Strategic Management Journal
Strat. Mgmt. J., **25**: 397–404 (2004)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/smj.385

RESEARCH NOTES AND COMMENTARIES

AN ASSESSMENT OF THE USE OF STRUCTURAL EQUATION MODELING IN STRATEGIC MANAGEMENT RESEARCH

CHRISTOPHER L. SHOOK,[1] DAVID J. KETCHEN, JR.,[2]* G. TOMAS M. HULT[3] and K. MICHELE KACMAR[2]
[1] *College of Business, Auburn University, Auburn, Alabama, U.S.A.*
[2] *College of Business, Florida State University, Tallahassee, Florida, U.S.A.*
[3] *Eli Broad College of Business, Michigan State University, East Lansing, Michigan, U.S.A.*

Structural equation modeling (SEM) is a powerful, yet complex, analytical technique. The use of SEM to examine strategic management phenomena has increased dramatically in recent years, suggesting that a critical evaluation of the technique's implementation is needed. We compared the use of SEM in 92 strategic management studies published in nine prominent journals from 1984 to 2002 to guidelines culled from methodological research. We found that the use and reporting of SEM often have been less than ideal, indicating that authors may be drawing erroneous conclusions about relationships among variables. Given these results, we offer suggestions for researchers on how to better deploy SEM within future inquiry.

While the roots of strategic management research can be traced at least to the early 1960s, the field's prominence grew dramatically following the publication of Schendel and Hofer (1979) and the emergence of the *Strategic Management Journal* (*SMJ*) in 1980. As the field's stature has developed, so too has its theoretical sophistication. The desire to test complex models has led some authors to embrace structural equation modeling (SEM). Although SEM was introduced into the strategy literature in 1984 (Farh, Hoffman, and Hegarty, 1984), the technique has only recently become popular. For example, only five studies using SEM were published in *SMJ* prior to 1995, while 27 such studies appeared between 1998 and 2002.

In its most general form, SEM consists of a set of linear equations that simultaneously test two or more relationships among directly observable and/or unmeasured latent variables. While SEM serves purposes similar to multiple regression, differences exist between these techniques. SEM has a unique ability to simultaneously examine a series of dependence relationships (where a dependent variable becomes an independent variable in subsequent relationships within the same analysis) while also simultaneously analyzing multiple dependent variables (Jöreskog *et al*., 1999).

Like any statistical tool, SEM's benefits are obtained only if properly applied. Because SEM

Key words: structural equation modeling; data analysis; statistical methods
*Correspondence to: David J. Ketchen, Jr., College of Business, Florida State University, Tallahassee, FL 32306-1110, U.S.A. E-mail: dketchen@cob.fsu.edu

Received 25 October 2001
Final revision received 2 September 2003

requires several complicated choices, insight and judgment are crucial elements of its use (Jöreskog and Sörbom, 1996). Missteps compromise results' validity, inhibiting researchers' ability to develop knowledge and inform managers. Reviews of SEM usage in the fields of organizational behavior (Brannick, 1995), management information systems (Chin, 1998), marketing (Steenkamp and van Trijp, 1991), and logistics (Garver and Mentzer, 1999) have unveiled serious flaws. Steiger (2001) found that SEM textbooks ignore many important issues, suggesting that strategy researchers may have inadequate guidance to draw upon. To date, however, no systematic assessment has been made of SEM use within strategic management. Given (a) the difficulties found in other fields, (b) the growing use of SEM within strategic management, (c) SEM's complexity, and (d) the need to develop valid insights about organizations, a critical examination of SEM usage seems timely and warranted.

CRITICAL ISSUES IN THE APPLICATION OF SEM

We focused on six critical issues discussed in other fields' SEM assessments: data characteristics, reliability and validity, evaluating model fit, model respecification, equivalent models, and reporting. Where appropriate, we culled 'best practices' from the literature on SEM methodology (e.g., Bollen, 1989; Fornell and Larcker, 1981; Jöreskog and Sörbom, 1996) to judge researchers' decisions. Some of these best practices rely on statistical tests, while others rely on 'rules of thumb.' For other issues, our goal was to describe trends in SEM usage.

