. Although stages 2 and 3 are equivalent in sample size requirements, stage 1 may not require as large a sample size contingent on which inner approximation is selected.

Overall, for an initial sense of the sample size required at stages 2 and 3, one simply looks at the model specification or, equivalently, the graphical model such as that depicted in Figure 1 and finds the largest of two possibilities: (1) the block with

the largest number of formative indicators (i.e., largest measurement equation) or (2) the dependent LV with the largest number of independent LVs influencing it (i.e., the largest structural equation). If one were to use a regression heuristic of 10 cases per predictor, the sample size requirement would be 10 times either (1) or (2), whichever is greater.

Ideally, for a more accurate assessment, one needs to specify the effect size for each regression analysis and look up the power tables provided by Cohen (1988) or Green's (1991) approximation to these tables. Using Figure 1 as an example, the only block with formative indicators consists of four indicators influencing ξ_2. The dependent LV with the largest number of independent LVs influencing it is η_2, with three paths going into it. Thus, the largest regression at any one time consists of four independent variables. Assuming a medium effect size as defined by Cohen (1988), a minimum sample size of 84 is needed to obtain a power of .80. With a large effect size, the sample requirement drops to 39.

For stage 1, the use of a path weighting scheme would result in the same sample requirements as necessary for stages 2 and 3. However, with the use of a factor or a centroid weighting scheme, only simple regressions between the LVs are performed in calculating the weights to be used for the inside approximation. In this situation, only the measurement model with formative indicators becomes the critical factor in sample size requirements. Had all latent variables been modeled as reflective (mode A), the use of either a factor or a centroid weighting scheme would entail only a series of simple regressions during the entire stage 1 process, resulting in minimum sample size requirements of 53 and 24 for medium and large effect sizes, respectively.

In fact, the minimum sample size required to assess component loadings for reflective indicators is likely even smaller. Given that the standard requirement for loadings is normally set at .60 or above, the effect size of component loadings is larger than what is considered large in regression power analysis (i.e., f^2 of .35, Cohen, 1988). For example, a .60 loading represents an f^2 effect size of .56 and requires a sample size of 15 to obtain a power of .80 for detection. This situation is demonstrated partly

in the Monte Carlo study to follow, wherein sample sizes of 20 could not detect structural paths of .40, but easily detected loadings of .60 and .80.

Model Evaluation

As noted, PLS makes no distributional assumption (other than predictor specification) in its procedure for estimating parameters. Thus, traditional parametric-based techniques for significance testing/evaluation would not be appropriate. Instead, Wold (1980, 1982) argued for tests consistent with the distribution-free/predictive approach of PLS. In other words, rather than based on covariance fit, evaluation of PLS models should apply prediction-oriented measures that are also nonparametric. To that extent, the R^2 for dependent LVs, the Stone-Geisser (Stone, 1974; Geisser, 1975) test for predictive relevance, and Fornell and Larcker's (1981) average variance extracted measure are used to assess predictiveness, whereas resampling procedures such as jackknifing and bootstrapping are used to examine the stability of estimates. Readers interested in more detail regarding statistical tests, with examples of their use, can consult Barclay, Higgins, and Thompson (1995), Chin (1998), Chin and Gopal (1995), and Mathieson, Peacock, and Chin (1996).

Partial Least Squares Estimates: The Issue of Consistency at Large

Although one of the benefits of the PLS procedure can be argued to be its ability to estimate LV case values, these scores can lead to biased parameter estimates. Essentially, the case values for the LVs are "inconsistent" relative to the CBSEM model due to the fact that they are aggregates of the observed variables, which in part include measurement error. This bias tends to manifest itself in higher estimates for component loadings (outer model relations) and lower estimates at the structural level (inner model

relations). The estimates will approach the "true" latent variable scores as both the number of indicators per block and the sample size increase. This limiting case is termed "consistency at large" (Wold, 1982, p. 25). Intuitively, the larger the number of indicators in a block, the more the "essence" of the LV is confirmed by the data. However, the sample size also needs to increase, as in the usual asymptotic notion of consistency, in order for the sample covariance matrix to become a better estimate of the population covariance matrix. Thus, in PLS, better estimates cannot simply be obtained by increasing the sample size. Both more indicators and more cases are needed. Furthermore, increasing the block size not only results in estimates that approach the "true" parameter scores, but also lowers the standard errors, which have been shown to vary inversely with the square root of the block size (Lyttkens, 1966, 1973).

