2008

There's SEM and "SEM": A Critique of the Use of PLS Regression in Information Systems Research

Anne Rouse
Deakin Business School, Deakin University Burwood, Victoria, Australia, anne.rouse@deakin.edu.au

Brian Corbitt
RMIT University Melbourne, Australia, brian.corbitt@rmit.edu.au

There's SEM and "SEM": A Critique of the Use of PLS Regression in Information Systems Research

Anne C. Rouse
Deakin Business School, Deakin University
Burwood, Victoria, Australia
Email: anne.rouse@deakin.edu.au

Brian Corbitt
RMIT University
Melbourne, Australia
Email: brian.corbitt@rmit.edu.au

Abstract

In disciplines other than IS, the use of covariance-based structural equation modelling (SEM) is the mainstream method for SEM analysis, and for confirmatory factor analysis (CFA). Yet a body of IS literature has developed arguing that PLS regression is a superior tool for these analyses, and for establishing reliability and validity. Despite these claims, the views underlying this PLS literature are not universally shared. In this paper the authors review the PLS and mainstream SEM literatures, and describe the key differences between the two classes of tools. The paper also canvasses why PLS regression is rarely used in management, marketing, organizational behaviour, and that branch of psychology concerned with good measurement – psychometrics. The paper offers some practical options to Australasian researchers seeking greater mastery of SEM, and also acts as a roadmap for readers who want to check for themselves what the mainstream SEM literature has to say.

Keywords
Structural equation modelling, PLS, CFA, reliability

INTRODUCTION

The last twenty-five years have seen a growing use in business research of "structural equation modelling" or SEM. These are a class of tools for analysing the structure of measures of theoretical (unobserved) constructs (such as abilities, attitudes, and judgements) and for analysing the relationships between these constructs. Most of the developments in SEM have come from studies in psychology, where the focus is on psychological constructs. Since much of information systems (IS) research is also concerned with psychological constructs, these tools offer the potential for development of better measures, and for identification of statistically-validated theories based on complex path relationships. Drawing on mainstream texts on SEM and prior SEM training, the authors set out to explain in layman's language what SEM does and why it has been embraced so widely for research involving psychological constructs. We canvass, in particular how it differs from a tool routinely described, in Information Systems, as SEM, known as "partial least squares" (PLS) regression. Marcoulides and Saunders (2006) have criticized the tendency by some IS researchers to treat PLS regression as a "silver bullet" (p iii), and this paper responds to their editorial by alerting IS academics that alternative views exist of SEM, and of PLS regression, outside of our discipline.

DEFINITIONS — BROAD AND PARTICULAR

Exactly what is "structural equation modelling" (SEM)? At its broadest, SEM involves a number of largely linear modelling techniques that are based on solving a set of structural equations using matrix algebra. Using this definition, SEM would include most of the multivariate statistical tools including multiple regression, path analysis, factor analysis, and principal components analysis (see SEMNET, 2008b). It is in this sense that the analytical tool used widely in information systems research – partial least squares regression or PLS – can be said to be a type of "structural equation modelling". However, outside of information systems, this broad use of the label SEM is rare (a claim that can be confirmed by consulting the range of mainstream SEM texts listed in the reference section).

The broad definition of SEM above would also embrace what is more typically described as "structural equation modelling" in the disciplines of psychology, marketing, education, and management. This introduces an element of confusion because the same term denotes quite different things. In those disciplines the term SEM is reserved to mean that particular class of analytical tools that are based on iterative analysis of covariances and calculation of a “discrepancy measure”. A covariance is the unstandardized (or “raw”) form of a correlation. SEM software (including LISREL, AMOS, EQS, but *not* PLS) uses algorithms and analytical methods initially developed by the psychologist Karl Joreskog in the 1970s.

In the class of modelling tools defined by SEM (the term used in the particular), a theoretical structural equation model is developed. This model predicts, or “implies” a particular structure for the data-based covariance matrix. Once the theoretical model's parameters have been estimated using iterative matrix algebra, the software attempts to produce an implied matrix based on the theoretical model that as closely as possible matches the data. The implied (or theoretical) covariance matrix can then be compared to the *actual* covariance matrix for the data that has been gathered. If the discrepancy between these two matrices is statistically small, then the structural equation model can be considered a statistically plausible explanation for relations between the measures. SEM is thus designed to maximize, then test, the degree of consistency between the theoretical model, and the actual data (Byrne 2001; Kline 2005; Rigdon 1996).

The development of this class of statistical tools was a major breakthrough in social and business research, because till they became available, the only way of establishing whether the gathered data matched theory was to test whether it was possible to predict one or more outcomes well in the particular sample (with the assumption that this prediction would be valid in other samples). However being able to *predict* a variable (say through regression) is less powerful than being able to *account for* (explain) the whole pattern of co-variation within the gathered data. In addition, by partitioning out the effects of measurement error (discussed below), SEM reduces the attenuation in regression paths that results from this error. In other words, with SEM, it is possible to obtain more accurate measures of the path coefficients, so revealing more of the underlying relationship between latent variables.

To avoid confusion, in this paper, the authors will use the more generally-accepted *particular* definition of SEM ─ a statistical technique based on analysis of covariance structures that explicitly models measurement errors and seeks to derive unbiased estimates for the relationships between latent constructs (SEMNET, 2008b). One developer of PLS software – Wynne Chin – is critical of this particularizing of the term SEM by social scientists (Chin 1998). However, since almost no business researchers typically describe other forms of regression or factor analysis as “structural equation modelling” it is Chin’s use of the term SEM to signify the broader meaning that is atypical. Furthermore, this atypical use can lead to confusion, and to beliefs that IS researchers have mastered “SEM” when in fact they have little knowledge of SEM developments in reference disciplines.

