Factor-based structural equation modeling with WarpPLS

Ned Kock

Full reference:
Kock, N. (2019). Factor-based structural equation modeling with WarpPLS. *Australasian Marketing Journal*, 27(1), 57-63.

Abstract

Structural equation modeling (SEM) is extensively used in marketing research. For various years now, there has been a somewhat heated debate among proponents and detractors of the use of the partial least squares (PLS) method for SEM. The classic PLS design, originally proposed by Herman Wold, has a number of advantages over covariance-based SEM; e.g., minimal model identification demands. However, that design does not base its model parameter recovery approach on the estimation of factors, but on composites, which are exact linear combinations of indicators. This leads to adverse consequences, primarily in the form of unacceptable levels of type I and II errors. Recently a new factor-based method for SEM has been developed, called PLSF, which we discuss in this paper. This method has the advantages of classic PLS, but without the problems inherent in the use of composites. For readers interested in trying it, the PLSF method is implemented in the SEM software WarpPLS.

KEYWORDS: Partial Least Squares; Structural Equation Modeling; Measurement Error; Type I Error; Type II Error; Variation Sharing

1. Introduction

Structural equation modeling (SEM) is extensively used in marketing research (Hayes et al., 2017). This sophisticated quantitative research method also finds broad application is many other areas of business research, as well as in research within the social and behavioral sciences. SEM employs latent variables (LVs), which are measured indirectly through “observed” or “manifest” variables. Observed or manifest variables associated with a specific LV are normally called “indicators”.

The measurement of LVs via indicators, typically obtained from the administration of questionnaires, includes error. In this context LVs refer to perception-based constructs, such as one’s satisfaction with one’s job. Indicators frequently store numeric answers on Likert-type scales (e.g., 1 = strongly disagree … 5 = strongly agree) to question-statements in questionnaires. Each set of question-statements is designed to refer to a specific LV, and expected to measure the LV with a certain degree of imprecision.

For various years now there has been a rather heated debate among proponents and detractors of the use of the partial least squares (PLS) method, developed by the Norwegian-born econometrician and statistician Herman Ole Andreas Wold, in the context of SEM. This debate stems from common factor model assumptions, which form the basis on which covariance-based SEM rests (Kline, 2010; Mueller, 1996). Fueling the debate is the fact that Wold’s classic PLS design for “soft” SEM has a number of advantages over covariance-based SEM, including: minimal model identification demands, practically no data or model parameter distribution assumptions, virtually universal convergence to solutions, and estimation of “pseudo-factors”. The latter can be used in a number of derivative tests.

In spite of the advantages above, there is a major problem with using PLS, as originally designed, for hypothesis-testing in the context of SEM. The original PLS design does not base its model parameter estimation process on the estimation of true factors. Estimation is based on “composites”, which are exact linear combinations of indicators, and are called above “pseudo-factors”. The composite estimates generated by the classic PLS design can be conceptually seen as factors minus their corresponding measurement residuals. Reliance on them leads to biased model parameter estimates, notably path coefficients and loadings, even as sample sizes grow to infinity.

