Strategic Management Journal
Strat. Mgmt. J., **20**: 195–204 (1999)

USE OF PARTIAL LEAST SQUARES (PLS) IN STRATEGIC MANAGEMENT RESEARCH: A REVIEW OF FOUR RECENT STUDIES

JOHN HULLAND*
Richard Ivey School of Business, The University of Western Ontario, London, Ontario, Canada

Advances in causal modeling techniques have made it possible for researchers to simultaneously examine theory and measures. However, researchers must use these new techniques appropriately. In addition to dealing with the methodological concerns associated with more traditional methods of analysis, researchers using causal modeling approaches must understand their underlying assumptions and limitations.

Most researchers are well equipped with a basic understanding of LISREL-type models. In contrast, current familiarity with PLS in the strategic management area is low. The current paper reviews four recent studies in the strategic management area which use PLS. The review notes that the technique has been applied inconsistently, and at times inappropriately, and suggests standards for evaluating future PLS applications. Copyright © 1999 John Wiley & Sons, Ltd.

Advances in causal modelling techniques have made it possible for researchers to simultaneously examine theory and measures. Such techniques can be thought of as superior to more traditional techniques (e.g., multidimensional scaling, factor analysis) in that they permit: (1) the explicit inclusion of measurement error, and (2) an ability to incorporate abstract and unobservable constructs (Fornell, 1982). Bagozzi (1980) suggests that causal models provide researchers with four key benefits: (1) they make the assumptions, constructs, and hypothesized relationships in a theory explicit; (2) they add a degree of precision to a theory, since they require clear definitions of constructs, operationalizations, and functional relationships; (3) they permit a more complete representation of complex theories; and (4) they provide a formal framework for constructing and testing both theories and measures.

The best-known causal modeling technique is LISREL (Jöreskog and Sörbom, 1989; Hagedoorn and Schakenraad, 1994). However, LISREL is poorly suited to deal with small data samples (Fornell, 1982), and can yield nonunique or otherwise improper solutions in some cases (Fornell and Bookstein, 1982). An alternative causal modeling approach known as Partial Least Squares (PLS) has been developed to avoid some of these limitations (Wold, 1974, 1985), although use of PLS requires its own set of assumptions. PLS has been used both in other business disciplines (e.g., Duxbury and Higgins, 1991; Hulland and Kleinmuntz, 1994; Smith and Barclay, 1997; Zinkhan, Joachimsthaler, and Kinnear, 1987) and in the strategic management area to examine risk–return outcomes (Cool, Dierickx, and Jemison, 1989), cooperative ventures (Fornell, Lorange,

Key words: causal modeling; measurement issues; partial least squares; strategic management research issues

* Correspondence to: John Hulland, Richard Ivey School of Business , The University of Western Ontario, London, Ontario N6A 3K7, Canada

CCC 0143–2095/99/020195–10 $17.50

Received 4 November 1996
Final revision received 7 April 1998

and Roos, 1990), global strategy (Johansson and Yip, 1994), and global integration (Birkinshaw, Morrison, and Hulland, 1995).[1]

While techniques such as LISREL and PLS can enrich existing methodological approaches to conducting strategic management research, they must be used appropriately. Most researchers are well equipped with a basic understanding of LISREL-type models. In contrast, current familiarity with PLS in the strategic management area is low, making it difficult for most researchers to properly evaluate its use by others. Exacerbating this problem of limited familiarity, existing applications of PLS in the strategic management area have used the technique inconsistently, and at times inappropriately.

To help shape future application of PLS in the strategy area, the following discussion reviews use of the technique in four recent studies (Birkinshaw *et al*., 1995; Cool *et al*., 1989; Fornell *et al*., 1990; Johansson and Yip, 1994). This discussion is organized around two general sets of issues: conceptual and methodological. As Bagozzi (1984) noted, theory and measurement are intimately intertwined, and both must therefore be considered in a causal modeling context.

CONCEPTUAL ISSUES

Although measures and theory are both important, the process of model specification necessarily begins by considering the theoretical model underlying a particular piece of research. That is, the causal modeling process begins at the conceptual level. Many of the conceptual issues which apply to all empirical research are also relevant in a causal modeling context. However, we focus here on three issues that play a particularly important role in causal modeling research: conceptual model specification, construct dimensionality, and the distinction between constructs and measures.

