Journal of Applied Structural Equation Modeling

eISSN: 2590-4221 Journal of Applied Structural Equation Modeling: 1(2),1-12, January 2018

SHOULD BOOTSTRAPPING BE USED IN PLS-SEM? TOWARD STABLE *P*-VALUE CALCULATION METHODS

Ned Kock

A.R. Sanchez, Jr. School of Business, Texas A&M International University, USA

nedkock@gmail.com

ABSTRACT

The use of the partial least squares (PLS) approach for structural equation modeling (SEM) has been experiencing explosive growth, particularly in the last few years. The calculation of p-values is extensively used for hypothesis testing in PLS-SEM. Such calculation typically relies on standard errors estimated via bootstrapping. This leads to unstable p-values and prohibitive computational demands when very large samples are analyzed. We discuss two calculation methods relying on exponential adjustment that generate stable standard errors and p-values, and that have minimal computational requirements. A Monte Carlo experiment shows that the methods yield estimates of the actual standard errors that are generally consistent with bootstrapping, and often more precise. The methods are implemented as part of the software WarpPLS, starting in version 5.0.

Keywords: *Partial Least Squares; Structural Equation Modeling; Standard Error; p-value; Path Bias; Monte Carlo Simulation.*

INTRODUCTION

The use of the partial least squares (PLS) method in the context of structural equation modeling (SEM) has been experiencing explosive growth among empirical researchers from a wide range of disciplines (Kock & Hadaya, 2018; Memon et al., 2017). This has happened in spite of PLS-SEM yielding approximations of model parameters that are asymptotically biased (Kock, 2015a). One reason for this is that biases decrease with the use of psychometrically sound measures (indicators). In PLS-SEM latent variables are approximated by composites, as opposed to factors, where the composites are exact linear combinations of indicators. The more indicators are available, and the higher their loadings on the latent variables, the less biased are the approximations of model coefficients yielded by PLS-SEM.

Estimation of path coefficients is an important element of empirical investigations employing PLS-SEM, since it provides the basis for hypothesis testing. Often each path coefficient will refer to a hypothesis, with each hypothesis being tested through the calculation of a *p*-value associated with the path coefficient. In the frequentist framework of statistical significance testing used in PLS-SEM, if a *p*-value is below a certain threshold then the corresponding hypothesis is assumed to be supported. The threshold is usually .05, used in conjunction with a one-tailed linear test of a directional hypothesis (Kock, 2015b).

Another approach employed for testing hypotheses associated with paths is to calculate a confidence interval for each path coefficient. If the value 0 (zero) does not fall within the interval, the hypothesis is supported. Otherwise the hypothesis is rejected. There has been much debate in the past over the possible benefits, if any, of using confidence intervals instead of *p*-values for hypothesis testing (Batterham & Hopkins, 2006; Newcombe, 1998; Poole, 1987; 2001). Recently Kock (2016) provided a detailed applied comparison of the two approaches in the context of PLS-SEM, showing that they lead to very similar results in tests of hypotheses.

In PLS-SEM the calculation of *p*-values for path coefficient estimates is normally conducted in three steps. First a standard error for the path coefficient estimate is calculated via resampling. By far the most widely used resampling technique in PLS-SEM is bootstrapping (Diaconis & Efron, 1983; Goodhue et al., 2012). The second step is to calculate a *t*-ratio, by dividing the path coefficient by the estimated standard error. The third and final step is the calculation of the *p*-value based on the *t*-ratio, which can be done with the incomplete beta function (Thompson et al., 1941). A common alternative for this third step is to use a table with pre-calculated *p*-value ranges associated with various *t*-ratio ranges.

As PLS-SEM use has grown, so has the availability of large datasets for analyses. The availability of large datasets is also often seen as having ushered in what is known as the "big data" era. Large datasets pose a serious challenge to resampling techniques such as bootstrapping. The reason is that resampling techniques create multiple replications of the original dataset. Frequently the number of replications is 100 or more. For a dataset with 100,000 cases, for example, the use of 100 replications would lead to the creation of 10 million cases of data for analysis. On standard personal computers this would typically lead to very long waiting periods, with a common outcome actually being a computer "crash" before any analysis results can be obtained.

Resampling techniques such as bootstrapping also present a different problem. They are inherently unstable. Let us illustrate this through an example, which will be further elaborated later through an empirical illustration. In an analysis of a sample of size 152 and employing bootstrapping with 500 replications, a path coefficient estimate of .217 yielded a standard error of .095 and a *p*-value of .012, while a path coefficient estimate of .194 yielded a standard error of .072 and a *p*-value of .004. Here the instability is reflected in a *higher p*-value for a *stronger* path coefficient (i.e., a path coefficient with greater absolute estimated value). One would expect a *lower p*-value to be yielded for a *stronger* path coefficient, because the chance probability of a path coefficient estimate decreases with its magnitude (Aczel & Sounderpandian, 2002; Kock, 2015b).

