ARTICLE IN PRESS

Journal of Operations Management xxx (2016) 1–19

Contents lists available at ScienceDirect

Journal of Operations Management

journal homepage: www.elsevier.com/locate/jom

Partial least squares path modeling: Time for some serious second thoughts

Mikko Rönkkö [a,*], Cameron N. McIntosh [b], John Antonakis [c], Jeffrey R. Edwards [d]

[a] Aalto University School of Science, PO Box 15500, FI-00076, Aalto, Finland
[b] Public Safety Canada, 340 Laurier Avenue West, Ottawa, Ontario, K1A 0P8, Canada
[c] Faculty of Business and Economics, University of Lausanne, Internef #618, CH-1015, Lausanne-Dorigny, Switzerland
[d] Kenan-Flagler Business School, University of North Carolina at Chapel Hill, Campus Box 3490, McColl Building, Chapel Hill, NC, 27599-3490, USA

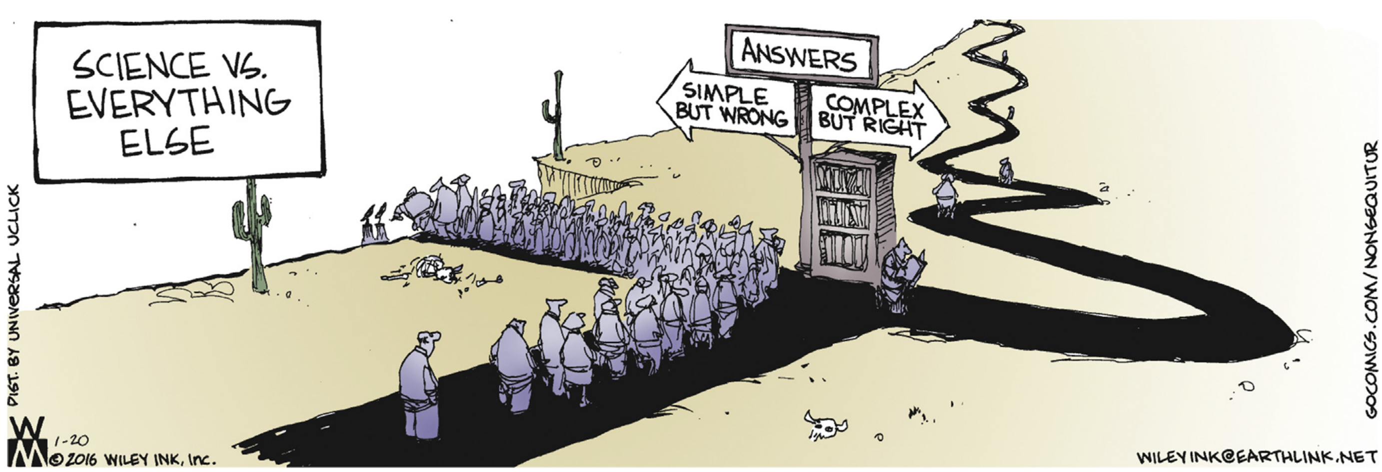
ARTICLE INFO

Article history:
Received 26 March 2016
Received in revised form
17 May 2016
Accepted 25 May 2016
Available online xxx

Keywords:
Partial least squares
Structural equation modeling
Statistical and methodological myths and urban legends

ABSTRACT

Partial least squares (PLS) path modeling is increasingly being promoted as a technique of choice for various analysis scenarios, despite the serious shortcomings of the method. The current lack of methodological justification for PLS prompted the editors of this journal to declare that research using this technique is likely to be deck-rejected (Guide and Ketokivi, 2015). To provide clarification on the inappropriateness of PLS for applied research, we provide a non-technical review and empirical demonstration of its inherent, intractable problems. We show that although the PLS technique is promoted as a structural equation modeling (SEM) technique, it is simply regression with scale scores and thus has very limited capabilities to handle the wide array of problems for which applied researchers use SEM. To that end, we explain why the use of PLS weights and many rules of thumb that are commonly employed with PLS are unjustifiable, followed by addressing why the touted advantages of the method are simply untenable.



1. Introduction

Partial least squares (PLS) has become one of the techniques of choice for theory testing in some academic disciplines, particularly marketing and information systems, and its uptake seems to be on the rise in operations management (OM) as well (Peng and Lai, 2012; Rönkkö, 2014b). The PLS technique is typically presented as an alternative to structural equation modeling (SEM) estimators

* Corresponding author.
E-mail address: mikko.ronkko@aalto.fi (M. Rönkkö).

http://dx.doi.org/10.1016/j.jom.2016.05.002
0272-6963/© 2016 Elsevier B.V. All rights reserved.

(e.g., maximum likelihood), over which it is presumed to offer several advantages (e.g., an enhanced ability to deal with small sample sizes and non-normal data).

