UNDERSTANDING STATISTICS, *3*(4), 283–297

A Beginner's Guide to Partial Least Squares Analysis

Michael Haenlein

Department of Department of Electronic Commerce
Otto Beisheim Graduate School of Management

Andreas M. Kaplan

Department of Media Management
University of Cologne

Since the introduction of covariance-based structural equation modeling (SEM) by Jöreskog in 1973, this technique has been received with considerable interest among empirical researchers. However, the predominance of LISREL, certainly the most well-known tool to perform this kind of analysis, has led to the fact that not all researchers are aware of alternative techniques for SEM, such as partial least squares (PLS) analysis. Therefore, the objective of this article is to provide an easily comprehensible introduction to this technique, which is particularly suited to situations in which constructs are measured by a very large number of indicators and where maximum likelihood covariance-based SEM tools reach their limit. Because this article is intended as a general introduction, it avoids mathematical details as far as possible and instead focuses on a presentation of PLS, which can be understood without an in-depth knowledge of SEM.

partial least squares, structural equation modeling, PLS, LISREL, SEM

First-generation techniques, such as regression-based approaches (e.g., multiple regression analysis, discriminant analysis, logistic regression, analysis of variance) and factor or cluster analysis, belong to the core set of statistical instruments which can be used to either identify or confirm theoretical hypothesis based on the analysis of empirical data. Many researchers in various disciplines have applied one of these

Requests for reprints should be sent to Michael Haenlein at e-mail: michael.haenlein@bain.com

methods to generate findings that have significantly shaped the way we see the world today, such as Spearman's (1904) work on general intelligence for psychology (factor analysis), Hofstede's (1983) publication on cross-cultural differences for sociology (factor and cluster analysis), and Altman's (1968) article on forecasting corporate bankruptcy for management research (discriminant analysis).

However, a common factor for all these methods is that they share three limitations, namely, (a) the postulation of a simple model structure (at least in the case of regression-based approaches); (b) the assumption that all variables can be considered as observable; and (c) the conjecture that all variables are measured without error, which may limit their applicability in some research situations.

Where the first assumption, the postulation of a simple model structure (i.e., one dependent and several independent variables) is concerned, Jacoby (1978) stated that "we live in a complex, multivariate world [and that] studying the impact of one or two variables in isolation, would seem … relatively artificial and inconsequential" (p. 91). Although model building always implies omitting some aspect of reality (Shugan, 2002), this assumption of regression-based approaches may be too limiting for an analysis of more complex and more realistic situations. This becomes, for example, especially obvious when one wants to investigate the potential effect of *mediating* or *moderating* variables (for a detailed definition of these two terms, see Baron & Kenny, 1986) on the relationship between one or more dependent and independent variables, which may result in some dependent variables influencing other dependent variables.

With respect to the second limitation, the assumption that all variables can be considered as observable, McDonald (1996) stressed that a variable can be called observable "if and only if its value can be obtained by means of a real-world sampling experiment" (p. 239). Therefore, any variable that does not correspond directly to anything observable must be considered as unobservable (Dijkstra, 1983). This definition makes it obvious that only a handful of relevant variables, such as age and gender, can be considered as observable, whereas "the effects and properties of molecules, processes, genes, viruses, and bacteria are usually observed only indirectly" (S. Wold, 1993, p. 138)

Regarding the conjecture of variables measured without error, one has to bear in mind that each observation of the real world is accompanied by a certain measurement error, which may comprise two parts (Bagozzi, Yi, & Philipps, 1991): (a) random error (e.g., caused by the order of items in a questionnaire or respondent fatigue; Heeler & Ray, 1972) and (b) systematic error, such as method variance (i.e., variance attributable to the measurement method rather than the construct of interest; Bagozzi et al., 1991). Because the observed score of an item is therefore always the sum of three parts, namely, the true score of the variable, random error, and systematic error (Churchill, 1979), first-generation techniques are, strictly speaking, only applicable when there is neither a systematic nor a random error component—a rare situation in reality.