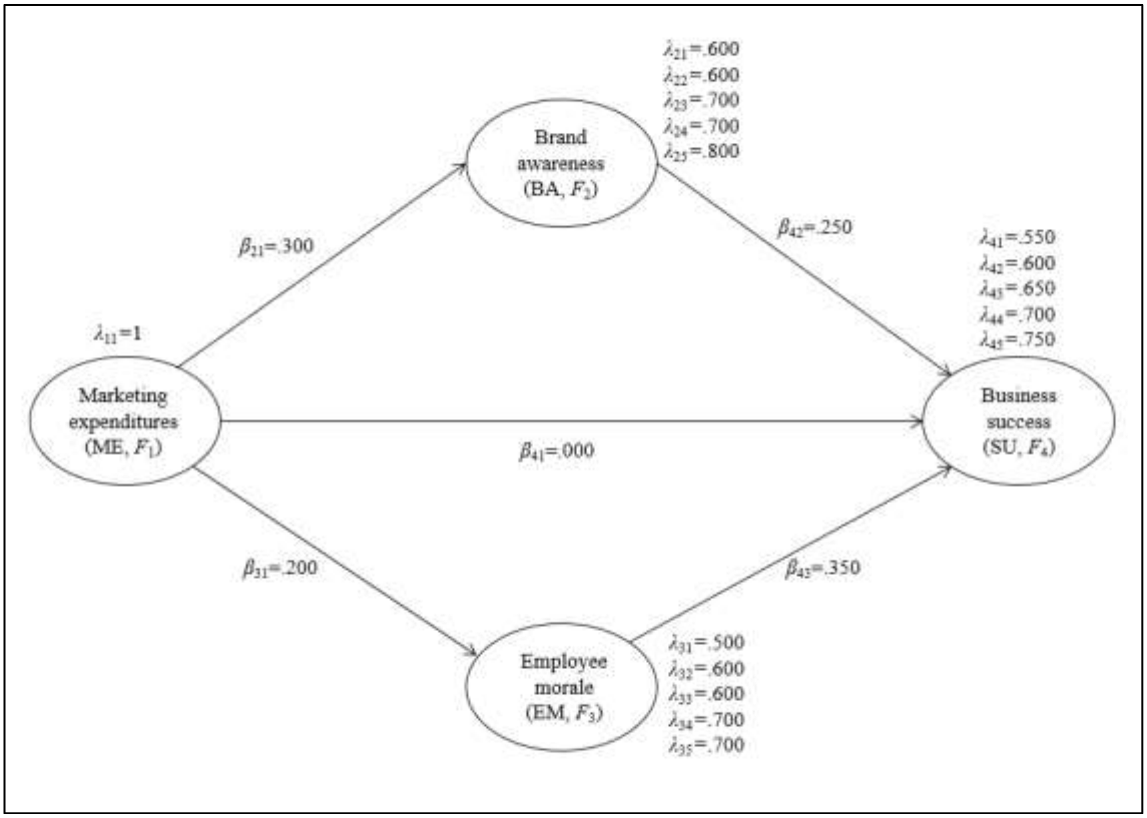
Recently a new factor-based method for SEM has been developed (Kock, forthcoming), which is discussed in this paper. Our discussion builds on an illustrative model, which is meant to make our exposition meaningful to those in the field of marketing and related fields. We created a large sample (*N*=10,000) of normally distributed data based on this illustrative model to exemplify the performance of the new method vis-à-vis other methods. These comparison methods are covariance-based SEM through full-information maximum likelihood, ordinary least squares regression with summed indicators, and PLS Mode A employing the path weighting scheme. The latter is the most widely used form of classic PLS.

Since this new method bears some resemblance to classic PLS, we refer to it as the “PLSF” method. The “F” in the PLSF acronym refers to the correlation-preserving factor estimation process that underlies the method. For readers interested in trying it, the PLSF method is implemented in WarpPLS (Kock, 2018). WarpPLS is an SEM software package that is unique not only because of its implementation of factor-based SEM, but also because it enables nonlinear analyses where best-fitting nonlinear functions are estimated for each pair of structurally linked variables in path models, and subsequently used (i.e., the nonlinear functions) to estimate path coefficients that take into account the nonlinearity.

2. Illustrative model

Our discussion is based on the illustrative model shown in Figure 1, which is used here to help us accomplish our goal of making our discussion meaningful to those in the field of marketing and related fields. The model contains four factors, associated with the following constructs: marketing expenditures (ME, F_1), brand awareness (BA, F_2), employee morale (EM, F_3), and business success (SU, F_4). These constructs are assumed to have been measured at the company level. The marketing expenditures construct is assumed to have been measured without error (e.g., in dollars per year), through a single indicator. The other three constructs are assumed to have been measured with error through Likert-type question-statements in a questionnaire, each through five indicators.

Figure 1. Illustrative model



In this illustrative model β_{ij} is the path coefficient for the link going from factor F_j to factor F_i, and λ_{ij} is the loading for the jth indicator of factor F_i. The general idea underlying this model is that a company's marketing expenditures positively influence business success but only indirectly, via positive influences on brand awareness and employee morale. For example, a solar panel manufacturing company, by spending money marketing its product, will make its solar panel brand better known to customers. It will also make its employees proud, by portraying the company positively in the media, and thus increase their morale. The combined result will be greater business success.

This is clearly a simplified model that is not meant to serve as the basis for future theoretical or empirical research. It is nevertheless a helpful model, as it helps us conduct a discussion that is not entirely conceptual or mathematical. From a methodological perspective, this model is also helpful because of the inclusion of mediating effects, and of the "zero" path from ME to SU. As it will be seen later in this paper, these model characteristics allow us to illustrate the tendency that composite-based SEM methods, such as classic PLS-based SEM, have to induce unacceptable levels of type I and II errors – a.k.a. false positives and false negatives, respectively.