Conceptual model specification

The studies by Birkinshaw *et al.*, Cool *et al.*, and Fornell *et al*. each present and test a single conceptual model. In contrast, Johansson and Yip specify (and estimate) eight different conceptual models which follow five seemingly different basic forms. There is nothing inherently wrong in making use of alternative models. Indeed, in the early stages of theory refinement such comparisons often play a critical role. However, the approach used by Johansson and Yip in presenting their various models is largely ad hoc—it is not clear why some models have been included while others have not. Furthermore, their eight models can all be represented as special cases of one of the two more general models shown in Figure 1.[2] Thus, it would be more appropriate to first estimate the two general models (i.e., 'unrestricted' Models A and B), and to then estimate each of the more specific variants (i.e., 'restricted' models). Although R^2 values and path coefficients from entirely different theoretical models are not directly comparable, when one model is a subset of (or 'nested within') a more general model the researcher can compare R^2 values on the key endogenous construct(s) using a Chow test. For Johansson and Yip's study, this approach would permit direct comparisons between Models 1, 2, 3, 4, and 6, and between Models 5, 7, and 8.

Construct dimensionality

The dimensionality of constructs is a critical consideration in the development of causal models. In many cases, researchers can correctly assume that their constructs are unidimensional. However, when a particular construct is more properly conceptualized as multidimensional (e.g., business strategy according to Venkatraman, 1989),

[1] For an excellent introduction to the use of PLS in practical applications, see Barclay, Higgins, and Thompson (1995). Barclay *et al*. also provide a thorough description of the objectives of PLS, the estimation process, sample size issues, the relative strengths and weaknesses of both LISREL and PLS, and the availability of software. For a good discussion of the model structure employed by PLS, see Cool *et al*. (1989).

[2] Specifically, Models 1, 2, 3, 4, and 6 in Johansson and Yip (1994) are all nested within Model A shown in the figure. For example, their Model 1 assumes that only paths 2, 7, 9, and 10 are nonzero, Model 2 assumes that only paths 3, 6, 8, and 10 are nonzero, and Model 3 assumes that only paths 2, 3, 6, 7, 8, 9, and 10 are nonzero. Model 4 assumes that paths 1, 3, 5, 6, 7, 8, and 10 are nonzero, while Model 6 assumes that paths 1, 3, 5, 6, 8, 10, 11, and 12 are nonzero.

Models 5, 7, and 8 are all nested within model B of the figure. Specifically, Model 5 assumes that paths 4, 5, and 6 are zero, Model 7 assumes that path 4 is zero, and Model 8 assumes that path 5 is zero.

