ELSEVIER

Long Range Planning 45 (2012) 341–358

LRP long range planning

http://www.elsevier.com/locate/lrp

Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods

Edward E. Rigdon

Several widely-cited weaknesses of Partial Least Squares (PLS) path modeling center on its character as a composite-based method rather than a factor-based method. Yet factor models as a framework for research may have been oversold. Insights from the forecasting literature suggest that PLS path modeling has strengths as a tool for prediction which have not been fully appreciated. PLS Mode A, typically thought of as "reflective measurement," is equivalent to the use of correlation weights, which deliver better prediction on out-of-sample data (data not used in estimating model parameters), while PLS Mode B is equivalent to the use of regression weights, which provide better in-sample prediction (prediction of data used to estimate model parameters). PLS path modeling can move forward by freeing itself entirely of its heritage as "something like but not quite factor analysis," by fleshing out inferential tools appropriate for a purely composite method, and by developing approaches for assessing measurement validity that properly recognize the distinction between theoretical concept and empirical proxy.

"'Two and two are four.'
'Sometimes, Winston. Sometimes two and two are five. Sometimes they are three. Sometimes they are all of them at once.'"

George Orwell, Nineteen Eighty-Four

Introduction

Partial least squares (PLS) path modeling was conceived as a composite-based alternative to factor-based structural equation modeling (SEM), employing the optimality properties of ordinary least squares regression to approximate results from factor-based SEM, along with more limited

0024-6301/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.lrp.2012.09.010

distributional assumptions and reduced computational demands. Both the method's originators and its critics have tended to evaluate PLS path modeling in terms of what it is not. However, research from multiple domains supports an argument for changing this thinking about PLS path modeling in two important ways. First, research in statistics and psychometrics challenges the factor-centric worldview which keeps PLS path modeling standing in the shadow of its slightly elder sibling. Second, research in forecasting points to advantages for PLS path modeling that are not widely understood. Taken together, these streams of research suggest that PLS path modeling users should, 1) discard their "factor envy" and recognize and evaluate PLS path modeling as a method strictly in terms of weighted composites of observed variables; 2) exploit the advantages which PLS path modeling offers; and 3) build the future of PLS path modeling in entirely factor-free terms. Combined, these arguments call for PLS path modeling to stand by itself as a method for modeling structural relations.

A brief review

Hair et al. (2012) provide a detailed overview of PLS path modeling, so familiarity with the method will be assumed here. Factor-based SEM and PLS path modeling have been bound together conceptually from their very beginnings. Both owe their origins to the same accident. Karl Jöreskog (1969), whose crucial innovations launched factor-based SEM as an inferential technique, was on his way to a career as a high school mathematics teacher when a coincidence changed his future and the future of multivariate analysis (Sörbom, 2001). While Jöreskog was visiting a graduate student friend, the friend received a telephone call from Herman Wold, the famed econometrician, who was seeking a research assistant. Jöreskog's friend connected Wold with Jöreskog. In time, Jöreskog's work under Wold's supervision led to the first of Jöreskog's many contributions to factor-based SEM, a technique involving structural relations between common factors. In turn, those contributions inspired Wold to imagine an analogous method based on weighted composites, just as principal components and canonical correlation analysis stand as composite-based alternatives to exploratory factor analysis (Dijkstra, 2010).

Factor-based SEM and PLS path modeling are described with similar graphical and symbolic languages. For example, imagine a simple model with six observed variables, y_1 — y_6. In factor-based SEM, y_1 — y_3 might be associated with one common factor, f_1, while y_4 — y_6 might be associated with a second common factor, f_2, according to a mathematical model such as:

$$y_i = \lambda_{i1} f_1 + \delta_i, \; i = 1, 2, 3$$

$$y_j = \lambda_{j2} f_2 + \delta_j, \; j = 4, 5, 6$$

where δ_i and δ_j are error terms in the factor model. The factors themselves may be related through a structural model such as:

$$f_2 = \beta_{21} f_1 + \zeta_2$$

As with ordinary least squares regression, error terms are constrained to be orthogonal with predictors in the same equation. Moreover, in factor-based SEM the covariance matrices of the factors and of the error terms are explicit parts of the factor model, whose structures may be (and typically are) constrained.

PLS path modeling with the same observed variables could be described similarly, but with key differences. First, in place of common factors f_1 and f_2, PLS path modeling proceeds with proxies f_1^* and f_2^*. These proxies — stand-ins for the latent variables of interest — are weighted composites of other variables in the model. These proxies are not assumed to be identical to the latent variables which they replace — they are recognized as approximations (Wold, 1982; Chin, 1998). Second, PLS path modeling allows relations between a set of observed variables and the associated proxy to take different forms. In one form, labeled Mode B, the proxy is modeled as a dependent variable, predicted by its associated observed variables, for example:

$$f_1^* = w_1 y_1 + w_2 y_2 + w_3 y_3 + d_1$$

Mode A, by contrast, maintains each observed variable as dependent upon its associated proxy, for example:

$$y_i = \lambda_{i1} f_1^* + \delta_i, \ i = 1, 2, 3$$

Third, and crucially, error terms in PLS path modeling are not explicit parts of the model, so relations among these error terms are unconstrained.

Whence the popularity of factor-based methods?

Factor-based SEM quickly became extremely popular as a multivariate technique. “Unlike many other developments in theoretical psychometrics, [SEM] spread from the methodology laboratory into the research laboratory with unusual rapidity” (Bentler, 1986, p. 35). “The development of the rigorous and generalized methods for testing hypotheses concerning underlying structures in covariance matrices is perhaps the most important and influential statistical revolution to have occurred in the social sciences” (Cliff, 1983, p. 115). By contrast, while PLS path modeling has come to enjoy a measure of popularity, the method has not been adopted with the same enthusiasm across so many research areas.

Factor analysis and the classical true-score theory of measurement are tied together even more closely than PLS path modeling and factor-based SEM. Both were pioneered by Charles Spearman, in two papers published in the same year (McDonald, 1999, p. 5). True score theory may be viewed as a special case of the common factor model (Steiger and Schönemann, 1978, p. 139). With true score theory coming to serve as a pillar of modern psychometrics, factor-based methods enjoyed a clear path toward widespread adoption among applied researchers in the social sciences.

Factor-based SEM’s chances were further boosted by the development of complete epistemological and psychometric frameworks which were themselves especially factor-friendly. Churchill’s (1979) influential framework for measure development specifically incorporated exploratory factor analysis — although Churchill clearly identified his recommendation as just one approach, not the only one. Churchill’s work was followed by later guidelines for measure development which were even more SEM-specific (e.g., Gerbing and Anderson, 1988). However, perhaps the most influential of these frameworks was the Holistic Construal (Bagozzi, 1984, 2011; Bagozzi and Phillips, 1982; Bagozzi and Yi, 2012). Google Scholar finds more than 1,200 references to Bagozzi and Phillips’s (1982) application of the approach to research problems in organization theory. The Holistic Construal describes an approach to measurement which is so like a factor-based structural equation model that the application of the statistical method to address the measurement problem is obvious and intuitive. “What is needed is a methodological paradigm that permits a direct assessment of the degree of correspondence between measurements and concepts and at the same time takes this relationship into account in the test of substantive hypotheses” (Bagozzi and Phillips, 1982, p. 460). The Holistic Construal was intended to be the tool which achieved this end, and factor-based SEM was the statistical engine which powered the Holistic Construal.