SEM IN INFORMATION SYSTEMS VS. OTHER BUSINESS DISCIPLINES

In psychology, and in the range of business disciplines (including IS) that gather psychological data, there has been a steady increase in the use of SEM to understand the underlying structure of unobservable psychological constructs that are reflected in quantitative measures (such as survey responses). A psychological construct is a theoretical, unobserved mental concept that is reflected in statements expressed by interviewees, or by responses to survey measures. In psychology, and those disciplines that call on psychology, such as management and marketing, SEM is increasingly replacing traditional tools like exploratory factor analysis for establishing the unidimensionality, reliability and construct validity — including discriminant and convergent validity — of quantitative measures (Ketchen and Bergh, 2006).

In the quantitative IS literature, too, there has been a growing reliance on what is termed “structural equation modelling” for these purposes. However, rather puzzlingly, in our discipline the tool of choice is increasingly no longer one of the commonly used SEM software tools (which tend to all output similar results if the same estimation method is used). Instead it is increasingly PLS regression — although this PLS analysis is invariably described as “structural equation modelling”. A perhaps-surprising observation is that outside of Information Systems, the use of PLS regression is rare in mainstream “Tier 1” research journals. This is illustrated in Table 1 below. This tabulates the numbers of PLS regression studies reported in the business literature between 2004 and 2006. Table 1 is based on searches within Proquest/ABI Inform Global and Business Source Premier/Academic (electronic databases) for articles reporting the use of PLS, or CFA based on PLS. The journals in which these appear were classified according to the recent ERA (Excellence in Research for Australia) listings using John Lamp’s web-based search engine (Lamp 2008). Not all Tier 1 IS journals are captured by these databases, though Tier 1 journals in marketing and management have good coverage. This means that the method employed will tend to *understate* the proportion of PLS articles that appears in the IS literature.

Table 1 illustrates that in disciplines other than Information Systems, PLS is rarely used in Tier 1 (A* and A level) journals. Instead, most PLS studies are reported in less well-established journals (Classified as B, C, or unclassified by the ERA). Table 1 suggests that, for journals where audiences (and reviewers) are typically required to undergo substantial quantitative methods training, PLS is generally not seen as an appropriate research tool for complex multivariate research. Goodhue et al (2006) reported a similar finding based on their study of earlier (pre 2004) articles relying on PLS, published in the three highest-ranked journals within each of IS, Marketing and Management. Those authors concluded that "the use of PLS is predominantly an MIS phenomenon" (p.2). Our Table 1 suggests this pattern is becoming even more marked.

Table 1: Outlets for PLS research by discipline and ranking of journals 2004-2006

Discipline in which article published:	Total articles found	A*	A	A* or A	% All A*or A	B, C or Unranked
Information Systems	50	34	9	43	74.1%	7
Marketing	18	2	3	5	8.6%	13
Management	30	1	3	4	6.9%	26
Engineering	10	2	1	3	5.2%	7
Accounting and Economics	7		3	3	5.2%	4
TOTAL				58	100%	57

The puzzle this paper seeks to explore is why only our discipline has so wholeheartedly adopted a relatively obscure and somewhat dated tool (PLS) at the expense of those statistical tools (and methods) used in more established reference disciplines. There are, after all, substantial benefits in using a "mainstream" analytical tool for path analysis involving unobserved variables. There are support groups such as SEMNET, which provides a listserv for SEM analysts including syllabi, references, and detailed coverage of debated issues (SEMNET, 2008). There are also many sophisticated textbooks (examples include Kline 2005; Schumacker and Lomax 2004; Loehlin 2004; Marcoulides and Moustaki 2002; Byrne 2001; Raykov and Marcoulides 2000; Maruyama 1998 and Hoyle 1995) explaining the theory and mathematics behind mainstream tools, and, importantly, the statistical assumptions that underlie them. SEMNET, for example, provides a list of over a dozen texts to 1997, including reviews, and these can be supplemented by more recent publications listed at amazon.com. Furthermore, there are usually advanced training programs available in the use of mainstream tools within most universities attended by researchers from a wide range of disciplines. None of these resources are available for PLS. Another benefit of using mainstream tools is that they are widely understood (and so can be critically reviewed) by members trained in allied disciplines. However, perhaps these older, more established disciplines have "missed the plot" in relation to PLS regression? Alternatively, perhaps in embracing PLS regression so wholeheartedly, the IS discipline has begun to behave like a little boy with a hammer – seeing things to "hit" with PLS everywhere.

A NON MATHEMATICAL DESCRIPTION AND HISTORY OF PLS REGRESSION

Partial least squares regression was originated in the sixties and seventies by Herman Wold (an econometrician who helped mentor Karl Joreskog), at a time when other SEM tools were still being developed (Wold 1989). It was promoted as a tool that overcame one of the key problems with alternative regression methods (at that time the dominant form of complex statistical analysis). This was the problem introduced by multicollinearity, where predictors of the DV are highly intercorrelated (Jagpal 1982). Multicollinearity leads to instability within regression. By creating, and regressing, orthogonal "components" (a smaller number of linear composites that preserved much of the variance of the original predictors), PLS regression could overcome the multicollinearity problems. Essentially PLS transforms a group of correlated individual measures into a group of orthogonal (uncorrelated) components which are used to predict the dependent variable(s) (DVs). Components are what are also sometimes described as "composite constructs" (MacCallum and Brown 1993) but are quite different from "factors" which have a precise mathematical definition.

