A note on how to conduct a factor-based PLS-SEM analysis

Ned Kock

Full reference:

Kock, N. (2015). A note on how to conduct a factor-based PLS-SEM analysis. *International Journal of e-Collaboration*, 11(3), 1-9.

Abstract

The composite-factor estimation dichotomy has been the epicenter of a long and ongoing debate among proponents and detractors of the use of the partial least squares (PLS) approach for structural equation modeling (SEM). In this brief research note we discuss the implementation of a new method to conduct factor-based PLS-SEM analyses, which could be a solid step in the resolution of this debate. This method generates estimates of both true composites and factors, in two stages, fully accounting for measurement error. Our discussion is based on an illustrative model in the field of e-collaboration. A Monte Carlo experiment suggests that model parameters generated by the method are asymptotically unbiased. The method is implemented as part of the software WarpPLS, starting in version 5.0. This note provides enough details for the method's implementation in other venues such as R and GNU Octave.

KEYWORDS: Partial Least Squares; Structural Equation Modeling; Measurement Error; Path Bias; Variation Sharing; Monte Carlo Simulation

Introduction

The debate among proponents and detractors of the use of Wold's partial least squares (PLS) method (Adelman & Lohmoller, 1994; Lohmöller, 1989; Wold, 1980) in the context of structural equation modeling (SEM) has been going on for a long time. So far, it shows no signs of resolution. It arises from common factor model assumptions, which form the basis on which covariance-based SEM (CB-SEM) rests (Kline, 2010; Mueller, 1996). The debate is centered around two main issues.

The first issue is that Wold's original PLS design for "soft" SEM has a number of advantages over CB-SEM, such as minimal model identification demands, practically no data or model parameter distribution assumptions, virtually universal convergence to solutions, and estimation of "pseudo-factors". The latter, "pseudo-factors", provide a partial solution to the factor indeterminacy problem of CB-SEM.

The second issue fueling the debate is that the original PLS design does not base its model parameter estimation methods on the estimation of true factors. Estimation is based on "composites", which are exact linear combinations of indicators, and are referred to above as "pseudo-factors". The composite estimates generated by the original PLS design can be seen as factors minus their corresponding measurement errors. Reliance on them leads to biased model parameter estimates (notably path coefficients and loadings) even as sample sizes grow to infinity (Kock, 2014b).

In this brief research note we describe what could be a solid step in the resolution of this debate, although it may open new avenues for debate on different issues. We show how researchers can implement what we refer to as "Factor-Based PLS-SEM" (PLSF-SEM). This new method generates estimates of both true composites and factors, in two stages, fully accounting for measurement error.

The PLSF-SEM method is implemented in version 5.0 of WarpPLS, which is under intensive internal testing and nearing beta release at the time of this writing. WarpPLS is an SEM software tool that is unique in that 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. Moreover, WarpPLS provides a comprehensive set of model fit and quality indices that are compatible with both composite-based and factor-based SEM.