The Holistic Construal, as developed over the years, is intricate and rich, drawing deeply from philosophy of science, but for the sake of brevity a simplified version will serve here (see Figure 1). The focus of scientific research is on concepts, such as “competitiveness,” which are not directly observed and which may not have one clear empirical expression. A conceptual definition (the triangle in Figure 1) specifies precisely what the researcher means by this concept. Concepts are also partly defined by their relations with other concepts within a “nomological network” (Cronbach and Meehl, 1955), a system of law-like relationships, discovered over time, which anchor each concept. Thus, depictions of the Holistic Construal typically include an antecedent concept and a consequent concept. Indicators (represented by rectangles) are observed variables associated with the concept. Between conceptual definition and indicators lies a construct (represented by an oval), something created from the empirical data which is intended to enable empirical testing of

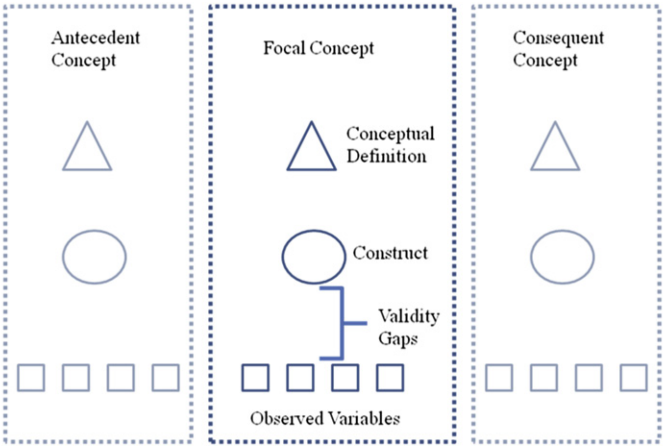

Figure 1. Simplified holistic construal

propositions regarding the concept. Indicators are linked only imperfectly to constructs, and thus are affected by error terms (not shown). The validity of measures relates to the tightness of the link between indicators and construct.

The advantage for factor-based SEM lay in the nature of the construct standing in the middle, between the conceptual definition and the indicators. Such a construct might be formed in many ways. Nunnally and Bernstein (1994, p. 159) asserted that "the heart of psychometric theory" is weighted combinations or linear composites of observed variables. Yet within the Holistic Construal, and across a broad swath of applied psychometric literature, the oval in the middle is understood to be a common factor. Moreover, the distinction between the concept itself and this construct is blurred (Maraun and Gabriel, 2013). In fact, there seems to be no distinction at all: "The central variables in a theory go by various names: latent variables, theoretical variables, constructs or theoretical constructs, unobservable variables, factors, traits or simply concepts" (Bagozzi and Yi, 2012, p. 9).

The implications are profound. If constructs and theoretical concepts are the same thing, and if the construct is a common factor extracted from the indicators, then common factors and theoretical concepts are also identical, so that factor-based modeling tools effectively allow the modeling of the unobservable theoretical concepts themselves. Factor-based SEM grants users the ability to directly assess the linkage between observed variable and theoretical concept using functions of the estimated parameters of the factor model, neatly solving the otherwise vexing problem of establishing measurement validity. Bollen's (1989) influential factor-based SEM text explicitly redefined the very notion of validity from the classical "the extent to which an indicator measures what it is supposed to measure" (p. 197) to "the magnitude of the direct structural relation between [a common factor] and [an indicator]" (p. 197). In addition, if the common factor is indeed the theoretical concept, then the error terms which the factor model assigns to the indicators are also measurement error terms. Accounting for those error terms analytically allows researchers to claim that their factor-based SEM parameter estimates are "adjusted for measurement error."

Given these understandings, the choice between factor-based SEM and PLS path modeling is really no choice at all. One method seems to offer direct, precise quantitative access to theoretical concepts, while the other can only approximate such results. Given wide and almost unchallenged acceptance of this argument, the logical course for those seeking to promote PLS path modeling was to make their method look as much like factor analysis as possible. And so the PLS path modeling literature came to describe Mode A estimation as "reflective measurement," just as the factor-based SEM literature refers to factor analysis as reflective measurement. This may be described as an article of faith among regular PLS path modeling users and among those who write about the method.

The unintended consequence, however, has been to help cement PLS path modeling's position in factor-based SEM's shadow, as a tool that does not quite do what its older sibling does. Beyond being limited to only approximating factor analytic results, and beyond working with weighted composite proxies instead of with the concept-equivalent constructs claimed by factor-based SEM, PLS path modeling lacks an inferential test statistic like factor-based SEM's χ^2 statistic. Thus, critics of PLS path modeling have argued that factor-based structural equation models are "testable," while PLS path models are not. In sum, PLS path modeling has been defined largely in terms of what it is not, with predictable consequences in terms of the popularity of the method.

Have the advantages of factor methods been oversold?

"As we have demonstrated, there is no need for factor theory at all" (Schönemann and Steiger, 1976, p. 188).

After dominating measurement in the social sciences for decades, this received view of measurement is under attack from many quarters. Rossiter (e.g., 2010, 2011) takes issue with reflective measurement and multiple indicators, and argues for the evaluation of validity in purely conceptual, not empirical, terms. Even while challenging Rossiter's approach, Rigdon et al. (2011) affirmed that the dominant measurement paradigm is fundamentally flawed, and called for researchers to look at a broader range of possible measurement models. Attacking conventional practice from several angles, Michell (e.g., 2008, 2013) has argued in part that conventional psychometric practice rests on doubtful assumptions, testable yet untested, while Maraun and colleagues (Maraun and Gabriel, 2013) argue that the dominant paradigm is actually built out of a confusion of key concepts. All of these challenges to conventional thinking raise questions about the appropriateness of granting factor analysis the role of empirical touchstone in measurement.

All along, there has been room to argue that the advantages of factor-based methods have been oversold, indeed that "…the advantages of factor analysis are largely illusory" (Steiger, 1990, p. 41). From the earliest days of factor-based SEM, the χ^2 statistic, which provides SEM's inferential advantage, has repeatedly demonstrated that factor models generally are not consistent with the data used to estimate them. "Factor analysis *is* a model. The probability that it fits any data set exactly…is essentially zero. Hence, for real data, the factor analysis model is an approximation. In many cases it is not a good approximation…" (Steiger, 1990, pp. 41-42). Despite this demonstrated inconsistency between model and data, researchers carry on interpreting results from factor-based structural equation models: "The status of factor analysis as a falsifiable scientific hypothesis has rarely been taken seriously since Thurstone popularized multiple factor analysis as a general research tool. It has not been taken seriously as a model by those theoreticians who urged its widespread use, and went to great lengths to ignore the theoretical problems attending to this model. It has not been taken seriously by most users, who were generally uninformed about the defects of the factor model" (Schönemann and Steiger, 1976, p. 188).

Early in the history of factor-based SEM, evaluation of the fit of these models shifted from the χ^2 statistic to alternative fit indices, which mostly lack an inferential character and whose interpretation depends largely on vague and changing rules of thumb. According to Sörbom (2001), Karl Jöreskog stated in a workshop that the Goodness of Fit Index (GFI) and Adjusted Goodness of Fit Index (AGFI) were added to the LISREL (Jöreskog and Sörbom, 1996) factor-based SEM package because, "Well, users had threatened us saying they would stop using LISREL if it always produces such large chi-squares. So we had to invent something to make people happy. GFI serves that purpose" (Sörbom, 2001, p. 10). In addition, users often engage in ex post model modification (Cliff, 1983), adding free parameters and/or discarding observed variables (sacrificing substantial amounts of data) in order to obtain values on alternative fit indices which meet rule-of-thumb criteria. Even if researchers confined their focus to the χ^2 statistic, these ex post modifications would badly complicate any inferential interpretation. The ability of factor-based SEM to produce unbiased factor model parameter estimates is predicated upon the estimated model being correct in the population, yet experience shows that this condition rarely holds in practice. This reality means that factor-based SEM's supposed advantages in terms of both its inferential nature and its unbiased parameter estimates are often or even typically not available to users of factor-based SEM.