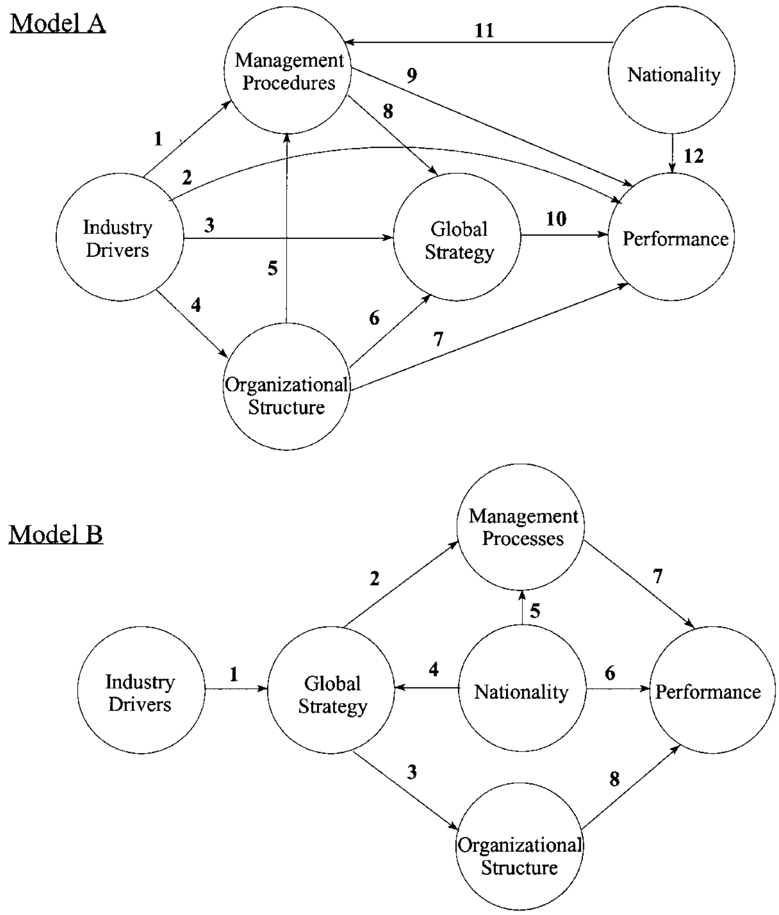

Figure 1. Two comprehensive conceptual models encompassing all eight models estimated in Johansson and Yip

researchers' causal models should include separate constructs representing each of these dimensions.

In general, the four studies reviewed here assume and use construct unidimensionality. However, an interesting contrast in approaches to this issue of dimensionality can be seen by comparing the treatments of the industry structure construct by Johansson and Yip and by Birkinshaw *et al*. Early discussion of their constructs by Johansson and Yip (1994: 580–582) strongly suggests that industry structure, global strategy, organization structure, management processes, and performance are all multidimensional. For example, they clearly identify four separate and distinct 'drivers' in their discussion of the industry structure construct (i.e., government drivers, cost drivers, market drivers, and competitive drivers). Intuitively, it seems likely that each of these four drivers could have a different impact on the other constructs (e.g., global strategy). However, in the empirical section of their paper, Johansson and Yip ignore these distinctions, using instead a single, higher order (but conceptually undefined) 'industry drivers' construct. In contrast, Birkinshaw *et al*. include three separate 'driver' constructs in their model in order to look at their individual effects on other strategic constructs. This latter approach is more consistent with existing theory while also yielding superior empirical results.

Constructs versus measures

Although related to one another, constructs and measures are distinct entities (Bagozzi, 1984). Much of the work in the four studies reviewed here properly separates the theoretical and empirical planes. However, there are two exceptions worthy of further discussion. First, although Birkinshaw *et al*. provide theoretical definitions for their constructs, two (economies of scale and differences in comparative advantage) are measured empirically by only single items. From a causal modeling perspective, these measures and their associated constructs must be viewed as interchangeable. Clearly, it is more appropriate to include multiple measures for each construct.

Second, Johansson and Yip are inconsistent in how they deal with the effects of 'nationality' on their other constructs. They appear to be ambivalent about whether nationality is more correctly viewed as 'an additional construct' (1994: 593) or as 'a possible method bias' (p. 586), employing both perspectives in their model estimations. Although such tentativeness is understandable in an exploratory study, it renders moot comparisons between models. Researchers need to clearly separate constructs from measures in order to properly test the nomological validity of both.

METHODOLOGICAL ISSUES

Three general sets of methodological considerations are relevant to the application of PLS in a management research context: (1) assessing the reliability and validity of measures; (2) determining the appropriate nature of the relationships between measures and constructs; and (3) interpreting path coefficients, determining model adequacy, and selecting a final model from the available set of alternatives. Each of these is dealt with below.

Reliability and validity

Although PLS estimates parameters for both the links between measures and constructs (i.e., loadings) and the links between different constructs (i.e., path coefficients) at the same time, a PLS model is usually analyzed and interpreted sequentially in two stages: (1) the assessment of the reliability and validity of the measurement model, followed by (2) the assessment of the structural model. This sequence ensures that the researcher has reliable and valid measures of constructs before attempting to draw conclusions about the nature of the construct relationships.

The adequacy of the measurement model can be assessed by looking at: (1) individual item reliabilities, (2) the convergent validity of the measures associated with individual constructs, and (3) discriminant validity.

Item reliability

In PLS, individual item reliability is assessed by examining the loadings (or simple correlations) of the measures with their respective construct. A rule of thumb employed by many researchers is to accept items with loadings of 0.7 or more, which implies that there is more shared variance between the construct and its measure than error variance (e.g., Carmines and Zeller, 1979). Since loadings are correlations, this implies that more than 50 percent of the variance in the observed variable (i.e., the square of the loading) is due to the construct.