Although closed-form solutions for estimating the amount of bias are not available for multiblock heterogeneous loading conditions, formulas for estimating the bias of PLS estimates relative to the parameter-oriented CBSEM ML estimates in the single- and two-block models have been provided (for derivations, see Lohmöller, 1989, pp. 207–212). In these situations, it was shown that the bias decreases as the loadings become more reliable, and the bias decreases as the number of indicators increases. For the simple two-block model, it was also demonstrated that the predicted correlation between indicators from different blocks is unbiased because the loading and correlation biases cancel each other out. The general proof that this canceling effect occurs even under conditions of unequal block sizes, weights, and loadings has been given by Areskoug (1982). Therefore, we again see the distinction between prediction orientation versus parameter orientation. The parameter estimation accuracy of the PLS procedure relative to the covariance-based ML procedure increases under consistency at large.

Yet, it can be argued that PLS estimates are consistent under the formal PLS model. The bias measures just discussed were calculated relative to CBSEM ML estimation, which presupposes that the underlying model that generates the data is covariance based. Schneeweiss (1990) has suggested that the consis-

tency at large notion is really a "justification for using PLS as an estimation method to estimate CBSEM parameters in cases where the number of manifest variables is large" (p. 38). Instead, Schneeweiss argues that PLS can be seen as a consistent estimator of parameters and latent variables as long as we determine which population parameters we are attempting to estimate. If we are estimating the parameters for the population model as defined by PLS, then we have the advantage of "treating PLS as a method for defining descriptive parameters in situations where blocks of manifest variables are related to each other" (p. 38), even if the data cannot be regarded as stemming from a CBSEM model. In this situation, the PLS estimation method will estimate the PLS parameters consistently. If, on the other hand, the data are generated from a covariance-based model, the PLS estimates will result in inconsistent estimates.

Therefore, while PLS can be used in a confirmatory sense following a covariance-based orientation, it can also be used for testing the appropriateness of a block of indicators in a predictive sense and for suggesting potential relations among blocks without necessarily making any assumptions regarding which LV model generated the data. As Wold (1980) noted, an initial PLS model is:

> ... usually tentative since the model construction is an evolutionary process. The empirical content of the model is extracted from the data, and the model is improved by interactions through the estimation procedure between the model and the data and the reactions of the researcher. Consequently, the researcher should begin with a generous number of observables-indicators in the various blocks. To use many observables makes for rich empirical content of the model and is favorable to the accuracy of the PLS estimation procedure. In the interaction between the data and the original model it will become apparent which indicators are relevant and which should be omitted (p. 70).