We searched the 10 empirical journals within MacMillan's (1994) 'forum for strategy research' from 1984 to 2002 for strategic management studies using SEM. As in Ketchen and Shook (1996), strategy studies were defined as studies examining relationships among the four broad constructs that represent the field according to Summer *et al.* (1990): strategy, environment, leadership/organization, and performance. Two coders had to agree that a study fit within this domain for the study to be included. We excluded studies using partial least squares (PLS). Although PLS is in some ways similar to SEM, it is not a covariance-based modeling technique and its application is very different from covariance-based methods. The 92 relevant studies we identified are listed in Table 1. Most studies were found in *SMJ* (34 studies, 37%), the *Academy of Management Journal* (24, 26%), and the *Journal of Management* (12, 13%). To assess whether the application of SEM has changed over time, we grouped the studies into two time periods: 1984 to 1995 (32 studies), and 1996 to 2002 (60 studies).

Two coders independently coded each article. One professor coded all studies. The other coding was split among another professor and two advanced PhD students. The percentage of coding agreement was 93 percent, which compares favorably to similar studies (e.g., 83%—Ford, MacCallum, and Tait, 1986; 93%—Ketchen and Shook, 1996). Disagreements were discussed until both coders agreed on the proper coding.

Data characteristics

The use of cross-sectional vs. longitudinal data is a basic issue in the application of SEM. Historically, the term 'causal modeling' has been used to describe SEM. The strongest inference of causality may be made only when the temporal ordering of variables is demonstrated (Kelloway, 1995). However, many strategy studies involve cross-sectional designs; in such cases, strong theoretical underpinnings are critical to causality inferences. We coded a study as cross-sectional if it involved one time period (69 of 92 studies, 75%) or as longitudinal if it involved two or more periods (23 studies, 25%). Cross-sectional designs were slightly, but not significantly ($z = 0.95$), more prevalent recently (22 of 32 studies, 69%, from 1984 to 1995 vs. 47 of 60, 78%, studies from 1996 to 2002).

Researchers should also ensure that their data meet the assumed distribution of their estimation approach. The common approaches to estimating structural equation models assume that indicator variables have multivariate normal distributions. Non-normal data may lead to inflated goodness-of-fit statistics and underestimated standard errors (MacCallum, Roznowski, and Necowitz, 1992). Thus, the uncritical use of non-normal data may hamper research progress by providing inaccurate findings. Despite these concerns, 75 studies (81%) did not note if the sample was normally distributed. The percentage of studies not mentioning sample distribution stayed constant (81% in 1984 to 1995; 82% in 1996 to 2002). Of the 17 studies that discussed normality, eight were based on normally

Table 1. Strategy articles that used SEM[a]