PLS regression was designed particularly to handle situations where very many measures (items) are gathered for a small number of cases (such as in chemometric studies). However, this is the opposite of the situation in business research, where typically only a handful of measures (or items) are reported for each case (or respondent). In business, researchers typically seek to understand the variation across cases (firms, or respondents).

In PLS regression the components are calculated using algorithms that maximise and account for both the variance of the component and the variance of the dependent variable at the same time (Garthwaite 1994; McDonald 1996; Haenlein and Kaplan 2004). These components, then, are a form of artificially-created measure whose purpose is to maximise prediction of the DVs and loadings in that particular sample, even though this prediction will probably not work as well in other samples. Furthermore, these algorithms are criticized by some statisticians as "not well understood" (Rigdon 1996; Tenenhaus 2008). This is in contrast to the algorithms underlying SEM which have been investigated and tested by dozens of researchers in the peer-reviewed literature. An important issue is that these components can be (and are) calculated even where data is missing (leading to erroneous statistics), McDonald (1996, p. 240) in an empirical exploration of the limits of PLS regression reported that:

> *The PLS methods are difficult to describe and extremely difficult to evaluate, partly because PLS constitutes a set of ad hoc algorithms that have generally not been formally analysed or shown to possess any clear global optimizing properties… and partly because … it can be difficult to determine what properties of latent variable models they possess, if any.*

The implicit assumption behind PLS is that all measured variance in the model is "useful variance", which should be explained (Chin, Marcolin and Newsted 2003). But this means that the components created by PLS account for the error variance in the sample as well as the variance of the underlying construct. It is to avoid this contamination that researchers in psychology are discouraged from using principal components analysis (PCA). "Best practice" is to instead use one of the "common factor" methods when undertaking exploratory factor analysis (Fabrigar *et al*. 1999; Gorsuch 1990).

PLS components are definitely *not* the same thing as the latent variables (or factors) calculated in SEM or exploratory factor analysis. As McDonald (1996 p 240) has observed: "there is no clear justification for [interpreting PLS components as if they were the same as latent variables or factors] since they cannot explain the covariation of their indicators except approximately". This was acknowledged by Wold (1989) who described the statistically less precise PLS regression as a form of "soft modelling". Despite this, the components calculated by PLS are routinely described, and treated, as "factors" in PLS-based journal articles. Even the AIS glossary of terms misleadingly defines PLS regression as "a second generation regression model that combines a factor analysis with linear regressions" (Straub, 2007). This imprecise terminology can be misleading to readers of research based on PLS regression, as it is often not clear that no actual factor analysis has been undertaken at all ─ meaning that the researcher does not have the rich picture of the inter-correlated relationships between dimensions that is revealed by the output of exploratory factor analysis.

In SEM, on the other hand, the latent variables are "common factors", not components. This means they are calculated based on an algebraic model that includes only the variance that is shared (or common) across all the measures that indicate the underlying construct. The distinction between *components* (essentially formative indicators that have no theoretical basis) and *common factors* –theoretical constructs that by definition are those that would account for the variance shared by indicators – is a fundamental one in psychometrics (Gorsuch 1990). However, the frequency with which these two forms of latent variables are confused in the IS literature suggests that this distinction is not well understood by IS researchers. This is revealed in Chin et al's (2003) view above that error variance is still "useful variance" that should be explained, a view that is strongly at odds with mainstream thinking in psychometrics (Gorsuch 1990).

Figure 1, below, illustrates the differences visually. Figure 1 also shows how, with factor analysis (but not components analysis), measurement error (e_1, etc) is separated out from the factor loadings. It was this capacity to separate out error variance that made SEM such a valuable tool within other reference disciplines (Kline, 2005).