We discuss two *p*-value calculation methods that address the problems above. Both methods employ stable exponential adjustments through the direct application of formulas. As such, unlike resampling methods, they do not generate sample replications. Because of that, they can be used with large datasets. Moreover, neither method makes any data or model parameter distribution assumption. Finally, a Monte Carlo experiment shows that the methods yield estimates of the actual standard errors that are consistent with those obtained via bootstrapping, in many cases yielding more precise estimates of the actual standard errors.

The two *p*-value calculation methods are implemented starting in version 5.0 of WarpPLS. 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. Starting in version 5.0 of WarpPLS, factor-based SEM algorithms have also become available (Kock, 2015a). Factor-based SEM algorithms conduct analyses fully accounting for measurement error, thus yielding parameters that are very similar to those generated by covariance-based SEM via full information maximum likelihood (Kock, 2017).

ILLUSTRATIVE MODEL

The illustrative model depicted in Figure 1 has been used in our Monte Carlo experiment and empirical illustration, which are discussed later. The model contains five latent variables, for which composites were estimated via PLS-SEM. The latent variables are communication flow orientation (C_1), usefulness in the development of information technology (IT) solutions (C_2), ease of understanding (C_3), accuracy (C_4), and impact on redesign success (C_5).

In this illustrative model β_{ij} is the path coefficient for the link going from composite C_j to composite C_i, λ_{ij} is the loading for the *j*th indicator of composite C_i, and ζ_i is the structural error associated with an endogenous composite C_i. With one exception, a set of indicators x_{ij} is used to measure each composite C_i, where each indicator is assumed to measure the composite with a certain degree of imprecision. The exception is the case in which only one indicator is used, communication flow orientation (C_1), where no imprecision is assumed.

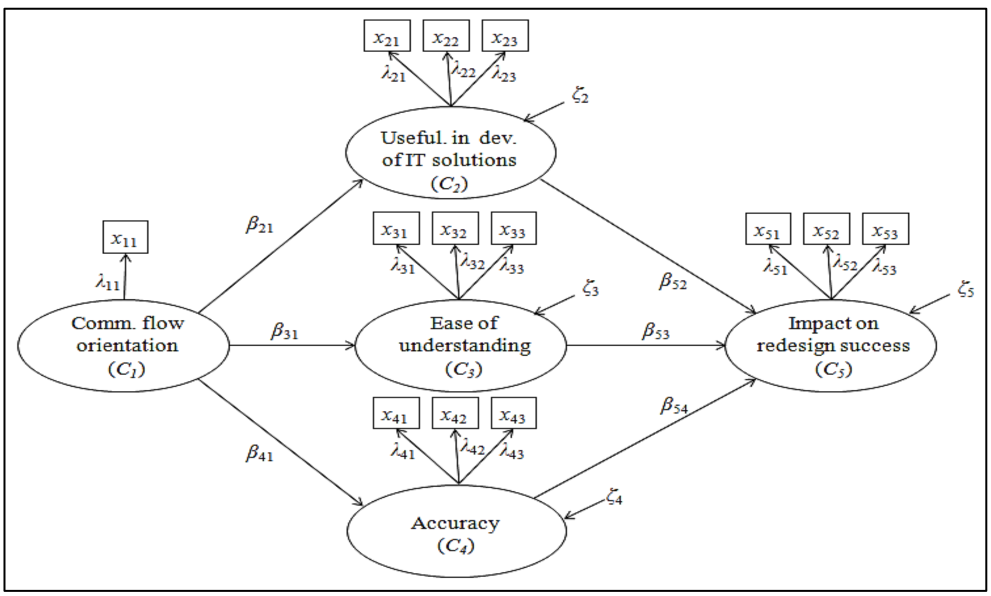


Figure 1. Illustrative Model

The model is based on communication flow optimization theory (Danesh-Pajou, 2005; Kock, 2003). Even though the theory is not the focus of our investigation, readers may be interested in knowing that the theory incorporates the expectation that business process redesign teams that place more emphasis on the improvement of the flow of communication in business processes would tend to achieve more successful results (in terms of business process redesign quality) than teams placing less emphasis on that improvement.

Communication flow orientation (C_1) is the degree to which a business process modeling approach explicitly shows how communication interactions take place in a business process. Usefulness in the development of IT solutions (C_2) is the degree to which a process modeling approach is useful in the development of a generic IT solution to automate the redesigned process. Ease of understanding (C_3) is the degree to which a process modeling approach is perceived to yield a process representation that is easy to understand. Accuracy (C_4) is the degree to which a process modeling approach is perceived to lead to an accurate representation of the process. Finally, impact on redesign success (C_5) is the degree to which the process modeling technique used is perceived to lead to an actual improvement of the targeted business process.