3. The PLSF method

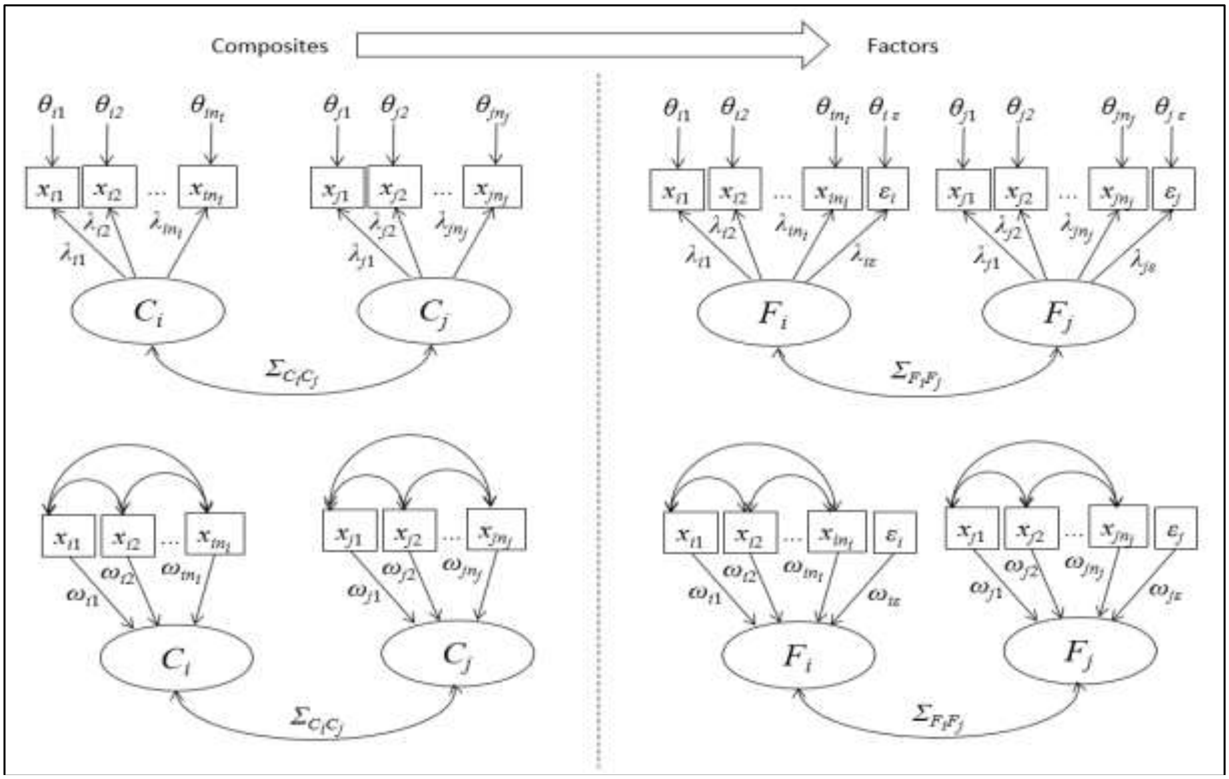
The PLSF method is comprised of four main stages. In its **first stage**, it estimates the reliabilities for each of the LVs in a model. Those reliabilities are then used to estimate composites in the **second stage**. These composites are different from the composites generated via classic PLS algorithms, such as PLS Mode A and PLS Mode B (Kock & Hadaya, 2018). A key difference is that they must satisfy a specific equation that employs the Moore–Penrose pseudoinverse transformation. In the **third stage**, factors are estimated based on the composites via a novel technique known as variation sharing. In the **fourth stage**, the final stage, all parameters are estimated based on the estimated factors and the original indicators. This final stage implements a two-stage least squares estimation to control for in-model endogeneity. Through these four stages, the PLSF method essentially produces composites and then "transforms" these composites into factors. This fundamental characteristic of the PLSF method is illustrated in Figure 2.

The figure shows two correlated composites on the left side, and their corresponding correlated factors on the right side. The factors are derived from the composites. The indicators are assumed to be imprecise measures, obtained via questionnaires, from their corresponding LVs. As such the indicators "reflect" their common factors. Even though this is the case (top part of the figure), the factors can also be seen as aggregations of their respective indicators and measurement residuals (bottom part of the figure). In each factor the measurement residual explains the variation in the factor that is not accounted for by the indicators. Therefore the measurement residual is uncorrelated with the factor's indicators. The percentage of the variance explained in a factor by its indicators, or the R^2 obtained when a factor is regressed on its indicators, equals the factor's reliability.

As we can see, reliability estimates need to be produced early in the PLSF method, and in fact are critical for the method. The Cronbach's alpha coefficients provide good estimates of the reliabilities when the degree of heterogeneity among the loadings in each LV is low. This appears to frequently be the case in practice. Two alternatives can be employed when this is not the case. One of these alternatives is to use the reliability estimates generated by Dijkstra's consistent PLS (a.k.a. PLSc) technique, which appear to be closer to the true reliabilities than Cronbach's alpha coefficients under high loading heterogeneity conditions. The other alternative is to use Cronbach's alpha coefficients as initial reliability estimates, and iterate across the second and third stages of the PLSF method. In each iteration the reliability estimates for each

LV would be adjusted to $1/2\,(\widehat{\omega}_i{}'\hat{\lambda}_i + \hat{\rho}_i)$, where for each LV indexed by i: $\widehat{\omega}_i$ is a column vector with weight estimates, $\hat{\lambda}_i$ is a column vector with loading estimates, and $\hat{\rho}_i$ is the composite reliability estimate.