In practice, it is common to find that at least several measurement items in an estimated model have loadings below the 0.7 threshold, particularly when new items or newly developed scales are employed. A low loading may be the result of: (1) a poorly worded item, (2) an inappropriate item, or (3) an improper transfer of an item from one context to another. The first problem leads to low reliability, the second to poor content (and construct) validity, and the last to nongeneralizability of the item across contexts and/or settings. Even when the researcher has a strong theoretical rationale for including such items in his or her model, items with extremely low loadings should be carefully reviewed, since they will add very little explanatory power to the model while attenuating (and therefore biasing) the estimates of the parameters linking constructs (Nunnally, 1978). In general, items with loadings of less than 0.4 (a threshold commonly used for factor analysis results) or 0.5 should be dropped.

The evidence of high individual item reliability across the four studies reviewed here is somewhat mixed. Although Birkinshaw *et al*. do not report individual item reliabilities, they note (1995: 647) that only items with 'individual factor loadings greater than 0.6 were retained, with most greater than 0.7.' Of the 18 measurement items included in the study by Cool *et al*., only one had a loading of less than 0.7, and it was retained for sound theoretical reasons. On the other hand, four of the 18 item loadings reported by Fornell *et al*. are less than 0.4, six are less than 0.5, and only six exceed the 0.7 threshold. Similarly, for the models reported in Johansson and Yip (1994), about one third of the loadings consistently fall below 0.4, whereas just over half exceed 0.7.[3]

Thus, while most of the measures used in these four studies appear to exhibit satisfactory individual item reliabilities, two of the four studies retained a significant number of low-reliability items in their final analyses. Since these low reliabilities can attenuate the estimated relationships between constructs, results based on the retention of low-reliability items must be interpreted with caution.

Convergent validity

When multiple measures are used for an individual construct, the researcher should be concerned not only with individual measurement item reliability, but also with the extent to which the measures demonstrate convergent validity. Traditionally, researchers using PLS have generally reported one or both of two measures of convergent validity (also referred to as composite reliability): Cronbach's alpha and the internal consistency measure developed by Fornell and Larcker (1981).[4] Fornell and Larcker argue that their measure is superior to alpha since it uses the item loadings obtained within the nomological network (or causal model). Nonetheless, the interpretation of the values obtained is similar, and the guidelines offered by Nunnally (1978) can be adopted for both. Specifically, Nunnally suggests 0.7 as a benchmark for 'modest' composite reliability, applicable in the early stages of research.[5]

Although it is not possible to calculate Cronbach's alphas for the individual constructs in any of the models reviewed here, internal consistency values were determined by using the reported loadings and Fornell and Larcker's formula. These values are reported in column one of Table 1 for three of these studies. (Birkinshaw *et al.* report internal consistency values in their Table 2).

Broadly speaking, the convergent validities of the constructs used by Birkinshaw *et al.*, Cool *et al.*, and Fornell *et al*. appear to be acceptable. In contrast, most of the constructs used by Johansson and Yip exhibit poor convergent validity. For Models 6 and 8 (their two most preferred models), only the measures used to assess performance demonstrate sufficient internal consistency to exceed Nunnally's 'modest' standard of 0.70. At the same time, two sets of measures (those representing organizational structure and global strategy) exhibit extremely low internal consistency (i.e., less than 0.5).

Low internal consistency can result from a variety of underlying causes, including poor construct definition and/or construct multidimensionality. In the first case, the faulty construct definition severely impairs the determination of relevant and appropriate measures for the construct. In the latter case, if the underlying construct is actually multidimensional, but it is measured using items which are assumed to be linked to a unidimensional construct, the measures as a group will demonstrate poor internal consistency. Furthermore, while some of the individual measurement items will have strong loadings linking them closely with the construct, others will have loadings which are close to zero, or even negative in sign. In such instances, the researcher should consider either splitting the original construct into new constructs (each with its own set of measures), or eliminating items until only a unidimensional construct remains. For example, the organizational structure construct—as measured by Johansson and Yip—appears to be multidimensional. To rectify this problem, they could either drop one of their measures or split the construct into two distinct subconstructs.

