Long Range Planning 47 (2014) 161–167

Contents lists available at ScienceDirect

Long Range Planning

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

Rethinking Partial Least Squares Path Modeling: Breaking Chains and Forging Ahead

Edward E. Rigdon

Rigdon (2012) argued that researchers should embrace PLS path modeling as an explicitly composite-based technique, and in connection with this proposal sketched a "concept-centric" approach to measurement as an alternative to the dominant "factor-centric" measurement paradigm. Bentler and Huang (2014) reassert the classical linkage between factor analysis and measurement—but this linkage depends on the implausible assumption that indeterminate factors are identical to the conceptual variable in researchers' theoretical models. Dijkstra (2014) correctly notes that assessing measurement validity within Rigdon's framework is difficult—but the ease of validity assessment within factor-centric measurement framework is an illusion. Sarstedt et al. (2014) endorse Rigdon's call to divorce PLS path modeling from factor analysis, and the current paper offers some further thoughts about the costs and challenges of embracing PLS path modeling as a composite-based technique.

> "Concepts that have proven useful in ordering things easily achieve such authority over us that we forget their earthly origins and accept them as unalterable givens. Thus, they might come to be stamped as 'necessities of thought,' 'a priori givens,' etc. The path of scientific progress is often made impassible for a long time by such errors. Therefore it is by no means an idle game if we become practiced in analyzing long-held commonplace concepts and showing the circumstances on which their justification and usefulness depend, and how they have grown up, individually, out of the givens of experience. Thus, their excessive authority will be broken." Albert Einstein (1916, p. 101), *Nachruf auf Ernst Mach*, *Physikalische Zeitschrift 17, p. 101*

Introduction

It is a pleasure and an honor to exchange views with such eminent scholars in the pages of this journal. Theo Dijkstra has been active in this area ever since he was a graduate student under Herman Wold's supervision at Wharton when PLS path modeling first emerged decades ago. Observations in Dijkstra (2010) played a central role in precipitating the stream of thought represented in Rigdon (2012). Peter Bentler is a past winner of the Distinguished Scientific Applications Award from the American Psychological Association (an award presented jointly to Dr. Bentler and to Karl Jöreskog for their lifetime contributions to factor-based structural equation modeling), and Wenjing Huang recently completed her dissertation under Dr. Bentler's supervision, with additional guidance from Dr. Dijkstra. Marko Sarstedt, Christian Ringle, Jörg Henseler and Joe Hair have made more contributions to PLS path modeling than I would be allowed to list here. I open with these remarks to assure readers that I know the stature of the scholars with whom I am disagreeing. I also recognize that *Long Range Planning* is not a quantitative methods journal, and its readers will be more likely to benefit if this reply emphasizes conceptual issues and practical challenges, while minimizing mathematical minutiae. That is my intent, granting that the topic is statistical.

To recap, the issue revolves around the methodology of multiple indicator modeling, where a model includes multiple observed variables for every conceptual variable. Between each conceptual variable and its associated multiple observed variables stands a proxy, a single empirical stand-in for the conceptual variable. Factor-based structural equation modeling (SEM) employs common factors in this role, while composite-based methods employ weighted sums of the observed variables.

http://dx.doi.org/10.1016/j.lrp.2014.02.003
0024-6301/© 2014 Elsevier Ltd. All rights reserved.

Key arguments, responses and rejoinders

The key arguments in Rigdon (2012) were:

1. PLS path modeling is often discussed and used as if it was a kind of factor analysis, but mathematically it is purely a composite-based method.
2. PLS path modeling is criticized for being a composite-based method, but this composite-based character is a strength, not a weakness.
3. When assumptions hold, including the assumption that the factor model is correct in the population, factor-based methods excel in the estimation of factor model parameters.
4. In practice, that crucial assumption of correctness is typically contradicted by data, so that the optimality of factor methods for parameter estimation does not necessarily hold.
5. Factor-based methods are fundamentally unsuitable for prediction, especially for prediction outside the dataset used to estimate the factor model, because of factor indeterminacy.
6. The dominant measurement paradigm in the social sciences is itself a factor model, so that factor-based methods seem to have the advantage in evaluating measurement validity, under that paradigm.
7. That dominant measurement paradigm depends implicitly upon the improbable assumption that an indeterminate common factor derived from data is identical to a fundamentally unobservable conceptual variable.
8. Measurement validity is better understood as the degree of correspondence between a conceptual variable and an empirical proxy (whether formed as a composite, as a common factor, or otherwise) than in terms of the correspondence between the proxy and the observed variables from which the proxy is formed.
9. The PLS path modeling community needs to embrace the method's character as a composite-based method, but that requires shedding the factor-based jargon, perspectives, evaluation tools and measurement framework of factor-based SEM and developing alternatives.