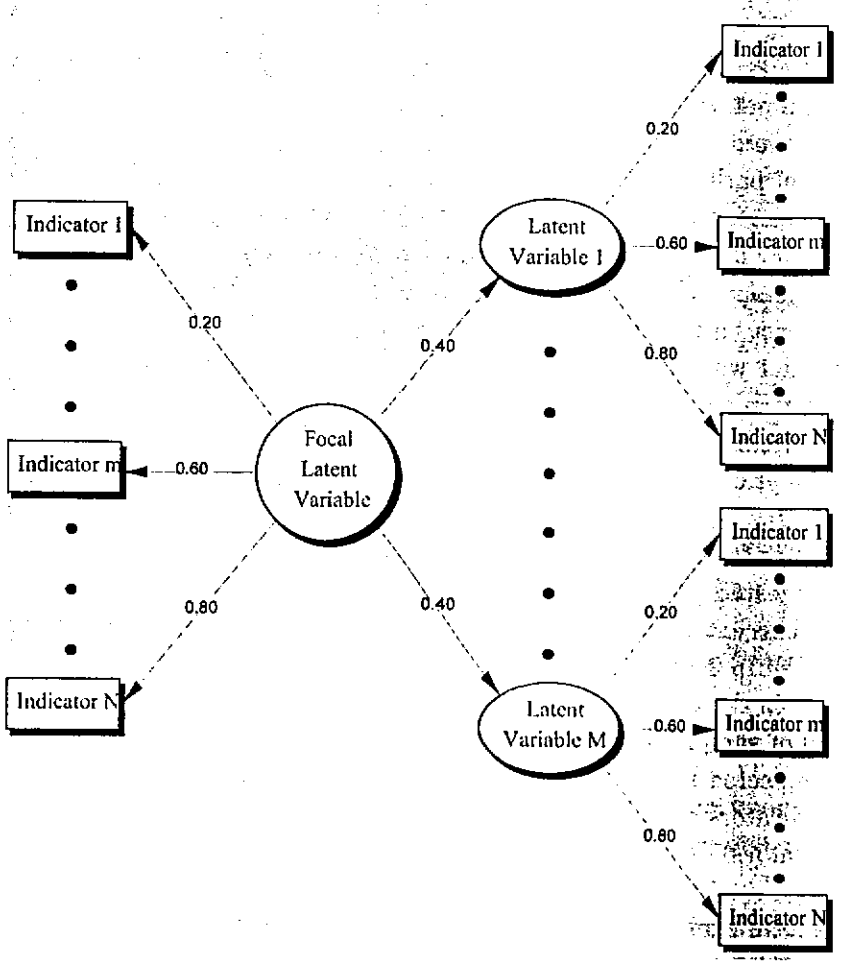
Monte Carlo Simulation

In this section, the results of two Monte Carlo studies are summarized.[1] The studies were designed to examine how well the PLS algorithm performs in recovering the "true" population parameters under varying numbers of latent variables, indicators, and sample cases. More specifically, the estimates for structural paths and component loadings were examined under conditions (as discussed earlier) that are normally out of the range of CBSEM analysis. Although the distinction in generating data conforming to a PLS model versus a CBSEM model and its impact on the notion of consistency was made, a CBSEM-based model was nonetheless applied in order to see how well PLS performs under the notion of consistency at large. In terms of structural path recovery, PLS was also compared to simple path-analytic regression using a simple summation of the indicators. No attempt was made in these studies to compare the PLS estimates to CBSEM estimates. The sample size and model complexity tested in these studies were, by and large, beyond the scope of current CBSEM programs. The data were generated using PRELIS 2.14 and tested using PLS-Graph, Version 2.91.

In the first Monte Carlo study, the three treatments consisted of sample sizes (20, 50, 100, 150, and 200 cases), number of connected latent variables (2, 4, 8, 12, and 16), and number of indicators attached to each latent variable (4, 8, 12, 16, or 32 indicators). For each of the 125 cells in the design, 100 replications were made. Figure 3 represents the model tested. In this model, a focal exogenous latent variable is connected to m endogenous latent variables with standardized paths of .40. As with the other m LVs, the focal LV has n indicators. Consistent with CBSEM models, all latent variables were modeled with reflective indicators. In all runs, the "true" loadings for each LV were kept heterogeneous with the first 25% of indicators set at standardized loadings of .20, the next 25% at .60, and the last 50% at 0.80.

[1] Due to page limitations, results of the Monte Carlo studies are not presented in tabular form. Tabled results can be obtained from either author, the editor, or on the World Wide Web at http://disc-nt.cba.uh.edu/chin/sage/appendix.htm.

Figure 3 Model Used to Generate Results for Study 1. All Structural Paths Connecting the Focal Latent Variable to the Other *m* Latent Variables Were Set at .40. For Each Latent Variable, Component Loadings Were Set at .20, .60, and .80 for 25%, 25%, and 50% of the *n* Indicators, Respectively



Essentially, this represents a situation in which one fourth of the indicators for each LV are useless, one fourth adequate, and one half considered good.