To overcome these limitations of first-generation techniques, more and more authors started using structural equation modeling (SEM) as an alternative. Compared to regression-based approaches, which analyze only one layer of linkages between independent and dependent variables at the same time, SEM, as a second-generation technique, allows the simultaneous modeling of relationships among multiple independent and dependent constructs (Gefen, Straub, & Boudreau, 2000). Therefore, one no longer differentiates between dependent and independent variables but

> distinguishes between the exogenous and endogenous latent variables, the former being variables which are not explained by the postulated model (i.e. act always as independent variables) and the latter being variables that are explained by the relationships contained in the model. (Diamantopoulos, 1994, pp. 108)

Additionally, SEM enables the researcher to construct unobservable variables measured by indicators (also called *items*, *manifest variables*, or *observed measures*) as well as to explicitly model measurement error for the observed variables (Chin, 1998a), and hence it overcomes the limitations of first-generation techniques described earlier and consequently gives the researcher the flexibility to "statistically test a priori substantive/theoretical and measurement assumptions against empirical data (i.e. confirmatory analysis)" (Chin, 1998a, p. vii).

In general, there are two approaches to estimating the parameters of an SEM, namely, the covariance-based approach and the variance-based (or components-based) approach. Covariance-based SEM, in particular, has received high prominence during the last few decades and, "to many social science researchers, the covariance-based procedure is tautologically synonymous with the term SEM" (Chin, 1998b, p. 295). Although there are several different tools that can be used to perform this kind of analysis, such as EQS, AMOS, SEPATH, and COSAN, the LISREL program developed by Jöreskog in 1975 became the most popular one and, consequently, the term *LISREL* is sometimes used as a synonym for covariance-based SEM.

The focus of this article is to give an introduction to the other side of the coin, variance-based SEM, and to present partial least squares (PLS) analysis as one technique from this group in more detail. In contrast to articles already published in this area (e.g., Cassel, Hackl, & Westlund, 1999; Dijkstra, 1983; Garthwaite, 1994), our focus is on an easily understandable presentation of this topic, accessible to beginners without extensive knowledge of statistics in general or SEM in particular. Additionally, we try to answer the question under which circumstances a researcher might want to prefer variance-based over covariance-based SEM, given the specific assumptions and limitations of each of these methods.

For this purpose, our article is structured as follows: In the next section, we provide a short introduction to theories, SEM, and the measurement of unobservable

variables as necessary background to understanding the fundamentals behind SEM. We then proceed to the main topic of this article, an introduction to PLS analysis, explaining the basics of this technique as well as statistical assumptions and limitations associated with it. The article concludes with some recommendations regarding the suitability of PLS for SEM, focusing particularly on situations in which covariance-based SEM might not be appropriate.

THEORIES, SEM, AND INDICATORS

As indicated in the beginning of this article, SEM can be (and often is) used to test (and consequently to either support or reject) theoretical assumptions with empirical data. It is therefore essential to have a sound understanding of the structure of theories to understand the different components of a structural equation model.

According to Bagozzi and Philipps (1982), a theory may contain three different types of concepts: (a) theoretical concepts that "are abstract, unobservable properties or attributes of a social unit of entity" (p. 465); (b) empirical concepts which "refer to properties or relations whose presence or absence in a given case can be inter-subjectively ascertained, under suitable circumstances, by direct observations" (Bagozzi & Philipps, 1982, p. 465); and (c) derived concepts, which are unobservable (like theoretical concepts) but "unlike theoretical concepts … must be tied directly to empirical concepts" (Bagozzi & Philipps, 1982, p. 465). Additionally, there are three possible types of relationship that link these concepts: (a) *nonobservational hypotheses*, which link theoretical concepts with other theoretical concepts; (b) *theoretical definitions*, which connect theoretical and derived concepts; and (c) *correspondence rules*, which link theoretical or derived to empirical concepts and serve "to provide empirical significance to theoretical terms" (Bagozzi, 1984, p. 17).

Using this framework, it is possible to construct a research model that represents a certain theory, simply by converting theoretical and derived concepts into unobservable (latent) variables, and empirical concepts into indicators, which are linked by a set of hypotheses (representing either nonobservational hypotheses, theoretical definitions, or correspondence rules). This model can then be represented graphically by a path diagram (also called an *arrow scheme*; see Figure 1), which shows how the various elements relate to one another (Diamantopoulos, 1994).

Based on the path diagram (and using Figure 1 as an example), it is then possible to set up three sets of equations, which can be used to describe the relationships between the different parameters of the research model.