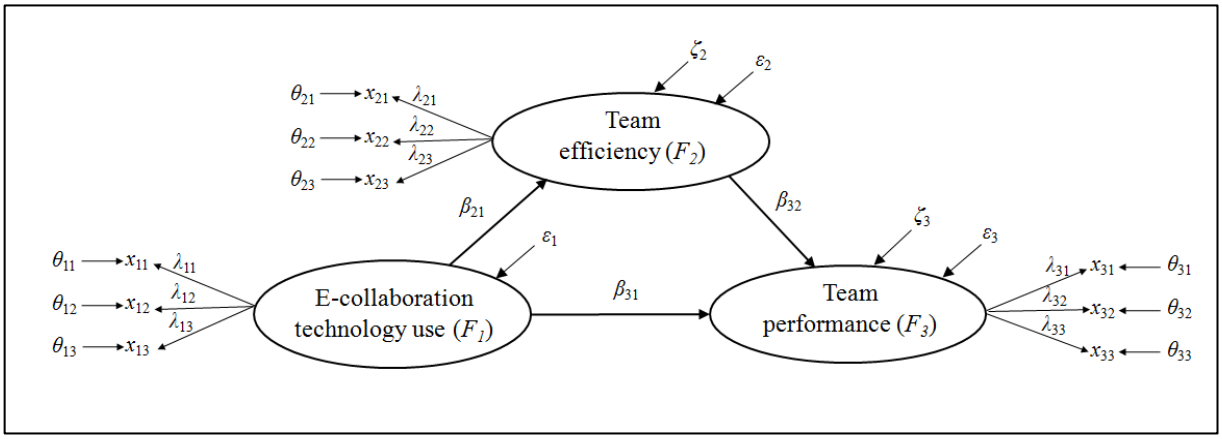
Illustrative model

Our discussion is based on the illustrative model depicted in Figure 1, which builds on an actual empirical study in the field of e-collaboration (Kock, 2005; 2008; Kock & Lynn, 2012). This illustrative model incorporates the belief that *e-collaboration technology use* (F_1) by teams of workers tasked with the development of new products in organizations (e.g., a new consulting service, a new car part) increases both *team efficiency* (F_2) and *team performance* (F_3). *Team efficiency* (F_2) is related to the speed and cost at which teams operate. *Team performance* (F_3) is related to how well the new products developed by teams perform in terms of sales and profits.

In this illustrative model β_{ij} is the path coefficient for the link going from factor F_j to factor F_i; λ_{ij} is the loading for the *j*th indicator of factor F_i; θ_{ij} is the indicator error for the *j*th indicator of factor F_i; ε_i is the measurement error associated with F_i; and and ζ_i is the structural error

associated with F_i, which exists only for endogenous factors. An endogenous factor has at least one other factor pointing at it in the model.

Figure 1. Illustrative model



Note that the measurement errors ε_i are not the same as the structural errors ζ_i. Measurement errors exist for any factors that are measured with a certain degree of imprecision, whether the factors are exogenous or endogenous. Structural errors exist only for endogenous factors. Analogously, the measurement errors ε_i should not be confused with the indicator errors θ_{ij}, even though these two types of errors are related. The former arise due to the existence of the latter, and can be seen as "extra" indicators that account for the explained variances in their respective factors that are not accounted for by the actual factor indicators.

PLSF-SEM's first stage: Composites

In our discussion in this and the following sections all variables are assumed to be standardized; i.e., scaled to have a mean of zero and standard deviation of one. This has no impact on the generality of the method or of the discussion. All standardized variables can be rescaled back to their original scales.

PLSF-SEM's first stage yields initial estimates of the composites. These estimates are used in the method's second stage, where factors and other model parameter estimates are produced. It starts by setting weights and loadings as 1 (reversed indicators must be properly adjusted), and initializing the composite estimates with a standardized vector of the summed indicators. Then measurement errors $\hat{\varepsilon}_i$, reliabilities $\hat{\alpha}_i$, measurement error weights $\hat{\omega}_{i\varepsilon}$, and composite weights $\hat{\omega}_{iC}$ are set as indicated in (1)-(4).

$$\hat{\varepsilon}_i := Stdz[Rnd(N)]. \tag{1}$$

$$\hat{\alpha}_i := \frac{n_i \bar{\Sigma}_{x_i x_i}}{[1+(n_i-1)\bar{\Sigma}_{x_i x_i}]}. \tag{2}$$

$$\widehat{\omega}_{i\varepsilon} \coloneqq \sqrt{1 - \hat{\alpha}_i}. \tag{3}$$

$$\widehat{\omega}_{iC} \coloneqq \sqrt{\hat{\alpha}_i}. \tag{4}$$

In these equations $Rnd(N)$ is a function that returns an independent and identically distributed (i.i.d.) variable with N rows, with N being the sample size; $Stdz(\cdot)$ is a function that returns a standardized column vector; n_i is the number of indicators of factor F_i; x_i is a matrix with N rows and with each column referring to one of the indicators associated with F_i; and $\bar{\Sigma}_{x_i x_i}$ is the mean of the non-redundant correlation coefficients among the column vectors that make up x_i (e.g., the mean of the lower triangular version of $\Sigma_{x_i x_i}$).