Discriminant validity

The traditional methodological complement to convergent validity is discriminant validity, which represents the extent to which measures of a given construct differ from measures of other constructs in the same model. In a PLS context, one criterion for adequate discriminant validity is that a construct should share more variance with its measures than it shares with other constructs in a given model. To assess discriminant validity, Fornell and Larcker (1981) suggest the use of

---

[3] In fact, one of the retained loadings reported by Johansson and Yip is as low as 0.08.
[4] Internal consistency = $((\Sigma\ \lambda_{yi})^2)\ /\ ((\Sigma\ \lambda_{yi})^2 + \Sigma\ \mathrm{var}\ (\epsilon_i))$.
[5] Strictly speaking, this discussion of convergent validity and the preceding discussion of item reliability can only be applied to measures that are reflective, rather than formative. This issue is explored more fully in the next section of this paper.

Table 1. Internal consistency and average variance extracted values, by construct, for three published studies

	Internal consistency	AVE	Root AVE
Cool, Dierickx, and Jemison (1989)			
Return	0.96	0.88	0.94
Risk	0.97	0.94	0.97
Organizational fit	0.68	0.56	0.75
Operational efficiency	0.96	0.93	0.96
Input factor payments	0.94	0.88	0.94
Product–market investments	0.88	0.72	0.85
Market share	0.98	0.96	0.98
Rivalry	0.99	0.98	0.99
Fornell, Lorange, and Roos (1990)			
Internal push	0.61	0.23	0.48
Stakeholder strength	0.70	0.55	0.74
Analytical scope	0.80	0.42	0.65
Performance	0.72	0.40	0.64
Johansson and Yip (1994)			
Model 6			
Industry drivers	0.55	0.43	0.66
Management processes	0.62	0.46	0.68
Organizational structure	0.24	0.42	0.65
Global strategy	0.41	0.23	0.48
Performance	0.85	0.75	0.86
Model 8			
Industry drivers	0.54	0.46	0.68
Global strategy	0.38	0.27	0.52
Management processes	0.62	0.45	0.67
Organizational structure	0.40	0.46	0.68
Performance	0.84	0.72	0.85

Average Variance Extracted (i.e., the average variance shared between a construct and its measures).[6]

This measure should be greater than the variance shared between the construct and other constructs in the model (i.e., the squared correlation between two constructs). This can be demonstrated in a correlation matrix which includes the correlations between different constructs in the lower left off-diagonal elements of the matrix, and the square roots of the average variance extracted values calculated for each of the constructs along the diagonal. For adequate discriminant validity, the diagonal elements should be significantly greater than the off-diagonal elements in the corresponding rows and columns.

Birkinshaw *et al*. report both average variance extracted (AVE) values and inter-construct correlations in their Table 3, providing clear evidence of discriminant validity. For the other three studies, the table reports AVE and root AVE values for each construct in columns two and three.[7] Although none of these studies report inter-construct correlations, it is possible to roughly assess the discriminant validity of the constructs using the estimated path coefficients. For example, the root AVE values reported for the Cool *et al*. study are all larger than the path coefficients they estimate. Likewise, the root AVE

[6] Average variance extracted = $\sum \lambda_{yi}^2 \Big/ \left(\sum \lambda_{yi}^2 + \sum \text{var}(\epsilon_i)\right)$.

[7] Cool *et al*. report AVE as 'convergent validity' in their Table 1. They also report 'discriminant validity' for each construct. However, this latter measure assesses the average squared correlation of a particular construct with all other constructs in the model. Since this value will vary considerably depending on the other constructs included, it is not really appropriate to compare such values to any fixed threshold. In general, the approach described here is preferred.

values found for the Fornell *et al*. study all exceed their reported path coefficients. Thus, it is possible to conclude that discriminant validity is adequate in both of these studies.