STANDARD ERRORS AND *P*-VALUES

The calculation of *p*-values is commonplace in PLS-SEM as a basis for hypothesis testing. First estimates of path coefficients and standard errors must be produced, which in PLS-SEM are used in the calculation of *t*-ratios, by dividing the estimated path coefficients by the estimated standard errors. The *t*-ratios are then used in the estimation of *p*-values. In this section we briefly describe the three standard error estimation methods we compared.

Method 1: Bootstrapping.

Bootstrapping (Diaconis & Efron, 1983) is the most widely used method for standard error estimation in PLS-SEM. Let $\mathcal{S}$ be a set of samples created based on an empirical dataset, where each sample in $\mathcal{S}$ is built by taking rows from the original dataset. Each row is taken at random and with replacement (i.e., the same row can be repeated), and each sample in $\mathcal{S}$ has the same size (i.e., number of rows) as the original dataset. Let N_S be the number of such samples. The standard error denoted as S_1, obtained via bootstrapping for a given path coefficient β, is calculated according to (1), where: β_i is the path coefficient estimate for sample i, and $\bar{\beta}$ is the mean path coefficient across all samples.

$$S_1 = \sqrt{\frac{1}{N_S}\sum_{i=1}^{N_S}\left(\beta_i - \bar{\beta}\right)^2}. \tag{1}$$

The bootstrapping approach to estimation of standard errors can be seen as a type of Monte Carlo simulation approach (Robert & Casella, 2005). Many samples are created based on the original sample, mimicking the sample creation process normally seen in Monte Carlo simulations employed in the context of comparative assessments of different SEM approaches and techniques (see, e.g., Goodhue et al., 2012). The key difference is that typically in Monte Carlo simulations employed in the context of SEM technique assessments, the samples are created based on a true population model.

Methods 2 and 3: Stable Exponential Adjustments

It follows from the central limit theorem (Aczel & Sounderpandian, 2002; Miller & Wichern, 1977) that the standard error of a path coefficient, which is (i.e. the path coefficient) a standardized quantity, should be proportional to the inverse square root of the sample size. We also know that the standard error is a function of the magnitude of the path coefficient (Petraitis et al., 1996). Finally, the standard error should converge to zero as the sample size grows to infinity. These criteria call for standard error estimation functions that apply exponential adjustments to the inverse square roots of the sample sizes (Kock & Hadaya, 2018). The functions should also be sensitive to path coefficients' magnitudes; magnitudes that can in turn be measured by the path coefficients' absolute values.

Our attempts to fit functions that meet the criteria above, to both simulated and empirical data, led us to equations (2) and (3). They are meant to be used for the estimation of the standard error of a path coefficient β in a PLS-SEM model, for a given empirical sample. In these equations, N is the sample size, and the symbol e denotes Euler's number (given by $\sum_{n=0}^{\infty} 1/n! \cong 2.71828$). As it will be seen shortly in our Monte Carlo experiment, the first equation yields a lower bound estimate of the standard error (S_2) associated with β, and the second an upper bound estimate (S_3).

$$S_2 = \frac{1}{\sqrt{N}} e^{-\left(\frac{e^{|\beta|}}{\sqrt{N}}\right)}. \quad (2)$$

$$S_3 = \frac{1}{\sqrt{N}} e^{-\left(\frac{e|\beta|}{\sqrt{N}}\right)}. \quad (3)$$

Function fitting exercises often lead to fractional constants. The use of Euler's number (e) in the equations above, as opposed to fractional constants that are close to e in value, reflects our attempt to achieve a certain measure of mathematical elegance in their formulation. While this may appear to be a rather subjective criterion, its importance has often been highlighted by pure and applied mathematicians (Osborne, 1984; Poincaré, 2000).

As it will also be seen shortly in our presentation of the results of a Monte Carlo experiment, the standard error estimates yielded by the equations above seem to be consistent with the estimates obtained via bootstrapping, and in many cases provide better approximations of the actual standard errors. Better approximations occurred more often with the upper bound estimate (S_3) than with the lower bound estimate (S_2).

MONTE CARLO EXPERIMENT

A Monte Carlo experiment based on the true population model show in Figure 2 was conducted to assess the performance of the three standard error estimation methods discussed in the previous section. Performance was assessed in terms of statistical power and closeness to the actual standard errors obtained through the analyses of simulated samples. This Monte Carlo experiment was conducted as part of extensive internal tests of version 5.0 of WarpPLS.