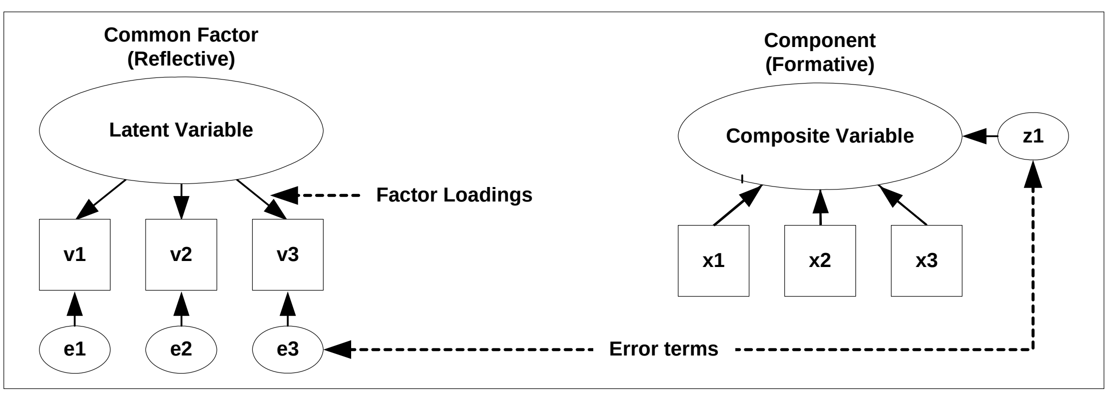

Figure 1: Structure of a common factor compared with that of a component

To illustrate the difference more pragmatically, the latent variable in the factor model shown in Figure 1 might be the construct "perceived life stress" which could be indicated by scores on items measuring such things as "perceived level of anxiety", "perceived stress at the moment" or a score for "there is too much going on in my life right now" (these items have been made up for this illustration). Here the latent variable is understood to *underlie* or cause the scores reported by the respondents. On the other hand, in the component model, the latent variable might also be a measure of "life stress" but in that case, the example items would typically be non psychological measures. They might be, for instance, scores of whether the respondent experienced job loss, divorce, or a recent accident (example taken from Bollen and Lennox 1991). In that case, there is no underlying psychological construct of "life stress" that exists apart from these scores, so what the component is really measuring is "life stress events". The supposed "latent variable" is formed, or caused, by the actual measures, and the scale that measures this is technically described not as a scale but an "index". A similar formative notion is "socio economic status" which is also a non- psychological *index* rather than a psychological scale.

(Common) factors are, by their nature, *reflective*, so, as shown in Figure 1, the arrows point from the underlying factor to the indicators which reflect, or are caused by, this factor. For example, items on a 7 point Likert measure of satisfaction would reflect the underlying level of "satisfaction" experienced by the respondent. In contrast, components, including those calculated by PLS, are *formative* – they are a linear function formed by (or caused by) a series of items, which do not need to have anything in common at all. One of the arguments for using PLS regression is that it supposedly handles formative measures better than SEM, though McDonald (1996) has demonstrated that SEM can handle formative indicators relatively easily by constraining the measurement error variances (eg z1 in Figure 1). In the IS literature, there have been puzzling examples of authors citing as a reason for using PLS regression the fact that it can handle formative indicators, even though the theoretical model reported is based largely on psychological (and hence reflective) constructs.

As (common) factor analysis (FA) is based on an algebraic model that "strips away" the error variance that is not shared amongst indicators, FA allows for more accurate measures of underlying constructs. In contrast, the loadings onto components are generally inflated by the statistical method used to create the component, giving the impression to readers that measures are more reliable at indicating an underlying construct than they really are (Gorsuch 1990). Despite these limitations, which are quite well known in psychology, in IS many researchers continue to treat (and describe) components as if they were factors (for an example, see Gefen and Straub 2005).

Because, with PLS, the "factor" loadings and the path coefficients reported by the program tend to be tightly bound to the sample from which they are derived (McDonald 1996), PLS models are less likely to reproduce well in other samples. This is a problem shared with other forms of regression, and for regression based analysis it is "best practice" to cross-validate the study with a new set of data (or a "hold out" sample) to make sure that the analysis is not capitalizing on chance (Smith and Albaum 2004, p. 671). However this is not often done in IS research.

In addition, because error is incorporated, the "factor" loadings that are reported for PLS are higher than those that would be found through exploratory factor analysis or SEM. In contrast, the size of the relationships (paths) between variables tends to be reduced because of this error (McDonald 1996; Anderson and Gerbing 1988). This means that research based on PLS, particularly when small samples are used, is likely to discount relationships that would show up with larger samples and more accurate statistical tools. When reporting non statistical findings researchers often seek for explanations, so the erroneous "explanations" can become part of the discipline's cumulative wisdom, even though the effect they seek to explain is really the result of tools and samples that do not have the power to reveal the relationship.

The overstated component loadings also have implications for the measurement model produced by PLS, that is, the relationship between indicators and the underlying constructs. Calculations of *average variance extracted*, and *reliability*, are based on these loadings, and so also overstated when PLS regression is used. The overstated loadings also have implications for measures of *discriminant validity*, which typically are calculated by comparing the average variance extracted with the (squared) between-variable correlation. In PLS the former is larger than it should be, while the latter is smaller. In sum, PLS regression will tend to make measures appear more reliable, and more discriminating, than would be the case if other mainstream tools (like exploratory factor analysis, or SEM-based confirmatory factor analysis) were used. Again, erroneous findings may become part of the discipline's cumulative research base. This overstatement should eventually be picked up in the literature if the research was replicated, but replication studies are rare in Information Systems.

A more fundamental problem with PLS is that unlike exploratory or confirmatory factor analysis, PLS regression does not reveal the underlying dimensionality of the data – so it does not test that the components it creates are unidimensional (where responses all tend to move in the same direction). Unidimensionality means there is only one construct underlying the data, not several. Since measures of reliability (like Cronbach alpha) rely on (and assume *a priori*) the fact that the data being analysed is unidimensional, this also means that the reliability scores reported from PLS regression may be incorrect. Unidimensionality is a core requirement for "good" psychometric measures because otherwise some items making up the scale might correlate (with other measures) in a different way to the rest of the scale. Just as validity cannot be established in the absence of reliability, reliability cannot be accurately established if a scale is not unidimensional. However, because of the nature of the algorithms used, unidimensionality cannot be measured with PLS regression; instead it is assumed to be there *a priori* (Gerbing and Anderson 1988)

Finally, although PLS regression is described as being a tool for “confirmatory factor analysis” (CFA) this too is based on a confusing use of terminology. CFA is an analytical technique using SEM software (such as AMOS, Lisrel or EQS) where the focus is only on the relationship between observed measures and their underlying unobserved construct – what SEM writers call the “measurement model”. CFA is based on the underlying “common factor” mathematical model. Using CFA it is possible to statistically test the fit between a factor model and the pattern of covariances (as discussed above) by considering various “goodness of fit” indicators based on the discrepancy between the theoretical model and the actual data. It is because of this statistical testing that the factor analysis is “confirmed”. PLS regression does not offer this statistical testing, so does no confirmation, and it cannot confirm “factors” as it does not calculate common factors.