The first set relates the indicators of the exogenous variables (x) to their associated measurement error (δ) and the latent exogenous variables (ξ):

$$x_1 = \lambda_{x11}\, \xi_1 + \delta_1$$

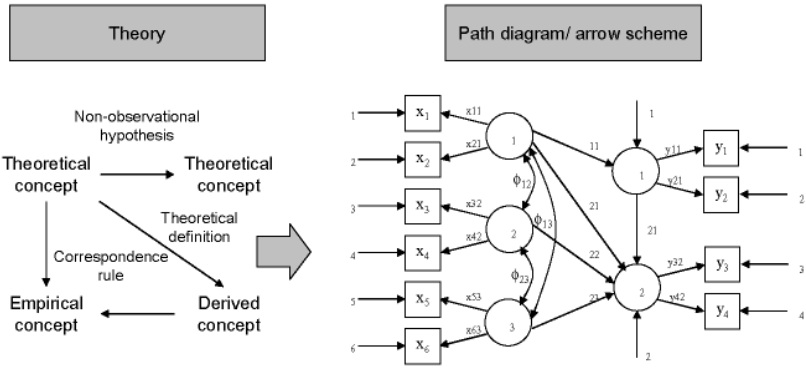

FIGURE 1 Relationship between theory and path diagram/arrow scheme (Bagozzi & Philipps, 1982; Diamantopoulos, 1994). η (eta) = latent endogenous variable; ξ (xi) = latent exogenous (i.e., independent) variable; ζ (zeta) = random disturbance term; γ (gamma) = path coefficient; ϕ (phi) noncausal relationship between two latent exogenous variables; y_i = indicators of endogenous variables; ε_i (epsilon) = measurement errors for indicators of endogenous variable; λ_{yi} (lambda *y*) = loadings of indicators of endogenous variable; x_i = indicators of endogenous variable; δ_i (delta) = measurment errors for indicators of exogenous variable; λ_{xi} = (lambda *x*) loadings of indicators of exogenous variable.

$$x_2 = \lambda_{x21}\,\xi_1 + \delta_2$$
$$x_3 = \lambda_{x32}\,\xi_2 + \delta_3$$
$$x_4 = \lambda_{x42}\,\xi_2 + \delta_4$$
$$x_5 = \lambda_{x53}\,\xi_3 + \delta_5$$
$$x_6 = \lambda_{x63}\,\xi_3 + \delta_6$$

The second set describes the relationship between the indicators of the endogenous variables (*y*), their associated measurement error (ε), and the latent endogenous variables (η):

$$y_1 = \lambda_{y11}\,\eta_1 + \varepsilon_1$$
$$y_2 = \lambda_{y21}\,\eta_1 + \varepsilon_2$$
$$y_3 = \lambda_{y32}\,\eta_2 + \varepsilon_3$$
$$y_4 = \lambda_{y42}\,\eta_2 + \varepsilon_4$$

Finally, the last set deals with the relationship between the latent endogenous (η) and exogenous (ξ) variables:

$$\eta_1 = \gamma_{11}\,\xi_1 + \zeta_1$$
$$\eta_2 = \beta_{21}\,\eta_1 + \gamma_{21}\,\xi_1 + \gamma_{22}\,\xi_2 + \gamma_{23}\,\xi_3 + \zeta_2.$$

In contrast to the equations just formulated, the random disturbance terms ζ do not reflect measurement error but are known as "errors in equations" and "reflect random disturbances (i.e. they indicate that the endogenous variables are not perfectly explained by the independent variables)" (Diamantopoulos, 1994, p. 110).

By using matrix algebra, these three sets of equations can also be written in the following way:

$$\mathbf{x} = \Lambda_x \boldsymbol{\xi} + \boldsymbol{\delta}, \tag{1}$$

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\epsilon}, \tag{2}$$

$$\boldsymbol{\eta} = \mathrm{B}\, \boldsymbol{\eta} + \Gamma\, \boldsymbol{\chi} + \boldsymbol{\zeta}. \tag{3}$$

This results in a set of theoretical equations (Equation 3), representing nonobservational hypotheses and theoretical definitions, and measurement equations (Equations 1 and 2), representing correspondence rules (Bagozzi & Philipps, 1982). The theoretical equations are then also referred to as the *structural model*, whereas the measurement equations build the *measurement model*, and both combined can be subsumed by the term *structural equation model*.

As Bagozzi (1984) emphasized, there are three different types of unobservable variables: (a) variables that are unobservable in principle (e.g., theoretical terms); (b) variables that are unobservable in principle but either imply empirical concepts or can be inferred from observations (e.g., attitudes, which might be reflected in evaluations); and (c) unobservable variables that are defined in terms of observables. Because none of these types can be measured directly, the researcher needs to measure indicators instead, which cover different facets of the unobservable variable. In general, indicators can be split into two groups: (a) reflective indicators that depend on the construct and (b) formative ones (also known as *cause measures*) that cause the formation of or changes in an unobservable variable (Bollen & Lennox, 1991).