Moreover, with real data, the practical difference between factor analysis and composite-based principal components analysis is less than it seems: "…when the same number of components or factors are extracted, the results from different types of component or factor analysis typically yield highly similar results. Discrepancies are rarely if ever of any practical importance in subsequent interpretations" (Velicer and Jackson, 1990, p. 5). Of course, one can construct data sets, simulated from populations based on factor models, where different methods produce predictably different results (McDonald, 1996a), but in practice the observed differences are not substantively important.

Finally, researchers have long warned that the factors derived in factor analysis are not necessarily equivalent to the theoretical concepts at the center of scientific research. Cliff (1983) warned of the nominalist fallacy, the unsubstantiated assumption that a common factor is equivalent to whatever label is assigned to it. "Even in what is called 'confirmatory' factor analysis, it is not the nature of the factors which is confirmed: the only thing which is confirmed is that the observed covariance matrix is not *inconsistent* with a certain pattern of parameters. It does not tell us what those parameters mean, and experience has shown that our belief that we do know what they mean is often ill-founded" (Cliff, 1983, p. 122).

If factor and concept are not equivalent, then neither are factor model error terms equivalent to measurement errors. Thus, the claim that factor-based SEM results are "adjusted for measurement error" must be moderated substantially. Both over-correction and under-correction can induce bias in parameter estimates. Moreover, composite methods with multiple indicators do achieve a degree of adjustment for unreliability, due to the nature of weighted composites. Suppose a set of p variables y_i ($i = 1$ to p) are summed with weights w_i ($i = 1$ to p) to form a composite C, according to:

$$C = \sum_i w_i y_i$$

Let σ_C^2 be the variance of the composite C, let σ_i^2 be the variance of y_i and let σ_{ij} ($i = 1$ to p, $j = 1$ to i, $i \neq j$) be the covariance between two different indicators y_i and y_j. Then the variance of the composite will be equal to (Mulaik, 2010, p. 83, equation 3.21):

$$\sigma_C^2 = \sum_{i=1}^{p} w_i^2 \sigma_i^2 + 2 \sum_{i=1}^{p} \sum_{j=1}^{i-1} w_i w_j \sigma_{ij}$$

On the surface, covariances count twice. However, the variances themselves can be divided into shared variance and unshared or unique variance. The unique variance of each indicator, then, is counted only once in the variance of the composite, while shared variance is counted three times — once within the variance term of the indicator and twice among the covariances. So unique variance, including random measurement error, is substantially underweighted in forming the composite. This effect can be accentuated by using differential weights that put more weight on more reliable indicators. In other words, the very act of creating a weighted sum partly achieves the goal of "accounting measurement error." Yes, a factor-based analysis could be used to perform a more thorough accounting…but the researcher cannot know whether that accounting will be correct, in the absence of heroic assumptions. Thus, it becomes entirely unclear whether the "adjustment for measurement error" induced by a factor model is "better" in any meaningful way than what happens when weighted composites are employed as proxies, as in PLS path modeling.

Instead of viewing factor analysis as a pathway to the unobservable, while viewing composite-based methods as mere approximations, it is more plausible to regard both factor analysis and composite-based methods like PLS path modeling as alternative methods for constructing empirical approximations to those underlying concepts:

> Latent variable models are not detectors of unobservable latent structures, properties/attributes, causal sources, or anything else. They are of a kind with component models…Both component

and latent variable models are really *replacement variate generators*, quantitative recipes for the construction of variates, $\boldsymbol{\theta}_k$, k = 1…m, that *replace* a set of variates X_j, j=1…p, in some particular, optimal fashion. (Maraun and Halpin, 2009, p. 115).

Both factor methods and composite methods take the indicators as input and form empirical representations of the theoretical concepts. "Principal components analysis has often been characterized as a convenient but inferior 'approximation' of factor analysis. Instead, it might be more realistic to view factor analysis as a computationally difficult, theoretically problematic, approximation of component analysis" (Steiger and Schönemann, 1978, p. 174). Given this, it behooves researchers to think carefully about which method — factor-based SEM or composite-based PLS path modeling — better serves their research purposes (Fornell and Bookstein, 1982).

But researchers may also need to reconsider the whole notion of measurement. If factor analysis and the received view of psychometrics are bound together, then a challenge to one must raise questions about the other: "If Spearman-Thurstone factor analysis should be discarded…we must also discard structural equation models with latent variables (obviously), and virtually all of classical test theory, internal consistency reliability, including Cronbach's alpha, and most of validity theory, since these developments rest explicitly or implicitly on the common factor model…" (McDonald, 1996b, p. 597). The Holistic Construal described measurement primarily in terms of theoretical concepts, constructs, and indicators or observed variables. More than one author has noted confusion regarding the use of terms like concept and construct. Within the organization research literature, Schwab (1980, p. 5) long ago discarded any distinction: "…a construct will be thought of as referring to a conceptual variable." Similarly, Rossiter (2010; 2011) identifies the term, "construct" with "a conceptual definition," suggesting equality between "concept" and "construct." Michell (2013) argues that the term "construct" has so many meanings, and is so loaded with excess baggage as a term, that it has become essentially useless. Yet there is need for a term for something constructed from a set of indicators, which stands in the place of an intrinsically theoretical, idealized entity. Therefore, it may be helpful to modify the Holistic Construal framework, by 1) explicitly representing the concept itself, distinct from any empirical representation; and 2) replacing the "construct," which lies between concept and indicator, with something more clearly identified as being an empirical stand-in for the theoretical concept. The most appropriate word might be "proxy" — the term used in PLS path modeling to describe this very thing. These modifications lead to an alternative measurement framework, which might be called the Concept Proxy Framework (see Figure 2).

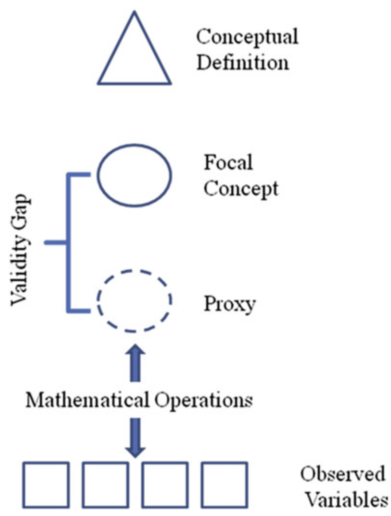

Figure 2. Concept proxy framework

Recognizing the distinction between concept and proxy means recognizing the actual steps and the actual difficulties in measuring theoretical concepts. The theoretical concept remains idealized and out of reach, yet scientific research is centered on these concepts (Schwab, 1980). Using any of a variety of means, the researcher constructs approximations to these theoretical concepts out of the raw material of observed variables. The question of measurement validity is not about the links between indicators and proxies — these are mere mathematical operations, with no causal significance. Instead, measurement validity questions center on the similarity in behavior between each proxy and the theoretical concept for which it stands. This is a tremendous challenge — but it is the real task that researchers should have been addressing all along.

Authors who have lately challenged the conventional view of measurement would surely object to aspects of this framework, as well. Rossiter (2010, 2011) denies the role of empirical evidence in measurement validation. Indeed, if there is no actual concept but only a theoretical definition, then validation cannot encompass anything more than an assessment of fidelity between the definition and the content of the measurement item. Rossiter also looks askance at multiple-item measurement, while the Concept Proxy Framework, informed by the forecasting community's experience, values multiple indicators. From another perspective, Michell (2008) would likely insist that this framework fails to address the key question of whether the concept is merely categorical/ordinal or is continuously valued. As Michell (2008) argues, measurement is only possible for concepts that are continuously valued.