The impact of consistency at large on loading estimates was clear. Specifically, increasing sample size alone does not provide a better approximation to the population value. Instead, the number of indicators also has to increase. As an example, for the .80 loading in the 150 sample size analysis, a substantial improve-

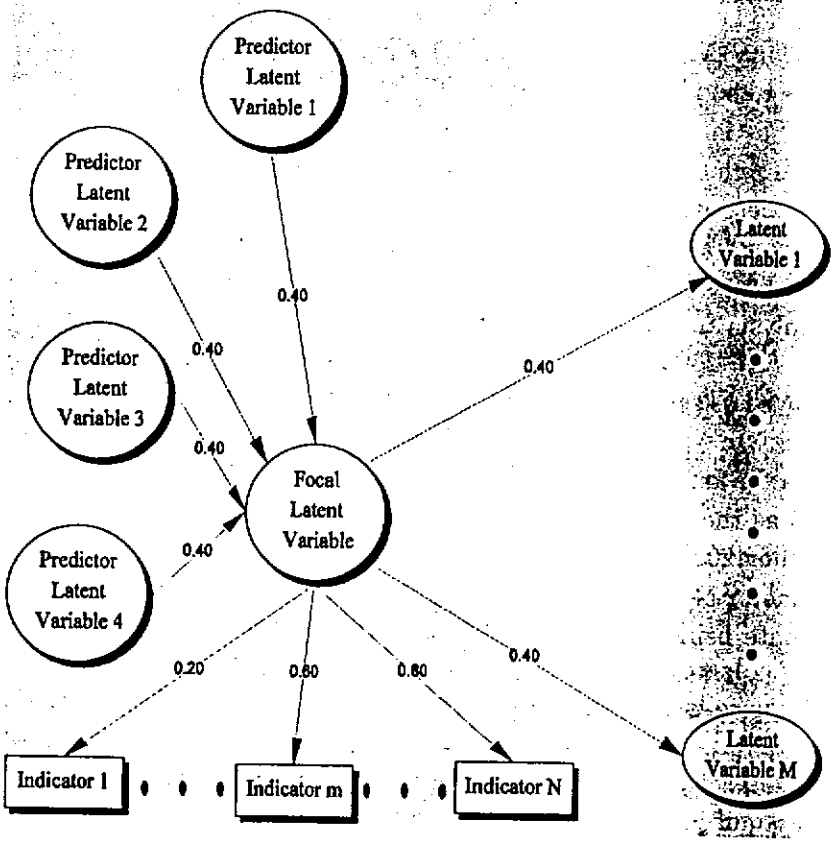
ment occurs when the number of indicators increases from four to eight, whereupon it starts to plateau. As expected, increasing sample size lowered standard errors. At the extreme, the very small loading of .20 was not detected until sample sizes of 150 and 200 were reached. However, as discussed earlier, it was possible to successfully estimate and detect path loadings of .60 and .80 at the small sample size of 20, albeit with reasonably large standard errors.

Interestingly, the number of connected latent variables did not seem to help the loading estimates. At best, the standard error dropped slightly as the number of LVs increased, but only when the number of indicators was at four or eight. For example, using the 100 sample size analysis with four indicators, the standard error for the .60 loading drops from .07 to .05 as the number of LVs increases from 2 to 16. Similarly, the standard error for the .80 loading goes from .04 to .02. It is possible that a more significant LV influence may occur if a structural path larger than the .40 had been specified.