Written in mathematical terms (see Figure 2), this difference becomes obvious. Whereas reflective indicators can be expressed as a function of their associated latent variables, such as

$$y_1 = \lambda_{y1}\, \eta + \varepsilon_1$$
$$y_2 = \lambda_{y2}\, \eta + \varepsilon_2$$
$$y_3 = \lambda_{y3}\, \eta + \varepsilon_3$$

or, by using matrix algebra,

$$\mathbf{y} = \Lambda_y \boldsymbol{\eta} + \boldsymbol{\epsilon},$$

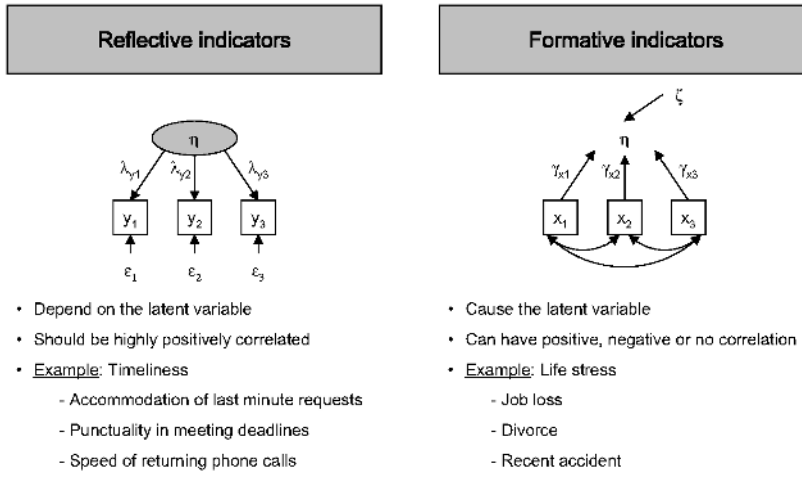

FIGURE 2 Reflective versus formative indicators.

formative indicators are not influenced by but influence the latent variables, so that

$$\eta = \gamma_{x1}\, x_1 + \gamma_{x2}\, x_2 + \gamma_{x3}\, x_3 + \zeta.$$

Consequently, if the unobservable can be considered as giving "rise to something observed," as is the case when, for example, the unobservable describes a personality trait or attitude, reflective indicators should be used. For example, Moorman, Deshpandeé, and Zaltman (1993) operationalized the unobservable variable "timeliness" by the following three reflective indicators: (a) accommodation of last-minute requests, (b) punctuality in meeting deadlines, and (c) speed of returning phone calls. In contrast, formative indicators are appropriate if constructs "are perceived as explanatory combinations of indicators" (Fornell & Bookstein, 1982, p. 442), such as the unobservable variable "life stress," which can be considered as a combination of formative indicators such as job loss, divorce, a recent accident, and death in the family (Chin & Newsted, 1999).

This leads to one major difference between formative and reflective indicators: Whereas reflective indicators should have a high correlation (as they are all dependent on the same unobservable variable), formative indicators of the same construct can have positive, negative, or zero correlation with one another (Hulland, 1999), which means that a change in one indicator does not necessarily imply a similar directional change in others (Chin, 1998a). In the examples given earlier, a person who is considered more timely than another is expected to accommodate last-minute requests more often *and* to be more punctual in meetings *and* to return phone calls more promptly. On the other hand, a higher degree of life stress does not imply that a

person has become unemployed, got divorced, and lost his or her parents all at the same time—one of these events alone may be sufficient to increase stress.

PLS ANALYSIS: BASIC IDEA AND UNDERLYING ASSUMPTIONS

As highlighted in the beginning of this article, there are two approaches to estimate the parameters of an SEM, that is, the covariance-based approach and the variance-based approach.

The covariance-based approach "attempts to minimize the difference between the sample covariances and those predicted by the theoretical model. … Therefore, the parameter estimation process attempts to reproduce the covariance matrix of the observed measures" (Chin & Newsted, 1999, p. 309). Because of the popularity of covariance-based SEM, there is a wide variety of articles that provide an introduction to this technique, and because a detailed description of that method goes beyond the scope of this article, the reader is referred to Diamantopoulos (1994) for an excellent and easily understandable presentation of covariance-based SEM and LISREL. We come back to some points related to this approach later, when we highlight the circumstances under which a researcher might want to prefer variance-based over covariance-based SEM.