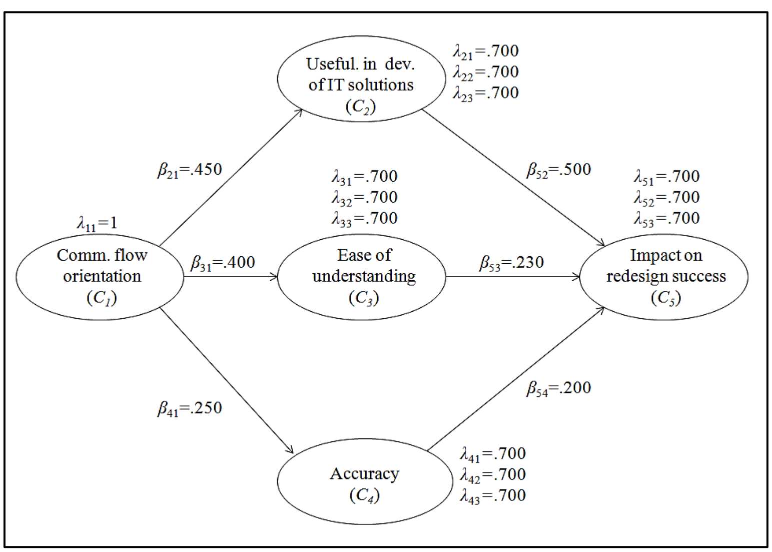


Figure 2. True population Model

We created an analyzed 1,000 samples for each of the following sample sizes: 50, 100, 200, 300, and 500. The PLS Mode A algorithm with the path weighting scheme (Lohmöller, 1989) was used in the analyses. These are the most widely used algorithm (PLS Mode A) and inner model estimation scheme (path weighting) in the context of PLS-SEM.

Table 1. Summarized Monte Carlo Experiment Results

Method	BOOT	STBL2	STBL3	BOOT	STBL2	STBL3
Sample size	50	50	50	300	300	300
CO>GT(TruePath)	.450	.450	.450	.450	.450	.450
CO>GT(AvgPath)	.383	.383	.383	.388	.388	.388
CO>GT(Power)	.905	.954	.946	1	1	1
CO>GT(SEPath)	.125	.125	.125	.076	.076	.076
CO>GT(EstSEPath)	.120	.115	.122	.047	.053	.054
CO>EU(TruePath)	.400	.400	.400	.400	.400	.400
CO>EU(AvgPath)	.347	.347	.347	.347	.347	.347
CO>EU(Power)	.781	.900	.867	1	1	1
CO>EU(SEPath)	.131	.131	.131	.072	.072	.072
CO>EU(EstSEPath)	.133	.116	.124	.049	.053	.055
CO>AC(TruePath)	.250	.250	.250	.250	.250	.250
CO>AC(AvgPath)	.224	.224	.224	.218	.218	.218
CO>AC(Power)	.419	.611	.559	.985	.995	.994
CO>AC(SEPath)	.141	.141	.141	.061	.061	.061
CO>AC(EstSEPath)	.166	.118	.129	.054	.054	.056
GT>SU(TruePath)	.500	.500	.500	.500	.500	.500
GT>SU(AvgPath)	.333	.333	.333	.347	.347	.347
GT>SU(Power)	.711	.863	.823	1	1	1
GT>SU(SEPath)	.206	.206	.206	.160	.160	.160
GT>SU(EstSEPath)	.146	.116	.125	.052	.053	.055
EU>SU(TruePath)	.230	.230	.230	.230	.230	.230
EU>SU(AvgPath)	.175	.175	.175	.163	.163	.163
EU>SU(Power)	.254	.410	.356	.917	.921	.906
EU>SU(SEPath)	.131	.131	.131	.085	.085	.085
EU>SU(EstSEPath)	.157	.119	.132	.054	.054	.056
AC>SU(TruePath)	.200	.200	.200	.200	.200	.200
AC>SU(AvgPath)	.159	.159	.159	.147	.147	.147
AC>SU(Power)	.240	.405	.335	.866	.868	.849
AC>SU(SEPath)	.137	.137	.137	.073	.073	.073
AC>SU(EstSEPath)	.165	.119	.132	.053	.054	.056

Notes: BOOT = Method 1: Bootstrapping; STBL2 = Method 2: Stable exponential adjustment (with $e^{|\beta|}$ in equation); STBL3 = Method 3: Stable exponential adjustment (with $e|\beta|$ in equation); XX>YY = link from composite XX to YY; CO = communication flow orientation (C_1); GT = usefulness in the development of IT solutions (C_2); EU = ease of understanding (C_3); AC = accuracy (C_4); SU = impact on redesign success (C_5); TruePath = true path coefficient; AvgPath = mean path coefficient estimate; Power = statistical power; SEPath = standard error of path coefficient estimate; EstSEPath = method-specific standard error of path coefficient estimate.

A summarized set of results is shown in Table 1, where we restrict ourselves to sample sizes 50 and 300. True path coefficients, mean path coefficient estimates, statistical power values, standard errors of path coefficient estimates, and method-specific standard errors of path coefficient estimates are shown next to one another. Full results, for all sample sizes included in the simulation, are available in Appendix A.

Path coefficient estimates are used in the estimation of standard errors based on the stable methods (STBL2 and STBL3); i.e., the method-specific standard errors. As we can see, the mean path coefficient estimates differ from the true path coefficients across different sample sizes, and generally underestimate the true path coefficients. This underestimation stems from the use of composites in PLS-SEM, which in turn leads to a composite *correlation* attenuation (Nunnally & Bernstein, 1994) that "propagates" to the path coefficients (Kock, 2015a).