Figure 2. From composites to factors


Composites
Factors
θ_{i1} θ_{i2} θ_{in_i} θ_{j1} θ_{j2} θ_{jn_j}
x_{i1} x_{i2} ... x_{in_i} x_{j1} x_{j2} ... x_{jn_j}
λ_{i1} λ_{i2} λ_{in_i} λ_{j1} λ_{j2} λ_{jn_j}
C_i C_j
$\Sigma_{C_iC_j}$
θ_{i1} θ_{i2} θ_{in_i} $\theta_{i\varepsilon}$ θ_{j1} θ_{j2} θ_{jn_j} $\theta_{j\varepsilon}$
x_{i1} x_{i2} ... x_{in_i} ε_i x_{j1} x_{j2} ... x_{jn_j} ε_j
λ_{i1} λ_{i2} λ_{in_i} $\lambda_{i\varepsilon}$ λ_{j1} λ_{j2} λ_{jn_j} $\lambda_{j\varepsilon}$
F_i F_j
$\Sigma_{F_iF_j}$
x_{i1} x_{i2} ... x_{in_i} x_{j1} x_{j2} ... x_{jn_j}
ω_{i1} ω_{i2} ω_{in_i} ω_{j1} ω_{j2} ω_{jn_j}
C_i C_j
$\Sigma_{C_iC_j}$
x_{i1} x_{i2} ... x_{in_i} ε_i x_{j1} x_{j2} ... x_{jn_j} ε_j
ω_{i1} ω_{i2} ω_{in_i} $\omega_{i\varepsilon}$ ω_{j1} ω_{j2} ω_{jn_j} $\omega_{j\varepsilon}$
F_i F_j
$\Sigma_{F_iF_j}$

To produce the composite estimates, weights and loadings must satisfy the equation below, where: $\Sigma_{x_i x_i}$ is the matrix of correlations among the indicators associated with the LV indexed by i, $diag(\Sigma_{x_i\hat{\theta}_i})$ is the diagonal matrix of *covariances* among the indicators and corresponding error terms, and the superscript $+$ denotes the Moore–Penrose pseudoinverse transformation. Note that $diag(\Sigma_{x_i\hat{\theta}_i})$ is a diagonal matrix because in the common factor model $\Sigma_{x_{ij}\theta_{ij}} = 0$ for all $i \neq j$. That is, the indicator error terms are correlated with their corresponding indicators, as they account for the variation in those indicators that is not caused by the corresponding factor, and at the same time are uncorrelated with other indicators in the same factor.

$$\widehat{\omega}_i = \Sigma_{x_i x_i}{}^{-1}\left[\Sigma_{x_i x_i} - diag\left(\Sigma_{x_i\widehat{\theta}_i}\right)\right]\hat{\lambda}_i{}'^{+}.$$

Iterations of the equation above together with a few other ancillary equations produce estimates of weights and loadings. With the estimates of weights, the PLSF method generates estimates of composites. Those are then used by the PLSF method to produce factor estimates, employing the variation sharing equations below as part of an iterative process. Iterations take place until the sum of the absolute differences $\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}$ falls below a small fraction, or until the sum of the absolute differences between successive estimates of $\Sigma_{\hat{F}_i\hat{F}_j}$ changes by less than a small fraction. The estimated matrix of correlations among factors $\hat{\Sigma}_{F_iF_j}$ is calculated as $\Sigma_{\hat{C}_i\hat{C}_j}/\sqrt{\hat{\rho}_i\hat{\rho}_j}$, where $\Sigma_{\hat{C}_i\hat{C}_j}$ is the corresponding element of the matrix of correlations among estimated composites. The matrix of correlations among estimated factors $\Sigma_{\hat{F}\hat{F}}$ varies across iterations.

$$\hat{\varepsilon}_i = Stdz\left(\hat{\varepsilon}_i + \left(\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}\right)\frac{\hat{\Sigma}_{F_iF_j}}{\widehat{\omega}_{i\varepsilon}}\left(\hat{C}_j\widehat{\omega}_{jC} + \hat{\varepsilon}_j\widehat{\omega}_{j\varepsilon}\right)\right),$$

$$\hat{F}_i = Stdz\left(\hat{F}_i + \left(\widehat{\omega}_{iC} - \Sigma_{\hat{F}_i\hat{C}_i}\right)\hat{C}_i\widehat{\omega}_{iC}\right),$$

$$\hat{\varepsilon}_i = Stdz\left(\hat{\varepsilon}_i - \Sigma_{\hat{C}_i\hat{\varepsilon}_i}\hat{C}_i\widehat{\omega}_{iC} + \left(\widehat{\omega}_{i\varepsilon} - \Sigma_{\hat{F}_i\hat{\varepsilon}_i}\right)\hat{F}_i\widehat{\omega}_{i\varepsilon}\right).$$