Unlike covariance-based SEM, PLS, first introduced by H. Wold (1975) under the name *NIPALS* (*nonlinear iterative partial least squares*), focuses on maximizing the variance of the dependent variables explained by the independent ones instead of reproducing the empirical covariance matrix. Like any SEM, a PLS model consists of a structural part, which reflects the relationships between the latent variables, and a measurement component, which shows how the latent variables and their indicators are related; but it also has a third component, the weight relations, which are used to estimate case values for the latent variables (Chin & Newsted, 1999).

In contrast to covariance-based SEM, which estimates first model parameters and then case values (i.e., estimated values for each latent variable in each data set) by regressing them onto the set of all indicators (Dijkstra, 1983), PLS starts by calculating case values. For this purpose, the "unobservable variables are estimated as exact linear combinations of their empirical indicators" (Fornell & Bookstein, 1982, p. 441), and PLS treats these estimated proxies as perfect substitutes for the latent variables (Dijkstra, 1983). The weights used to determine these case values are estimated so that the resulting case values capture most of the variance of the independent variables that is useful for predicting the dependent variable (Garthwaite, 1994). This is based on the implicit assumption that all measured variance of the variables in the model is useful variance that should be explained (Chin, Marcolin, & Newsted, 1996). Using these weights, it is then possible to de-

termine a value for each unobservable variable, simply by calculating a weighted average of its indicators. This results in a model in which all unobservable variables are approximated by a set of case values and that can, therefore, be estimated by a set of simple, first-generation, ordinary least squares regressions. Consequently, the basic idea of PLS is quite straightforward: First, the weight relations, which link the indicators to their respective unobservable variables, are estimated. Second, case values for each unobservable variable are calculated, based on a weighted average of its indicators, using the weight relations as an input. Finally, these case values are used in a set of regression equations to determine the parameters for the structural relations (Fornell & Bookstein, 1982).

This explanation makes it obvious that the most crucial part of a PLS analysis is the estimation of the weight relations. Of course, it would be easier simply to assume equal weights for all indicators, but this approach has two disadvantages: First, there is no theoretical rationale for all indicators to have the same weighting. Because it can be assumed that the resulting parameter estimates of the structural model depend on the type of weighting used, at least as long as the number of indicators is not excessively large (McDonald, 1996), the (exogenous) assumption of equal weights makes the results highly arbitrary. Second, as Chin, Marcolin, and Newsted (2003b) stressed, such a procedure does not take into account the fact that some indicators may be more reliable than others and should, therefore, receive higher weights.

Consequently, PLS uses a more complex, two-step estimation process to determine the weights (w_i): First, it starts with an outside approximation, in which case values for each latent variable (e.g. η_2 in Figure 1) are estimated, based on a weighted average of their respective indicators (e.g., $\eta_2' = w_1 y_3 + w_2 y_4$). The weights used to calculate this aggregation are determined in a manner similar to a principal-components analysis for reflective or regression analysis for formative indicators (Cassel, Hackl, & Westlund, 1999). In the next step, the inside approximation, improved case values are determined as a weighted average of neighboring latent variables (e.g., $\eta_2'' = w_3\eta_1' + w_4\xi_1' + w_5\xi_2' + w_6\xi_3'$). For this process, there are three different weighting schemes available (centroid, factor, and path weighting scheme; for a detailed description, see Lohmöller, 1989), but one can demonstrate that the choice between them has only a minor impact on the final results. Using this second estimate of the case values, the weight relations are modified (e.g., $\eta_2'' = w_1' y_3 + w_2' y_4$) and the process of inside and outside approximation starts from the beginning again and is repeated until convergence of the case values is achieved (Cassel et al., 1999).

Hence, being a limited information approach (Dijkstra, 1983), PLS has the advantage that it "involves no assumptions about the population or scale of measurement" (Fornell & Bookstein, 1982, p. 443) and consequently works without distributional assumptions and with nominal, ordinal, and interval scaled variables. However, one has to bear in mind that PLS, like any statistical technique, also re-

quires certain assumptions to be fulfilled. Beyond those known from the standard (i.e., Gaussian classical linear ordinary least squares) regression model (see, e.g., Gujarati, 1995), the most important assumption is predictor specification (Chin & Newsted, 1999). This requirement states that the systematic part of the linear regression must be equal to the conditional expectation of the dependent variable (for a mathematical formulation, see H. Wold, 1975) and can be considered as fulfilled in most cases. Furthermore, by using a Monte Carlo simulation, Cassel et al. (1999) showed that PLS is quite robust with regard to several inadequacies (e.g. skewness or multicollinearity of the indicators, misspecification of the structural model) and that the latent variable scores always conform to the true values.