Generally the method-specific standard error of path coefficient estimates obtained via STBL3 were the closest to the actual (or true) standard errors of path coefficient estimates. This suggests that standard errors estimated via STBL3 are not only stable when compared with those estimated via BOOT, but also more accurate. Moreover, both STBL2 and STBL3 led to greater statistical power than BOOT at small sample sizes. Power tends to be compromised the most with small sample sizes, and to invariably increase as sample sizes go up regardless of the standard error and *p*-value calculation method used.

Interestingly, our results suggest that the actual standard errors do not depend on the number of competing structural paths pointing at a composite. The results imply that the actual standard errors depend primarily on the sample size and magnitude of the estimated path coefficients. This goes counter to the suggestion by Goodhue et al. (2012) to use Cohen's (1988; 1992) power tables to estimate minimum sample sizes in PLS-SEM, since in those tables minimum required sample sizes increase with number of predictors. A more advisable strategy would be to use the equations provided here, and by Kock & Hadaya (2018), to estimate minimum sample sizes.

EMPIRICAL ILLUSTRATION

Table 2 summarizes the results of an empirical study. The study served as the basis for the development of the illustrative and true population models discussed earlier. Shown next to one another are estimated path coefficients, method-specific standard errors of path coefficient estimates, and *p*-values.

The data for this empirical study was collected from 156 individuals involved in business process redesign projects in Northeastern U.S.A. The participants employed one of two business process modeling approaches; one focused on the communication flow within business processes, and the other focused on the chronological flow of activities. These are depicted in Appendix B. The questionnaire used for data collection is provided in Appendix C.

Overall, all three *p*-value calculation methods yielded estimates consistent with communication flow optimization theory (Kock, 2003). The theory forms the underlying theoretical foundation for the model, and has been validated before through other empirical studies using different datasets (Danesh-Pajou, 2005; Danesh-Pajou & Kock, 2005; Kock et al., 2008). Given this, the empirical study results summarized above lend further "real data" validation of the two new *p*-value calculation methods.

Table 2. Empirical Study Results

Method	BOOT	STBL2	STBL3
CO>GT(EstPath)	.485	.485	.485
CO>GT(EstSEPath)	.057	.070	.072
CO>GT(*p*)	<.001	<.001	<.001
CO>EU(EstPath)	.362	.362	.362
CO>EU(EstSEPath)	.061	.071	.074
CO>EU(*p*)	<.001	<.001	<.001
CO>AC(EstPath)	.269	.269	.269
CO>AC(EstSEPath)	.075	.072	.076
CO>AC(*p*)	<.001	<.001	<.001
GT>SU(EstPath)	.506	.506	.506
GT>SU(EstSEPath)	.079	.070	.072
GT>SU(*p*)	<.001	<.001	<.001

EU>SU(EstPath)	.217	.217	.217
EU>SU(EstSEPath)	.095	.072	.076
EU>SU(*p*)	.012	.002	.003
AC>SU(EstPath)	.194	.194	.194
AC>SU(EstSEPath)	.072	.073	.077
AC>SU(*p*)	.004	.004	.006

Notes: *N* = 156; BOOT = Method 1: Bootstrapping; number of bootstrapping samples (replications) used = 500; STBL2 = Method 2: Stable exponential adjustment ($e^{|\beta|}$); STBL3 = Method 3: Stable exponential adjustment ($e|\beta|$); XX>YY = link from composite XX to composite YY; CO = communication flow orientation (C_1); GT = usefulness in the development of IT solutions (C_2); EU = ease of understanding (C_3); AC = accuracy (C_4); SU = impact on redesign success (C_5); EstPath = estimated path coefficient; EstSEPath = method-specific standard error of path coefficient estimate; *p* = *p*-value.

DISCUSSION AND CONCLUSION

The use of PLS-SEM has been experiencing explosive growth. Estimation of path coefficients plays a key role in empirical investigations employing PLS-SEM. Frequently each path coefficient will refer to a hypothesis, and each hypothesis will be tested through the calculation of a *p*-value associated with the path coefficient. The calculation of *p*-values is normally conducted in three steps: (1) a standard error for the path coefficient estimate is obtained via bootstrapping; (2) a *t*-ratio is obtained by dividing the path coefficient estimate by the standard error; and (3) the *p*-value is calculated through the incomplete beta function, or obtained from a table, based on the *t*-ratio.

Large datasets pose a serious challenge to bootstrapping in particular, and resampling techniques in general, because these techniques create multiple replications of the original dataset. Moreover, resampling techniques such as bootstrapping are inherently unstable; often leading to higher *p*-values for stronger path coefficients, and lower *p*-values for weaker path coefficients. This type of instability defies commonsense and is puzzling to empirical researchers, as one would expect *p*-values to reflect chance probabilities of path coefficients, and thus to go down as path coefficients become stronger.