Recent scrutiny suggests, however, that many of the purported advantages of PLS are not supported by statistical theory or empirical evidence, and that PLS actually has a number of disadvantages that are not widely understood (Goodhue et al., 2015; McIntosh et al., 2014; Rönkkö, 2014b; Rönkkö and Evermann, 2013; Rönkkö et al., 2015). As recently concluded by Henseler (2014): "[L]ike a hammer is a suboptimal tool to fix screws, PLS is a suboptimal tool to estimate common factor models", which are the kind of models OM researchers use (Guide and Ketokivi, 2015, p. vii). Unfortunately, whereas a person attempting to hammer a screw will quickly realize that the tool is ill-suited for that purpose, the shortcomings of PLS are much more insidious because they are not immediately apparent in the results of the statistical analysis. Although PLS promises simple solutions to complex problems and often produces plausible statistics that are seemingly supportive of research hypotheses, both the technical and applied literature on the technique seem to confound two distinct notions: (1) something can be done; and (2) doing so is methodologically valid (Westland, 2015, Chapter 3). As stated in a recent editorial by Guide and Ketokivi (2015): "Claiming that PLS fixes problems or overcomes shortcomings associated with other estimators is an indirect admission that one does not understand PLS" (p. vii). However, the editorial provides little material aimed at improving the understanding of PLS and its associated limitations.

Although there is no shortage of guidelines on how the PLS technique should be used, many of these are based on conventions, unproven assertions, and hearsay, rather than rigorous methodological support. Although OM researchers have followed these guidelines (Peng and Lai, 2012), such works do not help readers gain a solid and balanced understanding of the technique and its shortcomings. This state of affairs makes it difficult to justify the use of PLS, beyond arguing that someone has said that using the method would be a good idea in a particular research setting (Guide and Ketokivi, 2015, p. vii) Therefore, in order to mitigate common misunderstandings, we clarify issues concerning the usefulness of PLS in a non-technical manner for applied researchers. In light of these issues, it becomes apparent that the findings of studies employing PLS are ambiguous at best and at worst simply wrong, leading to the conclusion that PLS should be discontinued until the methodological problems explained in this article have been fully addressed.

2. What is PLS and what does it do?

A PLS analysis consists of two stages. First, observed variables are combined as weighted sums (composites); second, the composites are used in separate regression analyses, applying null hypothesis significance testing by comparing the ratio of a regression coefficient and its bootstrapped standard error against Student's *t* distribution. In a typical application, the observed variables are intended to measure theoretical constructs. In this type of analysis, the purpose of combining multiple indicators into composites is to produce aggregate measures that can be expected to be more reliable than any of their components, and can therefore be used as reasonable proxies for the constructs. Thus the only difference between PLS and more traditional regression analyses using summed scales, factor scores, or principal components, is how the indicators are weighted to create the composites. Moreover, instead of applying traditional factor analysis techniques, the quality of the measurement model is evaluated by inspecting the correlations between indicators and composites that they form, summarized as the composite reliability (CR) and average variance extracted (AVE) indices.

Although PLS is often marketed as a SEM method, a better way to understand what the technique actually does is to simply consider it as one of many indicator weighting systems. The broader methodological literature provides several different ways to construct composite variables. The simplest possible strategy is taking the unweighted sum of the scale items, with a refined version of this approach being the application of unit weights to standardized indicators (Cohen, 1990). The two most common empirical weighting systems are principal components, which retain maximal information from the original data, and factor scores that assume an underlying factor model (Widaman, 2007), with different calculation techniques producing scores with different qualities (Grice, 2001b). Commonly used prediction weights include regression, correlation, and equal weights (Dana and Dawes, 2004). Although not linear composites, different models based on item response theory produce scale scores that take into account both respondent ability and item difficulty (Reise and Revicki, 2014). Outside the context of research, many useful indices are composites, such as stock market indices that can weight individual stocks based on their price or market capitalization. Given the large number of available approaches for constructing composites variables, two key questions are: (1) Does PLS offer advantages over more well-established procedures?; and (2) What is the purpose of the PLS weights used to form the composites? We address these questions next.

2.1. On the "optimality" of PLS weights

Most introductory texts on PLS gloss over the purposes of the weights, arguing that PLS is SEM and therefore it must provide an advantage over regression with composites (e.g., Gefen et al., 2011); however, such works often do not explicitly point out that PLS itself is also simply regression with composites. Other authors suggest the weights are optimal (e.g., Henseler and Sarstedt, 2013, p. 566), but do not explain why and for which specific purpose. As noted by Krämer (2006): "In the literature on PLS, there is often a huge gap between the abstract model [...] and what is actually computed by the PLS path algorithms. Normally, the PLS algorithms are presented directly in connection with the PLS framework, insinuating that the algorithms produce optimal solutions of an obvious estimation problem attached to PLS. This estimation problem is however never defined" (p. 22).

The purpose of PLS weights remains ambiguous (Rönkkö et al., 2015, p. 77), as various rather *different* explanations abound (see Table 1 for some examples). However, none of these works (or their cited literature) provide mathematical proofs or simulation evidence to support their arguments. Perhaps the most common argument is that the indicator weights maximize the R^2 values of the regressions between the composites in the model (e.g. Hair et al., 2014, p. 16). However, this claim is problematic, for two main reasons: (a) why maximizing the R^2 values is a good optimization criterion is unclear; and (b) PLS has not been shown to be an optimal algorithm for maximizing R^2. In contrast, Rönkkö (2016a, sec. 2.3) demonstrates a scenario where optimizing indicator weights directly with respect to R^2 produces a 118% larger R^2 value than PLS, thus demonstrating that if the purpose of the analysis is to maximize R^2, the PLS algorithm is not an effective algorithm for this task.

Another common claim is that PLS weights reduce the impact of measurement error (e.g., Chin et al., 2003, p. 194; Gefen et al., 2011,