Through the variation sharing equations above, successive estimates of factors $\hat{F}_i$ and measurement residuals $\hat{\varepsilon}_i$ acquire or lose variation from correlated factors, composites, and measurement residuals (denoted as $\hat{F}_j$, $\hat{C}_j$ and $\hat{\varepsilon}_j$). This is done in such a way as to enforce the following constraints: $\hat{\Sigma}_{F_iF_j} = \Sigma_{\hat{F}_i\hat{F}_j}$, $\Sigma_{\hat{F}_i\hat{C}_i} = \widehat{\omega}_{iC}$, $\Sigma_{\hat{F}_i\hat{\varepsilon}_i} = \widehat{\omega}_{i\varepsilon}$, and $\Sigma_{\hat{C}_i\hat{\varepsilon}_i} = 0$. The first constraint, namely $\hat{\Sigma}_{F_iF_j} = \Sigma_{\hat{F}_i\hat{F}_j}$, drives the iterative convergence process. As we can see, the PLSF method is analogous to covariance-based SEM in that it maximizes the fit between covariance matrices. The key difference is that the PLSF method maximizes the fit between factor covariance matrices, whereas covariance-based SEM maximizes the fit between indicator covariance matrices.

4. Coefficient estimation accuracy

We created a large sample (*N*=10,000) of normally distributed data, based on the illustrative model described earlier, to exemplify the performance of the PLSF method vis-à-vis other methods. Such a large sample contains only a small amount of sampling error, which makes it a good choice for cross-method comparisons highlighting algorithmic outcomes. It also provides the basis for a straightforward discussion about the performance of different methods.

Here PLSF is compared against the following methods: covariance-based SEM through full-information maximum likelihood (FIML); ordinary least squares regression with summed indicators (OLS); and PLS Mode A employing the path weighting scheme (PLS). The latter is the most widely used form of classic PLS path modeling employed in the field of marketing.

We used WarpPLS 6.0 (Kock, 2018) for the implementation of the PLSF, OLS, and PLS methods. The following algorithms were selected in the WarpPLS software: Factor-Based PLS Type CFM2 (for the PLSF method), Robust Path Analysis (for OLS), and PLS Mode A (for PLS). For FIML, we used R 3.5.1 and the package lavaan 0.6-3 (Rosseel, 2012).

Table 1 lists the path coefficients and a summarized set of loadings; the latter only for SU, to avoid crowding, as the same loading estimation patterns repeat themselves across LVs. The columns labeled “True” list the true values for path coefficients and loadings. The “Est.” columns list the corresponding estimates employing each method. The “Diff.” columns list the differences between estimates and true values for each method. The rows labeled “RMSE” list root-mean-square errors associated with the differences among estimates and true values, calculated as the square roots of the averages of the squared differences.

Table 1. Path coefficients and loadings for large sample (*N*=10,000)

Path coefficients													
			PLSF			FIML			OLS			PLS	
	True		Est.	Diff.		Est.	Diff.		Est.	Diff.		Est.	Diff.
ME>BA	0.3000		0.2997	-0.0003		0.2987	-0.0013		0.2715	-0.0285		0.2725	-0.0275
ME>EM	0.2000		0.1868	-0.0132		0.1873	-0.0127		0.1634	-0.0366		0.1643	-0.0357
ME>SU	0.0000		-0.0060	-0.0060		-0.0071	-0.0071		0.0223	0.0223		0.0215	0.0215
BA>SU	0.2500		0.2468	-0.0032		0.2471	-0.0029		0.1948	-0.0552		0.1969	-0.0531
EM>SU	0.3500		0.3449	-0.0051		0.3447	-0.0053		0.2599	-0.0901		0.2653	-0.0847
RMSE				0.0070			0.0071			0.0526			0.0500
Loadings (for SU)													
			PLSF			FIML			OLS			PLS	
	True		Est.	Diff.		Est.	Diff.		Est.	Diff.		Est.	Diff.
SU1<SU	0.5500		0.5469	-0.0031		0.5504	0.0004		0.6800	0.1300		0.6489	0.0989
SU2<SU	0.6000		0.5989	-0.0011		0.5983	-0.0017		0.7055	0.1055		0.7001	0.1001
SU3<SU	0.6500		0.6551	0.0051		0.6454	-0.0046		0.7340	0.0840		0.7264	0.0764
SU4<SU	0.7000		0.7042	0.0042		0.7053	0.0053		0.7623	0.0623		0.7809	0.0809
SU5<SU	0.7500		0.7551	0.0051		0.7545	0.0045		0.7870	0.0370		0.8074	0.0574
RMSE				0.0040			0.0038			0.0898			0.0842