Partial least squares always performed better than the simple summed regression approach, although it did best at lower numbers of indicators. As discussed earlier, the minimum sample size for a medium effect size, which .40 represents, is 53. Thus, it was not surprising to find nonsignificant results for the $N = 20$ analysis and less accurate estimates for the $N = 50$ analysis. For the $N = 200$ analysis, both the PLS and regression estimate improved as the number of indicators increased. However, the PLS estimate was approximately .05 closer to the population parameter for four- and eight-indicator conditions. The gap dropped to .03 as the number of indicators increased. For sample sizes of 150 or 200, the mean PLS estimate yielded the population parameter at indicator levels of 16 and 32 with the regression estimate being quite close. Overall, with 16 or more indicators, the regression approach might be preferred because it obtains similar results at a simpler level of computation. At the more realistic level for social science research (i.e., four or eight indicators), however, the more accurate estimates would suggest using PLS.

In the second study, treatments consisted of sample sizes (50, 100, and 200), number of connected latent variables (2, 8, and 16), and number of indicators attached to each latent variable (4, 8, 16, or 32 indicators). Heterogeneous loadings were set in the same fashion as in the first study. In addition to restricting the scope of the treatments, the model was modified by adding four independent exogenous LVs impacting our focal LV (all with standardized paths of .40). Thus, the focal LV becomes a mediator between four exogenous LVs and m endogenous LVs. Figure 4 depicts the model analyzed.

Figure 4 Model Used to Generate Results for Study 2. All Structural Paths Connecting the Focal Latent Variable to the Other Latent Variables Were Set at .40. There Are Four Exogenous Latent Variables and m Endogenous Latent Variables Connected to the Focal Latent Variable. For Each Latent Variable, Component Loadings Were Set at .20, .60, and .80 for 25%, 25%, and 50% of the n Indicators, Respectively.



The results of Study 2 essentially corroborated the results obtained in the first study. Earlier, it was noted that a regression involving four independent variables and medium effect sizes would need a minimum sample size of 80. Interestingly, the analysis at sample size 50 still generated significant results. The standard errors, as expected, dropped as sample size increased. Though not as strong, the PLS estimates again performed best relative to the simple summed regression approach with four and eight indicators; the difference in the standardized beta was approximately .02 or .03. With greater numbers of indicators, the PLS and regression estimates were essentially the same.

Overall, the results show that the PLS approach can provide information about the appropriateness of indicators at sample size as low as 20. Furthermore, it performed better than the simple summed regression approach with four or eight indicators. Computationally, the most complex model in this study, which involved 672 indicators, 21 latent variables, and 200 cases, took approximately 1.5 minutes to run on a 166-MHz Pentium computer. Such an analysis would not be possible using CBSEM.

Summary

As mentioned at the beginning of this chapter, by far the most common approach for SEM has been covariance-based procedures. Yet, the PLS procedure may be more suitable under certain circumstances. The PLS model and estimation procedure covered in this chapter revealed many differences between CB-SEM and PLS.

Programs such as LISREL, EQS, AMOS, SEPath, CALIS, and RAMONA use covariance-based procedures with the objective of obtaining optimal parameter accuracy. The level of theoretical/substantive knowledge that the researcher brings to the study is a major factor inasmuch as any given model becomes the basis for explaining the covariances among all the indicators. To obtain consistent parameter estimates, the empirical conditions of the data require a multivariate normal distribution (under an

ML function) and independence of observations. Finally, indicators are typically required to be modeled as reflective and unique case values for LVs cannot be obtained (i.e., factor indeterminacy).

Partial least squares was developed as a counterpart to CB-SEM analysis. In general, it can be viewed as complementary to CBSEM because its main objective is prediction. The focus of PLS, under predictor specification, is on the variance of dependent variables and no assumptions are made regarding the joint distribution of the indicators or the independence of sample cases. Because of its prediction orientation, factors are determinate and unique case values for the LVs are estimated. Indicators can be modeled as either formative or reflective.

Sample size requirements under PLS can be quite minimal relative to a CBSEM analysis. Computationally, PLS is an order of magnitude faster given that its procedure involves only a series of least squares estimations. By virtue of the fact that, at any moment, only a subset of the parameters is being estimated, PLS can handle much larger/complex models with many LVs and indicators in each block. Models consisting of over 1,000 indicators can be easily executed. Furthermore, as the number of indicators per block increases along with the sample size, the PLS estimates tend to become more stable as they converge to the "true" parameter values.