We have discussed two *p*-value calculation methods that seem to successfully address both of the problems above. Both methods rely on stable exponential adjustments, which are obtained through the direct application of formulas. Neither method generates replications of the original dataset. Moreover, neither method makes data or model parameter distribution assumptions. As we have demonstrated through a Monte Carlo experiment, the methods generally yield compatible and often more precise estimates of the actual standard errors than bootstrapping. Additionally, we have shown that the methods yield estimates consistent with a theoretical framework that has been previously validated through various empirical studies.

Users of WarpPLS, starting in version 5.0, are able to test the methods for themselves. Also, we hope that this research 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

Aczel, A.D., & Sounderpandian, J. (2002). *Complete business statistics.* Boston, MA: McGraw-Hill/Irwin.

Batterham, A. M., & Hopkins, W. G. (2006). Making meaningful inferences about magnitudes. *International Journal of Sports Physiology and Performance*, 1(1), 50-57.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155-159.

Danesh-Pajou, A. (2005). *IT-enabled process redesign: Using communication flow optimization theory in an information intensive environment.* Doctoral Dissertation. Philadelphia, PA: Temple University.

Danesh-Pajou, A., & Kock, N. (2005). An experimental study of process representation approaches and their impact on perceived modeling quality and redesign success. *Business Process Management Journal*, 11(6), 724-735.

Diaconis, P., & Efron, B. (1983). Computer-intensive methods in statistics. *Scientific American*, 249(1), 116-130.

Goodhue, D.L., Lewis, W., and Thompson, R. (2012). Does PLS have advantages for small sample size or non-normal data? *MIS Quarterly*, 36(3), 981-1001.

Kock, N. (2003). Communication-focused business process redesign: Assessing a communication flow optimization model through an action research study at a defense contractor. *IEEE Transactions on Professional Communication*, 46(1), 35-54.

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

Kock, N. (2015b). One-tailed or two-tailed P values in PLS-SEM? *International Journal of e-Collaboration*, 11(2), 1-7.

Kock, N. (2016). Hypothesis testing with confidence intervals and P values in PLS-SEM. *International Journal of e-Collaboration*, 12(3), 1-6.

Kock, N. (2017). Structural equation modeling with factors and composites: A comparison of four methods. *International Journal of e-Collaboration*, 13(1), 1-9.

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., Danesh-Pajou, A., & Komiak, P. (2008). A discussion and test of a communication flow optimization approach for business process redesign. *Knowledge and Process Management*, 15(1), 72-85.

Lohmöller, J.-B. (1989). *Latent variable path modeling with partial least squares.* Heidelberg, Germany: Physica-Verlag.

Memon, M.A., Ting, H., Ramayah, T., Chuah, F., & Cheah, J.-H. (2017). A review of the methodological misconceptions and guidelines related to the application of structural equation modeling: A Malaysian scenario. *Journal of Applied Structural Equation Modeling*, 1(1), i-xiii.

Miller, R.B., & Wichern, D.W. (1977). *Intermediate business statistics: Analysis of variance, regression and time series.* New York, NY: Holt, Rihehart and Winston.

Newcombe, R. G. (1998). Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Statistics in Medicine*, 17(8), 857-872.

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory.* New York, NY: McGraw-Hill.

Osborne, H. (1984). Mathematical beauty and physical science. *The British Journal of Aesthetics*, 24(4), 291-300.

Petraitis, P.S., Dunham, A.E., & Niewiarowski, P.H. (1996). Inferring multiple causality: the limitations of path analysis. *Functional Ecology*, 10(4), 421-431.

Poincaré, H. (2000). Mathematical creation. *Resonance*, 5(2), 85-94.

Poole, C. (1987). Beyond the confidence interval. *American Journal of Public Health*, 77(2), 195-199.

Poole, C. (2001). Low P-values or narrow confidence intervals: Which are more durable? *Epidemiology*, 12(3), 291-294.

Robert, C.P., & Casella, G. (2005). *Monte Carlo statistical methods.* New York, NY: Springer.

Thompson, C.M., Pearson, E.S., Comrie, L.J., & Hartley, H.O. (1941). Tables of percentage points of the incomplete beta-function. *Biometrika*, 32(2), 151-181.