Figure 3. Differences (RMSEs) with respect to true values

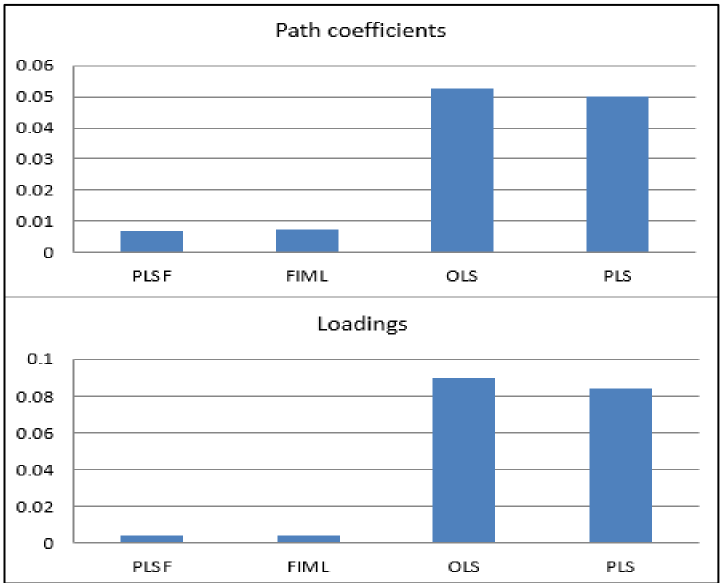


Figure 3 highlights the differences with respect to the true values for each of the methods. As noted earlier, these differences are presented as root-mean-square errors (RMSEs). Note that the large sample created has true values associated with coefficients that are somewhat different from the true population values shown earlier for the illustrative model (Kock & Moqbel, 2016). It is the true sample values that the methods try to recover, which is why PLSF and FIML seem to miss the marks by approximately the same margins – e.g., by estimating a 0.2000 path respectively as 0.1868 and 0.1873. In other words, the bar charts are somewhat misleading, possibly portraying the PLSF and FIML methods as marginally less precise than they truly are.

As can be inferred from the results, the performances of PLSF and FIML were very similar in terms of estimation of path coefficients and loadings. Also, PLSF and FIML performed significantly better than OLS and PLS, whose corresponding RMSEs with respect to the true values were multiple orders of magnitude higher.

In terms of the path coefficients, the OLS and PLS methods significantly *underestimated* the true values. For example, the path coefficient for the link BA>SU (true value = 0.2500) was estimated as 0.1948 by OLS and as 0.1969 by PLS. The OLS and PLS methods also significantly *overestimated* the "zero" path coefficient for the link ME>SU (true value = 0.0000). The OLS method overestimated it as 0.0223, and PLS as 0.0215.

5. False negatives and false positives

So what is the big deal about underestimating a path coefficient as 0.1969, as did the PLS method, when in reality (i.e., at the population level) the coefficient is 0.2500? The main problem with this type of underestimation is that it can have a significant impact on the probability of type II errors, or false negatives. The probability that a method will *avoid* false negatives, in the context of a specific model, is known as the statistical power of the method. A power value of 80 percent (or 0.8) or higher is generally considered acceptable in business research.

One of the "Explore" menu options available from the main window of WarpPLS, namely the "Explore statistical power and minimum sample size requirements" menu option, allows users to estimate statistical power and minimum sample size requirements. This feature helps us to form an idea about the loss in power that is caused by the path coefficient underestimations of OLS and PLS. Let us use the path coefficient for the link BA>SU (true value = 0.2500) as an example. It was estimated as 0.2468 by PLSF, which is very close to the estimate generated by FIML. Entering this value in WarpPLS we find that 80 percent power is achieved with a sample size of approximately 88 to 102.

However, if we enter the 0.1969 path coefficient estimate produced by PLS, we find that 80 percent power is achieved with a sample size of approximately 146 to 160. Incrementally changing the power settings in the "Explore statistical power and minimum sample size requirements" feature of WarpPLS, allows us to establish that a sample size in the range of 93 to 104 for this 0.1969 path coefficient would be associated with a power level of 64 percent, which is unacceptably low. And this range is slightly higher than the 88 to 102 range considered above. In other words, using the classic composite-based PLS method would significantly decrease the power of our analysis, when compared with the factor-based PLSF method. This essentially means that classic PLS leads to unacceptably high probabilities that type II errors will occur. The same is true also for OLS, with a comparable decrease in power and increase in type II error probability.

Another problem associated with the composite-based methods such as PLS occurs in connection with type I errors, or false positives. This problem is possibly even more severe than the problem regarding type II errors discussed above, because the acceptable probability of type I errors is normally much lower than that for type II errors. A value of 5 percent (or 0.05) or lower is generally considered acceptable in business research, for type I errors. This is $1/4$ of the 20 percent allowable for type II errors; which is equivalent to 80 percent power ($100 - 20 = 80$).