However, there is also another side of the coin, namely, the problem of consistency at large. In general, a consistent estimator can be described as "one that converges in probability to the value of the parameter being estimated as the sample size increases" (McDonald, 1996, p. 248). However, because the case values for the latent variables in PLS are aggregates of manifest variables that involve measurement error, they must be considered as inconsistent (Fornell & Cha, 1994). Therefore, "the path coefficients estimated through PLS converge on the parameters of the latent-variable model [only] as both the sample size and the number of indicators of each latent variable become infinite" (McDonald, 1996, p. 248)—a problem known under the term *consistency at large*. Hence in all real-life situations, in which both the number of cases in the sample and the number of indicators per latent variable will be finite, PLS tends to underestimate the correlations between the latent variables and overestimate the loadings (i.e., the parameters of the measurement model; Dijkstra, 1983). Only when the number of cases in the sample *and* the number of indicators per latent variable increase to infinity do the latent variable case values approach the true values and this problem disappears (Lohmöller, 1989).

CHOICE BETWEEN COVARIANCE- AND VARIANCE-BASED SEM

Given this problem of consistency at large, one might question the suitability of PLS and by right ask why a technique that cannot guarantee one of the key features of any statistical model, namely, consistency of estimators, should be presented to the reader of a statistics journal at all.

The answer to this question is that PLS comes into its own principally in situations in which covariance-based SEM tools reach their limit, namely, when the number of indicators per latent variable becomes excessively large.

As stated earlier, the goal of covariance-based SEM is to determine the matrix of model parameters Φ in such a way that the resulting covariance matrix predicted by the theoretical model $\Sigma(\Phi)$ is as close as possible to the sample covariance matrix S. For this, one needs to define a discrepancy function $F(S, \Sigma)$, which "takes on

a value of zero only when S = Σ and otherwise is positive, increasing as the difference between S and Σ increases" (MacCallum, Browne, & Sugawara, 1996, p. 132). Given that the sample covariance matrix is based on p measured indicators, the most commonly used such function is the normal-theory maximum likelihood function, defined as

$$F_{ML} = \ln |\Sigma| - \ln |S| + \mathrm{Tr}\,(S\Sigma^{-1}) - p$$

(MacCallum et al., 1996).

Regarding the number of indicators per latent variable p, the researcher should try to identify as many of them as possible, since "no serious work on path models … can be done without using a very large number of indicators for each attribute" (McDonald, 1996, p. 267). For example, the work of Marsh, Hau, Balla, and Grayson (1998) gave an indication that more indicators per latent variable lead to fewer improper solutions and more stable results, and similar findings were generated by Nasser and Wisenbaker (2003). Although in some areas, such as management research, researchers rarely have more than a handful of indicators per unobservable variable to their disposition (Baumgartner & Homburg, 1996), there are also disciplines in which this number can be very large, as many as 500 indicators per latent variable or more. For example, Bookstein, Sampson, Streissguth, and Barr (1996) analyzed the influence of prenatal alcohol exposure on neurobehavioural functioning by using 474 indicators to measure the latter construct. Another area in which this might be applicable is that of functional magnetic resonance imaging studies in which brain functions (e.g., remembrance, addiction) are analyzed by using a "surrogate" hemodynamic response as an indicator for neuronal response (for an introduction into functional magnetic resonance imaging, see Parry & Matthews, 2000).

However, in these cases, the sample covariance matrix can easily reach a size that is difficult to handle with conventional computer systems because based on p measured indicators the sample covariance matrix has $p(p + 1)/2$ distinct elements (i.e., p^2 elements in total, excluding $[p^2 - p]/2$ in the upper or lower triangular matrix).

Using the example of a simple SEM with $q = 5$ latent variables (similar to the one in Figure 1), each measured by 200 indicators, this results in

$$200q\,(200q + 1)/2,$$

or 500,500 distinct elements of the sample covariance matrix.

Additionally, and probably more seriously, the statistical power of such a model would be so large that it would in fact be impossible to apply any kind of fit test to judge overall model quality.