Table 2, below, summarises the key differences between PLS regression and mainstream SEM discussed above.

Table 2: Comparison of Structural Equation Modelling and PLS Regression

SEM	PLS
Technique used widely in many disciplines, substantial psychometric research undertaken since the seventies on underlying algorithms and their limitations	Technique largely ignored by psychometric literature as outdated and poorly understood; underlying algorithms not widely published nor tested to establish limitations
Calculation of unobserved variables is based on the well understood common factor statistical model	Calculation of unobserved variables based on creation of components (linear combinations) using obscure algorithms. Components designed to maximise fit with particular sample, so measurement and structural models may not generalize well.
Focus is on explaining the pattern of covariances in the data using unbiased statistical estimates and tests of goodness of fit (see below).	A form of *regression*, where the focus is on predicting the DVs within the particular sample (with the assumption that the predictions will hold in other samples). Shares the limitations of other forms of regression.
Allows statistical testing of the hypotheses that model paths are different from zero	Allows approximate testing (using bootstrap procedures) of the hypotheses that paths are different from zero
Statistical testing of the correspondence between theoretical model and sample covariances possible, producing a wide range of indicators of “goodness of fit”	No overall model fit testing possible. Only indicator of goodness of fit is R^2 predicted (which does not reveal whether model fits the data well).
Allows for confirmatory factor analysis where number of factors and factor structures are tested statistically	Does not create factors; does not test “factor” (i.e. component) numbers or loadings statistically.
Does not over inflate factor loadings. Statistical testing of the factor (measurement) model across sub-samples possible if sample large enough.	Inflates component loadings due to the inclusion of error variance Produces attenuated path loadings due to this inclusion Does not provide for testing generalisability of the measurement model
Does not overstate reliability and average variance extracted for unobserved factors	Tends to overstates these, suggesting measures are more reliable and valid than they may actually be Produces a poor measurement model (cf Hair 2006)
Allows for statistical confirmation of factor structure (CFA) and correlation between constructs, so allowing statistical testing of measure unidimensionality	Does not test for unidimensionality
Needs approximately 150-200 responses (depending on reliability) to produce findings that will generalise May produce non-significant findings for large samples (over approximately 400) but this can be addressed by splitting sample and testing robustness of model across samples.	Needs several hundred responses to produce path results that will generalise to other samples (cf Marcoulides and Saunders 2006) Note: Only structural model (paths) was tested by those authors

CRITICAL CONSUMPTION OF THE PLS LITERATURE

As part of a review of the PLS literature that is currently ongoing, the authors classified papers reporting, or encouraging, use of PLS regression into three categories — illustrated in Table 3. This classification can act as a tool to encourage more discriminating consumption of PLS regression journal articles. The first category of paper basically reports the *use* of PLS regression as an analytical tool for IS research, often with a short justification that is based on claims in papers from the second category. The second, *promoter* category includes articles that highlight the benefits of PLS as an analytical tool by asserting a number of attributes of PLS regression when compared to mainstream tools. Examples include articles by Chin, Gefen, Straub and colleagues, and, in marketing, Fornell and Bookstein. In seeking to promote PLS as an analytical tool, these articles often downplay the benefits of alternatives, or exaggerate problems with other SEM tools.

The third category of paper involves a *critique* of PLS regression, in the sense that these papers test the boundaries or limitations around claimed benefits of PLS (viz a viz alternatives). Most of theses critiques have targeted only the structural (path) model, rather than the quality of the measurement model.

Table 3: Classification scheme applied to typical journal papers reporting use of PLS

Typical Examples	Using PLS	Promoting PLS	Critiquing PLS
Van Slyke *et al* (2007); Saade and Bahli (2005). Enns *et al* (2003) – used PLS for measurement model; Meso *et al* (2005).	X		
Chin (1995; 1997, 1998) Fornell (1982; 1987); Fornell and Larckner (1981); Fornell and Bookstein (1982); Gefen *et al* (2000); Gefen and Straub (2003);		X	
Chin et al (2003); Goodhue et al (2006); Marcoulides and Saunders (2006); Hsu *et al* (2006).			X

Few of the large number of PLS articles within the IS literature belong to the third class (critique). This contrasts with the psychological-statistical literature (including marketing literature which draws on this material) where there have been a large number of peer-reviewed “critique” articles exploring and testing the limitations of tools like regression, factor analysis and SEM. Papers in the first category (*reports* of use of PLS regression) rely almost exclusively on quotes from other “use” papers or “promoter” papers as justification for their reliance on PLS. Given this, we suggest that authors not rely on these as justification for the use of PLS, as they tend to merely repeat what has been asserted in the “promoter” category of articles.