Academy of Management Journal	*Journal of Management Studies*	*Strategic Management Journal*
Singh, 1986	Venkatraman, 1990	Boyd, 1990
Walker and Weber, 1987	*Judge and Douglas, 1998	Hoskisson, Hitt, Johnson, and Moesel,
Keats and Hitt, 1988	*Golden, Dukerich, and Fabian, 2000	1993
Miller, Droge, and Toulouse, 1988		Boyd, 1994
*Heide and Miner, 1992	*Journal of Management*	Hagedoorn and Schakenraad, 1994
Judge and Zeithaml, 1992	Venkatraman and Ramanujam, 1987	Kotha and Vadlamani, 1995
*Hoskisson, Johnson, and Moesel,	*Bartunek and Franzak, 1988	Hopkins and Hopkins, 1997
1994	Fryxell, 1990	*Stimpert and Duhaime, 1997
Wally and Baum, 1994	Fryxell and Barton, 1990	Boyd and Reuning-Elliott, 1998
Phan and Hill, 1995	Keats, 1990	*Murtha, Lenway, and Bagozzi, 1998
Amason, 1996	Fryxell and Wang, 1994	Bensaou, Coyne, and Venkatraman,
*Hitt, Hoskisson, Johnson, and	Gerbing, Hamilton, and Freeman,	1999
Moesel, 1996	1994	Capron, 1999
*Simonin, 1997	Boyd and Fulk, 1996	Combs and Ketchen, 1999
Stimpert and Duhaime, 1997	Daily and Johnson, 1997	Holm, Eriksson, and Johanson, 1999
Daily, Johnson, Ellstrand, and Dalton,	Miller, Droge, and Vickery, 1997	Knight *et al.*, 1999
1998	Dooley, Fryxell, and Judge, 2000	Kroll, Wright, and Heiens, 1999
*Finkelstein and Boyd, 1988	*Lukas, Tan, and Hult, 2001	*McEvily and Zaheer, 1999
*Johnson and Greening, 1999		*Palmer and Wiseman, 1999
*Busenitz, Gomez, and Spencer, 2000	*Management Science*	*Simonin, 1999
*Isobe, Makino, and Montgomery,	Venkatraman and Ramanujam, 1987	Yeoh and Roth, 1999
2000	Venkatraman, 1989	Steensma and Lyles, 2000
Sharma, 2000	Miller, 1991	*Kale, Singh, and Perlmutter 2000
*Baum, Locke, and Smith, 2001	*Sethi and King, 1994	Capron, Mitchell, and Swaminathan,
Li and Atuahene-Gima, 2001	McGrath, Tsai, Venkatraman, and	2001
Song and Montoya-Weiss, 2001	MacMillan, 1996	Geletkanycz, Boyd, and Finkelstein,
*Hoskisson, Hitt, Johnson, and		2001
Grossman, 2002	*Organization Studies*	*Hult and Ketchen, 2001
*Hult, Ketchen, and Nichols, 2002	Wong and Birnbaum-Moore, 1994	Spanos and Lioukas, 2001
	*Ginsberg and Venkatraman, 1995	*Yli-Renko, Autio, and Sapienza, 2001
Administrative Science Quarterly		Andersson, Forsgren, and Holm, 2002
Walker and Poppo, 1991	*Organization Science*	Johnson, Korsgaard, and Sapienza,
*Gupta, Dirsmith, and Fogarty, 1994	Marcoulides and Heck, 1993	2002
*Dukerich, Golden, and Shortell, 2002	*Mjoen and Tallman, 1997	Koka and Prescott, 2002
	*Zaheer, McEvily, and Perrone, 1998	Li and Atuahene-Gima, 2002
Decision Science	Larsson and Finkelstein, 1999	Matusik, 2002
Farh, Hoffman, and Hegarty, 1984	Young-Ybarra and Wiersema, 1999	McEvily and Chakravarthy, 2002
Hult, 1998	Schilling and Steensma, 2002	Schroeder, Bates, and Junttila, 2002
Daily, Johnson, and Dalton, 1999		Worren, Moore, and Cardona, 2002

[a] Using MacCallum *et al.*'s (1996) guidelines, the 29 studies marked with an asterisk had adequate power; 43 other studies did not and we were unable to assess power for 20 studies.

distributed data and nine were not. In all of the nine latter studies, authors took corrective actions (e.g., transforming the data). Data transformation is not without problems, however. To the extent that a researcher has developed a strong theoretical foundation and believes in the original specification, transforming the data can provide an incorrect specification.[1] Thus, researchers must be mindful of the trade-offs inherent in transforming data (cf. Satorra, 2001). One alternative to transformation is to use an estimation approach available in EQS (i.e., ml,robust) that adjusts the model fit chi-square test statistic and standard errors of individual parameter estimates. Another approach is to select an estimation strategy, such as generalized least squares in LISREL, that does not assume multivariate normality (Hayduk, 1987: 334; cf. Jöreskog *et al.*, 1999).

Reliability and validity

Measures' reliability and validity should be assessed when using SEM. Initially, the estimated

[1] We thank an anonymous reviewer for offering this insight.

reflective loadings and their accompanying significance levels are examined. If the loadings are not statistically significant, the researcher should (in most cases) eliminate such indicator(s) (Anderson and Gerbing, 1988) or consider the possibility that the indicators may be formative (i.e., represent one aspect of a multidimensional construct) rather than reflective (i.e., represent attributes of a unidimensional construct) (Bollen, 1989). Following the assessment of loadings and t-values, reliability and validity should be examined.

Coefficient alpha, the most common measure of reliability, has several limitations. For example, coefficient alpha wrongly assumes that all items contribute equally to reliability (Bollen, 1989). A better choice is composite reliability, which draws on the standardized loadings and measurement error for each item. A popular rule of thumb is that 0.70 is an acceptable threshold for composite reliability, with each indicator reliability above 0.50 (Fornell and Larcker, 1981).

We coded whether a study described reliability and what measure(s) were reported. In total, reliability was examined in 56 studies (61%). Assessing reliability was slightly, but not significantly ($z = 1.12$), more prevalent recently (17 studies, 53%, early period vs. 39 studies, 65%, recent period). Overall, 34 studies (37%) relied on coefficient alpha. Composite reliability was reported in 18 studies (20%), and both composite reliability and coefficient alphas were reported in three studies (3%).