A forecasting perspective

The conventional view of measurement which merged "theoretical concept" with "construct" with "common factor" promised researchers easy answers. One needed only to develop and refine a set of multiple indicators and estimate a factor model, and then one had in hand an empirical representation which could be assumed identical to the theoretical concept. By contrast, recognizing concept and proxy as distinct, and recognizing that the validity of inferences will depend on the similarity between a never-observed concept and a constructed proxy, emphasizes the difficulty in making valid inference. Recognizing the concept/proxy distinction does not create the difficulty; rather, it allows researchers to finally acknowledge the true nature of the challenge that they face.

Rossiter (2010, 2011) argues that validity here is a conceptual, not empirical, question, and insists that expert judgment of content is the only means by which the validity of measures may be evaluated. Recognizing concepts as distinct from their proxies and as inherently unobservable could lead to the conclusion that conceptual evaluation is the best one might achieve.

However, just as researchers are challenged to make inferences about theoretical concepts like "competitiveness," forecasters are challenged to make predictions about observations which are unseen because they lie in the future. It may be helpful to view measurement as a forecasting problem, with individual observed variables as separate forecasts and the proxy as a combination of those individual forecasts. When not describing PLS path modeling as an approximation to factor-based SEM, the earliest authors in the PLS literature emphasized the method's strength in prediction, and researchers applying PLS path modeling often assert the "predictive" nature of their research, though researchers often seem to mean nothing more than "aiming to maximize R^2 for dependent variables." Pursuing this approach in more depth highlights differences between factor-based SEM and composite-based methods, and also points the way toward a better future for composite-based modeling.

Factor indeterminacy

Factor-based methods are fundamentally unsuited to prediction-oriented research. It is true that the factor model approach can produce R^2's at the factor level which are higher than R^2's achieved in regression with observed variables, because of the way the factor model partials variance (Rigdon, 1994). On the other hand, it is also well known that factor analysis is designed to explain covariance, while principal component analysis is better suited for explaining variance (Jöreskog, 1979), so that researchers will expect higher R^2 values for the observed variables from the composite-based approach, but a better reproduction of the covariance matrix of the observed variables from a factor analysis.

However, the most important problem in using factor methods for prediction has to do with factor indeterminacy (for a chronicle of the extensive controversy in this area, see Mulaik, 2010, ch. 13; Steiger, 1979). Factor analysis involves turning p observed variables into m common factors plus p errors or unique factors. Assuming that identification constraints are met, researchers can estimate the parameters of the factor model, yielding summary moments (means, variances and covariances) for the common factors. For each respondent or case, it is a small matter to generate a vector of case values (scores on each common factor for each respondent) that conform to a factor's moments. The problem is that this set of values is not unique — an infinite number of different sets of values can be created which will fit the model equally well (e.g., Schönemann, 1971). In the context of regression, Anscombe (1973) demonstrated how radically different sets of observations could produce the same summary moments and the same regression line, but with very different distributions and different implications for predicting new values from the observed data. Factor indeterminacy describes something very similar.

Factor indeterminacy is not a matter of scaling or the result of random sampling (Schönemann, 1971). Factor indeterminacy is a consequence of factors models' attempt to break a fundamental rule in mathematics, having to do with the rank of a matrix. While the formal definition of rank is technical (Mulaik, 2010, p. 47), in connection with a covariance matrix, rank can be thought of as the maximum number of non-redundant dimensions of information available in the matrix. For a covariance matrix with p observed variables, maximum rank is p, meaning that there are at most p dimensions of information. If any of the p observed variables are perfectly redundant with one or more of the others, then rank will be less than p, but rank can never be more than p.

Composite methods, such as principal components, render a covariance matrix of rank p into at most p components. Factor methods, by contrast, aim to extract m + p factors. This is the root cause of factor indeterminacy (Schönemann and Steiger, 1976). It is impossible for those m + p factors to all be fully determined, because there isn't enough information in the data.

As first demonstrated by Spearman (Steiger and Schönemann, 1978), to reflect this, each factor must be constructed partly as a function of its indicators (with the exact function defined by the model's parameters) but also partly as a function (similarly defined) of a random variable not in the model. The only limitation on the nature of this random variable is that it must be orthogonal with the factor's indicators. That leaves a great deal of room for uncertainty (Wilson, 1928, Guttman, 1955).

Within the realm of factor models, this phenomenon has "disturbing consequences" (Fornell and Bookstein, 1982, p. 449). A given factor model is consistent with infinitely many different vectors of case values (Schönemann and Steiger, 1976, Schönemann and Wang, 1972). Moreover, differences between these various sets of factor scores and the original factors can be substantial. The correlation between different sets of equally compatible factor scores can even be negative (Guttman, 1955; Schönemann and Wang, 1972), again depending on the random component a researcher happens to use in constructing the two sets of factor scores. The arbitrary nature of these factor scores can produce peculiar results for the researcher who tries to use a common factor to predict other variables not included in that factor model. Say a researcher uses data on a set of p observed variables to estimates a factor model of organizational innovativeness, and then attempts to use that factor to predict two organizational outcomes, outside of that factor model. Suppose that outcome A is highly correlated with the indicators of the innovativeness factor, while outcome B is uncorrelated with those indicators. It is entirely possible that the factor derived from those indicators may be more highly correlated with outcome B than with outcome A (Schönemann and Steiger, 1978). If the p error factors are used as predictors along with the common factor, it is always possible to predict any outside-the-model outcome perfectly, regardless of the correlation between the outcome and the indicators (Schönemann and Steiger, 1978). As Schönemann and Steiger (1978, p. 290) ask, "Why would anyone want to estimate factors with the absurd properties described…?" Factor indeterminacy fundamentally threatens the usefulness of factor methods in research (Guttman, 1955, p. 79).

Moreover, the tools for "estimating" factor scores, included in standard factor-based SEM software packages, actually produce none of these vectors of genuine factor scores. Instead, standard procedures, such as the "regression method of factor score estimation" (Bollen, 1989, p. 305) construct one set of "estimated factor scores" effectively by averaging across all of these diverse sets of genuine factor scores (Schönemann and Steiger, 1976). Researchers can expect that most of the "estimated factor score" methods available, including Jöreskog's (2000) "latent variable score" method, will indeed reproduce the covariance matrix of the observed variables when assumptions hold (Beauducel, 2007). However, researchers who attempt to use these "estimated factor scores" as stand-ins for the factors themselves and make predictions about the values of individual observations may be "…asking more from the factor score estimates than they can provide" (Bollen, 1989, p. 306). Reproducing model parameters is not the same thing as making valid predictions about individual observations.

Can factor model users "solve" the problem of factor indeterminacy? Yes. Recall that the root cause of indeterminacy is a problem of rank, trying to extract m + p common and unique factors from a rank p covariance matrix. So one approach to eliminating indeterminacy is to resolve this problem. Krijnen et al. (1989) showed that indeterminacy disappears if some of the unique factors are forced to have variances of 0 — in other words, if the researcher extracts fewer unique factors, thus resolving the rank problem. Indeterminacy also disappears at the limit if the number of indicators per factor approaches infinity, or if the correlation between common factors and their indicators approach 1. Indeed:

> In the final analysis, perhaps the simplest yet most telling messages common to the often-conflicting work on indeterminacy are that high-quality, well-planned measurement procedures can alleviate indeterminacy problems. Ironically, such measures are in fact *components*, since the only truly determinate 'factor' is one that is perfectly predictable from [its indicators]. It seems high-reliability linear composites are what the factor analytic community has been seeking all along (Steiger, 1996, p. 549-550).