Using the preceding example, the minimum sample size $N_{\min}$ to estimate such a model would be

$$N_{min} = 200q,$$

or 1,000, as it is a necessary condition that the number of cases at least exceeds the number of indicators, because otherwise the input matrix is not definite (Marsh et al., 1998). Additionally, the number of degrees of freedom can be determined as

$$df = [200q\ (200q + 1)/2] - 2 \times 200q - k,$$

where k is the number of parameters in the structural model to be estimated because in the measurement model each indicator is associated with two parameters: (a) measurement error (either δ or ε) and (b) a loading on the associated latent variable (either λ_x or λ_y). In the case of a model similar to the one shown in Figure 1, the structural model has $k = 13$ parameters to be estimated, resulting in $df = 498{,}487$. Hence, the statistical power for the test of model fit based on root-mean-square error of approximation using $\alpha = .05$, $\varepsilon_0 = 0.05$, and $\varepsilon_a = 0.08$ (see MacCallum et al., 1996, for details), would be 1.0000. Therefore, one would probably be able to detect rounding errors to the 10th place and the model would be likely to fail even with simulated data.

On the other hand, the problem of consistency at large would no longer be an issue in PLS because "with a sufficiently large number of indicators, the choice of weights ceases to have any influence on the parameters of the path model" (McDonald, 1996). The PLS composite model would be very close to the underlying factor model, and the distinction between linear composites and underlying factors would become vanishingly small. Therefore, the researcher would be well advised to use PLS instead of covariance-based SEM in such situations. Recapitulating these arguments by using the words of S. Wold (1993), H. Wold's son, one can say that "the natural domain for LV [latent variable] models such as PLS … is where the number of 'significant' LV's is small, much smaller than the number of measured variables … and than the number of observations." (p. 137).

SUMMARY AND OUTLOOK

There are many additional points regarding PLS that deserve mention, we discuss briefly two of them in the following paragraphs. For example, another situation in which PLS might be preferable to LISREL are cases in which constructs are measured primarily by formative indicators. This may be a common occurrence in managerial research, as the work of Jarvis, MacKenzie, and Podsakoff (2003) gave an indication that managerial constructs might be reflected better by formative than by reflective indicators. As MacCallum and Browne (1993) showed, the predominance of formative indicators may lead to severe (identification) problems, implied covariances of zero among some indicators, and/or the existence of equivalent models in covariance-based SEM. PLS, in contrast, does not create problems with re-

spect to analyzing formative indicators and can therefore be used for models with either reflective, formative, or both types of indicators (Fornell & Bookstein, 1982).

One more area in which PLS is typically recommended is that of situations in which the sample size is small. For covariance-based SEM, it is generally advisable that the "sample size should exceed 100 observations regardless of other data characteristics to avoid problematic solutions and obtain acceptable fit concurrently" (Nasser & Wisenbaker, 2003, p. 754), and many researchers even recommend a minimum sample size of 200 cases (e.g., Marsh et al., 1998) to avoid results that cannot be interpreted, such as negative variance estimates (i.e., Heywood cases) or correlations greater than one (i.e., improper solutions; Dillon, Kumar, & Mulani, 1987). PLS, on the other hand, is applicable even under conditions of very small sample sizes. Although a detailed discussion of this point can only be made for each individual model based on statistical power analysis, a Monte Carlo simulation performed by Chin and Newsted (1999) indicated that PLS can be performed with a sample size as low as 50, and H. Wold even "analysed 27 variables using two latent constructs with a data set consisting of ten cases" (Chin, Marcolin, & Newsted, 2003a, Appendix A, p. 5). However, given the problem of consistency at large, the question of whether such results are actually of any usability is very difficult to answer.

In summary, we hope that this article helps give readers a first impression about PLS analysis, the assumptions underlying this approach, and the advantages it has compared to traditional covariance-based SEM in certain situations. For additional information, readers are referred to H. Wold's (1975) first presentation of PLS, Lohmöller's (1989) extensive discussion of this approach, and MacDonald's (1996) article regarding path analysis with composite variables. Readers who plan to use PLS for their own research are referred to PLS Graph (Chin, 2001), a freely available software tool with an intuitive graphical user interface.

ACKNOWLEDGMENTS

This work has benefitted substantially from several comments we received through the structural equation modeling group SEMNET. Specifically, we thank A. Alexander Beaujean (University of Missouri–Columbia), Les Hayduk (University of Alberta) and Edward E. Rigdon (Georgia State University). Additionally, we acknowledge the helpful comments of two anonymous reviewers, which helped us to significantly improve an earlier version of this manuscript.

REFERENCES

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609.

Bagozzi, R. P. (1984). A prospectus for theory construction in marketing. *Journal of Marketing, 48*, 11–29.

Bagozzi, R. P., & Philipps, L. W. (1982). Representing and testing organizational theories: A holistic construal. *Administrative Science Quarterly, 27*, 459–489.