The means for assessing validity when using SEM are reviewed by Bollen (1989). These means rely on rules of thumb. At the most basic level, the measurement model itself offers evidence of convergent and discriminant validity (Anderson and Gerbing, 1988), assuming it is found to be acceptable (i.e., if it has significant factor loadings $\geq$0.70 and fit indices $\geq$0.90). Acceptable convergent validity is achieved when the average variance extracted is $\geq$50% (this measure roughly corresponds to the Eigenvalue in exploratory factor analysis).

We coded whether a study examined convergent validity and what measure(s) were reported. In total, convergent validity was examined in 36 studies (39%). Assessing convergent validity was slightly, but not significantly ($z = 1.59$), more prevalent recently (9, 28%, early period vs. 27, 45%, recent period). Nine studies (10%) reported the average variance extracted. Less stringent assessments of convergent validity (e.g., examination of factor loadings and the correlation matrix) were used in 24 studies (26%).

Calculating the shared variance between two constructs and verifying that it is lower than the average variances extracted for each individual construct are the most common means to assess discriminant validity in the measurement model (Fornell and Larcker, 1981). Another viable approach is to conduct 'pairwise tests' of all theoretically related constructs (Anderson, 1987; Bagozzi and Phillips, 1982). The pairwise analysis tests whether a confirmatory factor analysis model representing two measures with two factors fits the data significantly better than a one-factor model.

We coded whether a study examined discriminant validity and what measure(s) were reported. Discriminant validity was examined in 37 studies, or 40% (11 in 1984 to 1995, 34%; 26 in 1996 to 2002, 43%). Pairwise tests were reported in 21 studies (23%), nine studies (10%) reported average variance extracted compared to shared variance, seven studies (7%) cited other measures (e.g., examination of the correlation matrix).

Our results for reliability and validity offer cause for concern. Confidence in an SEM study's findings depends on a solid foundation of measures with known and rigorous properties. Yet, little attention has been paid to reliability and validity. When measures are unwittingly weak, the consequences may include inappropriate modifications to structural models and spurious findings. Low reliability and validity may cause relationships to appear non-significant, regardless of whether the links exist. As a result, promising research paths may have been prematurely cut short. To the extent that such paths could have generated insights transferable to managers, strategic management's mission to inform practice was not served. Thus, careful attention to measurement issues is needed in future studies.

Evaluating model fit

Assessing a model's fit is one of SEM's most controversial aspects. Before assessing individual parameters, one must assess the overall fit of the observed data to an *a priori* model (Jöreskog *et al.*, 1999). Unlike traditional methods, SEM relies on non-significance. Fit indices ascertain if the covariance matrix derived using the hypothesized model is different from the covariance matrix derived from the sample. A non-significant difference indicates that the errors are non-significant, lending support to the model. A chi-square test is the most common fit measure, but it is only recommended with moderate samples (e.g., 100 to 200; Tabachnick and Fidell, 1996). With large samples, trivial differences between the two matrices become significant. The test also is suspect when using small samples because some are not distributed as chi-square populations.

Because of these limitations, several 'comparative fit' indices that contrast the fit of one model with the fit of competing or baseline models have emerged. Rules of thumb, not significance tests, are used in determining acceptable fit levels because the underlying sampling distributions for these indices are unknown. These heuristics may or may not be appropriate for a specific data set (Brannick, 1995). Thus, until a definitive measure is developed, researchers should use multiple measures to provide evidence about their models (Breckler, 1990). The use of multiple indices assures readers that authors have not simply picked a supportive index. Research by Gerbing and Anderson (1992) suggests that among the most stable and robust fit indices are the DELTA2 index, the relative noncentrality index (RNI), and the comparative fit index (CFI) (cf. Hu and Bentler, 1999).

We coded how many and which fit indices were used. Studies reported between zero and nine fit measures. Multiple fit measures were used in 83 studies (90%). Four studies (4%) did not report any fit measures, and only one fit measure was reported in five (6%). The most popular measure was chi-square (79 studies, 86%), followed by the Goodness of Fit Index (44, 48%), the CFI (38, 41%), and the root mean square residual (32, 35%). Only one study used all three fit measures recommended by Gerbing and Anderson (1992). Encouragingly, the use of multiple indicators has increased over time. For example, from 1984 to 1995, 13 of 32 studies (41%) used four or more, while 42 of 60 later studies (70%) used four or more ($z = 2.70$, $p < 0.01$).