In summary, PLS should not be viewed as simply a distribution-free alternative to CBSEM. Rather, it represents a different approach to empirical modeling—a descriptive, prediction-oriented one. As Falk and Miller (1992) state, using the terms "soft" and "hard" modeling for PLS and CBSEM, respectively:

> A wide gulf exists between predictive relationships and causal relationships. While causal relationships are predictive, predictive relationships need not be even remotely causal. With soft modeling the researcher is forced to adopt a more predictive, therefore descriptive, stance.
>
> If hard modeling procedures cannot be used because of study limitations, then soft modeling procedures cannot be expected to answer hard modeling questions about

statistical inferences. On the other hand, we reject the notion that study limitations should prevent data from being analyzed. Our position is that many research questions can and should be answered without making causal inferences. In our view this is the role of soft modeling. (pp. xi–xii)

We agree that the PLS methodology will likely grow in usage in the future. The approach is congruent with a large percentage of social science research where:

- The objective is prediction, and/or
- The phenomenon in question is relatively new or changing and the theoretical model or measures are not well formed, and/or
- The model is relatively complex with large numbers of indicators and/or LVs, and/or
- There is an epistemic need to model the relationship between LVs and indicators in different modes (i.e., formative and reflective measures), and/or
- The data conditions relating to normal distribution, independence, and/or sample size are not met.

Author Note: The assistance of Robert Willis in running the Monte Carlo simulations reported in this chapter is gratefully acknowledged.

References

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factory analyses. *Psychometrika, 49*, 155–173.

Areskoug, B. (1982). The first canonical correlation: Theoretical PLS analysis and simulation experiments. In K. G. Jöreskog and H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part 2, pp. 95–118). Amsterdam: North-Holland.

Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to causal modeling: Personal computer adoption and use as an illustration (with commentaries). *Technology Studies, 2,* 285–323.

Bollen, K. A., & Lennox, R. D. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110,* 305–314.

Boomsma, A. (1983). *On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality.* Amsterdam: Sociometric Research Foundation.

Chin, W. W. (1998). The partial least squares approach for structural equation modeling. In G. A. Marcoulides (Ed.), *Modern methods for business research* (pp. 295–336). Mahwah, NJ: Erlbaum.

Chin, W. W., & Gopal, A. (1995). Adoption intention in GSS: Relative importance of beliefs. *The Data Base for Advances in Information Systems, 26,* 42–64.

Chou, C.-P., & Bentler, P. M. (1995). Estimates and tests in structural equation modeling. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 37–55). Newbury Park, CA: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, P., Cohen, J., Teresi, J., Marchi, M., & Velez, C. N. (1990). Problems in the measurement of latent variables in structural equations causal models. *Applied Psychological Measurement, 14,* 183–196.

Dillon, W. R., Kumar, A., & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin, 101,* 126–135.

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modeling.* Akron, OH: University of Akron Press.

Fornell, C. (1987). A second generation of multivariate analysis: Classification of methods and implications for marketing research. In M. J. Houston (Ed.), *Review of marketing* (pp. 407–450). Chicago: American Marketing Association.

Fornell, C. (1989). The blending of theoretical and empirical knowledge in structural equations with unobservables. In H. Wold (Ed.), *Theoretical empiricism* (pp. 153–173). New York: Paragon.

Fornell, C., & Bookstein, F. (1982). Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. *Journal of Marketing Research, 19*, 440–452.

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

Fornell, C., & Yi, Y. (1992). Assumption of the two-step approach to latent variable modeling. *Sociological Methods and Research, 20*, 291–320.

Geisser, S. (1975). The predictive sample reuse method with applications. *Journal of the American Statistical Association, 70*, 320–328.

Gerbing, D. W., & Anderson, J. C. (1987). Improper solutions in the analysis of covariance structures: Their interpretability and a comparison of alternative respecifications. *Psychometrika, 52*, 99–111.