For a “zero” path coefficient, the percentage of false positives equates the statistical power associated with that path, because the latter is the probability that the path will be found to be significantly greater than zero (which in reality it is not). Given this, we can use the same line of mathematical reasoning employed by Kock & Hadaya (2018) to obtain the equation below (see also: Kock et al., 2017), where: N is the sample size at which a certain percentage of false positives is achieved (FP); z_{CL} is the z-score associated with a given confidence level CL; z_{FP} is the z-score associated with FP; and $|\widehat{\beta_0}|$ is the absolute value of the estimate obtained for the “zero” path coefficient.

$$N = \left(\frac{z_{CL} + z_{FP}}{|\widehat{\beta_0}|}\right)^2.$$

As noted earlier, the PLS method overestimated the “zero” path coefficient for the link ME>SU (true value = 0.0000), as 0.0215. We can use the Excel function $\mathrm{NORMSINV}(\cdot)$ to obtain z-scores. In business research, normally the confidence level CL used is 95 percent. Thus, the value of N shown below is the approximate sample size at which the percentage of false positives will be twice the acceptable level of 5 percent (i.e., 10 percent) for a $|\widehat{\beta_0}|$ value of 0.0215.

$$N = \left(\frac{z_{0.95} + z_{0.10}}{0.0215}\right)^2 \rightarrow$$

$$N = \left(\frac{1.6449 - 1.2816}{0.0215}\right)^2 \rightarrow$$

$$N \cong 286.$$

That is, the PLS method yields 10 percent false positives in our model with a sample size of 286. Stated differently, the probability that the PLS method would estimate the “zero” path in our model as being statistically significant was found to be 10 percent at that sample size. This is an unacceptably high percentage of false positives, double the 5 percent that is generally considered the maximum allowable. The OLS method estimated the “zero” path coefficient as 0.0223, greater than the 0.0215 for PLS, so clearly it would yield 10 percent false positives at a lower sample size.

Note from the equation that the values of z_{FP} and N must increase together for the equation to be satisfied, other things being equal. And, for the value of z_{FP} to increase, the percentage of false positives FP has to increase. Therefore, it follows that the higher the sample size, the greater will be the percentage of false positives with the composite-based OLS and PLS methods. This is exactly the opposite of what one would expect from trustworthy methods.

These results mirror those obtained by Kock (forthcoming), who provides a more detailed discussion of the PLSF method and results of comparisons against the same methods used here. That study found that false positives associated with a "zero" path coefficient were already at 27.2 percent for OLS and 23.6 for PLS at a relatively small sample size of 300. At that sample size, both the PLSF and FIML methods yielded false positive percentages below 5. The study used a more complex model than the one employed here, which seemed to have amplified the negative effects of the poor coefficient estimation accuracy of OLS and PLS. The fundamental reason for these results is that, unlike factor-based methods, composite-based methods tend to "capitalize on error" when used to estimate coefficients associated with non-existent causal paths.

6. Discussion and conclusion

We have highlighted problematic characteristics of composite-based methods, such as OLS and PLS, when they are compared with factor-based methods such as PLSF and FIML. The PLSF method is implemented starting in version 5.0 of the software WarpPLS (Kock, 2018). While our discussion has focused on path coefficients, which are structural model coefficients, we should note that our results show that measurement model coefficients are also biased in composite-based methods. Notably, loadings tend to be overestimated. For example, the loading for the link SU1<SU, whose true value is 0.5500, was overestimated as 0.6800 by OLS and as 0.6489 by PLS.

The distortion in measurement model coefficients undermines the use of tried-and-true criteria for measurement model quality assessment in FIML. For example, a widely used criterion for convergent validity assessment is that loadings must be equal to or greater than 0.5 for all indicator-LV links in a model. This criterion is meant to help researchers assess whether the indicators associated with the loadings do in fact "belong" to the LVs to which they were assigned in the measurement model design stage. This type of assessment is an essential element of what is known as a "confirmatory factor analysis". But if loadings are inflated, it is much more likely that indicators that do not actually belong to a LV will be mistaken as belonging. In other words, it is difficult to conduct a confirmatory *factor* analysis with entities that are not factors, but composites.

The recognition of this and related problems has motivated methodological researchers to develop various ad-hoc tests for the classic composite-based PLS method. This in part explains the proliferation of such ad-hoc tests that have recently been proposed; e.g., the heterotrait-monotrait test. New tests may be seen as novel and useful developments, but they are not necessarily so if the new tests are proposed to address issues that are caused by more fundamental problems, particularly if the fundamental problems are not actually addressed. It is in this context that the PLSF method offers value. It addresses the fundamental problem plaguing composite-based methods in general, including classic PLS, which is that these methods do not explicitly account for measurement error.

Our results demonstrate that composite-based methods, such as OLS and PLS, are doubly problematic when it comes to the estimation of structural model parameters. Those methods tend to *underestimate* path coefficients that exist at the population level. And, at the same time, they tend to *overestimate* path coefficients that do not exist – i.e., that are "zero" at the population level. Added to these problems is the one related to loadings, which tend to be *overestimated*, leading to measurement model quality assessment difficulties; e.g., in convergent validity assessment (Kock, 2014). Neither PLSF nor FIML, the factor-based methods employed in our analyses, suffers from the same problems.