Adequate statistical power is essential to assessing model fit when using a chi-square test. More generally, adequate power makes it more likely that a study will accurately depict variables' relations. We attempted to code studies' power directly, but power was not described in 90 of the 92 studies. To assess whether the studies' power was adequate, we consulted MacCallum, Browne, and Sugawara (1996), which lists the minimum sample size needed for adequate power (i.e., 0.80) at various degrees of freedom to assess close model fit. While this approach is somewhat crude, more precise assessment methods require statistics authors seldom report (cf. Satorra and Saris, 1985). Using MacCallum *et al.*'s (1996) guidelines, 29 studies (31%) had adequate power, while 43 (47%) did not. We could not assess power for 20 studies (22%) that did not report degrees of freedom and one that did not clearly report sample size.

Our findings suggest that many extant studies have obtained significant results despite insufficient sample sizes. A potential explanation for the number of studies that found significant results despite lacking adequate statistical power is that model respecifications were conducted.[2] Regardless of the underlying explanation, our findings are troubling. Any author building on a past study should consider its power when interpreting the findings. With this issue in mind, we note with asterisks in Table 1 the studies that we found to have adequate power.

Future scholars have several tools to ensure adequate power. One is to perform the calculations outlined by Satorra and Saris (1985); these calculations are complex and require the derivation of various alternative models (Chin, 1998). Authors also can use a program listed in MacCallum *et al.* (1996) to calculate needed sample sizes. A final approach is to use the MacCallum *et al.* (1996) guidelines that we used. One caution is that those power estimates may be inflated when using small samples (MacCallum *et al.*, 1996).

Model respecification

Model respecification occurs when one tests a proposed model and then seeks to improve model fit, often through adding or removing paths among constructs. Respecification is common in the social sciences because *a priori* models often do not adequately fit the data, but respecification is also controversial. Anderson and Gerbing (1988) argue that respecifications should be based on theory and content considerations in order to avoid exploiting sampling error to achieve satisfactory goodness of fit. Chin (1998) and Kelloway (1995) assert

[2] We appreciate an anonymous reviewer for providing this idea.

that changes should not only be theoretically justified, but also validated with a new sample. Brannick (1995) argues that respecifications should not be done at all. Indeed, fit indices assume such searches have not been done. Thus, after a specification search, it is unclear how to interpret an index score.[3] If theoretical justification for modifications exists, alternative models should have been proposed *a priori* rather than making *a posteriori* changes.

We coded whether specification searches were reported. Among the 92 studies, 43 (47%) listed a respecification. For those 43 studies, we coded the following: (1) if authors explicitly stated that the specification searches were exploratory in nature (16 did, 17%); (2) if the changes in the specification searches were validated on a hold-out sample (2 were, 2%); and (3) if the authors cited theoretical reasons as justification for the changes (18 did, 20%). While these results were disconcerting, there was some good news. Authors reported performing specification searches less frequently in recent studies (22 of 60, 37% vs. 21 of 32, 66%; $z = 2.65$; $p < 0.01$), perhaps reflecting the increasing robustness of strategy theory over time.

Equivalent models

For any proposed structural model, other structural models tested using the same data may suggest different relationships among latent constructs, but still yield equivalent levels of fit. These alternative models may be radically different, leading MacCallum *et al.* (1993) to assert that the existence of equivalent models could call into question the conclusions drawn from most SEM studies. Thus, authors should acknowledge the possible existence of equivalent models as a limitation of results obtained via SEM (Breckler, 1990; MacCallum *et al.*, 1993).