The promoter group of articles can be further subdivided into three categories – firstly there were the early mathematical articles describing the technique (eg Wold 1989). Few IS readers would have the training to work through these. Secondly there are a handful of articles written in the marketing literature in the eighties and nineties about PLS regression and its reported advantages to business researchers (eg Fornell 1981; 1982). These largely form the “evidence base” that has been used to justify the benefits of PLS, however they were at odds with alternative views from the worlds of psychology and marketing (for examples, see Anderson and Gerbing 1988; McDonald 1996 and SEM textbooks listed in the References section). These alternatives had come to be seen as mainstream by the end of the nineties. More recently (and third) there have been a series of articles written largely by IS specialists about how PLS regression can be used to carry out the processes typically (in other disciplines) undertaken using a mix of exploratory factor analysis, reliability analysis, and SEM (eg Gefen and Straub 2005; Chin 1995; 1997; 1998). A key author in this third category is Wynne Chin of the University of Houston. Chin is the developer of PLS-GRAPH, one of the PLS regression software packages. Many of the PLS regression analyses undertaken within Information Systems uses a beta test version of PLS-Graph he has made available. In the main these articles can be seen as “proselytizing” in that they illustrate how traditional techniques for establishing reliability and validity and path relationships might be done within PLS instead. The label “proselytizing” is used because the more established methods tend to be ignored, or criticized, although many of the claims made for problems associated with traditional tools do not acknowledge well-known solutions that have been published in the psychometric literature. The existence of this literature poses a problem for the IS discipline, because many doctoral students will regard this “promoter” literature (and particularly the tutorials) as reflecting mainstream business research views, when this is not the case at all, if the wider reference disciplines making up “business” are considered.

We believe there is merit in encouraging consumers of PLS regression research to consider what class of article has been written, and to take note that many of the claims related to the tool are based essentially on those in the *user* or *promoter* categories, rather than on empirical evidence. In most cases these promoter articles consist largely of assertions, many of which are treated cautiously by those with deep SEM training — for an example, see the carefully worded treatment by Hair *et al*. (pp 878-880) in the brief discussion of PLS in their SEM chapter. We also encourage readers to compare the relatively small samples being reported in the PLS literature against the Monte Carlo studies reported by Marcoulides and Saunders (2006) and by Goodhue et al (2006) — both of whom found that PLS, like other mainstream tools, needs relatively large samples to produce viable results. Since these messages might seem to be at odds with several widely-read “promoter” articles (and tutorials), we urge readers who are not convinced by our own paper to pursue the details themselves using the citations provided.

CONCLUSIONS

Few IS researchers have exposure to formal psychometric training based in alternative reference disciplines, and few have deep (multi-year) training in SEM methods. Consequently, many of the claims made for PLS have, until Goodhue et al's (2006) and Marcoulides and Saunders' (2006) articles, remained unchallenged in the IS literature. Inappropriate use of terms like "structural equation modelling" or "factor" may have served to forestall challenge, as they give the appearance that PLS is well grounded in traditional psychometric concepts, even though their idiosyncratic use by PLS proponents can serve to add confusion. So one conclusion we have drawn is that the IS literature needs to be far more precise in its use of terms that have been taken from reference disciplines. PLS is a form of regression, and is not "SEM"; it calculates and regresses components, not "factors".

IS is increasingly now addressing concerns (such as those related to e-business, HCI and IT strategy) that overlap those addressed by allied disciplines like marketing, management or psychology. So IS researchers publishing in these areas may well find their research being reviewed by specialists with substantial mainstream SEM training. Given this, it makes sense for IS researchers to become aware of the way SEM is treated in those older, and in terms of statistical sophistication, more established disciplines. At present, few in those disciplines will be familiar with PLS regression, and particularly in the non-Tier 1 journals, tend to leave unchallenged those claims made in their own publications for PLS. These claims are usually justified on the basis of IS articles promoting the tool (Goodhue et al, 2006). However, this situation is unlikely to continue as more researchers (eg Hair *et al.* 2006) articulate the limitations of PLS regression.

On the surface, PLS regression is certainly an attractive option for IS researchers, as the technique makes few demands. According to proponents it is possible to do high quality theory-testing research using PLS. It is also relatively easy to "establish" reliability and validity using the tool without having to master the intricacies of principal components analysis, covariance, factor analysis, or multiple regression. A researcher with minimal education in multivariate statistics can begin getting output from PLS with only a day's training in use of PLS's syntax. In addition, there are tutorials available (Gefen and Straub 2005; Chin 2002) which, if followed specifically, will produce an academic paper that appears to use the conventions developed within psychology (and marketing) for establishing reliability and validity.

However, those with deep training in SEM will soon recognize that PLS-based tests of reliability and validity are not the mainstream methods used elsewhere. Furthermore, PLS regression has the potential to produce results which may not generalise to alternative settings, making it problematic when used to *test* theory (a problem that is exacerbated in a discipline that does not appear to encourage replication studies). PLS-based research may also present what are psychometrically quite poor measures as reliable and valid – leading later researchers to uncritically rely on these in their own studies. This has implications for the discipline's cumulative research tradition.

It is now well accepted in mainstream psychometric and marketing literature that SEM, and even regression (including PLS regression) will, with small samples, tend to produce estimates that will never generalise beyond that sample because the regression statistics (e.g. p values) are poor approximations for the true values. This makes small samples (under around 100 to 120 responses, depending on the variance) largely inadequate for testing theory (see Allison and Allison 1999, pp. 57-59). It is for this reason that researchers should be encouraged to replicate small-sample studies (even those published in Tier 1 journals) rather than to treat them as definitive. Unfortunately, papers still appear asserting that PLS can be used effectively with quite small numbers, despite the recent studies (Goodhue *et al.*,2006; Marcoulides and Saunders 2006; Hsu *et al*. 2006) that provide empirical evidence contradicting this. This is a worrying trend as IS doctoral students may proceed with analyses of small samples that might never result in publication, based on outdated advice and practice.