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research, 26*, 499–510.

Hu, L.-T., & Bentler, P. M. (1995). Evaluating model fit. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 76–99). Newbury Park, CA: Sage.

Hui, B. S. (1978). *The partial least squares approach to path models of indirectly observed variables with multiple indicators.* Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Hui, B. S. (1982). On building partial least squares models with interdependent inner relations. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part 2, pp. 249–272). Amsterdam: North-Holland.

Jöreskog, K. G. (1973). A general method for estimating a linear structural equation system. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 85–112). New York: Academic Press.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications*. Chicago: SPSS.

Keesling, J. W. (1972). *Maximum likelihood approaches to causal analysis*. Unpublished doctoral dissertation, University of Chicago.

Lohmöller, J.-B. (1984). *LVPLS program manual: Latent variables path analysis with partial least-squares estimation.*

Cologne, Germany: Zentralarchiv für empirische Sozialforschung.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares.* Heidelberg, Germany: Physica-Verlag.

Lyttkens, E. (1966). On the fix-point property of Wold's iterative estimation method for principal components. In P. R. Krishnaiah (Ed.), *Multivariate analysis: Proceedings of an international symposium held in Dayton, Ohio, June 14–19, 1965* (pp. 335–350). New York: Academic Press.

Lyttkens, E. (1973). The fixed-point method for estimating interdependent systems with the underlying model specification. *Journal of the Royal Statistical Society, A136,* 353–394.

MacCallum, R. C. (1986). Specification searches in covariance structure analyses. *Psychological Bulletin, 100,* 107–120.

MacCallum, R. C., & Browne, M. W. (1993). The use of causal indicators in covariance structure models: Some practical issues. *Psychological Bulletin, 114,* 533–541.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130–149.

Maraun, M. D. (1996). Metaphor taken as math: Indeterminacy in the factor analysis model. *Multivariate Behavioral Research, 31,* 517–538.

Mathieson, K., Peacock, E., & Chin, W. (1996). *Extending the technology acceptance model: The influence of perceived user resources.* Working Paper WP 96–18, Faculty of Management, University of Calgary, Alberta, Canada.

Mulaik, S. A., James, L. R., Van Alstine, J., Bennett, N., Lind, S., & Stillwell, C. D. (1989). An evaluation of goodness-of-fit indices for structural equation models. *Psychological Bulletin, 105,* 430–445.

Noonan, R., & Wold, H. (1982). PLS path modeling with indirectly observed variables: A comparison of alternative estimates for latent variables. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part 2, pp. 75–94). Amsterdam: North-Holland.

Rindskopf, D. (1984). Structural equation models: Empirical identification, Heywood cases, and related problems. *Sociological Methods and Research, 13,* 109–119.

Schneeweiss, H. (1990). Models with latent variables: LISREL versus PLS. *Contemporary Mathematics, 112,* 33–40.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions, *Journal of the Royal Statistical Society, B36*, 111–133.

van Driel, O. P. (1978). On various causes of improper solutions in maximum likelihood factor analysis. *Psychometrika, 43*, 225–243.

Wiley, D. E. (1973). The identification problem for structural equation models with unmeasured variables. In A. S. Goldberger & O. D. Duncan (Eds.), *Structural equation models in the social sciences* (pp. 69–83). New York: Academic Press.

Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce: Theory and application of partial least squares. In J. Kmenta & J. B. Ramsey (Eds.), *Evaluation of econometric models*. New York: Academic Press.

Wold, H. (1982). Soft modeling: The basic design and some extensions. In K. G. Jöreskog & H. Wold (Eds.), *Systems under indirect observations: Causality, structure, prediction* (Part 2, pp. 1–54). Amsterdam: North-Holland.

Wold, H. (1985). Partial least squares. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 6, pp. 581–591). New York: Wiley.

Wold, H. (1988). Specification, predictor. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (Vol. 8, pp. 587–599). New York: Wiley.