In other words, the best solution to the factor indeterminacy problem is to avoid the use of factors.

Regression weights

Research on forecasting long ago established that combinations of individual forecasts tend to yield more accurate predictions (Makridakis and Winkler, 1983). Plausibly, different methods for combining forecasts can yield differences in forecast accuracy. Indeed, typical PLS path modeling packages offer different approaches for combining indicators or proxies to achieve best predictions. In the outer model, Mode B involves a multiple ordinary least squares (OLS) regression of the proxy on its associated indicators. OLS regression produces weights which optimize R^2 (Cohen et al., 2003). Mode B estimation in PLS path modeling is known to produce higher R^2 values for proxies of dependent concepts, as compared to R^2 values using Mode A estimation (Dijkstra, 2010, pp. 35-36; Rigdon, in press). So researchers might well expect that using OLS regression with the data in hand might produce a set of weights that would provide the most accurate out-of-sample prediction (prediction of data not used to estimate the parameters of the prediction model).

This intuition appears to be incorrect according to the work of Dana and Dawes (2004), who recommend that "…regression coefficients should almost never be used for social science predictions" (Dana and Dawes, 2004, p. 328). Dana and Dawes's (2004) simulation results suggest that, if regression is used to estimate weights, which are then fixed and used to make predictions based on data not used to estimate those weights, then OLS regression provides optimum results only under very specialized circumstances. Dana and Dawes (2004) varied the degree of predictability in their data populations, the number of predictors (from three to ten), population collinearity among the predictors, and the sample size used to estimate weights. OLS weights were superior only when predictability was high (with an average validated R — multiple correlation based on prediction of the complete population, with parameters fixed at sample values — above .9) and sample size was large. As a rule of thumb for minimum sample size, instead of the often-heard ten observations per predictor (Chin, 1998, p. 311), Dana and Dawes (2004, p. 328) recommend OLS weights for prediction only with sample sizes larger than 100 observations per predictor. Even at that sample size, OLS weights were superior only when validated R was above .6.

Correlation weights

A substantial stream of research in the forecasting literature has examined alternatives to OLS regression weights. Some of these techniques, like ridge regression, are rather sophisticated, but many are even simpler than OLS regression. Dana and Dawes's (2004) study dealt mostly with the latter. One alternative to OLS regression weights is "correlation weights." Unlike regression weights, correlation weights ignore covariance among the predictors. In simple matrix terms, if X represents a vector of predictors, y represents a dependent variable, S_{XX} represents the covariance matrix of the predictors and S_{Xy} represents the covariances between the predictors and the dependent variable, then OLS regression weights B_r can be estimated as:

$$\widehat{B_r} = S_{XX^{-1}} S_{Xy}$$

Correlation weights B_c would be estimated as:

$$\widehat{B_c} = S_{Xy}$$

Accounting for correlation between predictors means that relatively redundant predictors will be underweighted, while more unique predictors will carry more weight. Again, intuitively, such an adjustment seems appropriate, likely to produce a set of weights that draws the most unduplicated predictive ability out of a given set of predictors. Yet, a stream of research has favored correlation weights over regression weights, at least under some circumstances (Goldberg, 1972; Waller and Jones, 2010). Dana and Dawes (2004) found correlation weights to be superior to OLS regression weights with validated R of .6 or less, and with somewhat higher validated R when sample size is "low," meaning fewer than 15 observations per predictor. Thus, across a wide range of practical values for degree of predictability and sample size, correlation weights yield the same or better out-of-sample prediction.

Conveniently for users, standard PLS path modeling packages include correlation weights as an option for the outer model, though not under that name. What the PLS literature calls "Mode A," and links with factor models and reflective measurement, is actually a correlation weighting scheme. Under Mode A, in a simple model with two concepts labeled X and Y, with individual indicators labeled x and y, the weight for each indicator of concept Y is estimated as (Dijkstra, 2010, pp. 32-33, adjusted for the two-concept case):

$$\widehat{w_y} \propto \text{sign}_{yx} S_{yx} \widehat{w_x}$$

where $\widehat{w}$ indicates an estimated weight, "sign" indicates the sign function, and S_{yx} is the cross-covariance matrix containing the covariances of the indicators of one concept with the indicators of the other concept, and summation is over all indicators in the other block. The operator "$\propto$" means "proportional to," and accounts for the standardization which is commonly applied in PLS path modeling. Under Mode B estimation, these weights would be estimated as:

$$\widehat{w_y} \propto \text{sign}_{yx} S_{xx}^{-1} S_{yx} \widehat{w_x}$$

The difference between the two cases is that Mode B estimation incorporates S_{xx}, the covariance matrix of the indicators of the X concept, just as regression weighting uses the covariance matrix of the predictors, while Mode A estimation omits this matrix, just as it is omitted when estimating correlation weights.

The equivalence between PLS Mode A and correlation weights suggests that PLS path modeling users can expect results from Mode A which mirror results from correlation weights. While Mode B's regression weights can be expected to produce higher R^2 values for the data used to estimate the weights, users can expect Mode A's correlation weights to provide superior out-of-sample prediction across a wide and practical range of circumstances. Instead of choosing Mode A estimation to mimic a questionable factor-based measurement model, PLS path modeling users should instead choose Mode A estimation if their focus is on optimizing out-of-sample prediction, under suitable conditions.

Correlation weights yield one additional advantage: no unexpected values. With regression weights, positive zero-order correlations can turn into negative weights, due to the effects of collinearity or suppression involving other predictors. Such sign reversals are not errors, but they do complicate interpretation. Since correlation weights ignore collinearity among the predictors, positive correlations between predictors and dependent variable yield positive weights, and negative correlations yield negative weights. This simplifies interpretation and avoids challenges in explaining results to unsophisticated audiences.

Unit or equal weights

Correlation weights are not the only alternative, however, and are certainly not the most simple approach one can imagine. Even simple averages of individual forecasts (Einhorn and Hogarth, 1975), using implicit equal or unit weights (weights fixed to 1), can perform remarkably well: "Empirical evidence indicates that simple combining schemes (i.e., using a simple arithmetic average to combine several/many methods) are as accurate as elaborate ones" (Makridakis, 1989, p. 602). When predictors are standardized, as they typically are in PLS path analysis, equal weights and unit weights produce the same result (Rozeboom, 1979). When using equal weights, one assumes that predictors are all scaled, or rescaled, so that their correlations, among themselves and with the dependent variable, are all positive. Wainer (1976) observed that, with regard to setting weights for a composite under such conditions, "It don't make no nevermind." Unit weights also offer a variety of advantages over regression weights. Because they are not estimated from data, they are not affected by collinearity, heterogeneity or outliers. Their error of estimation or standard error is always zero (Einhorn and Hogarth, 1975), and they can be used when sample size is too small to use anything else. In addition, simple averaging offer even greater transparency than correlation weights. This transparency is important not only in easing explanations for clients or other audiences but also in enabling other researchers to incorporate a given study's results into their own mental models: "Learning requires some form of transparency, which forecasters can best achieve when they understand what they are doing" (Goldstein and Gigerenzer, 2009, p. 770).

Dana and Dawes's (2004) results supported the out-of-sample predictive performance of unit weights. In their simulations, unit weights performed best — better than correlation weights and better than regression weights — when the level of predictability in the population was low (with average validated R below .4) and especially when sample size was low. In their simulations, when sample size was no more than ten times the number of predictors, unit weights outperformed all other options even when average validated R was as high as .6. With lower sample sizes, the advantage stayed with unit weights even when average validated R was higher than .6. Thus, when sample size is "low," and in situations where predictability is expected to be low, just summing together each concept's indicators to form one index for each concept is a valid and valuable alternative. Standard PLS packages do not offer a unit weight option, but researchers should certainly consider creating their own unit-weight composites. Unit-weight composites would at least provide a reality check, ensuring that researchers gain something in return for adopting the added complexity of more sophisticated methods.