Before adopting PLS regression as the "tool of choice" for multivariate analysis we suggest that IS researchers (particularly doctoral candidates) obtain training in mainstream methods, including multiple regression, factor analysis and SEM. They will then be in a position to realistically evaluate the alternative analytical methods available and to critically consume assertions made about PLS. Australasian research students are particularly well-resourced because here such training is becoming increasingly widespread. Many psychology, statistics and marketing departments deliver courses in these topics, and the highly-regarded ACSPRI consortium (Australian Consortium for Social and Political Research) provides intensive training programs in Melbourne, Canberra and Queensland several times a year. Alternatively, IS researchers could work in collaboration with those from disciplines who have obtained much deeper training in SEM and CFA than is typical for IS specialists.

In addition, we suggest that the IS discipline in Australasia encourage our researchers to become more critical consumers of multivariate (and particularly SEM) research. The Marcoulides and Saunders (2006) article suggests that this is beginning to happen in the US. We should also be encouraging a more sceptical attitude to assertions made even in the Tier 1 IS literature, and encouraging researchers to check these claims with their cross-discipline colleagues with greater psychometric training. As an aside, there is a substantial gulf between those who have mastered basic multivariate methods (like regression and PCA) and those who have mastered the psychometric evaluation of measures. Ideally, those from whom advice is sought would have this latter expertise.

As this paper has demonstrated, there is certainly *not* universal agreement about the benefits of PLS regression over other methods (like SEM) for studying psychological constructs. This is most tellingly illustrated by the absence of treatment of PLS regression in virtually all leading texts on SEM (many of which we have included in the references). Hair *et al* (2006) is one of the few multivariate texts to discuss PLS, but the authors conclude that it is not an appropriate alternative for SEM. Essentially the tool is promoted by only a relatively small group of business researchers (rather than those trained in psychometrics), particularly Chin, Gefen and Straub in Information Systems, and, in the earlier marketing literature, Fornell and Bookstein.

PLS regression may well be a useful tool for firm-level measures and econometric data, which often involve formative indicators and are not concerned with psychological constructs. PLS may also be appropriate to rescue a study when an SEM model fails to converge, though there are often alternative strategies that can be employed to avoid such failure. However, as Hair *et al* (2006) warn, if PLS is used, researchers should acknowledge its limitations; in particular, the recognized problems with the measurement model (the loadings, reliability and validity of measured constructs). PLS researchers also need to use adequately sized samples, and be precise in their use of statistical terms.

As Marcoulides and Saunders have suggested, IS researchers should stop unquestioningly treating PLS as a statistical "silver bullet" or, using our metaphor, as an "all purpose hammer". Those considering its use in their research projects are encouraged to use this paper as a roadmap to other, and generally more mainstream, literature on SEM and CFA, and to consult this literature themselves before relying on arguments (including ours) about the limits of statistical tools. Certainly, before relying on PLS for substantial research projects (like a doctoral thesis) we recommend mainstream SEM training, so that any decision to use PLS is a fully-informed one.

REFERENCES

Allison, P. D. and Allison, S. 1999. *Multiple Regression: A Primer*. Pine Forge Press.

Anderson, J. and Gerbing, D. 1988. "Structural Equation Modelling In Practice: A Review And Recommended Two-Step Approach", *Psychological Bulletin*, (103:3), pp. 411-23.

Bollen, K. and Lennox, R. 1991. "Conventional Wisdom on Measurement: A structural Equation Perspective". *Psychological Bulletin* (110:2), pp. 305-314.

Byrne, B. M. 2001. *Structural Equation Modelling with AMOS. Basic Concepts Applications and Programming*. Routledge.

Chin, W. W. 1995. "Partial Least Squares Is To LISREL As Principal Components Analysis Is To Common Factor Analysis". *Technology Studies, (*22), pp. 315-319.

Chin, W. W. 1997. "Overview of the PLS Method" Retrieved June 1, 2008, from http://disc-nt.cba.uh.edu/chin/PLSINTRO.HTM

Chin, W. W. 1998. "Issues and Opinions on Structural Equation Modelling". *MIS Quarterly*. (22:1). pp. vii – xvi.

Chin, W. W., Marcolin, B. L., and Newsted, P. R. 2003. " Partial Least Squares Latent Variable Modelling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and Electronic Mail Emotion/Adoption Study", *Information Systems Research (*14:2) pp. 189-217.

Enns, H. Huff, S. and Higgins, B. R. 2003."CIO influence behaviors: The Impact of Technical Background". *Information and Management*. (40**:**5)**.** pp. 467-485.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., and Strahan, E. J. 1999. "Evaluating the Use of Exploratory Factor Analysis in Psychological Research". *Psychological Methods, (4)*, pp. 272-299.

Fornell, C. 1987. "A Second Generation of Multivariate Analysis: Classification of Methods and Implications for Marketing Research", in *Review of Marketing*, M. J. Houston ed., Chicago: American Marketing Association, , IL, pp. 407-450.

Fornell, C. 1982. *A Second Generation of Multivariate Analysis, Vol. 1: Methods*. New York: Praeger.

Fornell, C., and Bookstein, F. 1982. "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory", *Journal of Marketing Research*, (19), pp. 440-452.

Fornell, C., and Larckner, D. F. 1981. "Structural Equation Models with Unobserved Variables and Measurement Error". *Journal of Marketing Research. (18).* pp. 39-50.