Bagozzi, R. P., Yi, Y., & Philipps, L. W. (1991). Assessing construct validity in organizational research. *Administrative Science Quarterly, 36*, 421–458.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Baumgartner, H., & Homburg, C. (1996). Applications of structural equation modeling in marketing and consumer research: A review. *International Journal of Research in Marketing, 13*, 139–161.

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

Bookstein, F. L., Sampson, P. D., Streissguth, A. P., & Barr, H. M. (1996). Exploiting redundant measurement of dose and developmental outcome: New methods from the behavioural teratology of alcohol. *Developmental Psychology, 32*, 404–415.

Cassel, C. M., Hackl, P., & Westlund, A. H. (1999). Robustness of partial least-squares method for estimating latent variable quality structures. *Journal of Applied Statistics, 26*, 435–446.

Chin, W. W. (1998a). Issues and opinion on structural equation modelling. *MIS Quarterly, 22*(1), vii–xvi.

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

Chin, W. W. (2001). *PLS-Graph user's guide, Version 3.0.* Houston, TX: Soft Modelling.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (1996). *A partial least squares latent variable modelling approach for measuring interaction effects: Results from a Monte Carlo simulation study and voice mail emotion/adoption study.* Paper presented at the 17th International Conference on Information Systems, Cleveland, OH.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003a). A partial least squares latent variable modelling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study—Supplement A. *Information Systems Research, 14*. Available at http://www.informs.org/pubs/Supplements/ISR/1526-5536-2003-1502-SupplA.pdf

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003b). A partial least squares latent variable modelling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research, 14*, 189–217.

Chin, W. W., & Newsted, P. R. (1999). Structural equation modelling analysis with small samples using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small sample research* (pp. 307–341). Thousand Oaks, CA: Sage.

Churchill, G. A. J. (1979). A paradigm for developing better measures of marketing constructs. *Journal of Marketing Research, 16*, 64–73.

Diamantopoulos, A. (1994). Modelling with LISREL: A guide for the uninitiated. *Journal of Marketing Management, 10*, 105–136.

Dijkstra, T. (1983). Some comments on maximum likelihood and partial least squares methods. *Journal of Econometrics, 22*, 67–90.

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.

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

Fornell, C., & Cha, J. (1994). Partial least squares. In R. P. Bagozzi (Ed.), *Advanced methods of marketing research* (pp. 52–78). Cambridge, England: Blackwell.

Garthwaite, P. H. (1994). An interpretation of partial least squares. *Journal of the American Statistical Association, 89*(425), 122–127.

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural equation modelling and regression: Guidelines for research practice. *Communications of the Association for Information Systems, 4*, 1–79.

Gujarati, D. N. (1995). *Basic econometrics*. New York: McGraw-Hill.

Heeler, R. M., & Ray, M. L. (1972). Measure validation in marketing. *Journal of Marketing Research, 9*, 361–370.

Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of International Business Studies, 14*, 75–89.

Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: A review of four recent studies. *Strategic Management Journal, 20*, 195–204.

Jacoby, J. (1978). Consumer research: A state of the art review. *Journal of Marketing, 42*, 87–96.

Jarvis, C. B., MacKenzie, S. B., & Podsakoff, P. M. (2003). A critical review of construct indicators and measurement model misspecification in marketing and consumer research. *Journal of Consumer Research, 30*, 199–218.

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: Seminar.

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

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.

Marsh, H. W., Hau, K.-T., Balla, J. R., & Grayson, D. (1998). Is more ever too much? The number of indicators per factor in confirmatory factor analysis. *Multivariate Behavioral Research, 33*, 181–220.

McDonald, R. P. (1996). Path analysis with composite variables. *Multivariate Behavioral Research, 31*, 239–270.

Moorman, C., Deshpandé, R., & Zaltman, G. (1993). Factors affecting trust in market research relationships. *Journal of Marketing, 57*, 81–101.

Nasser, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement, 63*, 729–757.

Parry, A., & Matthews, P. M. (2002). *Functional magnetic resonance imaging (fMRI): A "window" into the brain*. Oxford, England: Oxford University Press.

Shugan, S. M. (2002). Marketing science, models, monopoly models, and why we need them. *Marketing Science, 21*, 223–228.

Spearman, C. (1904). "General intelligence," objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Wold, H. (1975). Path models with latent variables: The NIPALS approach. In H. M. Blalock, A. Aganbegian, F. M. Borodkin, R. Boudon, & V. Capecchi (Eds.), *Quantitative sociology: International perspectives on mathematical and statistical modeling* (pp. 307–357). New York: Academic.

Wold, S. (1993). Discussion: PLS in chemical practice. *Technometrics, 35*, 136–139.