Single indicators

A substantial though controversial literature backs the value of simple heuristics in decision-making and forecasting in the face of uncertainty (Gigerenzer et al., 2011). Even more basic than forming a simple average, some researchers have advocated using just a single indicator to represent a concept (Rossiter, 2011), and the heuristics literature demonstrates the utility and even superiority of relying on single cues under particular conditions (Goldstein and Gigerenzer, 2009). Still, foundational papers in the forecasting literature (Clemen, 1989), and recent work in marketing (Diamantopoulos et al., 2012) point to the predictive superiority of the most basic combinations of forecasts: "Consider what we have learned about the combination of forecasts over the last twenty years…The results have been virtually unanimous: combining multiple forecasts leads to increased forecast accuracy. This has been the result whether the forecasts are judgmental or statistical, econometric or extrapolation" (Clemen, 1989, p. 559). Dana and Dawes (2004) simulations uncovered no circumstances where a single indicator strategy was the best approach. In the interest of making things as simple as possible, but no simpler (Einstein, 1934, p. 165), — researchers appear to be well advised to favor multiple indicators.

Fungible weights

Researchers with only a basic grounding in regression may find it surprising that simplified alternatives can defeat ordinary least squares regression. Scholarly response to the research demonstrating the superiority of simple methods in forecasting was hostile and dismissive (Goldstein and Gigerenzer, 2009). Basic statistics texts assure students of the optimality of OLS regression when assumptions hold. It is important to remember that optimality, even in terms of within-sample R^2, does not mean "optimal by a substantial margin." The advantage for OLS regression can be so small as to be effectively irrelevant. Waller (2008) demonstrated that, for a given dataset, many sets of alternate weights exist which yield only a slightly smaller R^2 despite carrying very different substantive implications. Waller (2008) labeled these "fungible weights."

As an example, Waller (2008, pp. 697-698) described a regression model with three predictors where OLS regression yielded within-sample $R^2 = .176$. Waller developed a method to derive alternative sets of weights, with each set yielding $R^2 = .171$, a reduction of only .005. With that slight reduction (less than 3% of the OLS regression R^2 value), Waller produced alternate sets of weights with dramatically different interpretations. Different sets of weights might assign the same predictor a positive weight or a negative weight, and might assign that predictor the largest standardized weight of all predictors, or might assign that predictor a weight of zero. The point for the current discussion is that the advantage of OLS regression weights, even for within-sample R^2, may not be very large. Other, very different weights can yield almost the same within-sample R^2 — and may yield higher out-of-sample R^2.

Conclusions and recommendations

There are three main implications to be drawn from this paper. First, PLS path modeling can and should separate itself from factor-based SEM and renounce entirely all mechanisms, frameworks and jargon associated with factor models. Instead of emphasizing how it is "like" factor-based SEM but with advantages and disadvantages across different conditions, PLS path modeling should celebrate its status as a purely composite-based method. PLS path modeling gains nothing from its "factor-like" pretensions, but has much to gain from leaving this legacy thinking behind. Obvious as it is, this will not be easy. It has been noted before that the partisanship of factor model supporters approaches religious zeal (Steiger, 1990). Factor analysis pioneer Charles Spearman had more than modest aspirations for the general intelligence factor which be believed he had derived using his method: "It seems not altogether chimaeric to look forward to the time when citizens, instead of choosing their career at almost blind hazard, will undertake just the professions really suited to their capacities. One can even conceive the establishment of a minimum index to qualify for parliamentary vote, and, above all, for the right to have offspring" (Hart and Spearman, 1912, pp. 78-79). Arguments based upon logic, basic analysis, and decades of experience are more than sufficient to justify favoring composite methods over factor methods in practical research.

Second, the PLS path modeling community should work to complete and validate a purely composite-based approach to evaluating modeling results. Rejecting the factor model does not mean rejecting rigor, but it does mean defining rigor in composite terms. There are clearly bad measures and bad models out there — measures which do not closely coincide with conceptual definitions, and models that are inconsistent with the data and that have weak explanatory power. PLS path modeling users should decide, within their own research situations, which variety of predictive power (within-sample or out-of-sample) is more important. Moreover, PLS path modeling can follow the example of logistic regression and develop an approach to model assessment that places results from a particular analysis within a framework of worst-case and best-base alternative models (Chin, 2010 offered a similar suggestion). The equal weights approach represents an excellent worst case — not just because this is a commonly used naïve strategy, but because, as Dijkstra and Henseler (2011) and the forecasting literature broadly show, equal weights will often perform quite well. One candidate for best case is Cohen's (1982) set correlation (Rigdon and Gefen, 2011). PLS path modeling typically forms just one composite for each set of indicators, constraining the relationship between each set of predictors to be unidimensional. Cohen's (1982) set correlation accounts for the explanatory power of relationships between sets of predictors across multiple dimensions. Comparing a particular model with a set correlation analysis would show just how much predictive ability a researcher was "leaving on the table" for the sake of specifying unidimensional relationships. If the gap is large, a researcher would have to question the unidimensionality assumption. PLS path modeling certainly does include the ability to extract and model multiple dimensions from given sets of indicators (Apel and Wold, 1982), but researchers rarely make use of this possibility, and, as with equal weights, standard PLS packages do not even provide this functionality.

Third, given that the weaknesses of the factor model are built into standard approaches to evaluating the quality of measures, walking away from the factor model means that researchers must develop an entirely different approach to measure validation. In line with PLS path modeling's emphasis on prediction, a logical candidate for an alternative measurement framework is one that is based on forecasting theoretical concepts using observed variable indicators as individual forecasts, with proxies, standing between the concepts and the indicators, as combinations of forecasts. Each proxy/combination should be expected to perform better than each individual indicator/forecast, but evaluating validity will chiefly mean determining the magnitude of the gap between proxy and concept. The forecasting literature recognizes feedback as the crucial element in improving forecast accuracy over time (e.g., Arkes, 2001). Collecting a body of feedback means building databases of research results over time — a thankless but crucial task. Instead of dutifully reporting a collection of statistics that address the fit of factor models or the magnitude of factor model parameters but don't address the question of measurement validity (Cliff, 1983, p. 122), researchers must commit themselves to actually nailing down the behavior of important theoretical concepts and the laws which link them. Social science disciplines must also commit themselves to replication, to build the databases of results that will enable the feedback that will improve forecast accuracy/measurement validity over time. In truth, this burden is not newly invented but only newly acknowledged, once the methodological haze of factor models is cleared away.

Future research questions

This paper certainly points to many areas needing further investigation. While this paper has leaned on results from the forecasting literature, PLS path modeling users will still need to ensure that all of the implications drawn from that literature truly apply in the PLS path modeling context. PLS path modeling uses weighting at two different stages in its estimation process. This paper has focused on what are known as "outer weights," where indicators are combined to predict a composite. PLS path modeling also uses "inner weights," where composites are combined to predict other composites. Standard PLS packages offer different approaches for estimating these inner weights. Applying results from the forecasting literature here again leads to certain expectations. Are those expectations born out in practice? And does the interaction of inner weighting schemes and outer weighting schemes yield the expected results? Users will certainly look forward to definitive results and guidance on limiting conditions.