Garthwaite, P. H., 1994. "An Interpretation of Partial Least Squares", *Journal of the American Statistical Association*, (89:425), pp. 122-27

Gefen, D and Straub, D. 2005. "A Practical Guide to Factorial Validity Using PLS-Graph:Tutorial and annotated example". *Communications of the AIS*. (16). pp. 91-109.

Gefen, D., Straub, D. and Boudreau, M. 2000. "Structural Equation Modelling Techniques and Regression: Guidelines for Research Practice", *Communications of the Association for Information Systems (*7:7), pp. 1-78

Goodhue, D. Lewis, W. and Thompson, R. 2006. "PLS small Sample Size and Statistical Power in MIS Research". *Proceedings of 39th HICSS Conference:* Hawaii. January. Gorsuch, R. L. 1990, "Common Factor Analysis Versus Component Analysis: Some Well and Little Known Facts", *Multivariate Behavioral Research*, (25:1)`, pp. 33-39.

Haenlein, M. and Kaplan, A. M. 2004. "A Beginers' Guide to Partial Least Squares Analysis". *Understanding Statistics*. (3:4). pp. 283-297.

Hair, J. F., Black, W. C., Babin, B. J. Anderson, R. E, and Tatham, R. L. 2006. *Multivariate Data Analysis*. 6th ed. Pearson International.

Hoyle, R. H. 1995. *Structural Equation Modelling: Concepts, Issues and Applications*. Newbury Park: Sage.

Hsu, S-H, Chen, W-H, and Hsieh, M-J. 2006. "Robustness Ttesting of PLS, LISREL, EQS and ANN-based SEM for Measuring Customer Satisfaction". *Total Qualty Management*. (173), pp. 355-371.

Jagpal, H. S. 1982. "Multicollinearity in Structural Equation Models with Unobserved Variables", *Journal of Marketing Research*, (194), pp. 431-439

Ketchen, D. J. and Bergh D. D. 2006. *Research Methodology in Strategy and Management, Vol 3*. Emerald.

Kline, R. B. 2005. *Principles and Practice of Structural Equation Modelling*. Guilford Press.

Lamp, J. 2008. "ERA Journal Access". Retrieved June 28, 2008, from http://lamp.infosys.deakin.edu.au/era/.

Loehlin, J. C. 2004. *Latent Variable Models: An Introduction to Factor, Path and Structural Equation Analysis. 4th Edition*. Lawrence Erlbaum Associates.

MacCallum, R. C., and Browne, M. W. 1993. "The Use of Causal Indicators in Covariance Structure Models: Some Practical Issues". *Psychological Bulletin,(11:4*3), pp. 533-541

Marcoulides, G. A. and Moustaki, I. 2002. *Latent Variable and Latent Structure Models*. Mahwah, N.J. Lawrence Erlbaum Associates.

Marcoulides, G. A. and Saunders, C. 2006. "PLS: A Silver Bullet?", *MIS Quarterly, (30*:2), pp. iii-ix.

Maruyama, G. M. 1998. *Basics of Structural Equation Modeling*. Newbury Park: Sage.

McDonald R.P. 1996." Path Analysis with Composite Variables", *Multivariate Behavioral Research*, (*31*:2), pp. 239-270.

Meso, P., Musa, P., and Mbarika, V. 2005."Towards a Model of Consumer Use of Information and Communication Technologies in LDES: The Case of sub-Saharan Africa". *Information Systems Research*. (15)., pp. 119-146.

Raykov, T.and Marcoulides, G. A. 2000. *A First Course in Structural Equation Modelling*. Lawrence Erlbaum Associates.

Rigdon, E. 1996."Methodological Alternatives to SEM/CFA. SEMNET FAQ", Retrieved June 1, 2008, from.http://www2.gsu.edu/~mkteer/relmeth.html

Saade, R. and Bahli, B. 2005. "The impact of Cognitive Absorption On Perceived Usefulness and Perceived Ease of Use in On-Line Learning: An Extension of the Technology Acceptance Model". *Information and Management*. (422). pp. 317-327.

Schumacker, R. E. and Lomax, R. G. 2004 *A Beginner's Guide to Structural Equation Modelling*. Routledge

SEMNET. 2008 "The Structural Equation Modelling Discussion Network". Retrieved June 1, 2008, from http://www2.gsu.edu/~mkteer/semnet.html.

SEMNET. 2008b. "SEMNET FAQ". http://www2.gsu.edu/~mkteer/semfaq.html [accessed 1 June, 2008]

Smith, S. M. and Albaum, G. S. 2004 *Fundamentals of Marketing Research* Newbury Park: Sage.

Straub, D. 2007. "AIS Dictionary of Terms".Retrieved June 1, 2008, from http://dstraub.cis.gsu.edu:88/quant/7glossary.asp

Tenenhaus, M. 2008. *Components Based Structural Equation Modelling. HEC Working Paper 887/2008*. HEC Paris, Jouy-en-Josas. Retrieved 25 June, 2008, from

http://www.hec.fr/hec/eng/professeurs_recherche/r_cahiers.php

Van Slyk, C., Ilie, V., Lou, H. and Stafford, T. 2007. "Perceived Critical Mass and the Adoption of a Communication Technology". *European Journal of Information Systems*. (16). pp. 270-283.

Wold, H. 1989. "Introduction to the Second Generation of Multivariate Analysis", in H. Wold Ed., *Theoretical Empiricism*. New York: Paragon House, pp. vii-xi.

COPYRIGHT