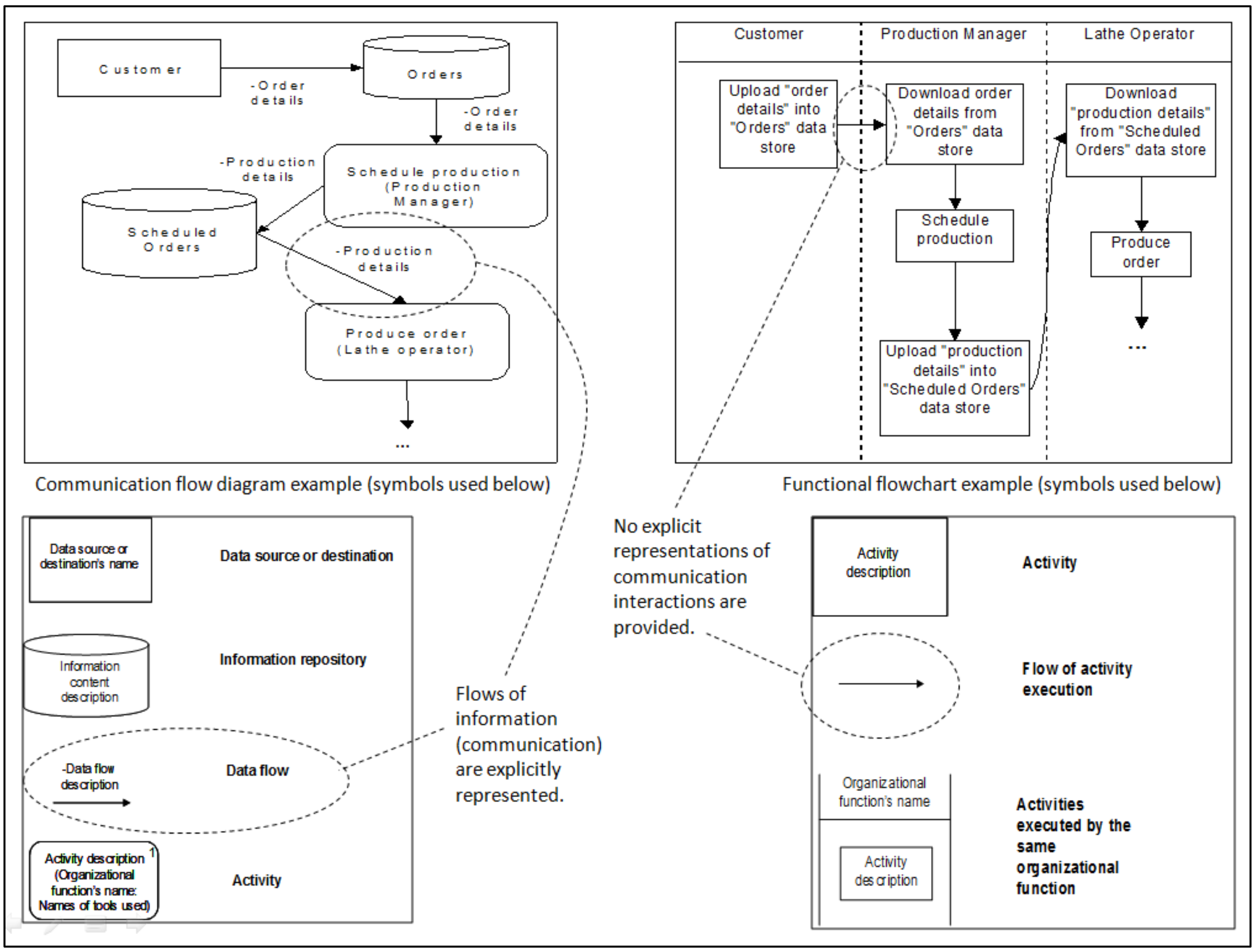
APPENDIX A: FULL MONTE CARLO EXPERIMENT RESULTS

The full Monte Carlo experiment results are provided in the table below. Notes: BOOT = Method 1 (bootstrapping); STBL2 = Method 2 (stable exponential adjustment with $e^{|\beta|}$ in equation); STBL3 = Method 3 (stable exponential adjustment with $e|\beta|$ in equation); XX>YY = link from composite XX to YY; CO = communication flow orientation (C_1); GT = usefulness in the development of IT solutions (C_2); EU = ease of understanding (C_3); AC = accuracy (C_4); SU = impact on redesign success (C_5); TruePath = true path coefficient; AvgPath = mean path coefficient estimate; Power = statistical power; SEPath = standard error of path coefficient estimate; EstSEPath = method-specific standard error of path estimate.