In 24 of the 92 studies, SEM was used only for measurement validation via confirmatory factor analysis (CFA). Eleven studies were focused solely on measurement validation. Thirteen studies followed the CFA with other techniques (e.g., regression) to examine relationships. In the studies where SEM usage was limited to measurement validation, no relationships among latent variables were examined using SEM; thus, we did not evaluate the authors' treatment of equivalent models (MacCallum *et al.*, 1993). Only one of the remaining 68 studies (2%) explicitly acknowledged that the possible existence of equivalent structural models is problematic. Beyond acknowledging equivalent models, other responses are available. For example, Anderson and Gerbing (1988) outline a two-step approach for identifying and evaluating equivalent models that is popular in the psychology and marketing literatures. Also, simultaneous relations models can be tested wherein, for example, variable A influences variable B, and B influences A. Parameters can be identified if additional variables influence A that do not influence B or vice versa.[4]

Reporting

Ideally, articles would provide enough detail to permit replication by others because replication serves as a safeguard against the acceptance of erroneous findings. Chin (1998: 8) suggests that the input matrix, the software and version used, starting values, number of iterations, the models tested, computational options used, and anomalies encountered during the analytic process should be described. No single study listed all of these items.

Readers can still gain a basic understanding of the analysis performed if the input matrix and software are listed. Although most statistical packages assume the use of the covariance matrix (Steiger, 2001), in practice both covariance and correlation matrices are used as the input for SEM. The majority of studies ($n = 75$, 81%) did not identify whether a correlation or covariance matrix was used. Correlation matrices were used in eight studies (9%) and covariance matrices were used in nine studies (10%). Whether the input matrix was provided was also coded. The input matrix was reproduced in eight studies (9%) and not reproduced in 84 studies (91%).

Different software packages and versions of the same software package sometimes have different default parameters. Thus, authors should provide the name and version of the software. The software package used was reported in 81 studies (88%). LISREL was used in 66 studies (72%), EQS was used in nine studies (10%), and CALIS, AMOS, and Mplus were each used in two studies (2%). The version of the software used was reported in 51 studies (55%), but not in 41 studies (45%). Only

[3] We thank an anonymous reviewer for this insight.

[4] We thank an anonymous reviewer for this insight.

five studies (5%) reported both the input matrix and the software used. Overall, the reporting practices to date inhibit strategic management scholars' ability to replicate and build on past findings.

CONCLUSIONS

One limitation of our study is that we could only assess the application of SEM based on what published articles reported. In some cases, authors may have made appropriate decisions and discussed these decisions with referees during the review process, but not included this material in the article. Yet, because the discipline desires knowledge accumulation, adequate information about statistical procedures should be made available. Unfortunately, our results show that such information is often not provided. In Table 2, we offer reviewers and editors a suggested checklist for assessing SEM studies. To enhance knowledge development, gatekeepers should require that authors address all of the issues listed. Recognizing the limitations on journal space, we also recommend that journals provide vehicles for communicating space-consuming details of statistical analyses (such as input matrices) to interested readers. For example, *Management Science* currently offers authors space on a journal-sponsored Internet web site.

Table 2. Checklist of issues that should be reported in SEM studies

1. Sample issues
   (a) General description
   (b) Number of observations
   (c) Distribution of sample
   (d) Statistical power
2. Measurement issues
   (a) Reliability of measures
   (b) Measures of discriminant validity
   (c) Measures of convergent validity
3. Reproduceability issues
   (a) Input matrix
   (b) Name and version of software package used
   (c) Starting values
   (d) Computational options used
   (e) Analytical anomalies encountered
4. Equivalent models issues
   (a) Potential existence acknowledged as a limitation
5. Respecification issues
   (a) Changes cross-validated
   (b) Respecified models not given status of hypothesized model

In closing, we believe that significant opportunities remain for SEM to generate insights within strategic management. For example, the field's core constructs (e.g., strategy, performance) are multidimensional and the relationships among them are complex (Summer *et al.*, 1990). Given SEM's ability to map and assess a web of relationships, it offers vast potential as a tool to diagnose key links. Also, SEM's ability to tap intangible latent variables (cf. Hult and Ketchen, 2001) might help unveil the 'unobservable' constructs that are central to the resource-based view, transactions costs economics, and agency theory (Godfrey and Hill, 1995). These outcomes can only be realized, however, if the technique is used prudently. Our hope is that future scholars will use the ideas presented here in order to maximize SEM's utility.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the help of James Combs, Karl Jöreskog, Bruce Lamont, Robert MacCallum, Timothy Palmer, and Randy Settoon with this paper.

REFERENCES

Anderson JC. 1987. An approach for confirmatory measurement and structural equation modeling of organizational properties. *Management Science* **33**(April): 525–541.

Anderson JC, Gerbing DW. 1988. Structural equation modeling in practice: a review and recommended two-step approach. *Psychological Bulletin* **103**(3): 411–423.