Technical readers will notice that the reliability estimate $\hat{\alpha}_i$ above is the Cronbach's alpha coefficient (Cronbach, 1951; Kline, 2010). We are aware that serious questions have been raised regarding Cronbach's alpha's psychometric properties. However, while the PLSF-SEM method uses the Cronbach's alpha coefficient as a basis for the estimation of measurement error and composite weights, it makes no assumptions about the coefficient's main purported psychometric properties that have been the target of criticism (Sijtsma, 2009). This is an important distinction in light of measurement error theory (Nunnally & Bernstein, 1994).

Moreover, we developed and tested a number of experimental versions of the PLSF-SEM method prior to writing this note, using various reliability estimates. The versions employing the Cronbach's alpha coefficient tended to yield the best results. Sijtsma (2009) notes that the Cronbach's alpha coefficient may have limited uses, even in the face of its psychometric limitations. It is quite possible that we found one, although this is an issue that clearly merits further investigation.

PLSF-SEM's first stage then proceeds by iteratively estimating factors $\hat{F}_i$, matrices $\hat{\theta}_i$ with N rows and with each column storing one of the indicator error terms, column vectors of weights $\widehat{\omega}_i$, composites $\hat{C}_i$, and column vectors of loadings $\hat{\lambda}_i$ according to (5)-(9). Convergence is achieved when the sum of the absolute differences between successive estimates of the matrix of loadings for the entire model $\hat{\lambda}$ changes by less than a small fraction.

$$\hat{F}_i \coloneqq Stdz\left(\hat{C}_i \widehat{\omega}_{iC} + \hat{\varepsilon}_i \widehat{\omega}_{i\varepsilon}\right). \tag{5}$$

$$\hat{\theta}_i \coloneqq x_i - \hat{F}_i \hat{\lambda}_i{}'. \tag{6}$$

$$\widehat{\omega}_i \coloneqq \Sigma_{x_i x_i}{}^{-1}\left[\Sigma_{x_i x_i} - diag\left(\Sigma_{x_i \hat{\theta}_i}\right)\right]\hat{\lambda}_i{}'^{+}. \tag{7}$$

$$\hat{C}_i \coloneqq \frac{1}{\widehat{\omega}_{iC}}(x_i \widehat{\omega}_i). \tag{8}$$

$$\hat{\lambda}_i \coloneqq \left(\hat{C}_i{}^{+} x_i\right)' \widehat{\omega}_{iC}. \tag{9}$$

In these equations $diag\left(\Sigma_{x_i\widehat{\theta}_i}\right)$ is the diagonal matrix of *covariances* among the indicators and corresponding error terms, and the superscript $+$ denotes the Moore–Penrose pseudoinverse transformation. It is useful to observe that $diag\left(\Sigma_{x_i\widehat{\theta}_i}\right)$ is a diagonal matrix because in the common factor model $\Sigma_{x_{ij}\theta_{ij}} = 0$ for all $i \neq j$. That is, in the common factor model indicator error terms are correlated with their corresponding indicators and uncorrelated with other indicators in the same factor.

Researchers familiar with the mathematics underlying PLS will see that the estimation steps above differ significantly from those employed in Wold's original PLS design (Adelman & Lohmoller, 1994; Lohmöller, 1989; Wold, 1980). Particularly noteworthy is that the estimation steps above incorporate significantly more information in defining the relationships among weights and loadings, chiefly information about the relationships among indicators and their error terms.