The forecasting literature has shown that the performance of different weighting schemes depends in part on the degree of predictability in the data (Dana and Dawes, 2004). How predictable are the concepts at the core of strategic management theory? It has been argued that phenomena in general, including in the business world, are substantially unpredictable (Makridakis and Taleb, 2009; Taleb, 2007). If the true level of predictability for strategic management concepts is actually fairly low, and if findings from the forecasting literature apply, this would provide further impetus toward the use of correlation weights (Mode A) or unit weights. Then again, a combination of unpredictability in the theoretical concepts and the errors associated with research methods could provide a grim prognosis for the future of academic research in an area.

A body of experience supports the use of composite-based methods for modeling structural relations, but PLS path modeling is only one of the methods in this class. While the focus on factor-based methods may have limited adoption of PLS path modeling, at the same time PLS path modeling may have been advantaged over other composite-based methods by its seeming similarity to factor-based SEM. In the headlong rush to adopt factor-based methods, a number of promising approaches have been ignored or left under-explored. The list of such methods includes, for starters, Shönemann and Steiger's (1976) regression component analysis and Cohen's (1982) set correlation. Including PLS path modeling, Hwang and Takane's (2004) generalized structured component analysis, and Buckler's (2010) nonlinear composite technique, composite-based methods offer a wide field for methodological research — once it is agreed that composite-based methods are worth the research attention. How truly different are they? How much overlap do they offer, and under what conditions? Makridakis et al. (2009, pp. 794-795) observe broadly that, "Statistically sophisticated, or complex, models fit past data well, but do not necessarily predict the future accurately…'Simple' models do not necessarily fit past data well, but predict the future better than complex or sophisticated statistical models." If researchers truly aim at prediction, then they must test and compare methods against this specific criterion.

Bringing composite-based methods including PLS path modeling to full maturity will mean developing (1) a complete suite of inferential tools for these methods, and (2) a complete and consistent approach to measurement which is factor-free. With more researcher attention devoted to this area, the field will surely see not just convenient and plausible suggestions but well-argued and thoroughly studied techniques that expand the range of inferential meaning which users can draw from composite analysis. The greater challenge lies in capturing and quantifying the gap between composite proxy and theoretical concept. As noted previously, forecasters have the advantage that the future becomes known over time, providing feedback which improves forecast accuracy, while theoretical concepts remain unobserved. Those researchers seeking a big problem to solve need look no further.

Conclusion

From the very beginning of the two methods, there have been reasons to question the dominance of factor-based SEM over composite-based PLS path modeling, and reasons to favor the latter. Recently, researchers have begun anew to question foundational principles upon which the dominance of factor methods stands, while composite methods have continued to evolve. Combined, these arguments point to a new future where composite methods, including PLS path modeling, stand as multivariate methods of choice for researchers not only in strategic management but across the social sciences.

Acknowledgements

The author gratefully acknowledges excellent critical feedback from two anonymous reviewers, from Marko Sarstedt, and from the marketing faculty at the University of Nebraska-Lincoln, especially A. Dwayne Ball and James Gentry, and the kind support of the editors of the special issue.

References

Anscombe, F.J., 1973. Graphs in statistical analysis. Amer. Stat. 27, 17—21.

Apel, H., Wold, H., 1982. Soft modeling with latent variables in two or more dimensions: PLS estimation and testing for predictive relevance. In: Jöreskog, K.G., Wold, H. (Eds.), Systems under Indirect Observation: Causality, Structure, Prediction (Part II). North-Holland, Amsterdam, pp. 209—248.

Arkes, H.R., 2001. Overconfidence in judgmental forecasting. In: Armstrong, J.S. (Ed.), Principles of Forecasting. Springer, New York, pp. 443—515.
Bagozzi, R.P., 1984. A prospectus for theory construction in marketing. J. Mark. Res. 48, 11—29.
Bagozzi, R.P., 2011. Measurement and meaning in information systems and organizational research. MIS Quar. 35, 261—292.
Bagozzi, R.P., Phillips, L.W., 1982. Representing and testing organizational theories: the holistic construal. Admin. Sci. Quar. 27, 459—489.
Bagozzi, R.P., Yi, Y., 2012. Specification, evaluation, and interpretation of structural equation modeling. J. Acad. Mark. Sci. 40, 8—34.
Beauducel, A., 2007. In spite of indeterminacy many common factor score estimates yield an identical reproduced covariance matrix. Psychometrika 72, 437—441.
Bentler, P.M., 1986. Structural modeling and Psychometrika: an historical perspective on growth and achievements. Psychometrika 51, 31—51.
Bollen, K.A., 1989. Structural Equations with Latent Variables. John Wiley and Sons, New York.
Buckler, F., 2010. NEUSREL 3.2 Universal Structural Modeling Software. Neusrel Causal Analytics, Cologne.
Chin, W.W., 1998. The partial least squares approach to structural equation modeling. In: Marcoulides, G.A. (Ed.), Modern Methods for Business Research. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 295—336.
Chin, W.W., 2010. Bootstrap cross-validation indices for PLS path model assessment. In: Esposito Vinzi, V., Chin, W.W., Henseler, J., Wang, H. (Eds.), Handbook of Partial Least Squares: Concepts, Methods and Applications. Springer-Verlag, Berlin, pp. 83—97.
Churchill, G.W., 1979. A paradigm for developing better measures of marketing constructs. J. Marketing Res. 16, 64—73.
Clemen, R.T., 1989. Combining forecasts: a review and annotated bibliography. Int. J. Forecasting 5, 559—583.
Cliff, N., 1983. Some cautions concerning the application of causal modeling methods. Multivar. Behav. Res. 18, 115—126.
Cohen, J., 1982. Set correlation as a general multivariate data-analytic method. Multivar. Behav. Res. 17, 301—341.
Cohen, J., Cohen, P., West, S.G., Aiken, L.S., 2003. Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, third. ed. Lawrence Erlbaum Associates, Mahwah, NJ.
Cronbach, L.J., Meehl, P.E., 1955. Construct validity in psychological tests. Psych. Bull. 52, 281—302.
Dana, J., Dawes, R.M., 2004. The superiority of simple alternatives to regression for social science prediction. J. Educ. Behav. Stat. 29, 317—331.
Diamantopoulos, A., Sarstedt, M., Fuchs, C., Wilczynski, P., Kaiser, S., 2012. Guidelines for choosing between multi-item and single-item scales in construct measurement: a predictive validity perspective. J. Acad. Mark. Sci. 40, 434—449.
Dijkstra, T.K., 2010. Latent variables and indices: Herman Wold's basic design and partial least squares. In: Esposito Vinzi, V., Chin, W.W., Henseler, J., Wang, H. (Eds.), Handbook of Partial Least Squares: Concepts, Methods and Applications. Springer-Verlag, Berlin, pp. 23—46.
Dijkstra, T.K., Henseler, J., 2011. Linear indices in nonlinear structural equation models: best fitting proper indices and other composites. Qual. Quant. 45, 1505—1518.
Einhorn, H.J., Hogarth, R.M., 1975. Unit weighting schemes for decision making. Organ. Behav. Hum. Perf. 13, 171—192.
Einstein, A., 1934. On the method of theoretical physics. Phil. Sci. 1, 163—169.
Fornell, C., Bookstein, F.L., 1982. Two structural equation models: LISREL and PLS applied to exit-voice theory. J. Mark. Res. 19, 440—452.
Gerbing, D.W., Anderson, J.C., 1988. An updated paradigm for scale development incorporating unidimensionality and its assessment. J. Mark. Res. 25, 186—192.
Gigerenzer, G., Hertwig, R., Pachur, R. (Eds.), 2011. Heuristics: The Foundations of Adaptive Behavior. Oxford University Press, Oxford, UK.
Goldberg, L.R., 1972. Parameters of personality inventory construction and utilization: a comparison of prediction strategies and tactics. Multivar. Behav. Res. M. 72—82.
Goldstein, D.G., Gigerenzer, G., 2009. Fast and frugal forecasting. Int. J. Forecasting 25, 760—772.
Guttman, L., 1955. The determinacy of factor score matrices with implications for five other basic problems of common-factor theory. Brit. J. Stat. Psych. 8, 65—81.