It may be helpful to the reader if I organize my response to the comments in terms of these nine arguments.

Argument 1: A composite method used like a factor method. The very first sentence of the abstract of Bentler and Huang (2014) states, "Rigdon (2012) suggests that partial least squares (PLS) can be improved by killing it, that is, by making it into a different methodology based on components." The sentence at least implies the view that PLS path modeling is a factor-based technique. This is a rather remarkable implication. At the end of every PLS path modeling estimation, the proxies are expressed as exact weighted sums of their observed variable indicators — in other words, as composites. There is no approximation or indeterminacy, so PLS path modeling cannot be a factor-based method, even though, as Dijkstra (2014) notes, the emerging methodology of confirmatory factor analysis was heavy on the mind of Herman Wold when PLS path modeling was born.

Then again, Bentler and Huang (2014) and Dijkstra (2014) both note that PLS path modeling is not a "latent variable" method. "Latent variable" is an unfortunate phrase with multiple meanings (which is why the term should be avoided). By "latent variable," the authors might mean "common factor" or they might mean "conceptual variable." Sarstedt et al. (2014) uses the phrase in connection with PLS path modeling, and that paper is similarly unclear about the phrase's meaning. If Bentler and Huang's (2014) "latent variable" means "common factor," then their insistence that PLS path modeling is not a "latent variable" method contradicts the abstract's implication that PLS path modeling somehow is a common factor method.

Perhaps, just as Wold was greatly impressed by Karl Jöreskog's factor modeling contributions and may have been inclined to see factors where the mathematics described composites, in the same way Bentler and Huang are thinking of Dijkstra's (1981) proposal, now labeled "PLSc," which takes parameter estimates from PLS path modeling and transforms them into unbiased estimates of factor model parameters. I think that "PLSc" is best understood not as a form of PLS path modeling but as a variation on what Gerbing and Hamilton (1994) calls iterated centroid estimation or "ICE," a simpler-to-calculate alternative to maximum likelihood estimation ("MLE") for factor models. Gerbing and Hamilton (1994) traces the history of ICE back through Hunter and Cohen's (1969) Fortran program, through Harman (1954) and back to the very beginnings of factor analysis.

Under ICE, each proxy with its observed variables is estimated separately. ICE starts by forming its proxies as unweighted sums of the associated observed variables. ICE calculates the communality of each observed variable as the squared covariance between the observed variable and its proxy. Then the covariance matrix of the observed variables is adjusted by replacing the diagonal values — initially, the variances of the observed variables — with each variable's communality. The updating of communalities and replacement of diagonal entries continues until convergence. The replacement of empirical variance with communality turns the proxies, which began as simple composites, into common factors. Gerbing and Hamilton (1994) reports that ICE results are highly comparable to MLE results for the same data, when the factor model is correct, though ICE runs in a small fraction of the time required by MLE.

There is a great deal of similarity between ICE and "PLSc". Both approaches start out with a standardized unweighted sum of the indicators as a proxy. ICE applies an adjustment to the covariance matrix at each iteration, while "PLSc" applies a single adjustment at the end of estimation. In ICE, the adjustment explicitly involves the communalities of the observed variables. Papers on "PLSc" do not make explicit reference to communalities, but the adjustment appears to achieve the same effect. The essential point is that, as with ICE, the proxies in "PLSc" start out as composites but end up as common factors.

To summarize, "PLSc" appears to be a factor-based approach that uses PLS path modeling parameter estimates as starting values. It does not contradict the argument that PLS path modeling is a composite-based method.