In Wold's original PLS design and its variants weights and loadings are typically assumed to be proportional to one another, and thus linearly related. At the population level, our simulations suggest that usually they are not (i.e., the relationship between any factor's weights and loadings is usually nonlinear). As a result, the original PLS design does not yield estimates of the true composites. It is our contention that there is one unique true composite associated with each factor, and that the PLSF-SEM method yields estimates of the true composites.

PLSF-SEM's second stage: Factors

PLSF-SEM's second stage starts with the estimation of the elements $\hat{\Sigma}_{F_iF_j}$ of a target correlation matrix $\hat{\Sigma}_{FF}$ via (10), which follows from the correlation attenuation notion of measurement error theory (Nunnally & Bernstein, 1994). In this equation $\Sigma_{\hat{C}_i\hat{C}_j}$ is the correlation between composites estimated in the first stage, corresponding to the pair of factors F_i and F_j. Here $\hat{\Sigma}_{F_iF_j}$ are the elements of the matrix of estimated correlations among factors $\hat{\Sigma}_{FF}$, which can be seen as a population matrix estimate.

$$\hat{\Sigma}_{F_iF_j} \coloneqq \frac{\Sigma_{\hat{C}_i\hat{C}_j}}{\sqrt{\hat{\alpha}_i\hat{\alpha}_j}}. \tag{10}$$

In this second stage the PLSF-SEM method will fit the matrix of correlations among estimated factors $\Sigma_{\hat{F}\hat{F}}$, which can be seen as a model-implied matrix estimate, to $\hat{\Sigma}_{FF}$. To that end, the method proceeds by initializing factors as indicated in (14) and iteratively performing the assignments in (11)-(15). Since factors and measurement errors are re-estimated in each iteration, so must the correlation matrix elements $\Sigma_{\hat{F}_i\hat{F}_j}$, $\Sigma_{\hat{F}_i\hat{C}_i}$ and $\Sigma_{\hat{F}_i\hat{\varepsilon}_i}$. These are the elements of the correlation matrices among factors, factors and composites, and factors and measurement errors, respectively.

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\hat{\varepsilon}_i + \left(\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}\right)\frac{\hat{\Sigma}_{F_iF_j}}{\widehat{\omega}_{i\varepsilon}}\left(\hat{C}_j\widehat{\omega}_{jC} + \hat{\varepsilon}_j\widehat{\omega}_{j\varepsilon}\right)\right]. \tag{11}$$

$$\hat{F}_i \coloneqq Stdz\left[\hat{F}_i + \left(\widehat{\omega}_{iC} - \Sigma_{\hat{F}_i\hat{C}_i}\right)\hat{C}_i\widehat{\omega}_{iC}\right]. \tag{12}$$

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\hat{\varepsilon}_i - \Sigma_{\hat{C}_i\hat{\varepsilon}_i}\hat{C}_i\widehat{\omega}_{iC} + \left(\widehat{\omega}_{i\varepsilon} - \Sigma_{\hat{F}_i\hat{\varepsilon}_i}\right)\hat{F}_i\widehat{\omega}_{i\varepsilon}\right]. \tag{13}$$

$$\hat{F}_i \coloneqq Stdz\left(\hat{C}_i\widehat{\omega}_{iC} + \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right). \tag{14}$$

$$\hat{\varepsilon}_i \coloneqq Stdz\left[\frac{1}{\widehat{\omega}_{i\varepsilon}}\left(\hat{F}_i - \hat{C}_i\widehat{\omega}_{iC}\right)\right]. \tag{15}$$

The assignments in (11)-(13) are called “variation sharing” assignments, and constitute a critical ingredient of the PLSF-SEM method. As a group they are akin to a “soft” version of the classic expectation-maximization algorithm (Dempster et al., 1977) used in maximum likelihood estimation, but with apparently faster convergence and nonparametric properties. Through these assignments factors and measurement errors obtain variation that they did not have at the end of PLSF-SEM’s first stage, but that is an integral part of the true measurement errors and factors. Ultimately all of this variation emanates from the true composites.