Hair Jr., J., Sarstedt, M., Ringle, C.M., Pieper, T., 2012. Use of partial least squares in strategic management research: a review of past practices and recommendations for future applications. Long Range Plann 45, 320–340.

Hart, B., Spearman, C., 1912. General ability, its existence and nature. Brit. J. Psych. 5, 51–84.

Hwang, H., Takane, Y., 2004. Generalized structured component analysis. Psychometrika 69, 81–99.

Jöreskog, K.G., 1969. A general approach to confirmatory maximum likelihood factor analysis. Psychometrika 34, 183–202.

Jöreskog, K.G., 1979. Basic ideas of factor and component analysis. In: Jöreskog, K.G., Sörbom, D. (Eds.), Advances in Factor Analysis and Structural Equation Models. Abt Books, Cambridge, MA, pp. 5–20.

Jöreskog, K.G., 2000. Latent variable scores and their uses. http://www.ssicentral.com/lisrel/techdocs/lvscores.pdf (accessed 11.10.12).

Jöreskog, K.G., Sörbom, D., 1996. LISREL 8: User's Reference Guide. Scientific Software, Chicago.

Krijnen, W.P., Dijkstra, T.K., Gill, R.D., 1998. Conditions for factor (in)determinacy in factor analysis. Psychometrika 63, 359–367.

Makridakis, S., 1989. Why combining works? Int. J. Forecasting 5, 601–603.

Makridakis, S., Hogarth, R.M., Gaba, A., 2009. Forecasting and uncertainty in the economic and business world. Int. J. Forecasting 25, 794–812.

Makridakis, S., Taleb, N., 2009. Decision-making and planning under low levels of predictability. Int. J. Forecasting 25, 716–733.

Makridakis, S., Winkler, R.L., 1983. Averages of forecasts: some empirical results. Man. Sci. 29, 987–996.

Maraun, M.D., Gabriel, S.G., 2013. Illegitimate concept equating in the partial fusion of construct validation theory and latent variable modeling. New Ideas Psych. 31, 32–42.

Maraun, M.D., Halpin, P.D., 2009. Manifest and latent variables. Measurement 6, 113–117.

McDonald, R.P., 1996a. Path analysis with composite variables. Multivar. Behav. Res. 31, 239–270.

McDonald, R.P., 1996b. Latent traits and the possibility of motion. Multivar. Behav. Res. 31, 593–601.

McDonald, R.P., 1999. Test Theory: A Unified Treatment. LEA, Mahwah, NJ.

Michell, J., 2008. Is psychometrics pathological science? Measurement 6, 7–24.

Michell, J., 2013. Constructs, inferences and mental measurement. New Ideas Psych. 31, 13–21.

Mulaik, S.A., 2010. Foundations of Factor Analysis, second ed. Taylor and Francis Group, Boca Raton, FL.

Nunnally, J.C., Bernstein, I.H., 1994. Psychometric Theory, third ed. McGraw-Hill, New York.

Rigdon, E.E., 1994. Demonstrating the effects of unmodeled random measurement error. Struct. Equ. Model. 1, 375–380.

Rigdon, E.E. Partial least squares path modeling, in: Hancock, G.R.; Mueller, R.D. (Eds.), *Structural Equation Modeling: A Second Course*, second ed. Information Age, Charlotte, NC, in press.

Rigdon, E.E., Gefen, D., 2011. Questioning Some Claims Regarding Partial Least Squares Path Modeling, Presented at the American Marketing Association Summer Educators Conference, Chicago.

Rigdon, E.E., Preacher, K.J., Lee, N., Howell, R.D., Franke, G.R., Borsboom, D., 2011. Avoiding measurement dogma: a response to Rossiter. Eur. J. Mark. 45, 1589–1600.

Rossiter, J.R., 2010. Measurement for the Social Sciences: The C-OAR-SE Method and Why it Must Replace Psychometrics. Springer, New York.

Rossiter, J.R., 2011. Marketing measurement revolution: the C-OAR-SE method and why it must replace psychometrics. Eur. J. Mark. 45, 1561–1588.

Rozeboom, W.W., 1979. Sensitivity of a linear composite of predictor items to differential item weighting. Psychometrika 44, 289–296.

Schönemann, P.H., 1971. The minimum average correlation between equivalent sets of uncorrelated factors. Psychometrika 36, 21–30.

Schönemann, P.H., Steiger, J.H., 1976. Regression component analysis. Brit. J. Math. Stat. Psych. 29, 175–189.

Schönemann, P.H., Steiger, J.H., 1978. On the validity of indeterminate factor scores. Bull. Psychonomic Soc. 12, 287–290.

Schönemann, P.H., Wang, M.-W., 1972. Some new results on factor indeterminacy. Psychometrika 37, 61–91.

Schwab, D.P., 1980. Construct validity in organizational behavior. In: Staw, B.M., Cummings, L.L. (Eds.), 1980. Research in Organizational Behavior, vol. 2. JAI Press, Greenwich, CT, pp. 3–43.

Sörbom, D., 2001. Karl Jöreskog and LISREL: a personal story. In: Cudeck, R., du Toit, S., Sörbom, D. (Eds.), Structural Equation Modeling: Present and Future. Scientific Software, Lincolnwood, IL, pp. 3–10.

Steiger, J.H., 1979. Factor indeterminacy in the 1930s and the 1970s: some interesting parallels. Psychometrika 44, 157—167.
Steiger, J.H., 1990. Some additional thoughts on components, factors and factor indeterminacy. Multivar. Behav. Res. 25, 41—45.
Steiger, J.H., 1996. Dispelling some myths about factor indeterminacy. Multivar. Behav. Res. 31, 539—550.
Steiger, J.H., Schönemann, P.H., 1978. A history of factor indeterminacy. In: Shye, S. (Ed.), Theory Construction and Data Analysis in the Social Sciences. Jossey-Bass, San Francisco, pp. 136—178.
Taleb, N., 2007. The Black Swan: The Impact of the Highly Improbable. Random House, New York.
Velicer, W.F., Jackson, D.N., 1990. Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. Multivar. Behav. Res. 25, 1—28.
Wainer, H., 1976. Estimating coefficients in linear models: it don't make no nevermind. Psych. Bull. 83, 213—217.
Waller, N.G., 2008. Fungible weights in multiple regression. Psychometrika 73, 691—703.
Waller, N.G., Jones, J.A., 2010. Correlation weights in multiple regression. Psychometrika 75, 58—69.
Wilson, E.B., 1928. Review of 'The abilities of man' by C. Spearman. Science 67, 244—248.
Wold, H.O., 1982. Soft modeling: the basic design and some extensions. In: Jöreskog, K.G., Wold, H. (Eds.), Systems under Indirect Observation: Causality, Structure, Prediction (Part II). North-Holland, Amsterdam, pp. 1—54.

Biography

Edward E. Rigdon is Professor of Marketing in the J. Mack Robinson College of Business at Georgia State University. His research focuses on structural equation modeling and related methods. His research has been published in journals including *Journal of Marketing Research*, *Journal of Consumer Research*, *MIS Quarterly*, *Structural Equation Modeling* and *Multivariate Behavioral Research*. He is also list owner/moderator for SEMNET, an email discussion list devoted to structural equation modeling (https://listserv.ua.edu/archives/semnet.html).
E-mail: erigdon@gsu.edu