Such is not the case for the study by Johansson and Yip. Consider the path coefficient between global strategy and management processes in their Model 8 (0.70). If this is taken as a rough proxy of the correlation between the two constructs, then the off-diagonal correlation between global strategy and management processes is greater than the root AVE value for each of the constructs considered on its own (i.e., 0.48, 0.68). This implies that the constructs and their measures cannot be adequately discriminated, and it is therefore entirely inappropriate to view them as distinct and separate theoretical entities.[8]

Construct–measurement relationships

The nature of the links between constructs and measures are referred to as epistemic relationships, or 'rules of correspondence' (Bagozzi, 1984; Fornell, 1982). Two basic types of epistemic relationships are relevant to causal modeling: reflective indicators and formative indicators. In the first case, indicators (measures) are believed to reflect the unobserved, underlying construct, with the construct giving rise to (or 'causing') the observed measures. In contrast, formative indicators define (or 'cause') the construct. A defined construct is completely determined by a linear combination of its indicators.

The discussion to this point has assumed that all indicators are reflective. When unobservable, underlying constructs are viewed as giving rise to associated measures, it is appropriate to talk about item reliability and convergent validity. However, this is not necessarily true for formative indicators (Bollen and Lennox, 1991; Cohen *et al*., 1990; MacCallum and Browne, 1993). In fact, formative indicators of the same construct 'can have positive, negative, or no correlation' with one another (Bollen and Lennox, 1991: 307). Consequently, observed correlations among the measures associated with a construct may not be meaningful, rendering irrelevant traditional assessments of individual item reliability and convergent validity.

This does not give the researcher *carte blanche* to arbitrarily link sets of measures to constructs, however. Cohen *et al*. (1990) suggest that when the relationship is formative, researchers must be careful to employ strong theory (which helps to identify appropriate measures) and multiple measures (to ensure acceptable content validity). Bollen and Lennox (1991: 307) expand on this latter point by emphasizing that researchers 'need a census of indicators, not a sample. That is, all constructs that form [the underlying construct] should be included.'

The choice between using formative or reflective indicators for a particular construct can at times be a difficult one to make. The researcher needs to think carefully about whether it is more correct to think of the underlying construct as 'causing' the observed measures (i.e., a reflective relationship) or of the measures as 'causing' (or defining) the construct (i.e., a formative relationship). An example of the former might be performance, whereas an example of the latter might be social economic status (SES). From a statistical perspective, use of formative indicators tends to increase the R^2 value for the endogenous (i.e., predicted) constructs, although this effect is usually not large. Thus, use of formative indicators tends to eliminate the need for the exogenous constructs, since all explanation is 'pushed' towards the endogenous variables.

Whether researchers use formative or reflective relationships in their models, their choice of a particular form of epistemic relationship should be both justified clearly and applied consistently. The studies by both Birkinshaw *et al*. and Cool *et al*. fail to describe the epistemic relationships between measures and constructs. However, it appears quite likely that reflective relationships were used exclusively in both studies. Consequently, the earlier discussions of reliability and validity should hold in both cases.

Fornell *et al*. use both formative and reflective relationships in their model. They provide a clear argument for choosing one form of epistemic relationship over the other for each of their four constructs. Although it is possible to question whether or not Fornell *et al*.'s choice of formative indicators is sufficiently complete for all of their constructs, the *natures* of the relationships studied

[8] In fact, a quick look at the two items used by Johansson and Yip to measure management processes—global budgeting and global group meetings—suggests that they are likely to be closely related to the three items used to measure global strategy—standardized products, integrated competitive moves, and overall global strategy.

are well defended. Consequently, discussions about reliability and validity for the three formative indicator constructs they employ are less relevant.[9]

In contrast, Johansson and Yip provide an incomplete justification for considering the industry drivers, global strategy, organization structure, and management process constructs as formative in nature, while performance is reflective. They suggest (1994: 587) that the first four of these constructs 'combine into broad factors,' whereas performance 'is more of an abstract perceptual construct.' Such arguments are neither compelling nor complete. Furthermore, Johansson and Yip do not consistently employ the same relationship form in all of their models. For example, in Models 1, 2, 3, 4, and 6 organization structure is viewed as a formative construct (i.e., it is *exactly* defined empirically by its measures). In contrast, in Models 5, 7, and 8 organization structure is reflective (i.e., it has surplus meaning beyond the measures employed). Such arbitrary shifts between relationship forms, without justification, are simply unacceptable.