Bagozzi RP, Phillips LW. 1982. Representing and testing organizational theories: a holistic construal. *Administrative Science Quarterly* **27**(September): 459–489.

Bollen KA. 1989. *Structural Equations with Latent Variables*. Wiley: New York.

Brannick MT. 1995. Critical comments on applying covariance structure modeling introduction. *Journal of Organizational Behavior* **16**: 201–213.

Breckler SJ. 1990. Applications of covariance structure modeling in psychology: cause for concern. *Psychological Bulletin* **107**(2): 260–273.

Chin W. 1998. Issues and opinion on structural equation modeling. *MIS Quarterly* **22**(1): 7–16.

Farh JL, Hoffman RC, Hegarty WH. 1984. Assessing environmental scanning at the subunit level: a multitrait–multimethod analysis. *Decision Sciences* **15**: 197–219.

Ford JK, MacCallum RC, Tait M. 1986. The application of exploratory factor analysis in applied psychology:

a critical review and analysis. *Personnel Psychology* **39**: 291–314.
Fornell C, Larcker DG. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research* **18**: 39–50.
Garver MS, Mentzer JT. 1999. Logistics research methods: employing structural equation modeling to test for construct validity. *Journal of Business Logistics* **20**(1): 33–58.
Gerbing DA, Anderson JC. 1992. Monte Carlo evaluations of goodness of fit indices for structural equation models. *Sociological Methods and Research* **21**(2): 132–160.
Godfrey PC, Hill CWL. 1995. The problem of unobservables in strategic management research. *Strategic Management Journal* **16**(7): 519–534.
Hayduk LA. 1987. *Structural Equation Modeling with LISREL: Essentials and Advances*. Johns Hopkins Press: Baltimore, MD.
Hu L, Bentler PM. 1999. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling* **6**(1): 1–55.
Hult GTM, Ketchen DJ. 2001. Does market orientation matter? A test of the relationship between positional advantage and performance. *Strategic Management Journal* **22**(9): 899–906.
Jöreskog KG, Sörbom D. 1996. *LISREL 8: User's Reference Guide*. Scientific Software International: Chicago, IL.
Jöreskog KG, Sörbom D, du Toit S, du Toit M. 1999. *LISREL 8: New Statistical Features*. Scientific Software International: Chicago, IL.
Kelloway EK. 1995. Structural equation modeling in perspective introduction. *Journal of Organizational Behavior* **16**: 215–224.
Ketchen DJ, Shook CL. 1996. The application of cluster analysis in strategic management research: an analysis and critique. *Strategic Management Journal* **17**(6): 441–458.
MacCallum RC, Browne MW, Sugawara HM. 1996. Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods* **1**(2): 130–149.
MacCallum RC, Roznowski M, Necowitz L. 1992. Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychological Bulletin* **111**(3): 490–504.
MacCallum RC, Wegner DT, Uchino BN, Farbigar LR. 1993. The problem of equivalent models in applications of covariance structure analysis. *Psychological Bulletin* **114**(1): 185–199.
MacMillan I. 1994. The emerging forum for business policy scholars. *Journal of Business Venturing* **9**(2): 85–89.
Satorra A. 2001. Goodness of fit testing of structural equation models with multiple group data and nonnormality. In *Structural Equation Modeling: Present and Future*, Cudeck RC, du Toit S, Sörbom D (eds). Scientific Software International: Lincolnwood, IL; 231–256.
Satorra A, Saris WE. 1985. Power of the maximum likelihood ratio test in covariance structure analysis. *Psychometrika* **50**(1): 83–90.
Schendel DE, Hofer CW. 1979. *Strategic Management: A New View of Business Policy and Planning*. Little, Brown: Boston, MA.
Summer CE, Bettis RA, Duhaime IH, Grant JH, Hambrick DC, Snow CC, Zeithaml CP. 1990. Doctoral education in the field of business policy and strategy. *Journal of Management* **16**: 361–398.
Steenkamp JB, van Trijp H. 1991. The use of LISREL in validating marketing constructs. *International Journal of Research in Marketing* **8**: 283–299.
Steiger J. 2001. Driving fast in reverse. *Journal of American Statistical Association* **96**: 331–338.
Tabachnick BG, Fidell LS. 1996. *Using Multivariate Statistics*. Harper Collins: New York.