The assignments above are only carried out for the variables indexed by i where $\widehat{\omega}_{i\varepsilon} > 0$. That is, factors and measurement errors are only adjusted in those cases where measurement error is assumed to exist, which are also cases where $\hat{\alpha}_i < 1$. Convergence is achieved when the sum of the absolute differences $\hat{\Sigma}_{F_iF_j} - \Sigma_{\hat{F}_i\hat{F}_j}$ falls below a small fraction, or when the sum of the absolute differences between successive estimates of $\Sigma_{\hat{F}_i\hat{F}_j}$ changes by less than a small fraction. Once convergence is achieved, final estimates of the composites, weights and loadings are generated through (16)-(18).

$$\hat{C}_i \coloneqq Stdz\left[\frac{1}{\widehat{\omega}_{iC}}\left(\hat{F}_i - \hat{\varepsilon}_i\widehat{\omega}_{i\varepsilon}\right)\right]. \tag{16}$$

$$\widehat{\omega}_i \coloneqq x_i{}^{+}\hat{C}_i\widehat{\omega}_{iC}. \tag{17}$$

$$\hat{\lambda}_i \coloneqq x_i{}'\hat{F}_i{}'^{+}. \tag{18}$$

The PLSF-SEM method then lastly proceeds to estimate path coefficients through a standard path analysis (Mueller, 1996; Wright, 1934; 1960) using the factor estimates. Standard errors for path coefficients and any other model parameter can be estimated via resampling or stable *P* value calculation methods (Kock, 2013; 2014), as is usually done in the original PLS design. The standard errors can subsequently be used to obtain chance probability estimates for hypothesis testing (Kock, 2014c), for any model parameter.

Monte Carlo experiment

We conducted a Monte Carlo experiment (Paxton et al., 2001) based on the true population model depicted in Figure 2, whereby 300 samples were created for each of the following sample sizes: 50, 100, and 300. This Monte Carlo experiment was conducted as part of extensive internal tests of version 5.0 of WarpPLS.

Figure 2. True population model

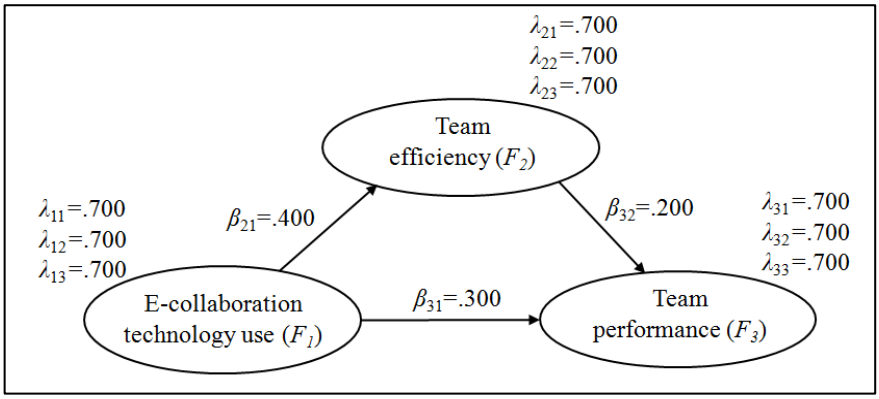


Table 1. Summarized Monte Carlo experiment results

SEM method	PLSA	PLSF	PLSA	PLSF	PLSA	PLSF
Sample size	50	50	100	100	300	300
EU>TE(TruePath)	.400	.400	.400	.400	.400	.400
EU>TE(AvgPath)	.339	.380	.309	.385	.303	.394
EU>TE(SEPath)	.125	.161	.128	.127	.110	.070
EU>TP(TruePath)	.300	.300	.300	.300	.300	.300
EU>TP(AvgPath)	.260	.301	.248	.294	.234	.297
EU>TP(SEPath)	.135	.157	.108	.133	.085	.079
TE>TP(TruePath)	.200	.200	.200	.200	.200	.200
TE>TP(AvgPath)	.201	.234	.189	.225	.174	.203
TE>TP(SEPath)	.144	.163	.098	.132	.061	.079
EU3<EU(TrueLoad)	.700	.700	.700	.700	.700	.700
EU3<EU(AvgLoad)	.793	.692	.802	.695	.808	.699
EU3<EU(SELoad)	.129	.108	.113	.077	.112	.049