Given that FIML is a fairly established method, why not simply use it instead of PLSF? Why do we need PLSF at all? There are several reasons why we need PLSF, of which three are particularly noteworthy. Firstly, the PLSF method is computationally much simpler than FIML. Most implementations of FIML require the calculation of matrices of second-order partial derivatives (Hessian matrices) and their inversion, which is often impossible or leads to unacceptable results (Kline, 2010; Mueller, 1996). The more complex the model, the more likely it is that unacceptable results will be produced.

Secondly, the PLSF method provides estimates of factor scores, which can subsequently be used in a variety of other tests. Factor scores can also be obtained via FIML, but they tend to be rather poor approximations of the true factors; see Kock (forthcoming) for several results and a discussion of this issue. Among tests that employ factor scores from the PLSF method are two that have been widely used in a variety of fields since their publication: full collinearity tests, which assess lateral and vertical collinearity among factors at the same time (Kock & Lynn, 2012), and can be used in common method bias assessments (Kock & Lynn, 2012; Kock, 2015); and factor-factor nonlinearity tests, whereby best-fitting nonlinear functions are built for each pair of causally linked factors, and then used in the calculation of modified path coefficients that take nonlinearity into consideration (Guo et al., 2011; Kock, 2010; Moqbel et al., 2013).

The “double-trouble” structural model parameter estimation situation associated with classic PLS and other composite-based methods has led to a new line of prediction-oriented research, based on the argument that classic PLS is as good for prediction as factor-based methods, if not better (Shmueli et al., 2016). Let us assume, for instance, that one wants to build a model of customer purchases at a supermarket, where purchases of products are modeled as influencing one another. The argument is that a model built based on parameters obtained via classic PLS would be quite successful at predicting purchases in the future (e.g., next month) based on past purchases (e.g., last month).

However, the above prediction-oriented type of application above is significantly different from hypothesis-testing in the context of SEM. The latter can greatly benefit from factor-based methods such as PLSF, because SEM-based tests of hypotheses rely heavily on parameter estimation accuracy. Given that the lack of estimation accuracy can lead to both type I and II errors, which is a very troubling scenario in the context of SEM, we tend to favor the PLSF method for testing hypotheses. In prediction-oriented scenarios, the advantages of PLSF are yet to be determined. This appears to be a promising area for future research.

References

Guo, K.H., Yuan, Y., Archer, N.P., & Connelly, C.E. (2011). Understanding nonmalicious security violations in the workplace: A composite behavior model. *Journal of Management Information Systems*, 28(2), 203-236.

Hayes, A. F., Montoya, A. K., & Rockwood, N. J. (2017). The analysis of mechanisms and their contingencies: PROCESS versus structural equation modeling. *Australasian Marketing Journal*, 25(1), 76-81.

Kline, R.B. (2010). *Principles and practice of structural equation modeling*. New York, NY: The Guilford Press.

Kock, N. (2010). Using WarpPLS in e-collaboration studies: An overview of five main analysis steps. *International Journal of e-Collaboration*, 6(4), 1-11.

Kock, N. (2014). Advanced mediating effects tests, multi-group analyses, and measurement model assessments in PLS-based SEM. *International Journal of e-Collaboration*, 10(3), 1-13.

Kock, N. (2015). Common method bias in PLS-SEM: A full collinearity assessment approach. *International Journal of e-Collaboration*, 11(4), 1-10.

Kock, N. (2018). *WarpPLS User Manual: Version 6.0*. Laredo, TX: ScriptWarp Systems.

Kock, N. (forthcoming). From composites to factors: Bridging the gap between PLS and covariance-based structural equation modeling. *Information Systems Journal*.

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal*, 28(1), 227–261.

Kock, N., & Lynn, G.S. (2012). Lateral collinearity and misleading results in variance-based SEM: An illustration and recommendations. *Journal of the Association for Information Systems*, 13(7), 546-580.

Kock, N., & Moqbel, M. (2016). Statistical power with respect to true sample and true population paths: A PLS-based SEM illustration. *International Journal of Data Analysis Techniques and Strategies*, 8(4), 316-331.

Kock, N., Avison, D., & Malaurent, J. (2017). Positivist information systems action research: Methodological issues. *Journal of Management Information Systems*, 34(3), 754-767.

Moqbel, M., Nevo, S., & Kock, N. (2013). Organizational members’ use of social networking sites and job performance: An exploratory study. *Information Technology & People*, 26(3), 240-264.

Mueller, R.O. (1996). *Basic principles of structural equation modeling*. New York, NY: Springer.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1-36.

Shmueli, G., Ray, S., Estrada, J. M. V., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models. *Journal of Business Research*, 69(10), 4552-4564.