Model goodness-of-fit

LISREL and other covariance structure analysis modeling approaches involve parameter estimation procedures which seek to reproduce as closely as possible the observed covariance matrix. In contrast, PLS has as its primary objective the minimization of error (or, equivalently, the maximization of variance explained) in all endogenous constructs. The degree to which any particular PLS model accomplishes this objective can be determined by examining the R^2 values for the dependent (endogenous) constructs.

One consequence of this difference in objectives between LISREL and PLS is that no proper overall goodness-of-fit measures exist for models estimated using the latter. Although existing PLS algorithms report goodness-of-fit statistics such as the Bentler–Bonett normed fit index (Bentler and Bonett, 1980), these statistics are meaningless since they are based on the assumption that the estimated model parameters are chosen in an attempt to minimize the difference between the observed and the reproduced covariance matrices (with the latter determined using the estimated parameter values)—an assumption that is not warranted for PLS.

Three of the four studies examined here report R^2 values for the endogenous constructs studied. These values range from a low of 12 percent (Birkinshaw *et al*., 1995) to a high of 64 percent (Cool *et al*., 1989). Although Fornell *et al*. (1990) include three endogenous constructs in their model, they only report an R^2 value for one of these (performance). In contrast, Johansson and Yip do not report R^2 values. Instead, they claim (1994: 595) that '. . . the model choice can focus on goodness-of-fit measures of the "inner" model . . . the Bentler–Bonett statistic . . . and the percent explanation of covariation among the latent variables'. Such a claim is simply incorrect.[10] In general, researchers employing PLS should report R^2 values for all endogenous constructs included in their models.

CONCLUSION

By combining and confronting theory with data (Fornell, 1982), and by forcing researchers to be explicit about both their measurement and theoretical assumptions (Bagozzi, 1980, 1984), causal models such as PLS can help strategic management researchers to achieve new insights.

[9] Only the performance construct is modeled as reflective by Fornell *et al*. As the earlier discussion noted, this construct exhibits acceptable convergent and discriminant validity. However, one of the retained items has a very low loading (0.38), and should arguably be dropped. The content validity of some of Fornell *et al*.'s formative constructs (particularly stakeholder strength, which includes only two rather abstract measures) can also be called into question.

[10] Even if the standards used to evaluate covariance structure analysis (e.g., LISREL) models *could* be appropriately applied to models estimated using PLS, the results presented by Johansson and Yip are questionable in at least two respects. First, the choice of the Bentler and Bonett (1980) normed fit index (NFI) as the sole arbiter of model fit is a poor one, since NFI is easily influenced by sample size and therefore often biased. After carefully reviewing a wide variety of goodness-of-fit measures, Gerbing and Anderson (1993) recommended use of several alternative measures, but not NFI. Second, although exact standards for evaluating goodness-of-fit measures have not yet been formally established, various pragmatic rules of thumb do exist. For example, for the NFI measure, researchers often suggest that a value of 0.95 or greater indicates a strong model, a value between 0.90 and 0.949 indicates an adequate model, and a value of less than 0.90 indicates poor model fit (Hulland, Chow, and Lam, 1996). Johansson and Yip report NFI values ranging from 0.26 (Model 1) to 0.46 (Models 7 and 8), suggesting that *none* of their models adequately fit the observed data.

Although PLS demands a level of rigor and clarity beyond that required by more traditional methodological approaches, as the field of strategic management continues to mature researchers need to increasingly rise to the challenge of meeting such demands.

However, researchers must use these new techniques appropriately. In addition to dealing with the methodological concerns associated with more traditional methods of analysis, researchers using causal modeling approaches such as LISREL and PLS must understand their underlying assumptions and limitations. As the preceding review of four recent uses of PLS in the strategic management area shows, the technique has been applied with considerable variability. The study by Cool *et al*. provides an excellent example for future applications of PLS. In contrast, some aspects of Johansson and Yip's use of PLS are seriously flawed, providing an important contrast to the work by Cool *et al*.

ACKNOWLEDGEMENTS

The author gratefully acknowledges the financial support provided by the Social Sciences and Humanities Research Council of Canada, the Ivey Business School, and the Barford Family. He also appreciates the thoughtful comments of two reviewers and the editor on an earlier version of this paper.

REFERENCES

Bagozzi, R. P. (1980). *Causal Models in Marketing*. Wiley, New York.