Notes: XX>YY = link from factor XX to YY; EU = e-collaboration technology use; TE = team efficiency; TP = team performance; XX1 … XXn = indicators associated with factor XX; TruePath = true path coefficient; AvgPath = mean path coefficient estimate; SEPath = standard error of path coefficient estimate; TrueLoad = true loading; AvgLoad = mean loading estimate; SELoad = standard error of loading estimate.

A summarized set of results based on the analyses of simulated samples is shown in Table 1. True values, mean parameter estimates, and standard errors are shown next to one another. Results obtained through the PLSF-SEM method (under the "PLSF" columns) are contrasted with results obtained through the PLS Mode A algorithm (under the "PLSA" columns). PLS Mode A with the "path weighting" scheme was employed, the most widely used in analyses employing the original PLS design. We show results for all of the structural paths in the model, but restrict ourselves to loadings for one indicator in one factor since all loadings are the same in

the true population model used. This is also done to avoid repetition, as the same general pattern of results for loadings repeats itself for all indicators in all factors.

As we can see from the summarized results, the PLSF-SEM method yielded virtually unbiased estimates at $N = 300$, whereas PLS Mode A yielded significantly biased estimates at that same sample size. One of the reasons for these significantly biased estimates with PLS Mode A are the relatively low loadings in the true population model ($\lambda_{ij} = .7$, for all i and j), which tend to be a challenge for algorithms based on the original PLS design.

The relatively low loadings in the true population model apparently had little effect on PLSF-SEM's asymptotic convergence to the true values of the model parameters, although those loadings probably slowed down that convergence somewhat as sample sizes increased. In other simulations we conducted with higher loadings, convergence was achieved at smaller sample sizes. Generally speaking, high loadings are to be expected based on the common factor model, as they imply the use of psychometrically sound measurement instruments.

For several of the path coefficients and loadings the PLSF-SEM method yielded lower standard errors, particularly as sample sizes increased. This is noteworthy because the PLSF-SEM method is clearly more computationally complex than PLS Mode A, and thus could have been expected to have a greater "cost" in terms of standard errors.

However, standard errors yielded at $N = 50$ were generally higher for the PLSF-SEM method. Apparently the difference was enough to have a negative effect on power, as the ratios of path coefficients to standard errors indicate. That is, at $N = 50$ one could argue based on the results that PLS Mode A has greater power than the PLSF-SEM method for this particular model, although the ratios of path coefficients to standard errors suggest that both methods may struggle to avoid type II errors at this small sample size, particularly for the paths whose true coefficients were lower than $.400$ (the path with the highest strength).

Discussion and conclusion

While Wold's original PLS design offers several advantages over CB-SEM, it is largely incompatible with the common factor model (Kline, 2010; Mueller, 1996). Arguably the common factor model is the core foundation of CB-SEM. Given this, in Monte Carlo simulations where data is created based on common factor model assumptions, the original PLS design yields biased model parameters. Generally path coefficients are underestimated, and loadings are overestimated. This "advantages-with-costs" scenario has led to much debate over the years among proponents and detractors of the original PLS design.

In this brief research note we discussed what could be a solid step in the resolution of this debate. We showed how researchers can implement Factor-Based PLS-SEM (PLSF-SEM), a new method that generates estimates of both true composites and factors. The method does so in two stages, and fully accounts for measurement error. Since it generates estimates of both true composites and factors, the PLSF-SEM method can potentially place researchers in a position where they can subsequently estimate any model parameter imaginable.