Method	**BOOT**	**STBL2**	**STBL3**	**BOOT**	**STBL2**	**STBL3**	**BOOT**	**STBL2**	**STBL3**	**BOOT**	**STBL2**	**STBL3**	**BOOT**	**STBL2**	**STBL3**
Sample size	**50**	**50**	**50**	**100**	**100**	**100**	**200**	**200**	**200**	**300**	**300**	**300**	**500**	**500**	**500**
CO>GT(TruePath)	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450	.450
CO>GT(AvgPath)	.383	.383	.383	.387	.387	.387	.389	.389	.389	.388	.388	.388	.388	.388	.388
CO>GT(Power)	.905	.954	.946	.999	.999	.999	1	1	1	1	1	1	1	1	1
CO>GT(SEPath)	.125	.125	.125	.101	.101	.101	.082	.082	.082	.076	.076	.076	.071	.071	.071
CO>GT(EstSEPath)	.120	.115	.122	.081	.086	.090	.058	.064	.066	.047	.053	.054	.037	.042	.043
CO>EU(TruePath)	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400	.400
CO>EU(AvgPath)	.347	.347	.347	.354	.354	.354	.347	.347	.347	.347	.347	.347	.346	.346	.346
CO>EU(Power)	.781	.900	.867	.978	.992	.988	1	1	1	1	1	1	1	1	1
CO>EU(SEPath)	.131	.131	.131	.091	.091	.091	.078	.078	.078	.072	.072	.072	.065	.065	.065
CO>EU(EstSEPath)	.133	.116	.124	.084	.087	.091	.060	.064	.066	.049	.053	.055	.038	.042	.043
CO>AC(TruePath)	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250	.250
CO>AC(AvgPath)	.224	.224	.224	.223	.223	.223	.219	.219	.219	.218	.218	.218	.215	.215	.215
CO>AC(Power)	.419	.611	.559	.687	.806	.777	.908	.944	.936	.985	.995	.994	.997	.998	.998
CO>AC(SEPath)	.141	.141	.141	.098	.098	.098	.073	.073	.073	.061	.061	.061	.055	.055	.055
CO>AC(EstSEPath)	.166	.118	.129	.103	.088	.094	.068	.065	.068	.054	.054	.056	.042	.042	.044
GT>SU(TruePath)	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500	.500
GT>SU(AvgPath)	.333	.333	.333	.343	.343	.343	.343	.343	.343	.347	.347	.347	.350	.350	.350
GT>SU(Power)	.711	.863	.823	.972	.985	.979	1	1	1	1	1	1	1	1	1
GT>SU(SEPath)	.206	.206	.206	.177	.177	.177	.168	.168	.168	.160	.160	.160	.154	.154	.154
GT>SU(EstSEPath)	.146	.116	.125	.092	.087	.091	.064	.064	.066	.052	.053	.055	.040	.042	.043
EU>SU(TruePath)	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230	.230
EU>SU(AvgPath)	.175	.175	.175	.165	.165	.165	.163	.163	.163	.163	.163	.163	.164	.164	.164
EU>SU(Power)	.254	.410	.356	.503	.564	.524	.794	.810	.784	.917	.921	.906	.989	.992	.987
EU>SU(SEPath)	.131	.131	.131	.106	.106	.106	.091	.091	.091	.085	.085	.085	.078	.078	.078
EU>SU(EstSEPath)	.157	.119	.132	.098	.089	.096	.067	.065	.069	.054	.054	.056	.041	.042	.044
AC>SU(TruePath)	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200	.200
AC>SU(AvgPath)	.159	.159	.159	.159	.159	.159	.150	.150	.150	.147	.147	.147	.148	.148	.148
AC>SU(Power)	.240	.405	.335	.482	.565	.503	.702	.751	.703	.866	.868	.849	.979	.971	.965
AC>SU(SEPath)	.137	.137	.137	.095	.095	.095	.079	.079	.079	.073	.073	.073	.065	.065	.065
AC>SU(EstSEPath)	.165	.119	.132	.101	.089	.096	.066	.065	.069	.053	.054	.056	.040	.042	.044

APPENDIX B: BUSINESS PROCESS MODELING APPROACHES USED

The two types of business process representations used in the empirical study, referring to high a low communication flow orientations, are exemplified in the figure below. In our data analyses, the representation on the left was coded as 1, as it refers to a high communication flow orientation; and the one on the right as 0, as it refers to a low communication flow orientation.



APPENDIX C: QUESTIONNAIRE USED IN EMPIRICAL STUDY

The latent constructs associated with the composites in the illustrative study were measured through the question-statements below. With one exception, all question-statements were answered on 7-point Likert-type scales. The exception was communication flow orientation (C_1), coded as either 1 or 0.

Communication flow orientation (C_1)

- C_{11}: Coded as either 1 or 0, corresponding to high or low communication flow orientation of the business process modeling approach used.

Usefulness in the development of IT solutions (C_2)

- C_{21}: This process modeling approach is useful in the development of a generic IT solution to automate the redesigned process.
- C_{22}: Creating a generic IT solution to enable the redesigned process is easy based on this process modeling approach.
- C_{23}: Graphical process representations using this approach facilitate the generation of a generic IT solution to automate the redesigned process.

Ease of understanding (C_3)

- C_{31}: Processes modeled using this approach are easy to understand.
- C_{32}: Graphical representations of processes using this approach are clear.
- C_{33}: This process modeling approach leads to graphical models that are easy to understand.

Accuracy (C_4)

- C_{41}: This process modeling approach leads to accurate process representations.
- C_{42}: Models created using this approach are correct representations of a process.
- C_{43}: Graphical representations using this approach clearly reflect the real process.

Impact on redesign success (C_5)

- C_{51}: Using this process modeling approach is likely to contribute to the success of a process redesign project.
- C_{52}: Success chances are improved if this process modeling approach is used.
- C_{53}: Using the graphical process representations in this approach is likely to make process redesign projects more successful.