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

Barclay, D. W., C. Higgins and R. Thompson (1995). 'The partial least squares (PLS) approach to causal modeling: Personal computer adaptation and use as an illustration', *Technology Studies*, **2**(2), pp. 285–309.

Bentler, P. M. and D. G. Bonett (1980). 'Significance tests and goodness of fit in the analysis of covariance structures', *Psychological Bulletin*, **88**(3), pp. 588–606.

Birkinshaw, J., A. Morrison and J. Hulland (1995). 'Structural and competitive determinants of a global integration strategy', *Strategic Management Journal*, **16**(8), pp. 637–655.

Bollen, K. and R. Lennox (1991). 'Conventional wisdom on measurement: A structural equation perspective', *Psychological Bulletin*, **110** (2), pp. 305–314.

Carmines, E. G. and R. A. Zeller, (1979). 'Reliability and validity assessment'. *Sage University Paper Series on Quantitative Applications in the Social Sciences*, No. 07-017, Sage, Beverly Hills, CA.

Cohen, P., J. Cohen, J. Teresi, M. Marchi and C. N. Velez (1990). 'Problems in the measurement of latent variables in structural equations causal models', *Applied Psychological Measurement*, **14**(2), pp. 183–196.

Cool, K., I. Dierickx, and D. Jemison (1989). 'Business strategy, market structure and risk–return relationships: A structural approach', *Strategic Management Journal*, **10**(6), pp. 507–522.

Duxbury, L. E. and C. A. Higgins (1991). 'Gender differences in work-family conflict', *Journal of Applied Psychology*, **76**, pp. 60–74.

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

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

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

Fornell, C., P. Lorange, and J. Roos (1990). 'The cooperative venture formation process: A latent variable structural modeling approach', *Management Science*, **36**(10), pp. 1246–1255.

Gerbing, D. W. and J. C. Anderson (1993). 'Monte Carlo evaluations of goodness-of-fit indices for structural equation models'. In K. A. Bollen and J. S. Long (eds.), *Testing Structural Equation Models*. Sage, Newbury Park, CA. pp. 40–65.

Hagedoorn, J. and J. Schakenraad, (1994). 'The effect of strategic technology alliances on company performance', *Strategic Management Journal*, **15**(4), pp. 291–309.

Hulland, J. S., Y. H. Chow and S. Lam (1996). 'Use of causal models in marketing research: A review', *International Journal of Research in Marketing*, **13**(2), pp. 181–197.

Hulland, J. S. and D. N. Kleinmuntz (1994). 'Factors influencing the use of internal summary evaluations versus external information in choice', *Journal of Behavioral Decision Making*, **7**(2), pp. 79–102.

Johansson, J.K. and G. S. Yip (1994). 'Exploiting globalization potential: U.S. and Japanese strategies', *Strategic Management Journal*, **15**(8), pp. 579–601.

Jöreskog, K. A. and D. Sörbom (1989). *LISREL 7 User's Reference Guide* (1st ed.). Scientific Software, Mooresville, IN.

MacCallum, R. C. and M. W. Browne (1993). 'The use of causal indicators in covariance structure models: Some practical issues', *Psychological Bulletin*, **114**(3), pp. 533–541.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). McGraw-Hill, New York.

Smith, J. B. and D. W. Barclay (January 1997). 'The effects of organizational differences and trust on the effectiveness of selling partner relationships', *Journal of Marketing*, **61**, pp. 3–21.

Venkatraman, N. (1989). 'Strategic orientation of business enterprises: The construct, dimensionality, and measurement', *Management Science*, **35**(8), pp. 942–962.

Wold, H. (1974). 'Causal flows with latent variables', *European Economic Review*, **5**, pp. 67–86.

Wold, H. (1985). 'Systems analysis by partial least squares'. In P. Nijkamp, L. Leitner, and N. Wrigley (eds.), *Measuring the Unmeasurable*. Marinus Nijhoff, Dordrecht, pp. 221–251.

Zinkhan, G. M., E. Joachimsthaler, and T. C. Kinnear (May 1987). 'Individual differences and marketing decision support system usage and satisfaction', *Journal of Marketing Research*, **24**, pp. 208–214.