At this point the reader may ask a reasonable question. Given that confirmatory factor analyses and hypothesis testing require primarily estimates of loadings and path coefficients, why would one want to generate factor estimates? The answer is that there are certain types of analyses that require factor scores, and more will likely be developed in the future as estimates of true factor scores become available to methodological researchers.

For example, the recently developed full collinearity test concurrently assesses lateral and vertical collinearity among factors (Kock & Lynn, 2012), providing the basis on which a number of methodological issues can be addressed (e.g., common method bias), but cannot be properly conducted without factor scores. Also, factor scores enable nonlinear analyses where best-fitting nonlinear functions are estimated for each pair of linked factors, and subsequently used to estimate path coefficients that take into account the nonlinearity (Guo et al., 2011; Kock, 2013; Moqbel et al., 2013).

It is our belief that the PLSF-SEM method is a solid step in the legitimization of modified versions of Wold's original "soft" PLS techniques for confirmatory factor and full-blown SEM analyses that are consistent with the common factor model. However, commonsense suggests that the PLSF-SEM method has weaknesses that will be uncovered as time goes by. It is very unlikely that any new method will be problem-free.

As the PLSF-SEM method is refined and improved, it may serve as the basis for the development of novel statistical tests that could lead to new insights in the context of SEM. Users of WarpPLS, starting in version 5.0, will be able to test the PLSF-SEM method and variations for themselves. Also, we hope that this brief note will provide enough details for implementations, in numerical programming environments such as R and GNU Octave, to be developed and tested under various conditions. We welcome comments, suggestions, and corrections.

Acknowledgments

The author is the developer of the software WarpPLS, which has over 7,000 users in more than 33 different countries at the time of this writing, and moderator of the PLS-SEM e-mail distribution list. He is grateful to those users, and to the members of the PLS-SEM e-mail distribution list, for questions, comments, and discussions on topics related to SEM and to the use of WarpPLS.

References

Adelman, I., & Lohmoller, J.-B. (1994). Institutions and development in the nineteenth century: A latent variable regression model. *Structural Change and Economic Dynamics*, 5(2), 329-359.

Cronbach, L.J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.

Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 (1), 1-38.

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.

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

Kock, N. (2005). What is e-collaboration. *International Journal of E-collaboration*, 1(1), 1-7.

Kock, N. (2008). E-collaboration and e-commerce in virtual worlds: The potential of Second Life and World of Warcraft. *International Journal of e-Collaboration*, 4(3), 1-13.

Kock, N. (2013). *WarpPLS 4.0 User Manual*. Laredo, TX: ScriptWarp Systems.

Kock, N. (2014). Using data labels to discover moderating effects in PLS-based structural equation modeling. *International Journal of e-Collaboration*, 10(4), 1-14.
Kock, N. (2014b). *One-tailed or two-tailed P values in PLS-SEM?* Laredo, TX: ScriptWarp Systems.
Kock, N. (2014c). 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., & 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.
Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares*. Heidelberg, Germany: Physica-Verlag.
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.
Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.
Paxton, P., Curran, P.J., Bollen, K.A., Kirby, J., & Chen, F. (2001). Monte Carlo experiments: Design and implementation. *Structural Equation Modeling*, 8(2), 287-312.
Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74(1), 107-120.
Wold, H. (1980). Model construction and evaluation when theoretical knowledge is scarce. In J. Kmenta and J. B. Ramsey (Eds.), *Evaluation of econometric models* (pp. 47-74). Waltham, MA: Academic Press.
Wright, S. (1934). The method of path coefficients. *The Annals of Mathematical Statistics*, 5(3), 161-215.
Wright, S. (1960). Path coefficients and path regressions: Alternative or complementary concepts? *Biometrics*, 16(2), 189-202.